EPIGENETIC MECHANISMS AND THEIR INVOLVEMENT IN RARE DISEASES

EDITED BY: Mojgan Rastegar and Dag H. Yasui
PUBLISHED IN: Frontiers in Genetics and
Frontiers in Cell and Developmental Biology

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-514-5
DOI 10.3389/978-2-88971-514-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

EPIGENETIC MECHANISMS AND THEIR INVOLVEMENT IN RARE DISEASES

Topic Editors:
Mojgan Rastegar, University of Manitoba, Canada
Dag H. Yasui, University of California, United States

Citation: Rastegar, M., Yasui, D. H., eds. (2021). Epigenetic Mechanisms and Their Involvement in Rare Diseases. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-514-5

Table of Contents

frontiers in Genetics

EDITORIAL
published: 01 September 2021
doi: 10.3389/fgene.2021.755076

Editorial: Epigenetic Mechanisms and Their Involvement in Rare Diseases

Mojgan Rastegar[1]* *and Dag H. Yasui*[2]*

[1] Department of Biochemistry and Medical Genetics, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada, [2] Department of Medical Microbiology and Immunology, University of California Davis School of Medicine, Davis, CA, United States

Keywords: epigenetics and rare diseases, MeCP2 isoforms and rett syndrome (RTT), DNA methylation and histone modifications, ATRX and gene regulatory mechanisms, activity dependent neuroprotective protein (ADNP) and chromatin remodeling, Beckwith-Wiedemann Syndrome (BWS) and Prader-Willi Syndrome (PWS), O-linked-D-N-acetylglucosamine (O-GlcNAc), MYCN-related epigenetic factors and non-coding regulatory RNAs

Editorial on the Research Topic

Epigenetic Mechanisms and Their Involvement in Rare Diseases

Epigenetic mechanisms are diverse modes of gene regulation, acting independent of genetic sequences. Epigenetics involves an array of "readers," "writers," and "erasers," with key roles in development, health, and disease. One important aspect of epigenetics is involvement in rare diseases. This special topic covers a series of original research and review articles that further our knowledge about epigenetic mechanisms in rare diseases.

One well-studied example of rare diseases caused by genetic mutations is Rett Syndrome (RTT). RTT is due to *de novo* mutations in the X-linked Methyl CpG Binding Protein 2 (*MECP2*) gene. The multi-functional "MeCP2" protein plays important roles in neuronal maturation and brain development.

Focusing on RTT, Sharifi and Yasui, provide an overview about MeCP2 protein biology and its functional relevance to RTT. The authors explain how *MECP2* mutations contribute to disease mechanisms, describing lessons learnt from RTT mice and model systems. They describe how *MECP2* expression is distributed among different organs, using helpful schematics. They further discuss MeCP2 DNA binding activities, and its association not only with CpG dinucleotide methylation, but also with CpH methylation in the context of CpA, CpC, or CpT. The authors elaborate on MeCP2 function as a dual transcriptional regulator, as an activator or effective suppressor of gene transcription. The authors also discuss MeCP2 splice variants; MeCP2E1 and MeCP2E2, MeCP2 role in liquid phase separation, and potential therapeutic strategies for RTT.

Complementing the first paper, Good et al., offer a timely review entitled "MeCP2: the genetic driver of Rett Syndrome epigenetics." The authors discuss how RTT-associated *MECP2* gene mutations can modify its DNA binding activities, and chromatin bundling capabilities, while altering MeCP2 protein stability. They further discuss the role of MeCP2 in alternative splicing and micro-RNA processing. Other aspects of MeCP2 function, diverse protein domains, and different mutations are also well-discussed. Interestingly, the authors explain the impact of proteasomal degradation through MeCP2 PEST sequences, and circadian-dependent dynamics of MeCP2 isoforms. Finally, the authors highlight the complexity of RTT pathology with differential relevance of MeCP2 isoforms.

The third paper on RTT is an original research article by Pejhan et al. The authors studied MeCP2 homeostasis regulatory network in the frontal cerebrum, hippocampus, amygdala, and cerebellum of post-mortem brain tissues from RTT patients and non-RTT controls. The authors

OPEN ACCESS

Edited and reviewed by:
Hehuang Xie,
Virginia Tech, United States

***Correspondence:**
Mojgan Rastegar
mojgan.rastegar@umanitoba.ca
Dag H. Yasui
dhyasui@ucdavis.edu

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *07 August 2021*
Accepted: *11 August 2021*
Published: *01 September 2021*

Citation:
Rastegar M and Yasui DH (2021) Editorial: Epigenetic Mechanisms and Their Involvement in Rare Diseases. Front. Genet. 12:755076. doi: 10.3389/fgene.2021.755076

establish that in humans, the MeCP2 homeostasis network (MeCP2E1/E2-BDNF-*miR132*) is brain-region specific. Their correlational studies suggested that cerebellum is the one tested brain region that may involve all these regulatory components. Their findings highlighted that MeCP2 isoform-specific protein levels do not fully follow their corresponding transcript levels. The authors further show that components of this regulatory network are significantly compromised in all tested brain regions, with reduced levels in RTT patients, except for *miR132* in the cerebellum. Among the two complementary *miR132-3p* and *miR132-5p* strands, the former appeared as the dominant *miR132* in these brain regions. This study is among limited research studies to examine human brain tissues for this rare disease.

In the section "Epigenetics in Prader-Willi syndrome," Mendiola and LaSalle present the emerging evidence of how loss of the small, non-coding RNA, *SNORD116,* contributes to neurological defects. The authors present evidence supporting Prader-Willi syndrome (PWS) as a metabolic disorder. They explain that patients present with poor feeding behavior as neonates, then progress to extreme hunger and obesity by early childhood. PWS is an imprinting disorder as *SNORD116* in neurons is expressed solely by the paternal allele. Loss of *SNORD116* expression by deletion, uniparental disomy, or abnormal DNA methylation of the imprinting control region commonly causes the phenotype. What is intriguing is the evolution of *SNORD116* function for sleep regulation and crosstalk with other key imprinted genes.

Similar to Prader-Willi Syndrome, Beckwith-Wiedemann Syndrome (BWS) is a rare imprinting disorder as paternally expressed *IGF2* and maternally expressed *H19* are altered. In their review, Papulino et al. present the complexity of the *IGF2/H19/CDKN1C* locus and how epigenetic DNA methylation regulates gene expression. As *IGF2* encodes a growth factor and *CDKN1C* encodes a cell cycle regulator, their emphasis is on examining the link with organ overgrowth and elevated cancer risk in BWS patients. In addition, valuable information is provided on the potential of repurposed "epi drugs" for treatment of BWS.

In contrast to DNA methylation and non-coding RNA, epigenetic defects in PWS and BWS, Fallah et al. describe histone methylation and acetylation defects underlying selected neurodevelopmental disorders such as Kabuki syndrome and Rubestein-Taybi syndrome, respectively. One provided key piece of information is the comparison of the broad range of tissues affected by these disorders. Of particular interest is the correlation between intellectual disability and dysmorphic body features. Although this article focuses on basic epigenetic mechanisms, treatments targeting histone defects underlying these disorders are now in use in clinics for patient treatments.

In a review by Timpano and Picketts, defective chromatin remodeling is discussed in the context of abnormal gene expression, and impacts on fundamental cellular activities, such as DNA replication, DNA repair, cellular proliferation and differentiation. Members of ATP-dependent chromatin remodeling factors include SWI/SNF, ISWI, INO80, CHD, and ATRX. One example of their association with rare disorder is the ATR-X Syndrome, a congenital disorder in males. These patients display heterogeneous phenotypes, although the disease is caused by single *ATRX* gene mutation(s). The authors provide an overview of ATRX interacting partners and biochemical functions. Further discussions include phenotypic and behavioral outcomes in different model systems of disease, transgenic mice, and potential ATRX roles in forebrain development.

Sun et al. study an epigenetic disorder in their original research. Activity Dependent Neuroprotective Protein (ADNP) is associated with a recently described disease; Helsmoortel-Van der Aa syndrome. The authors identify ANDP physical interaction with chromatin remodeling factors BRG1 and CHD4 and co-occupancy of these factors at intergenic sites. Interestingly, 75% of differentially expressed genes (DEG) from deletion or reduction in *Adnp* expression in embryonic stem cells (ESC) were upregulated. Genome wide chromatin accessibility was elevated along with altered histone marks due to ADNP loss, suggesting regulation of developmental genes through chromatin compaction. A connection between Helsmoortel-Van der Aa syndrome and other neurodevelopmental disorders such as Prader-Willi and Rubenstein-Taybi syndromes was suggested by common symptoms of intellectual disability and dysmorphic features.

The review by Konzman et al. establishes that O-GlcNAc transferase (OGT) enzyme regulation of post-translational modification O-linked-D-N-acetylglucosamine (O-GlcNAc) is involved in key cellular processes. The authors provide detailed examples of how environmental signals integrate into epigenetic DNA regulation of the human nervous system. The insights into OGT/O-GlcNAc function in human neurologic processes were obtained by studies of OGA/OGT function in model organisms such as *Drosophila* over the past several years.

Another original research on neuroblastoma is contributed by Krstic et al. Indeed, neuroblastoma is a rare disease, although it constitutes a high fraction of cancer deaths in children. The authors applied a chemo-genomic approach to study the global MYCN-epigenetic interactions. Among many epigenetic factors that they identified as MYCN targets, there are HDAC2, CBP, and CBX8. The authors reported that expression level of most MYCN-related epigenetic factors was associated with predictive patient outcomes. They completed a compound library screen for epigenetic proteins that showed wide response of neuroblastoma cells to epigenetic modulators. A particular chemical was C646, and the identified susceptibility of these cells correlated with MYCN expression.

Chen et al., provide a review entitled "A comprehensive genomic analysis constructs miRNA-mRNA interaction network in hepatoblastoma." While hepatoblastoma is rare, it is considered as the most common type of hepatic tumor in pediatric patients. By comparing hepatoblastoma and normal liver tissues, the authors identified 580 differentially expressed upregulated transcripts, and 790 downregulated transcripts. Among these, they noticed differentially expressed miRNAs. Analysis of protein-protein interaction network was further completed. The authors concluded that certain microRNAs, transcription factors, and hub genes act as potential regulators in hepatoblastoma.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

MR lab is supported by funding from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant #2016-06035), Ontario Rett Syndrome Association (ORSA) Hope Fund, University of Manitoba Research Grant program (URGP), University of Manitoba Collaborative Research Program (UCRP), Children's Hospital Research Institute of Manitoba (CHRIM), University of Manitoba CIHR Tri-Council Bridge funding, and Rady Innovation Fund. DY is supported by funding from the National Institutes of Health (NIH) grants; NIAA1R01AA027075 and NICHD5R01HD098038.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 01 July 2020
doi: 10.3389/fcell.2020.00553

ADNP Controls Gene Expression Through Local Chromatin Architecture by Association With BRG1 and CHD4

XiaoYun Sun[1†], WenJun Yu[3†], Li Li[2] and YuHua Sun[1,3*]*

[1] The Key Laboratory of Aquatic Biodiversity and Conservation, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China, [2] Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan, China, [3] The Innovation of Seed Design, Chinese Academy of Sciences, Wuhan, China

ADNP (Activity Dependent Neuroprotective Protein) is proposed as a neuroprotective protein whose aberrant expression has been frequently linked to rare neural developmental disorders and cancers, including the recently described neurodevelopmental Helsmoortel-Van der Aa syndrome. Recent studies have suggested that ADNP functions as an important chromatin regulator. However, how ADNP-regulated chromatin mechanisms control gene expression and stem cell fate commitment remains unclear. Here we show that ADNP interacts with two chromatin remodelers, BRG1 and CHD4. ADNP is required for proper establishment of chromatin accessibility, nucleosome configuration, and bivalent histone modifications of developmental genes. Loss of ADNP leads to enhancer over-activation and increased ratio of H3K4me3/H3K27me3 at key primitive endoderm (PrE) gene promoters, resulting in prominent up-regulation of these genes and priming ES cell differentiation toward endodermal cell types. Thus, our work revealed a key role of ADNP in the establishment of local chromatin landscape and structure of developmental genes by association with BRG1 and CHD4. These findings provide further insights into the role of ADNP in the pathology of the Helsmoortel-Van der Aa syndrome.

Keywords: ADNP, embryonic stem cells, chromatin, lineage-specifying genes, BRG1, CHD4

OPEN ACCESS

Edited by:
Mojgan Rastegar, University of Manitoba, Canada

Reviewed by:
Yong Xu, Nanjing Medical University, China
David D. Eisenstat, University of Alberta, Canada
Frank Kooy, University of Antwerp, Belgium

***Correspondence:**
Li Li
li.li@mail.hzau.edu.cn
YuHua Sun
sunyh@ihb.ac.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Cell and Developmental Biology

Received: *12 December 2019*
Accepted: *10 June 2020*
Published: *01 July 2020*

Citation:
Sun X, Yu W, Li L and Sun Y (2020) ADNP Controls Gene Expression Through Local Chromatin Architecture by Association With BRG1 and CHD4. Front. Cell Dev. Biol. 8:553. doi: 10.3389/fcell.2020.00553

INTRODUCTION

Embryonic stem cells (ESCs) possess an epigenome and chromatin structures that are required for the maintenance of self-renewal and pluripotency. The ES-specific chromatin state is directly or indirectly regulated by various factors, including epigenetic regulators and signaling molecules (Gifford et al., 2013). Chromatin remodelers are epigenetic regulators that use ATPase activity for nucleosome assembly and organization, chromatin access and nucleosome editing (Chen and Dent, 2014; Clapier et al., 2017). Great progress has been made in understanding the biochemical composition of the chromatin remodeler complexes and their role in ES cell self-renewal and pluripotency has been firmly established (Ho et al., 2008; Kidder et al., 2009; Lu and Roberts, 2013; Zhang et al., 2014; O'Shaughnessy-Kirwan et al., 2015; Zhao et al., 2017).

Activity Dependent Neuroprotective Protein was first described as a neuroprotective protein and has been implicated in various rare neural developmental disorders and

cancers, including the Helsmoortel-Van der Aa syndrome, gastric and colorectal cancers (Pinhasov et al., 2003; Vandeweyer et al., 2014). The Helsmoortel-Van der Aa syndrome is characterized by global developmental delay, intellectual disability, dysmorphic features, hypotonia and autism (Helsmoortel et al., 2014). However, the molecular mechanism underlying the syndrome remains poorly understood. ADNP contains nine zinc fingers and a homeobox domain, suggesting that it functions as a transcription factor. Consistently, ADNP deficiency in pluripotent P19 cells leads to aberrant gene activity, functioning as both transcriptional activator and repressor (Gozes et al., 2015). A growing body of research has shown that ADNP functions as an important chromatin regulator by physical association with chromatin remodelers. For instance, ADNP was shown to interact with core sub-units of the SWI/SNF chromatin remodeling complex such as BRG1 and BAF250 (Mandel and Gozes, 2007). By association with the chromatin regulator HP1, ADNP localizes to pericentromeric heterochromatin regions where it silences major satellite repeat elements (Mosch et al., 2011). ADNP forms a stable tripartite complex with CHD4 and HP1 (called the ChAHP) to control lineage gene expression in ESCs (Ostapcuk et al., 2018). Recently, it has been shown that ADNP regulates local chromatin architecture by competing for binding with CTCF, a master genome architecture protein (Phillips-Cremins et al., 2013; Kaaij et al., 2019).

Although approximately 15,000 ADNP bound sites were identified in ESCs, most ADNP ChIP-seq peaks are not localized at gene promoters (Ostapcuk et al., 2018). In addition, many genes bound by ADNP are not deregulated in the absence of ADNP (this work and Kaaij et al., 2019). Thus, the mechanism by which ADNP regulates gene expression remains unclear. In ES cells, most developmental transcription factors are in bivalent state which is characterized by the presence of both H3K4me3 and H3K27me3 at gene promoters. The bivalent domains are proposed to silence developmental genes in ES cells while keeping them poised for later activation (Ku et al., 2008). The enhancers of developmental genes are usually in a "poised" state, premarked by H3K4me1/H3K27me3; while the enhancers of pluripotency-related genes are marked by H3K27ac, a mark associated with active enhancers (Calo and Wysocka, 2013).

In this work, we hypothesize that the ADNP-regulated chromatin-remodeling mechanism contributes to ES cell gene expression state by modulating bivalent histone modifications and chromatin accessibility. We show that ADNP functions as a key chromatin regulator- this is potentially linked to its interaction with the chromatin remodelers, BRG1 and CHD4. ADNP is required for proper establishment of local chromatin accessibility, nucleosome configuration, and bivalent modifications of developmental genes. Loss of ADNP leads to enhancer over-activation and increased ratio of H3K4me3/H3K27me3 at key PrE gene promoters, resulting in prominent up-regulation of these genes and priming ES cell differentiation toward endodermal cell types. These findings provide further insights into the role of ADNP in the maintenance of ES cell phenotype and the pathology of the Helsmoortel-Van der Aa syndrome.

MATERIALS AND METHODS

ES Cell Culture

Mouse embryonic stem cells (mESCs) R1 were maintained in Dulbecco's Modified Eagle Medium (DMEM, BI, 01-052-1ACS) high glucose media containing 10% fetal bovine serum (FBS, Gibco, 10099141), 10% knockout serum replacement (KSR, Gibco, 10828028), 1 mM sodium pyruvate (Sigma, S8636), 2 mM L-Glutamine (Sigma, G7513), 1,000 U/ml leukemia inhibitory factor (LIF, Millipore, ESG1107) and penicillin/streptomycin (Gibco, 15140-122) at 37°C with 5% CO2.

The 2i culture condition was used as described previously (Chappell et al., 2013). The commercial ESGRO-2i Medium (Merck-Millipore, SF-016-200) was also used when necessary. We found that in 2i medium, *Adnp-/-* ESCs adopted morphology indistinguishable to that of control ESCs, and maintain self-renewal capacity for more than 20 passages that we tested.

Embryoid Body (EB) Formation

Embryonic stem cells differentiation into embryoid bodies was performed in attachment or suspension culture in medium lacking LIF or knockout serum replacement (KSR), as described in our previous report (Chappell et al., 2013).

Adnp shRNA Knockdown

The shRNA plasmids for *Adnp* (TRCN0000081670; TRCN0000081671), and the *gfp* control (RHS4459) were purchased from Dharmacon (United States). To make lentivirus, shRNA plasmids and *Trans*-lenti shRNA packaging plasmids were co-transfected into H293T cells according to the kit manual (Open Biosystems, TLP4615). After determining the virus titer, mESCs were transduced at a multiplicity of infection of 5:1. Puromycin selection (1 μg/ml) was applied for 4 days to select cells with stable viral integration. Quantitative PCR (qPCR) and Western blot were used to assess the knockdown of *Adnp*.

Generation of *Adnp-/-* ESCs

Adnp-/- mESCs were generated by CRISPR/Cas9 technology. Briefly, we designed two sgRNAs on exon 4 of the *Adnp* gene by using the online website http://crispr.mit.edu/. The sgRNAs sequences are: sgRNA1: 5′-CCCTTCTCTTACGAAAAATCAGG-3′; sgRNA2: 5′-CTACTTGGTGCGCTGGAGTTTGG-3′. SgRNAs were cloned into the pUC57-U6 expression vector with G418 resistance. The plasmids containing sgRNA and hCas9 were co-transfected into mESCs using Lipofectamine 2000 (Gibco). After 48 h, mESCs were selected with 500 μg/ml G418 for 7 days. Then the cells were re-seeded on 10 cm dishes coated with 0.1% gelatin to form colonies. The single colony was picked up and trypsinized for passage. DNA from the passaged cells was extracted and used for genotyping. At least two mutant ES cell lines were established in the lab.

Generation of 3 × FLAG Tagged *Adnp*-/- mESC Cell Lines

The full-length *Adnp* cDNA (NM_009628.3) was amplified by PCR and then cloned into pCMV-3 × Flag vector. The full-length *Adnp* cDNA sequence containing N-terminal 3 × Flag sequence was subcloned into the pCAG-IRES-Puro vector. To make stable transgenic cells, *Adnp*-/- mESCs were transfected with pCAG-IRES-Puro-3 × FLAG-Adnp vector using Lipofectamine 2000 (Gibco). 48 h later, cells were selected by 1 μg/ml puromycin. After 4–5 days drug selection, cells were expanded and passaged. Western blot assays were performed to confirm the transgenic cell line using FLAG antibodies.

Inducible transgenic cell lines were established according to the manual of the Tet-Express inducible expression systems (Clontech, 631169). Briefly, *Adnp*-/- ESCs were transfected with 2 μg pTRE3G-3 × FLAG-*Adnp* with linear 100 ng puromycin marker using Lipofectamine 2000 transfection reagent. 96 h later, 1 μg/ml puromycin was added and drug selection was performed for 2 weeks to establish the stable transgenic cell line. To induce target gene expression, 3×10^6 transgenic cells were plated in 6-well plates. The next day, the Tet-Express transactivator (Clontech, 631178) was added (3 μl Tet-Express to a final 100 μl total volume according to the kit manual) for 1 h in serum-free medium to induce target gene expression. Then cells were allowed to grow in complete medium for an additional 12–24 h before assaying for target protein induction. Western blotting was used to assess target protein expression levels using FLAG antibodies. In the absence of Tet-Express transactivator, pTRE3G provides very low background expression, whereas addition of Tet-Express proteins strongly transactivates target genes.

RNA Preparation, RT-qPCR and RNA-Seq

Total RNA from mESCs was extracted with a Total RNA isolation kit (Omega, United States). 1 μg RNA was reverse transcribed into cDNA with TransScript All-in-One First-Strand cDNA synthesis Supermix (TransGen Biotech, China). Quantitative real-time PCR (RT-qPCR) was performed on a Bio-Rad qPCR instrument using Hieff qPCR SYBR Green Master Mix (Yeasen, China). The primers used for RT-qPCR are listed in **Tables 2**, **3**. All experiments were repeated for three times. The relative gene expression levels were calculated based on the $2^{-\Delta\Delta Ct}$ method. Data are shown as means ± S.D. The Student's *t* test was used for the statistical analysis. The significance is indicated as follows: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

For RNA-Seq, mESCs were collected and treated with Trizol for RNA extraction. The isolated RNAs were quantified by a NanoDrop instrument, and sent to BGI Shenzhen (Wuhan, China) for whole RNA-Seq libraries and deep sequencing. RNA-Seq experiments were repeated for three times. Differentially expressed genes (DEGs) were defined by FDR < 0.05 and a Log_2 fold change > 1 was deemed to be DEGs.

Protein Extraction, and Western Blot Analysis

For protein extraction, ES cells and EBs were harvested and lysed in TEN buffer (50 mM Tris–HCl, 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.5% Na-Deoxycholate, with Roche cOmplete Protease Inhibitor). The lysates were quantified by the Bradford method and used for Western blot assay. Antibodies used for WB were ADNP (R&D Systems, AF5919, 1:500), FLAG (F3165, Sigma, 1:1000), HA (66006-1-Ig, Proteintech, 1:1000), BRG1 (21634-1-AP, Proteintech, 1:1000), CHD4 (ab181370, Abcam, 1:1000), SOX17 (24903-1-AP, Proteintech, 1:1000), GATA4 (19530-1-AP, Proteintech, 1:1000) and GATA6 (55435-1-AP, Proteintech, 1:1000). WB assay was performed as described previously. Briefly, the proteins were separated by 10% SDS-PAGE and transferred to a PVDF membrane. After blocking with 5% (w/v) non-fat milk for 1 h at room temperature, the membrane was incubated overnight at 4°C with the primary antibodies. Then the membranes were incubated with a HRP-conjugated goat anti-rabbit IgG (GtxRb-003-DHRPX, ImmunoReagents, 1:5000), a HRP-linked anti-mouse IgG (7076S, Cell Signaling Technology, 1:5000) for 1 h at room temperature. The GE ImageQuant LAS4000 mini luminescent image analyzer was used for photography. Western blots were repeated at least two times. Quantification of WB band intensity was performed by use of ImageJ software.

Co-immunoprecipitation (Co-IP)

Co-immunoprecipitation was performed for either ESCs or HEK293T cells as described in the text. Before performing co-IP, stable or transgenic cell lines were established as described above. For making transgenic cells, the full length or partial cDNAs of *Chd4* (geneID: 107932), *Brg1* (*Smarca4*, geneID: 20586) and *Adnp* genes were amplified by PCR and then cloned into the pCAG vector. The primers used for PCR are listed in **Table 1**. The constructs were verified by DNA sequencing. Co-IP experiments were performed with Dynabeads Protein G (Life Technologies, United States) according to the manufacturer's instructions. Briefly, 1.5 mg Dynabeads was conjugated with 10 μg IgG, or 10 μg anti-ADNP antibody, or 10 μg anti-FLAG antibody, or 10 μg anti-HA antibody, or 10 μg anti-BRG1 antibody, or 10 μg anti-CHD4 antibody. The whole cell lysates from cells were incubated with antibody-coupled Dynabeads overnight at 4°C. The next day, the beads were washed with PBST and boiled with loading buffer for 5 min. The protein samples were run on a SDS-PAGE gel and transferred to a PVDF membrane. The membrane was blocked with 5% (w/v) non-fat milk for 1 h at room temperature (RT), and followed overnight at 4°C with antibodies against ADNP (R&D Systems, AF5919, 1:500), FLAG (F3165, Sigma, 1:1000), HA (66006-1-Ig, Proteintech, 1:1000), BRG1 (21634-1-AP, Proteintech, 1:1000), CHD4 (ab181370, Abcam, 1:1000). Next day, the membranes were incubated with secondary antibodies (HRP-conjugated goat anti-rabbit IgG (GtxRb-003-DHRPX, ImmunoReagents, 1:5000), or HRP-linked anti-mouse IgG (7076S, Cell Signaling Technology, 1:5000) for 1 h at room temperature. After three times wash with PBST, the ECL substrate (Pierce, #32109) was applied for detection of signals. The GE ImageQuant LAS4000 mini luminescent image analyzer was used for photography.

Mapping experiments were performed in HEK293T cells. 2×10^7 cells were seeded in 10 cm dishes without antibiotics in DMEM medium containing 10% FBS at 37°C with 5% CO2. 24 h

TABLE 1 | The primers for qRT-qPCR.

Mouse genes	Forward (5′-3′)	Reverse (5′-3′)
β-*actin*	AGAGGGAAATCGTGCGTGAC	CAATAGTGATGACCTGGCCGT
Nanog	ACCCAACTTGGAACAACCAG	CGTAAGGCTGCAGAAAGTCC
Pou5f1	CGTTCTCTTTGGAAAGGTGTTC	GAACCATACTCGAACCACATCC
Pax6	AGTGAATGGGCGGAGTTATG	ACTTGGACGGGAACTGACAC
Nestin	CCCTGAAGTCGAGGAGCTG	CCCTGAAGTCGAGGAGCTG
Gsc	GCACCATCTTCACCGATGAG	AGGAGGATCGCTTCTGTCGT
Brachyury/T	CTGGGAGCTCAGTTCTTTCG	CCCCTTCATACATCGGAGAA
Gata4	TCTCACTATGGGCACAGCAG	GCGATGTCTGAGTGACAGGA
Gata6	CAAAAGCTTGCTCCGGTAAC	TGAGGTGGTCGCTTGTGTAG
Sox17	GCTTCTCTGCCAAGGTCAAC	CTCGGGGATGTAAAGGTGAA

TABLE 2 | The primers for ChIP-qPCR.

Mouse genes	Forward (5′-3′)	Reverse (5′-3′)
Park2 P1	CTGGGATCCGAGGCTAGAGT	ACCAGCGTTTCTGTCAGGTT
Sox17 P2	ACTAGTCTTGGGAAAGCGCC	AGAAAGAAAGCCCGGGGATG
Gata4 P1	CTAACGGGCCTGGTGTTCTT	CCCACTCACAGGGTGACTTC
Gata6 P1	TTTAGGGCTCGGTGAGTCCA	GAGGAAACAACCGAACCTCG
Nanog P1	CATCACGTCGGACTGCTTCT	CAGGGTTTCTCGTCCTTTCCT
pou5f1 P2	TGGAGACTTTGCAGCCTGAG	TTCTAGTCCACACTGCGTCG
Pax6 P1	ACGACGAAAGAGAGGATGCC	GGGCTTTCGCTGGAAGTAGA
Sox1 P1	GGCTGAGCTGAGTGCAAAGT	GGGTCGTGTTTAAATGCGCT

later, the plasmids containing a gene of interest were transfected into HEK293T cells using Lipofectamine 2000 (Gibco) according to the manufacturer's instructions. And 48 h later, the cells were harvested for the co-IP experiments.

Sequential Immunoprecipitation

3 × Flag-Tagged-ADNP Adnp-/- mESCs were seeded in 10 cm dishes and allowed to grow to 80–90% confluence. The cells were treated with 10 μM MG132 for 3 h and then harvested with a cell scraper. The lysate was prepared with lysis buffer containing 50 mM Tris–HCl (PH7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 and 1 × ROCHE protease inhibitor. Sequential IP was carried out as follows: 1.5 mg Dynabeads was conjugated with 10 μg anti-FLAG antibody (F3165, Sigma) at room temperature for 2 h, then the lysates were incubated with antibody coupled Dynabeads overnight at 4°C with rotation. After washing with IP wash buffer (50 mM Tris–HCl PH 7.4 and 150 mM NaCl) 3 times, 0.5 mg/mL 3 × FLAG peptides (F4799, Sigma) was added and incubated with the washed Dynabeads overnight at 4°C with rotation. Next day, the supernatants were collected by a magnetic stand and used for second round IP. 50 μl supernatants were saved as input. The remainder of the supernatants were incubated with the Dynabeads pre-coupled with anti-CHD4 antibody (21634-1-AP, Proteintech) overnight at 4°C with rotation. After extensive washes, the Dynabeads were resuspended with 5 × loading buffer. Then the mixture was boiled at 95°C for 5 min, followed by Western blot assay using the anti-BRG1 antibody.

TABLE 3 | The primers for plasmid constructions.

Mouse genes	Forward (5′-3′)	Reverse (5′-3′)
Adnp	ATGTTCCAACTTCCTGT CAACAATC	GCATATGGGCCGT GTTGCATC
Adnp-N	ATGTTCCAACTTCCTG TCAACAATC	TCACAATGTCAAA TCAAAGCTCAAAG
Adnp-C	ATGGTTCATATTGATG AAGAGATGG	GCATATGGGCCGTGT TGCATC
Adnp (1–735)	ATGTTCCAACT TCCTGTCAACAATC	TCATTCATGGTCCTC AATGACATGCT
Adnp (733–1473)	ATGGAACGGATA GGCTATCAGGTC	TCAGAGGCATTTG CTAGTAAAATTGTG
Adnp (1450–2055)	ATGCACAATTTTA CTAGCAAATGCCTC	TCAGTGGACTAG ATGCAGAGTGAT
Adnp (2035–2451)	ATGATCACTCTGC ATCTAGTCCAC	TCAGTACTTTTC ACAGTCGCGGAC
Adnp (2430–3371)	ATGGTCCGCGACT GTGAAAAGTAC	GCATATGGGC CGTGTTGCATC

Chromatin Immunoprecipitation (ChIP) and ChIP-seq

Chromatin Immunoprecipitation experiments were performed according to the Agilent Mammalian ChIP-on-chip manual as described (Singh et al., 2015). Briefly, 1×10^8 ES cells were fixed with 1% formaldehyde for 10 min at room temperature. Then the reactions were stopped by 0.125 M Glycine for 5 min with rotating. The fixed chromatin were sonicated to an average of 200–500 bp (for ChIP-Seq) or 500–1,000 bp (for ChIP-qPCR) using the S2 Covaris Sonication System (United States) according to the manual. Then Triton X-100 was added to the sonicated chromatin solutions to a final concentration of 0.1%. After centrifugation, 50 μl of supernatants were saved as input. The remainder of the chromatin solution was incubated with Dynabeads previously coupled with 10 μg ChIP grade antibodies (ADNP, R&D Systems, AF5919; H3K4me3, Abcam, ab1012; H3K27me3, Abcam, ab192985) overnight at 4°C with rotation. Next day, after 7 times washing with the wash buffer, the complexes were reverse cross-linked overnight at 65°C. DNAs were extracted by hydroxybenzene-chloroform:isoamylalcohol and purified by a Phase Lock Gel (Tiangen, China).

For ChIP-PCR, the ChIPed DNA were dissolved in 100 μl distilled water. Quantitative real-time PCR (qRT-PCR) was performed using a Bio-Rad qPCR instrument. The enrichment was calculated relative to the amount of input as described. All experiments were repeated at least for three times. The relative gene expression levels were calculated based on the $2^{-\Delta\Delta Ct}$ method. Data were shown as means ± S.D. The Student's t test was used for the statistical analysis. The significance is indicated as follows: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

For ChIP-seq, the ChIPed DNA were dissolved in 15 μl distilled water. Library constructions and deep sequencing were done by the BGI Shenzhen (Wuhan, China). All ChIP-seq experiments were repeated two times.

Calculation of H3K4me3/H3K27me3 Ratio at Gene Promoters

The promoter chromatin state was calculated as the relative ratio of the signal derived from the number of H3K4me3 and H3K27me3 sequence reads across a window between −3 and +3 kb of the annotated TSS. The relationship between H3K4me3/H3K27me3 ratio and expression was calculated by averaging of the H3K4me3/H3K27me3 ratio within a sliding window 100 observations wide, incrementing by 1, using a Spearman rank correlation. The ratio for *Adnp-/-* ESCs was relative to that of control ESCs. The calculation was based on the two ChIP-seq replicates. Data were shown as means ± S.D. The Student's *t* test was used for the statistical analysis. The significance is indicated as follows: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

ATAC-seq Assay

A 50,000 control ESCs and *Adnp-/-* ESCs in LIF-KSR medium were used for ATAC-seq assay. The experiment was performed in biological replicates using two independent isogenic cell lines for each genotype. Library preparation and ATAC-seq experiments were done by the BGI company (Wuhan, China). Libraries were paired-end sequenced (2 × 75 bp) using an Illumina NextSeq 500 device.

Immunoprecipitation in Combination With Mass Spectrometry

For IP-Mass spectrometry, the IP samples (previously immunoprecipitated by IgG or ADNP antibody) were run on SDS-PAGE gels and stained with Coomassie Brilliant Blue. Next, the entire lanes for each IP samples were cut out and transferred into a 15 ml tube containing 1 ml deionized water. Further sample treatment and the Mass Spectrometry analysis were done by the GeneCreate Biological Engineering Company (Wuhan, China).

Immunofluorescence Assay

Cells previously seeded onto glass slides were fixed with 4% paraformaldehyde for 10 min at room temperature. Then cells were washed with ice-cold PBST three times. Following the incubation with blocking buffer (5% normal horse serum, 0.1% Triton X-100, in PBS) at room temperature for 1 h, cells were incubated with primary antibodies anti-OCT3/4 (N-19) (sc-8628, Santa Cruz) at 4°C overnight. After three-times washing with PBST, the cells were incubated with the secondary antibodies (1: 500 dilution in antibody buffer, Alexa Fluor, Thermo Fisher) at room temperature for 1 h in the dark. The nuclei were stained with DAPI (D9542, Sigma, 1:1000). After washing with PBS twice, the slides were mounted with 100% glycerol on histological slides. Images were taken by a Leica SP8 laser scanning confocal microscope (Wetzlar, Germany).

Protein-Protein Interaction Assay Using a Rabbit Reticulocyte Lysate System

The Protein-Protein Interaction Assay using the Rabbit Reticulocyte Lysate System has been described (Sun et al., 2018). FLAG or HA tagged-ADNP, BRG1 or CHD4 proteins were synthesized using the TNT coupled reticulocyte lysate system according to the manual (L5020, Promega, United States). Briefly, 1 μg of a circular PCS2-version of plasmid DNA was added directly to the TNT lysate and incubated for 1.5 h at 30°C. 1 μl of the reaction products was subjected to WB assay to evaluate the synthesized protein. For protein-protein interaction assay, 5–10 μl of the synthesized HA or FLAG tagged proteins were mixed in a 1.5 mL tube loaded with TEN buffer, and the mixture was shaken for 30 min at room temperature. Next, IP or pull-down assay was performed using Dynabeads protein G coupled with FLAG or HA antibodies as described above.

Alkaline Phosphatase (AP) Staining

Alkaline phosphatase activity of mESCs was performed with a Leukocyte Alkaline Phosphatase Kit (Sigma, 86C-1KT) according to the manufacturer's instructions as described previously (Chappell et al., 2013).

BIOINFORMATICS ANALYSIS

ChIP-seq Analysis

ChIP-seq data were aligned in Bowtie2 (version 2.2.5) with default settings. Non-aligning and multiple mappers were filtered out. Peaks were called on replicates using the corresponding inputs as background. MACS2 (version 2.1.1) was run with the default parameters. Peaks detected in at least two out of three replicates were kept for further analysis. BigWig files displaying the full length for uniquely mapping reads were generated using the bedGraphToBigWig (UCSC binary utilities). To investigate the co-occupancy of ADNP, BRG1 and CHD4, we consulted previously published ChIP-seq data sets for BRG1 (GSE87820) and CHD4 (GSE64825) (Dieuleveult et al., 2016). To investigate the co-occupancy of ADNP with H3K4me1 and H3K27ac, we consulted previously published ChIP-seq data sets for H3K4me1 (GSM2575694) and H3K27ac (GSM2575695) (Dieuleveult et al., 2016).

RNA-seq Analysis

All sequencing reads were aligned to the 9 mm mouse genome assembly from the UCSC genome browser. Data were aligned using Bowtie2 with the default settings. Aligned and sorted reads were indexed using SAMtools (version 1.2). Reads were counted over exons using the R summarize Overlaps function and collapsed to yield one value per gene. The read counting is performed for exonic gene regions in a non-strand-specific manner while ignoring overlaps among different genes. Subsequently, the expression count values were normalized by Reads Per Kilobase per Million mapped reads (RPKM). The count table was used for differential expression calling with the EdgeR package. FDR < 0.05 and $\log_2$ fold change > 1 was deemed to be a differentially expressed gene. For comparative transcriptome analysis in the presence and absence of ADNP, BRG1 and CHD4, we consulted the published RNA-seq data sets for *Brg1* KO (GSE87821) and *Chd4* KO (GSE80280) (King and Klose, 2017).

ATAC-seq Analysis

Paired-end reads were aligned using Bowtie2 using default parameters. Only uniquely mapping reads were kept for further analysis. These uniquely mapping reads were used to generate bigwig genome coverage files similar to ChIP–seq. Heat maps were generated using deeptools2. For the meta-profiles, the average fragment count per 10-bp bin was normalized to the mean fragment count in the first and last five bins, which ensures that the background signal is set to one for all experiments. Merged ATAC-seq datasets were used to extract signal corresponding to nucleosome occupancy information with NucleoATAC. For comparison analysis of ADNP, BRG1 and CHD4 ATAC-seq signals, we consulted previously published ATAC-seq data sets for *Brg1* KD (GSM1941485-6) and *Chd4* KD (GSM1941483-4) (Dieuleveult et al., 2016).

Differential Binding and Gene Expression Analysis

Significant changes in ATAC-seq were identified using the DiffBind package, a FDR < 0.05 and $\log_2$ fold change > 1 was deemed to be a significant change. Gene ontology (GO) analysis for differentially regulated genes, and heat maps were generated from averaged replicates using the command line version of deepTools2. Peak centers were calculated based on the peak regions identified by MACS (see above).

Quantification and Statistical Analysis

Data are presented as mean values ± SD unless otherwise stated. Data were analyzed using Student's *t* test analysis. Error bars represent s.e.m. Differences in means were statistically significance when $p < 0.05$. Significant levels are: $^{*}p < 0.05$; $^{**}P < 0.01$; $^{***}p < 0.001$.

Data Availability

All RNA-seq, ATAC-seq and ChIP-seq data have been deposited into the database at https://bigd.big.ac.cn/. The accession numbers are CRA001624 and CRA002148. All other related data will be available upon reasonable request.

RESULTS

Adnp Ablation Leads to Significant Up-Regulation of PrE Genes

To understand the role of ADNP, we generated *Adnp* mutant ESCs by using CRISPR/Cas9 technology. gRNAs were designed to target the 3′ end of exon 4 of the mouse *Adnp* gene (**Figure 1A**). We have successfully generated 4 *Adnp* mutant alleles. The mutant ESCs we used in this work has the combination of 4- and 5-bp deletions in exon 4 of *Adnp*, as revealed by DNA genotyping around the CRISPR targeting site (**Figure 1B**). ADNP protein was almost undetectable in *Adnp*-/- ESCs by Western blot using ADNP antibodies from different resources (**Figure 1C** and **Supplementary Figure 1A**), which strongly supported that the mutant alleles are functional nulls.

In the traditional self-renewal medium containing LIF-KSR plus FBS, the newly established *Adnp*-/- ESC colonies overall exhibited typical ESC-like morphology, and abundantly expressed the core pluripotency factor OCT4 (**Figures 1D,E**). To understand how ADNP deficiency affects ES cell phenotype, comparative transcriptome analysis for control and early passage *Adnp*-/- ESCs was performed by RNA-sequencing (RNA-seq). A total of 1,026 differentially expressed genes (DEGs) ($\log_2$ fold change > 1 and $p < 0.05$) were identified based on two RNA-seq experiments (each has two replicates). Of which, an average of 766 genes were up-regulated and 260 genes were down-regulated (**Supplementary Figure 1B**). GO (gene ontology) analysis revealed that the majority of deregulated genes were enriched for DNA binding and catalytic activity (**Figure 1F**). In the absence of ADNP, the expression of key mesoderm specifying genes such as *Gsc* and *T* ($\log_2$ fold change < 1; RPKM < 1 in ESCs), and neuroectoderm specifying genes such as *Fgf5*, *Nestin* and *Olig2* ($\log_2$ fold change < 1; RPKM < 2), as well as pluripotency genes such as *Pou5f1* and *Nanog* ($\log_2$ fold change < 1) was barely changed. Remarkably, genes implicated in extraembryonic primitive endoderm (PrE) development such as *Gata4*, *Gata6*, *Sox7*, *Krt18*, *Sparc*, *Cited1*, *Dab2*, and *Cubn* ($\log_2$ fold change > 1; RPKM > 5 in mutant ESCs) were significantly up-regulated. The qRT-PCR assay confirmed the RNA-seq results (**Figure 1G**). These data suggested that ADNP performs an important role in repressing PrE genes in ESCs. Although PrE genes were up-regulated, *Adnp*-/- ESCs can maintain self-renewal capacity for many generations before eventually adopting a flattened morphology and exhibiting reduced alkaline phosphatase activity (**Figure 1H**). Thus, our results indicated that acute ADNP depletion in ESCs does not result in sudden and complete loss of self-renewal, while prolonged ADNP depletion may cause ESC differentiation toward endodermal cell types, likely due to up-regulation of the key endoderm-specifying genes.

It has been shown that loss of ADNP disrupted the differentiation potential of ESCs (Ostapcuk et al., 2018). Similar results were obtained in our hand by performing embryoid body (EB) formation of mutant and control ESCs (**Figure 1I**). In day 6 EBs derived from control ESCs, neural genes *Nestin* and *Pax6* as well as PrE genes *Gata6* and *Sox17* were induced as expected. In day 6 EBs derived from *Adnp*-/- ESCs, however, the PrE genes were abnormally up-regulated, at the expense of neural genes. WB analysis confirmed that GATA6 and GATA4 levels were higher in *Adnp*-/- ESC-derived EBs than in control ESC-derived EBs (**Figure 1J**). When FLAG-ADNP was re-introduced into mutant ESCs (**Supplementary Figure 1C**), the defective gene expression and the alkaline phosphatase activity were largely rescued (**Figures 1H,I**). This data demonstrated that the observed phenotypes were specifically due to the loss of ADNP.

ADNP Associates With Chromatin Remodelers BRG1 and CHD4

To understand the role of ADNP in ESCs, we sought to identify its interacting proteins by performing immunoprecipitation (IP) in combination with mass spectrometry (Mass Spec) assay using commercial ADNP antibodies (**Figure 2A**). The commercial

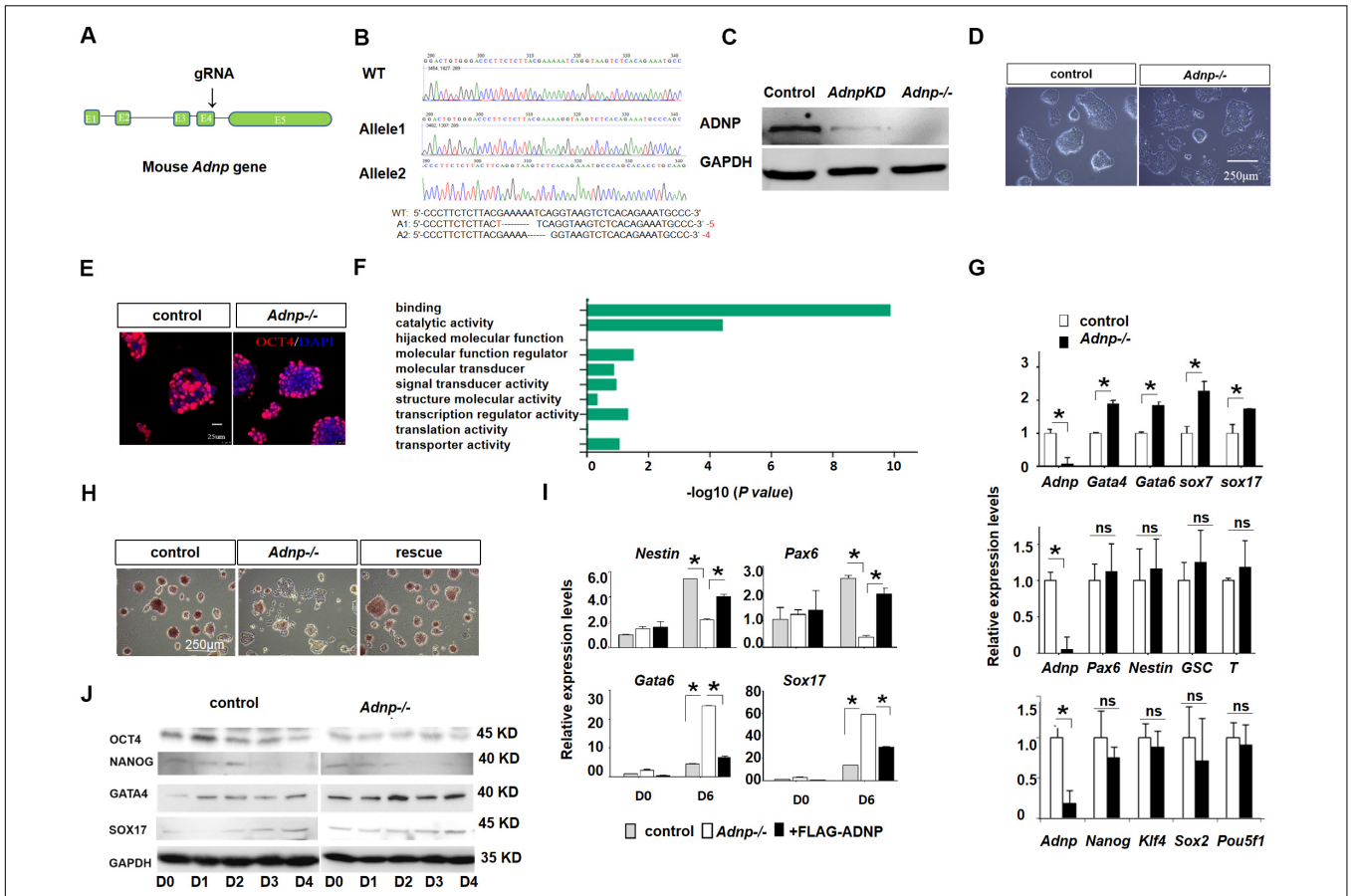

FIGURE 1 | *Adnp* ablation leads to significant up-regulation of PrE genes. **(A)** Schematic representation of the mouse *Adnp* gene, depicting the location of the CRISPR/Cas9 targeting site. **(B)** Genotyping results of two mutant alleles. **(C)** WB assay of ADNP for *Adnp* knockdown and *Adnp-/-* ESCs. **(D)** Phenotypes of wild type and *Adnp-/-* ESCs. **(E)** Immunofluorescence assay of OCT4 for control and *Adnp-/-* ESCs. **(F)** GO analysis of DEGs in the absence of ADNP. **(G)** qRT-PCR assay of pluripotency, endodermal, neuroectodermal and mesodermal genes. **(H)** Alkaline phosphatase stain of wild type, *Adnp-/-* and FLAG-ADNP rescued ESCs. **(I)** Expression of key PrE and neural genes of day 0 and day 6 EBs derived from control and mutant ESCs. **(J)** WB showing the abnormal induction of SOX17 and GATA4 levels in the absence of ADNP. The RNA-seq, qRT-PCR and WB assays were based on three replicates. Differences in means were statistically significant when $p < 0.05$. Significant levels are: $^*p < 0.05$.

ADNP antibodies we used are specific, as demonstrated by the fact that a clean band around 150 kD which is the predicted size of ADNP was detected by WB analysis, and this band became barely detected in *Adnp-/-* ESCs (**Figure 1C**). A total of 180 ADNP-interacting candidate proteins were identified, which included the known ADNP interactors HP1γ and CHD4 (Ostapcuk et al., 2018), which further supported the specificity of the antibodies used. We confirmed that the N-terminal fragment of ADNP binds to CHD4 which was in line with the recent report by Ostapcuk and co-workers (data not shown). Further mapping experiments revealed that the N-terminal fragment of ADNP binds to the C-terminal but not N-terminal fragment of CHD4 (**Figures 2B,C** and **Supplementary Figure 2A**).

Activity Dependent Neuroprotective Protein was previously shown to interact with BRG1 and BAF250, core sub-units of BAF ATP-dependent chromatin remodeling complexes in HEK293 cells (Mandel and Gozes, 2007). BRG1 and BAF250 are conserved components of the ES cell-specific BAF complex called esBAF (Ho et al., 2008). Although no esBAF components were identified in our Mass Spec assay, *Adnp-/-* ESCs resemble BRG1 or BAF250a deficient ESCs not only in gene expression but also in morphological aspects (Gao et al., 2008; Kidder et al., 2009). We therefore performed co-IP experiments to examine whether ADNP interacts with BRG1 or BAF250a in ESCs. Our co-IP results showed that endogenous BRG1 but not BAF250a was able to pull down ADNP (**Figure 2D**). Further mapping experiments showed that the C-terminal fragment of ADNP interacts with the N-terminal but not the C-terminal of BRG1 (**Figure 2E** and **Supplementary Figure 2B**). To investigate whether ADNP physically associates with BRG1, we used a reticulate lysate system to synthesize the FLAG-tagged C-terminal fragment of ADNP and the HA-tagged N-terminal fragment of BRG1. When they were mixed together, anti-FLAG antibodies could readily pull down HA-BRG1-N (**Figure 2F**).

Although endogenous BRG1 and CHD4 interact with each other in mouse embryos, it is not known whether this was the case in ESCs (Shimono et al., 2003; Singh et al., 2016). By performing co-IP experiments, we found that endogenous BRG1 pulled down

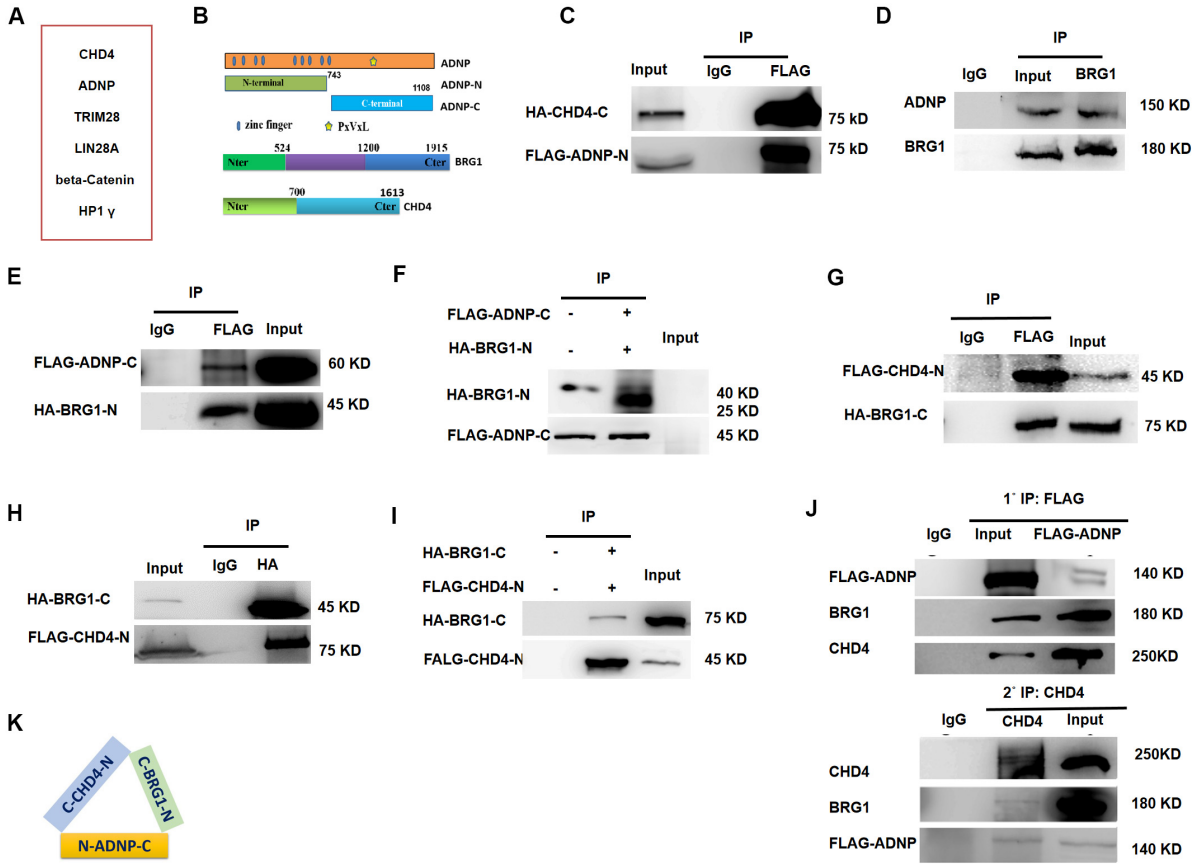

FIGURE 2 | ADNP forms complexes with BRG1 and CHD4. **(A)** A table showing representative ADNP interacting proteins by IP followed mass-spectrometry assay. **(B)** Schematic representation of wild type ADNP, BRG1 and CHD4, and their truncated mutant forms. **(C)** FLAG-tagged ADNP-N pulled down HA-tagged CHD4-C in 293T. **(D)** Endogenous BRG1 pulled down ADNP in ESCs. **(E)** FLAG-tagged ADNP-C pulled down HA-BRG1-N in 293T. **(F)** FLAG-tagged ADNP-C pulled down HA-BRG1-N using a TNT system. **(G)** FLAG-CHD4-C pulled down HA-BRG1-N in 293T cells. **(H)** HA-BRG1-N pulled down FLAG-CHD4-C in 293T. **(I)** FLAG-tagged pulled down HA-BRG1-N using a TNT system. **(J)** Sequential IP showing that ADNP, CHD4 and BRG1 could form a triplex in ESCs. **(K)** A cartoon showing the interaction of all three factors, based on the mapping results. All WB and IPs were repeated at least two times. Shown are representative images.

CHD4. Further mapping revealed that the C-terminal fragment of BRG1 strongly associates with the N-terminal fragment of CHD4 (**Figures 2G,H**). Using a reticulate lysate system, we further confirmed that they interact physically (**Figure 2I**).

Based on the above mapping results, we speculated that ADNP, CHD4, and BRG1 may form a triplex *in vivo*. To this end, we performed sequential immunoprecipitation experiments using a transgenic *Adnp*-/- ES cell line where a 3 × FLAG-tagged version of ADNP could be induced by the addition of the Tet-Express protein. We confirmed that in the presence of the Tet-Express protein, 3 × FLAG-ADNP levels in *Adnp*-/- ESCs were similar to endogenous ADNP in control ESCs (**Supplementary Figure 2C**). In the first round IP, FLAG antibodies easily pulled down endogenous BRG1 or CHD4. Next, FLAG antibody-bound protein complexes were eluted with excessive 3 × FLAG peptide, and were subjected to the second round of IP using CHD4 antibodies. As shown in **Figure 2J**, CHD4 antibodies could pull down both FLAG-ADNP and BRG1. Thus, our sequential IP data supported that ADNP, BRG1 and CHD4 could form a tripartite complex (ABC triplex) in ESCs (**Figure 2K**), although it is possible that this triplex is a part of large uncharacterized multiprotein complexes.

ADNP, BRG1, and CHD4 Co-occupy Target Genes Genome-Wide

To better understand the role of ADNP in the maintenance of ESCs, we sought to determine its direct targets and genome-wide binding profile by chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) analysis. A total of 10,642 sites were bound by ADNP compared to the input, and 838 target genes were identified. Of which, 1,632 peaks were found in promoter proximal regions, 5,951 peaks were found in gene bodies, and the majority of the remainder were localized to intergenic regions (**Figures 3A,B** and **Supplementary Figures 3A,B**). Thus, most of ADNP peaks were localized to intergenic or promoter-distal regions, which was similar to that from the recent published FLAG-ADNP ChIP-seq results (Ostapcuk et al., 2018). Gene ontology analysis revealed that ADNP targets are enriched for genes involved in metabolic processes and

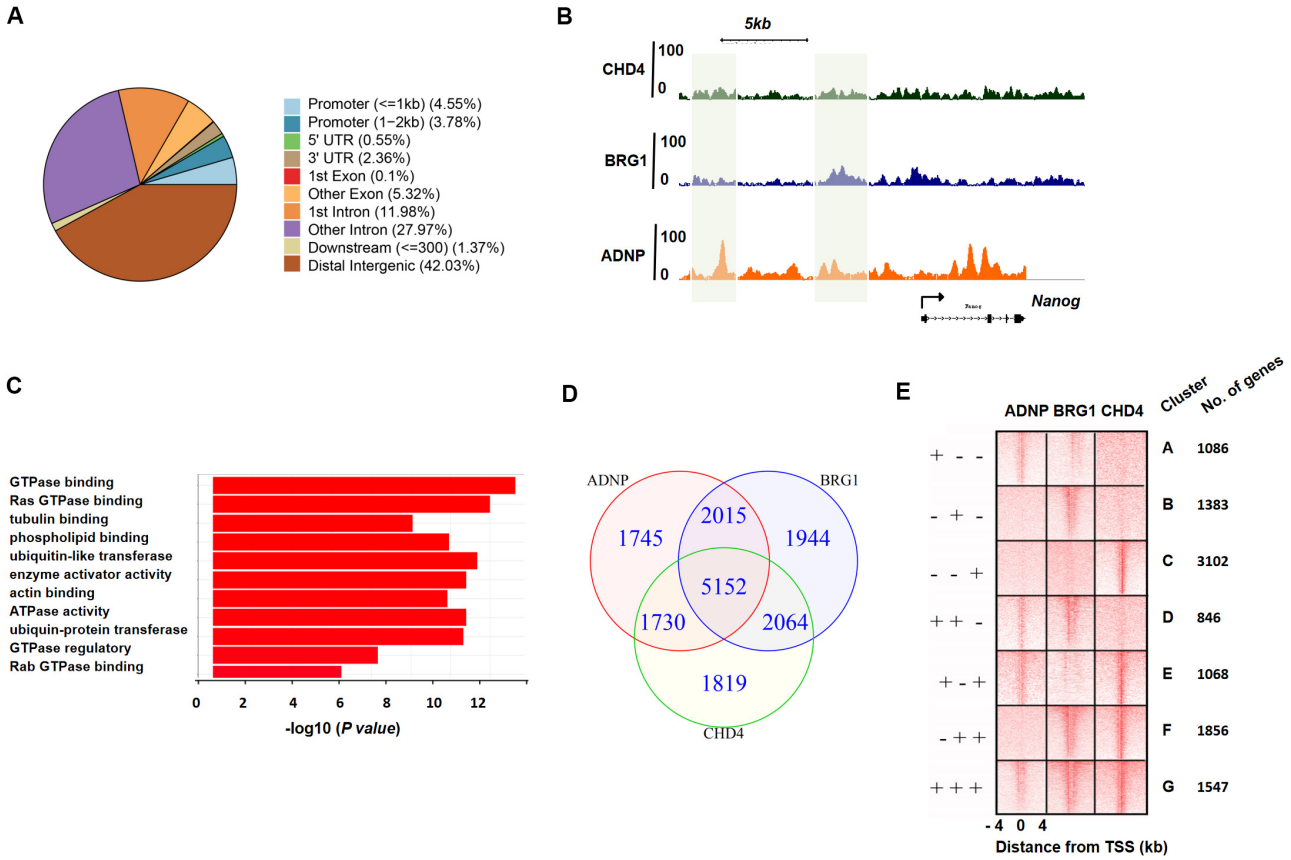

FIGURE 3 | ADNP, BRG1 and CHD4 co-occupy target genes genome-wide. **(A)** Annotation of ADNP peaks genome-wide. **(B)** A snapshot of ChIP-seq genome-browser view of ADNP, BRG1 and CHD4 occupancy at the *Nanog* locus. The co-localized loci by all three factors were highlighted in green. **(C)** GO analysis of ADNP target genes. **(D)** A Venn diagram analysis of ChIP-seq peaks for ADNP, BRG1 and CHD4. **(E)** A heat map illustrating distribution of ADNP, BRG1 and CHD4 in ± 4 kb genomic regions around transcription start sites (TSS), with gene cluster assignment and the number of each clusters.

cell signaling such as GTPase binding, G-protein signaling and cell adhesion (**Figure 3C**). As PrE genes were significantly deregulated in the absence of ADNP, we examined ADNP ChIP-seq peaks at these genes. Surprisingly, no significantly enriched ChIP-seq peaks were found at key PrE genes except *Gata4* (**Supplementary Figure 3B**). By contrast, pluripotency genes such as *Nanog* and *Pou5f1* were extensively bound by ADNP (**Figure 3A**). By ChIP-PCR, we confirmed that *Gata6*, *Sox7* and *Sox17* were barely bound by ADNP at gene promoter regions (**Supplementary Figure 3C**).

The association of ADNP with BRG1 and/or CHD4 prompted us to determine whether ADNP binding sites were co-occupied by the two chromatin remodeler factors. Unfortunately, the ChIP experiments using commercial BRG1 or CHD4 antibodies were not successful. We therefore consulted the published CHD4 or BRG1 ChIP-seq data, and revealed that 10,765 and 11,175 sites were significantly enriched for CHD4 and BRG1, respectively (Dieuleveult et al., 2016; King and Klose, 2017). CHD4 or BRG1 ChIP-seq peaks were localized to proximal promoter, gene body and intergenic regions, analogous to that of ADNP. Bioinformatics analysis were performed to examine the overlap among ADNP, BRG1 and CHD4 ChIP-seq peaks. When we compared ADNP and BRG1 sites, 67% (7,167/10,642) of ADNP peaks overlapped with 64% (7,167/11,175) of BRG1 peaks; when comparing ADNP with CHD4 sites, 65% (6,882/10,642) of ADNP peaks overlapped with 64% of CHD4 peaks (6,882/10,765). When comparing the binding of all three proteins, 31% (5,152/16,469) were co-bound by ADNP, BRG1 and CHD4 (**Figures 3B,D** and **Supplementary Figure 3A**).

We plotted ADNP, CHD4, and BRG1 ChIP-seq reads in a ± 4 kb region surrounding TSS and divided ADNP- or CHD4- or BRG1-bound genes into 7 categories (cluster A: ADNP+BRG1-CHD4-, cluster B: ADNP-BRG1+ CHD4-, cluster C: ADNP-BRG1-CHD4+, Cluster D: ADNP+BRG1+ CHD4-, cluster E: ADNP+BRG1-CHD4+, cluster F: ADNP-BRG1+CHD4+, and cluster G: ADNP+ CHD4+ BRG1+) (**Figure 3E**). We examined the effects of loss of ADNP on the expression of each cluster of genes. Interestingly, compared to all genes, loss of ADNP had a minimal effects on gene expression of all clusters except for cluster D (**Supplementary Figure 3D**). Loss of ADNP led to a significant down-regulation of cluster D genes ($p < 0.05$). GO analysis of cluster D revealed the enrichment of terms such as regulation of transcription, positive regulation of neural differentiation, cell cycle and metabolic

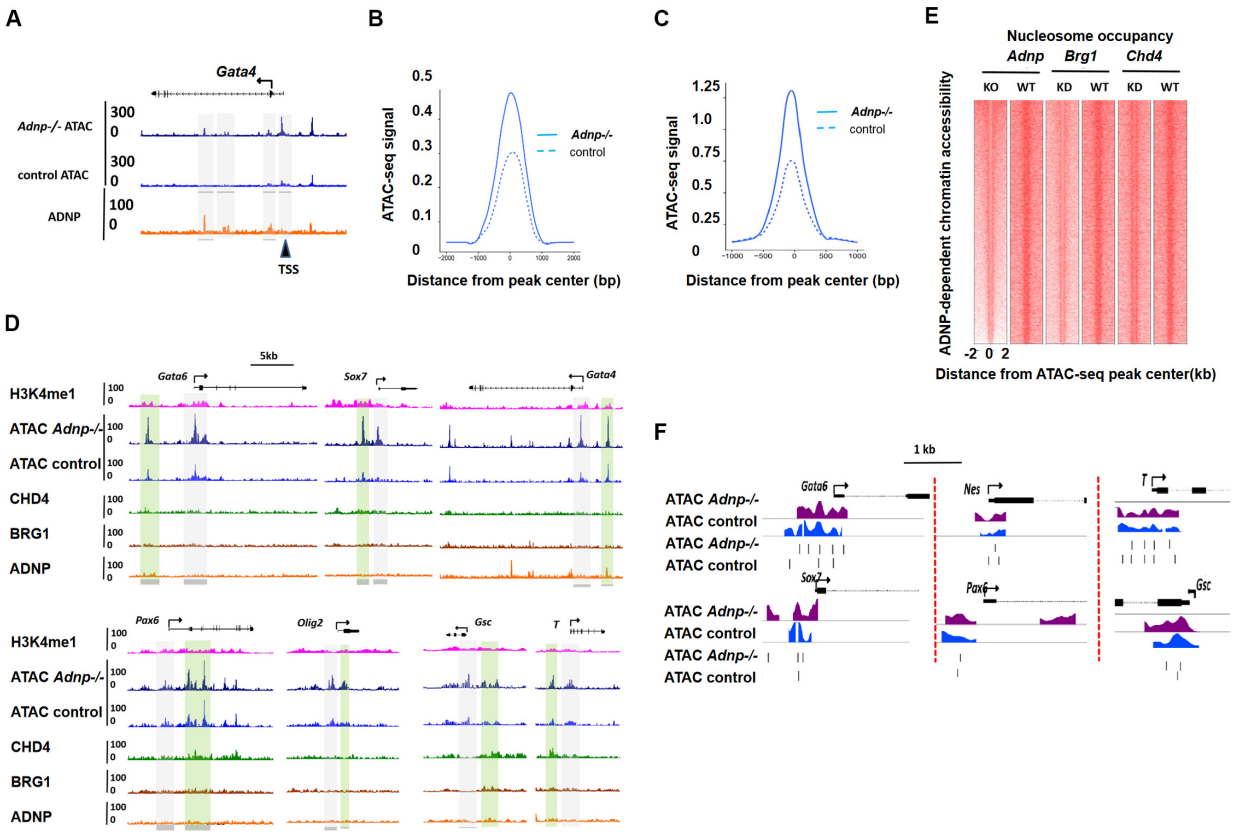

FIGURE 4 | ADNP depletion leads to chromatin accessibility and nucleosome configuration change. **(A)** A genomic snapshot of ATAC-seq and ADNP ChIP-seq signal at the *Gata4* locus in control and *Adnp-/-* ESCs. Gray color highlighting the ADNP-dependent ATAC-seq peak loci. Black triangle indicating the significant increase of ATAC-seq signal at TSS region where no ADNP ChIP-seq signal was detected. **(B)** A metaplot of average ADNP-dependent ATAC-seq signal in control and *Adnp-/-* ESCs. **(C)** A metaplot of average ATAC-seq signal at TSS regions in control and *Adnp-/-* ESCs. Similar results were obtained at enhancer regions (not shown). **(D)** A genomic snapshot of ADNP, BRG1, CHD4 and H3k4me1 ChIP-seq and ATAC-seq at lineage specifying genes in control and *Adnp-/-* ESCs. H3K4me1 was used for showing the poised enhancer regions. Gray color highlighting TSS-proximal regions, and green color highlighting enhancer regions. Except for *Gata4*, no key lineage specifying genes were bound by ADNP. Chromatin accessibility was increased at TSS-proximal and enhancer regions. **(E)** Nucleosome occupancy around ATAC-seq peak center in control, *Adnp* knockout, *Brg1* knockdown and *Chd4* knockdown ESCs. **(F)** A snapshot of nucleosome configuration at TSS region of representative lineage-specifying genes in control and *Adnp-/-* ESCs. ATAC-seq experiments were repeated two times for control and *Adnp-/-* ESCs. ATAC-Seq data for BRG1 and CHD4 were downloaded as described in the text. All analysis was based on the two replicates.

process (data not shown). This was in line with that loss of ADNP leads to compromised ESC pluripotency, particularly differentiation toward the neuronal lineage (Ostapcuk et al., 2018). Why the cluster D genes were most sensitive to loss of ADNP remains unclear.

ADNP Depletion Leads to Local Chromatin Accessibility and Nucleosome Configuration Change, and PrE Genes Appear Most Sensitive to Loss of *Adnp*

Activity Dependent Neuroprotective Protein interacting chromatin remodelers CHD4 and BRG1 have well-known functions for regulating chromatin accessibility and nucleosome configuration in ESCs (Tolstorukov et al., 2013; Lei et al., 2015). To understand how loss of ADNP affected gene expression, we performed transposase-accessible chromatin with massively parallel sequencing (ATAC-seq) for control and *Adnp-/-* ESCs. In control ESCs, the majority of ADNP-bound loci were largely devoid of ATAC-seq signals, suggesting that ADNP was bound to inaccessible chromatin. In the absence of ADNP, these sites became accessible as they showed significant ATAC-seq signals (**Supplementary Figures 4A,B**). The *Gata4* gene is shown here for individual representation (**Figure 4A**). This observation was in line with the recent report that ADNP may render local chromatin inaccessible by directly binding to these loci (Ostapcuk et al., 2018). Remarkably, we found that loss of ADNP also caused a widespread increase of ATAC-seq signals at genome loci where weak or no ADNP ChIP-seq signals were observed, primarily at gene enhancer and proximal-TSS regions (**Figures 4A–D**). This observation suggested that ADNP functions to restrict chromatin accessibility at gene regulatory regions, through a mechanism independent of its DNA binding activities. Alternatively, the chromatin accessibility at gene regulatory regions is very sensitive to loss of ADNP.

Specifically, we compared chromatin accessibility for endoderm, mesoderm and neuroectoderm specifying genes in the presence and absence of ADNP. In the absence of ADNP, a substantial increase of chromatin accessibility at both proximal-TSS and poised enhancer regions was observed for key endoderm specifying genes such as *Gata6* and *Sox7* (**Figure 4D**). Chromatin accessibility was also changed for key mesoderm and neuroectoderm specifying genes.

Next, we asked whether ADNP regulates nucleosome configuration in ESCs. Globally, nucleosome occupancy was significantly reduced in the absence of ADNP (**Figure 4E**). When examining the key lineage-specifying genes, we found that nucleosome positioning, phasing and occupancy were all significantly altered in the absence of ADNP (**Figure 4F**). It appeared that loss of ADNP had greater effects on nucleosome configuration for the PrE genes than mesoderm and neuroectoderm specifying genes: ADNP depletion led to a significant nucleosome occupancy increase around the TSS of PrE genes. Of note, the nucleosome configuration of ADNP-bound pluripotency genes was barely altered in the absence of ADNP (**Supplementary Figure 4C**).

ADNP-Regulated Chromatin Mechanism Is Linked With BRG1 and CHD4

Chromatin remodelers are well-known for their role in the regulation of chromatin structure (Musselman et al., 2012; Clapier et al., 2017; Wang et al., 2017). Based on the observation that ADNP, CHD4, and BRG1 could form complexes and co-occupy target genes, we reasoned that an ADNP-regulated chromatin mechanism might be linked with BRG1 and CHD4.

To explore this, we plotted ADNP ATAC-seq with ADNP, BRG1, and CHD4 ChIP-seq data sets, and asked whether ADNP-dependent ATAC hypersensitive peaks overlapped with BRG1 or CHD4 ChIP-seq peaks. Indeed, we found that ADNP, CHD4 or BRG1 ChIP-seq peaks partially overlapped with ADNP-dependent ATAC hypersensitive loci (**Figure 5A**). Of note, ADNP ChIP-seq signals were stronger at ADNP-dependent than ADNP-independent ATAC-seq peak loci. This was not the case for BRG1 and CHD4 ChIP-seq signals. This observation suggested that there is an inherent functional link among ADNP, BRG1 and CHD4 in shaping chromatin accessibility, and that ADNP may use both CHD4 and BRG1 to regulate chromatin accessibility by binding to the local genomic loci.

To further investigate the above hypothesis, we consulted the previously published BRG1 or CHD4 ATAC-seq data sets (King and Klose, 2017). We plotted BRG1, CHD4, and ADNP ATAC-seq data reads, and examined the co-localization of ATAC hypersensitive peaks in the absence of each factor. We found that about 10% of BRG1- or 17% of CHD4-dependent ATAC hypersensitive peaks overlapped with ADNP-dependent ATAC hypersensitive peaks (**Figures 5B,C**). Correlation analysis of loci with overlapping ATAC-seq signals revealed a high degree of co-localization of BRG1-, CHD4- and ADNP-dependent ATAC hypersensitive peaks (**Figure 5D**). Interestingly, ADNP depletion exhibited a much stronger effect on ATAC-seq signal than either BRG1 or CHD4 depletion at these loci (**Figure 5E**). These observations suggested that a co-dependency of BRG1 and CHD4 mediated by ADNP may be utilized to regulate chromatin architecture.

To determine the potential contribution of the two distinct chromatin remodelers, we compared the change of ATAC-seq signal in the absence of ADNP, BRG1 and CHD4, for developmental genes, especially the PrE-related genes. We observed that chromatin accessibility at a substantial fraction of genome loci was affected by all three factors (**Supplementary Figure 5A**). For PrE genes such as *Foxa2* and *Sparc*, either ADNP- or BRG1- or CHD4-depletion led to increased ATAC-seq signal, suggesting that BRG1 and CHD4 activities are synergistically required to maintain a closed chromatin architecture. For PrE genes such as *Sox7* and *Gata4*, chromatin accessibility was predominantly affected by ADNP and CHD4 as ATAC-seq signal was not altered by BRG1 depletion. For neuroectodermal genes such as *Fgf5* and *Nestin*, loss of BRG1 led to a reduction of ATAC-seq signal while CHD4 or ADNP depletion led to an increase of ATAC-seq signal, suggesting that BRG1 and CHD4 act antagonistically.

Loss of ADNP Caused Significant Change of Bivalent Histone Modifications for Developmental Genes

It has been known that ADNP-interacting chromatin remodelers BRG1 and CHD4 contribute to the establishment of bivalent histone modifications (Tolstorukov et al., 2013; Lei et al., 2015). We asked whether loss of ADNP led to the alteration of bivalent histone modifications for developmental genes in ESCs. To investigate this, we performed ChIP-seq analysis for H3K4me3 and H3K27me3 of control and *Adnp*-/- ESCs. Bioinformatics analysis of the ChIP-seq data showed that the levels of both H3K4me3 and H3K27me3 were changed by loss of ADNP (**Figures 6A,B**). We grouped gene promoters into three categories: H3K4me3 only, H3K27me3 only, and both H3K4me3 and H3K27me3, and asked how the histone marks changed in each category in the absence of ADNP. Bioinformatics analysis revealed that loss of ADNP caused a significant increase of both H3K4me3 (around the TSS) and H3K27me3 (0.5–4 kb upstream the TSS), resulting in a slightly increased number of all three cluster of promoters (**Figure 6C** and **Supplementary Figures 6A,B**).

To understand why loss of ADNP was associated with an up-regulation of PrE genes, we examined bivalent histone modifications for lineage-specifying genes in control and *Adnp*-/- ESCs. It seems that loss of ADNP had different effects on bivalent histone modifications depending on the lineage-specifying genes. For instance, at mesodermal genes such as *T* and *Gsc*, and neuroectodermal genes such as *Olig2*, *Pax6* and *Nestin*, a slight increase of both H3K4me3 and H3K27me3 levels was observed in *Adnp*-/- ESCs compared with control ESCs. At PrE specifying genes such as *Gata6*, *Gata4*, *Sox17*, *Sox7* and *Foxa2,* a substantial increase of H3K4me3 levels was observed, while H3K27me3 levels were slightly increased (**Figure 6D**). We confirmed this by ChIP-PCR (**Figure 6E**). It is known that the levels of H3K4me3 correlate with

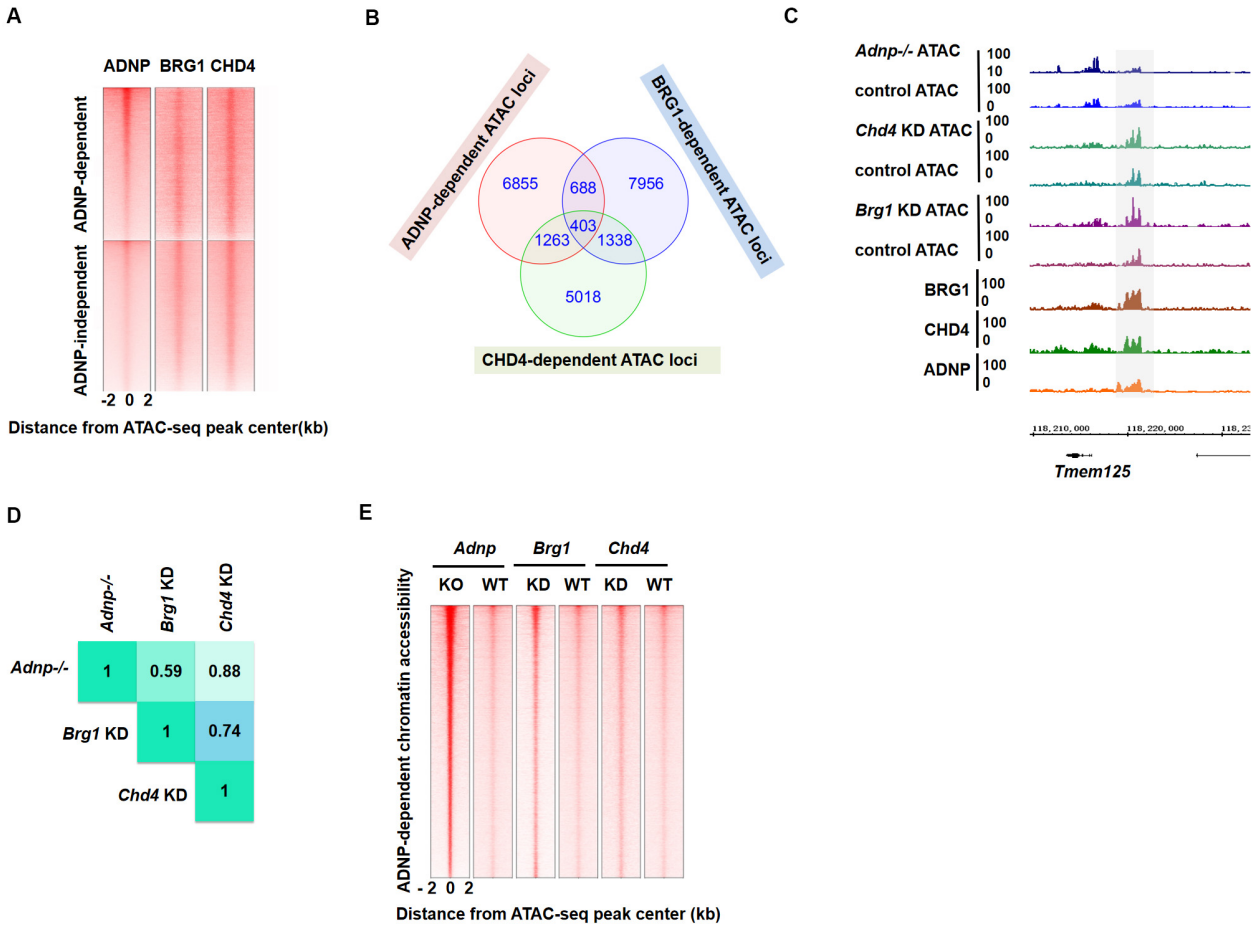

FIGURE 5 | BRG1 and CHD4 are closely-linked with ADNP-regulated chromatin accessibility. **(A)** Heatmap showing ADNP, BRG1 and CHD4 ChIP-seq enrichment around ± 2 kb ADNP-dependent and ADNP-independent ATAC-seq peak centers. **(B)** A Venn diagram analysis of BRG1-, CHD4- and ADNP-dependent ATAC hypersensitive peaks. **(C)** A snapshot showing overlapping BRG1-, CHD4- and ADNP-dependent ATAC-seq peaks at the *Tmem125* locus. **(D)** Correlation analysis of loci with overlapping ATAC-seq signals revealed a high degree of co-localization of BRG1-, CHD4- and ADNP-dependent ATAC hypersensitive peaks. **(E)** A heat map of ATAC-seq signal around the ATAC peak center, in the presence or absence of each factor. Sites are ranked by the increase of ATAC-seq signal following loss of ADNP activity.

gene activation, and the levels of H3K27me3 correlate with gene repression. And there is a positive correlation between transcript levels and H3K4me3/H3K27me3 ratio for bivalent genes in pluripotent stem cells (De Gobbi et al., 2011; Singh et al., 2015). We therefore compared the H3K4me3/H3K27me3 ratio for key lineage-specifying genes in *Adnp-/-* and control ESCs. A significant increase of H3K4me3/H3K27me3 ratio was observed at promoters of PrE genes such as *Gata6* and *Sox7* whose expression were prominently up-regulated in the absence of ADNP. For genes such as *T* and *Gsc* whose expression was barely changed in the absence of ADNP, the H3K4me3/H3K27me3 ratio in mutant ESCs was comparable to that of control ESCs (**Figure 6F**). Thus, loss of *Adnp* caused a significant increase of the H3K4me3/H3K27me3 ratio for key PrE specifying genes but not for mesodermal or neuroectodermal genes.

It is well-known that the accurate execution of gene expression programs requires 2 types of regulatory DNA elements in higher eukaryotes: promoters and enhancers. We previously showed that promoter-enhancer interactions play important roles in 3D genome organization and the control of gene expression in ESCs (Phillips-Cremins et al., 2013; Singh et al., 2015). To further understand the role of ADNP in the regulation of gene expression, we investigated how loss of ADNP affected enhancer activities of key lineage specifying genes. We found that there was a substantial increase of H3K4me3 at poised enhancer regions of PrE but not mesodermal or neuroectodermal genes in the absence of ADNP (**Figure 6G**). It has been suggested that enhancer over-activation correlates with increased H3K4me3 and decreased H3K4me1 levels (Shen et al., 2016). Thus, our data suggested that ADNP is required to maintain poised enhancers for PrE developmental genes, and loss of ADNP leads to enhancer over-activation.

MLL2 is the core component of the MLL complex that deposits H3K4me3, and EZH2 is the core component of the PRC2 complex that deposits the H3K27me3 mark at bivalent promoters

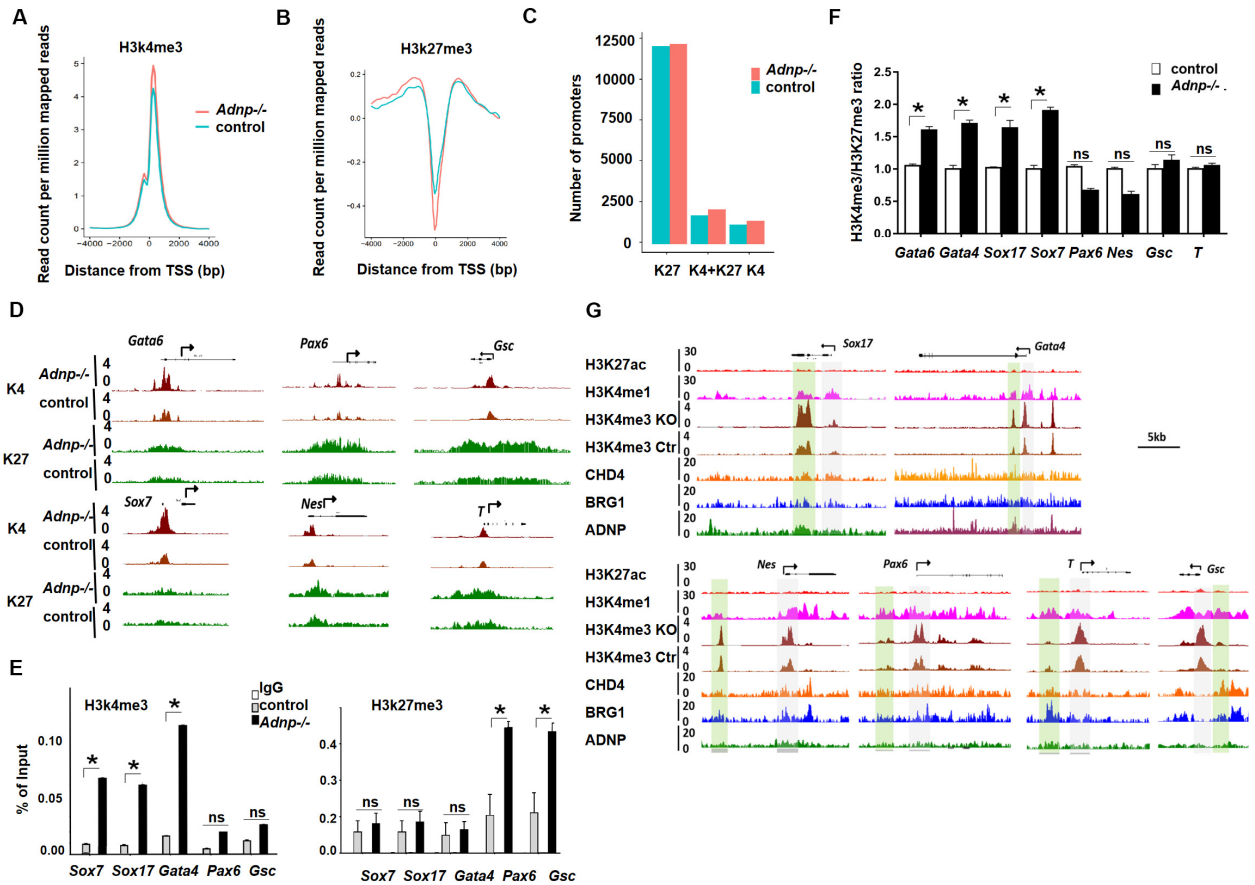

FIGURE 6 | Loss of ADNP caused histone modification change genome-wide. **(A)** A metaplot analysis of H3K4me3 occupancy at the TSS regions of all genes in control and *Adnp-/-* ESCs. **(B)** A metaplot analysis of H3K27me3 occupancy at TSS region of all genes in control and *Adnp-/-* ESCs. **(C)** Change of gene numbers of three gene clusters in the absence of ADNP. **(D)** A genomic snapshot of H3K4me3 and H3K27me3 ChIP-seq peaks at the indicating loci. **(E)** ChIP-PCR assay of H3K4me3 and H3K27me3 at the indicated gene promoters. **(F)** A quantitation of H3K4me3/H3K27me3 ratio at indicated gene promoters based on two replicates of H3K4me3 and H3K27me3 ChIP-seq data. **(G)** Genome browser view at the lineage specifying genes in the control and *Adnp-/-* ESCs. H3K4me1 ChIP signal was used for showing the poised enhancers. H3K27ac ChIP signal was used for the active enhancers. Gray color highlighting TSS-proximal regions, and green color highlighting enhancer regions. Note, H3K4me3 levels were substantially elevated at enhancer region of PrE but not for mesodermal and ectodermal genes. H3K4me3 and H3K27me3 ChIP-seq were repeated two times. Differences in means were statistically significant when $p < 0.05$. Significant levels are: $^{*}p < 0.05$.

(Ku et al., 2008). We investigated whether ADNP depletion affected MLL2 or EZH2 binding at gene promoters by performing MLL2 or EZH2 ChIP-PCR experiments. We found that MLL2 levels at *Sox7*, *Gata4* and *Gata6* promoters were significantly elevated in *Adnp-/-* ESCs compared with control ESCs, while EZH2 enrichment was significantly enhanced at *Nestin* and *Pax6* promoters (**Supplementary Figures 6C,D**). Consistently, RNA polymerase II (Pol II) was significantly elevated at *Sox17* and *Gata6* but not at *Gsc* and *Pax6* genes (**Supplementary Figure 6E**).

DISCUSSION

In this work, we show that ADNP functions as an important chromatin regulator or genome organizer by association with two distinct chromatin regulators, BRG1 and CHD4. ADNP, BRG1 and CHD4 are extensively co-localized genome-wide and they cooperatively control chromatin accessibility and nucleosome configuration. Loss of ADNP expression leads to significant change of nucleosome landscape, bivalent histone modifications and enhancer activities of PrE genes, resulting in de-repression of these genes and priming ESCs differentiation into endodermal cell types.

While this work was ongoing, Ostapcuk et al. (2018) reported a similar study showing that ADNP controls lineage-specifying genes by forming complex with HP1 and CHD4. In their work, it appeared that loss of ADNP had immediate effects on ES cell phenotype. This was demonstrated by a grossly abnormal ESC morphology, reduced alkaline phosphatase activities, deregulation of lineage-specifying genes and reduced expression of pluripotency genes of *Adnp-/-* ESCs. However, in our hands, acute ADNP depletion in ESCs does not result in sudden and complete loss of self-renewal, and our *Adnp-/-* ESCs exhibit a milder phenotype compared to the counterpart ESCs described by Ostapcuk et al. First, our newly established *Adnp-/-* ESCs exhibited an ESC-like morphology and strong

alkaline phosphatase activities. They could be passaged for many generations in the LIF/KSR medium. Only prolonged depletion of ADNP resulted in loss of ESC phenotype. Second, the RNA-sequencing analysis showed that the expression of pluripotency-related genes was barely changed in the newly established *Adnp*-/- ESCs. Third, the lineage-specifying genes were deregulated to a much lower extent when compared to that by Ostapcuk. For instance, the expression of *Igfbp4* and *Gsc* was not changed in the absence of ADNP in our study. The up-regulation of PrE genes was within 2–3-fold range, while this was over 5 times more in mutant cells by Ostapcuk. We think that the discrepancy could be due to the nature of the *Adnp* mutant alleles that were generated by Ostapcuk et al. and our group. In Ostapcuk's work, a very large fragment of the *Adnp* gene (including exons 3 and 4, most of exon 5 as well as introns 3 and 4) was deleted. By looking up the UCSC genome browser, it is likely that there are putative enhancers in the deleted region of the *Adnp* gene from Ostapcuk et al. (2018) in ESCs, but this requires validation. In our work, only 4 or 5 bp deletion in exon 4 of *Adnp* gene was introduced, which should only disrupt *Adnp* gene function. Importantly, our rescue experiments showed that FLAG-tagged ADNP could largely restore the phenotypes of *Adnp*-/- ESCs which was not reported by Ostapcuk et al. (2018). As our CRISPR-Cas9 mediated base pair deletions resemble the human *ADNP* mutation (ID number 64, c.190dupA) in patients with HVDAS syndrome (Van Dijck et al., 2019), this work may help to explain the pleiotropic phenotypes (other than neurodevelopmental defects) observed in patients.

In this work, we reported that ADNP could form a ABC triplex with BRG1 and CHD4 in ESCs. It has been shown that ADNP forms a stable ChAHP triplex with CHD4 and HP1 (Ostapcuk et al., 2018). In addition, ADNP was shown to associate with components of the SWI/SNF complex in HEK293 cells (Mandel and Gozes, 2007). Thus, it appears that ADNP could form different complexes with certain factors depending on its cellular functions. We propose that the ChAHP and ABC triplexes in ESCs are not exclusive, and that ADNP may control local genome structure or chromatin accessibility by recruiting different chromatin remodelers or regulators.

An interesting observation is that loss of ADNP leads to significant up-regulation of a panel of PrE genes, indicating that ADNP is required to robustly repress PrE genes in an undifferentiated ESCs. It was known that the balance between SOX17/GATA6 and NANOG/OCT4 maintains ESC in undifferentiated state (Niakan et al., 2010; Wamaitha et al., 2015). Although ADNP binds to pluripotency genes such as *Nanog* and *Pou5f1*, its loss had little effect on the expression of these genes. Thus, ADNP cannot regulate the expression of PrE genes through repressing *Nanog*. Signaling pathways, such as FGF/Erk signaling, play key role in the expression of PrE genes (Chappell et al., 2013). However, the expression of key components of the signaling pathways was barely altered in the absence of ADNP. Based on these observations, we propose that ADNP contributes to gene expression primarily by regulating local chromatin structure. Several lines of evidence supports the notion. First, ADNP is known to maintain proper local chromatin architecture in ESCs (Kaaij et al., 2019). Loss of ADNP may directly or indirectly alter promoter-enhancer interaction frequencies and affect gene expression. Second, ADNP is important for the proper bivalent histone modifications at developmental gene promoters. An increased ratio of H3K4me3/H3K27me3 at key PrE gene promoters was observed in the absence of ADNP. Third, loss of ADNP leads to prominent enhancer over-activation of key PrE genes by increasing H3K4me3. Fourth, ADNP regulates nucleosome configuration genome-wide. In the absence of ADNP, nucleosome positioning, phasing and occupancy were all changed at a greater extent in PrE than mesodermal and neuroectodermal genes. A recent study showed that the proper nucleosome landscape plays an important role in the control of gene expression (King et al., 2019). Taken together, we propose that loss of ADNP leads to both enhancer over-activation and increased ratio of H3K4me3/H3K27me3 at gene promoters of PrE genes which may explain why *Adnp*-/- ESCs exhibited significant up-regulation of PrE genes.

Another intriguing observation is that although the majority of ADNP bound sites are associated with protein-coding genes, most ADNP ChIP-seq signals are not found at promoter-proximal regions (Kaaij et al., 2019). In addition, many genes bound by ADNP are not deregulated in the absence of ADNP. It appears that gene expression changes with the loss of ADNP are not predicted by ADNP binding. This further implies that ADNP controls gene expression by controlling local chromatin architecture, which is likely mediated by BRG1, CHD4 and CTCF (Lei et al., 2015; O'Shaughnessy-Kirwan et al., 2015; Kaaij et al., 2019). Thus, we propose that ADNP may control gene expression by binding to gene regulatory regions (as a transcription factor) and by association with BRG1, CHD4 and CTCF (as a genome organizer).

A previous study has shown that BRG1 and CHD4 co-occupy distal sites corresponding to increased ESC master TF binding, and that co-dependency of BRG1 and CHD7 exists to support pluripotency network in ESCs (Yang et al., 2017). This study suggested that concerted activities of multiple chromatin remodelers are utilized to support ES cell pluripotency. To our knowledge, whether and how distinct chromatin remodelers work cooperatively to modulate chromatin architecture to regulate lineage-specifying genes is not clear. In this work, we showed that a co-dependency of SWI/SNF-BRG1 and CHD4 may underlie for robust chromatin regulation for developmental genes. Thus, our work extends previous studies by showing that chromatin remodelers are cooperatively used not only for supporting core pluripotency genes, but also for silencing developmental genes while keeping them poised for activation.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the CRA002148.

AUTHOR CONTRIBUTIONS

XS: collection of data and molecular biology experiments. WY: collection of data and bioinformatics

analysis. LL: bioinformatics analysis and data interpretation. YS: conception and design, collection of data, manuscript writing, financial support, and final approval of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Key Research and Development Program (#2016YFA0101100), National Natural Science Foundation of China (#31671526), and Hundred-Talent Programs (from CAS and IHB) (#Y623041501) to YS. LL was supported by NSFC (#31771430) and Huazhong Agricultural University Scientific & Technological Self-innovation Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.00553/full#supplementary-material

FIGURE S1 | (A) shRNA knockdown of *Adnp* in ESCs. **(B)** Volcano plot showing the number of up- and down-regulated DEGs. **(C)** WB showing the FLAG-ADNP levels in *Adnp-/-* ESCs. All data were based on two experimental repeats.

FIGURE S2 | (A) Synthesized FLAG-tagged ADNP-N failed to pull down MYC-tagged CHD4-C. **(B)** HA-BRG1-N pulled down full-length ADNP and FLAG-ADNP-C but not FLAG-ADNP-N in 293T cells. **(C)** WB showing that in the presence of Tet-Express protein, 3 × FLAG-tagged ADNP could be induced in *Adnp-/-* ESCs.

FIGURE S3 | (A) A snapshot of ChIP-seq genome-browser view of ADNP, BRG1 and CHD4 occupancy around the *Vamp5* and *Nodal* loci, showing the co-localization of all three factors in green. **(B)** A snapshot of ChIP-seq genome-browser view of ADNP occupancy around the *Gata4/Gata6/Sox17* loci. **(C)** Enrichment of ADNP at the indicated gene promoters by ChIP-PCR assay using ADNP antibodies (IgG as negative control) based on two repeats. ns: no significance. **(D)** A piano plot showing the expression change of the indicated gene clusters in the absence of ADNP, based on two replicates of RNA-seq data.

FIGURE S4 | (A) ChIP-seq genome browser view of ADNP ChIP-seq and ATAC-seq signals at part of chromosome 5 in control and *Adnp-/-* ESCs. Note the widespread increased ATAC signals in the absence of ADNP. **(B)** A example of ChIP-seq genome browser view of ADNP ChIP-seq and ATAC-seq signal at the *Dnahc8* locus. Gray: ATAC hypersensitive signal peaks were co-localized with ADNP ChIP-seq peaks. **(C)** Nucleosome configuration at promoter of the indicated pluripotency genes was not significantly altered in the absence of ADNP.

FIGURE S5 | (A) A snapshot of ATAC-seq signal change at the indicating loci in the absence of each factor. ATAC-seq experiments were repeated two times for control and *Adnp-/-* ESCs. ATAC-Seq data for BRG1 and CHD4 were downloaded as described in the text.

FIGURE S6 | (A) A metaplot analysis of H3K4me3 occupancy at TSS region of H3K4me3 only, H3K27me3 only and bivalent genes in control and *Adnp-/-* ESCs. **(B)** A metaplot analysis of H3K27me3 occupancy at TSS region of H3K4me3 only, H3K27me3 only and bivalent genes in control and *Adnp-/-* ESCs. The results were based on two replicates of H3K4me3 and H3K27me3 CHIP-seq experiments. **(C)** MLL2 enrichment at the indicated genes by ChIP-PCR assay. **(D)** EZH2 enrichment at the indicated genes by ChIP-PCR assay. **(E)** Pol II enrichment at indicated genes by ChIP-PCR analysis. All data were based on three repeat experiments. Differences in means were statistically significant when $p < 0.05$. Significant levels are: $^{*}p < 0.05$; $^{**}P < 0.01$; $^{***}p < 0.001$.

REFERENCES

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? *Mol. Cell.* 49, 825–837. doi: 10.1016/j.molcel.2013.01.038

Chappell, J., Sun, Y., Singh, A., and Dalton, S. (2013). MYC/MAX control ERK signaling and pluripotency by regulation of dual-specificity phosphatases 2 and 7. *Gene. Dev.* 27, 725–733. doi: 10.1101/gad.211300.112

Chen, T. P., and Dent, S. (2014). Chromatin modifiers: regulators of cellular differentiation. *Nat. Rev. Genet.* 15, 93–106. doi: 10.1038/nrg3607

Clapier, C. R., Iwasa, J., Cairns, B. R., and Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell. Biol.* 18, 407–422. doi: 10.1038/nrm.2017.26

De Gobbi, M., Garrick, D., Lynch, M., Vernimmen, D., Hughes, J. R., Goardon, N., et al. (2011). Generation of bivalent chromatin domains during cell fate decisions. *Epigenet. Chromatin* 4, 9–15.

Dieuleveult, M., Yen, K. Y., Hmitou, I., Depaux, A., Boussouar, F., Dargham, D. B., et al. (2016). Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* 530, 113–116. doi: 10.1038/nature16505

Gao, X. L., Tate, P., Hu, P., Tjian, R., Skarnes, W. C., and Wang, Z. (2008). ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6656–6661. doi: 10.1073/pnas.0801802105

Gifford, C. A., Ziller, M. J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., et al. (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* 153, 1149–1163.

Gozes, I., Yeheskel, A., and Pasmanik-Chor, M. (2015). Activity-dependent neuroprotective protein (ADNP): a case study for highly conserved chordata-specific genes shaping the brain and mutated in cancer. *J. of Alzheimer's Dis.* 45, 57–73. doi: 10.3233/jad-142490

Helsmoortel, C., Vultovan Silfhout, A. T., Coe, B. P., Vandeweyer, G., Rooms, L., van den Ende, J., et al. (2014). A SWI/SNF-related autism syndrome caused by de novo mutations in ADNP. *Nat. Genet.* 46, 380–384. doi: 10.1038/ng.2899

Ho, L., Ronan, J., Wu, J., Staahl, B. T., Chen, L., Kuo, A., et al. (2008). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for ESC self-renewal and pluripotency. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5181–5186. doi: 10.1073/pnas.0812889106

Kaaij, L., Mohn, F., van der Weide, R., de Wit, E., and Bühler, M. (2019). The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse. *Cell* 178, 1437–1451.

Kidder, B., Palmer, S., and Knott, J. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. *Stem. Cells* 27, 317–328. doi: 10.1634/stemcells.2008-0710

King, H. W., Fursova, N. A., Blackledge, N. P., and Klose, R. J. (2019). Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes. *Genome Res.* 28, 1494–1507. doi: 10.1101/gr.237180.118

King, H. W., and Klose, R. J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* 6:e22631. doi: 10.7554/eLife.22631

Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PloS Genet.* 4:e1000242. doi: 10.1371/journal.pgen.1000242

Lei, L., West, J., Yan, Z. J., Gao, X., Fang, P., Dennis, J. H., et al. (2015). BAF250a protein regulates nucleosome occupancy and histone modifications in priming embryonic stem cell differentiation. *J. Biol. Chem.* 290, 19343–19352. doi: 10.1074/jbc.m115.637389

Lu, P., and Roberts, C. W. (2013). The SWI /SNF tumor suppressor complex regulation of promoter nucleosomes and beyond. *Nucleus* 4, 374–378. doi: 10.4161/nucl.26654

Mandel, S., and Gozes, I. (2007). Activity-dependent neuroprotective protein constitutes a novel element in the SWI/SNF chromatin remodeling complex. *J. Biol. Chem.* 282, 34448–34456. doi: 10.1074/jbc.m704756200

Mosch, K., Franz, H., Soeroes, S., Singh, P. B., and Fischle, W. (2011). HP1 recruits activity-dependent neuroprotective protein to H3K9me3 marked pericentromeric heterochromatin for silencing of major satellite repeats. *PLoS One* 6:e15894. doi: 10.1371/journal.pone.0015894

Musselman, C. A., Ramírez, J., and Sims, J. K. (2012). Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. *Proc. Natl. Acad. Sci. U.S.A.* 109, 787–792.

Niakan, K. K., Ji, H. K., Maehr, R., Vokes, S. A., Rodolfa, K. T., Sherwood, R. I., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Gene. Dev.* 24, 312–326. doi: 10.1101/gad.1833510

O'Shaughnessy-Kirwan, A., Signolet, J., Costello, I., Gharbi, S., and Hendrich, B. (2015). Constraint of gene expression by the chromatin remodelling protein CHD4 facilitates lineage specification. *Development* 142, 2586–2597. doi: 10.1242/dev.125450

Ostapcuk, V., Mohn, F., Carl, S., Basters, A., Hess, D., Iesmantavicius, V., et al. (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* 557, 739–743. doi: 10.1038/s41586-018-0153-8

Phillips-Cremins, J. E., Sauria, M., Sanyal, A., Gerasimova, T. I., Lajoie, B. R., Bell, J. S. K., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1–15.

Pinhasov, A., Mandel, S., Torchinsky, A., Giladi, E., Pittel, Z., Goldsweig, A. M., et al. (2003). Activity-dependent neuroprotective protein: a novel gene essential for brain formation. *Brain. Res. Dev. Brain. Res.* 144, 83–90. doi: 10.1016/s0165-3806(03)00162-7

Shen, W., Xu, R., Guo, B., Rong, B., Gu, L., Wang, Z., et al. (2016). Suppression of enhancer overactivation by a RACK-Histone demethylase complex. *Cell* 165, 331–342. doi: 10.1016/j.cell.2016.02.064

Shimono, Y., Murakami, H., Kawai, K., Wade, P. A., Shimokata, K., and Takahashi, M. (2003). Mi-2 beta associates with BRG1 and RET finger protein at the distinct regions with transcriptional activating and repressing abilities. *J. Biol. Chem.* 278, 51638–51645. doi: 10.1074/jbc.m309198200

Singh, A., Sun, Y. H., Li, L., Zhang, W., Wu, T., Zhao, S., et al. (2015). Cell-cycle control of bivalent epigenetic domains regulates the exit from pluripotency. *Stem. Cell Rep.* 3, 323–330.

Singh, A. P., Foley, J. F., Rubino, M., Boyle, M. C., Tandon, A., Shah, R., et al. (2016). Brg1 enables rapid growth of the early embryo by suppressing genes that regulate apoptosis and cell growth arrest. *Mol. Cell. Biol.* 36, 1990–2010. doi: 10.1128/mcb.01101-15

Sun, X. Y., Chen, J., Zhang, Y. Y., Munisha, M., Dougan, S., and Sun, Y. H. (2018). Mga modulates bmpr1a activity by antagonizing bs69 in zebrafish. *Front. Cell Dev. Biol.* 6:126. doi: 10.3389/fcell.2018.00126

Tolstorukov, M. Y., Sansam, C. G., Lu, P., Koellhoffer, E. C., Helming, K. C., Alver, B. H., et al. (2013). Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10165–10170. doi: 10.1073/pnas.1302209110

Van Dijck, A., Vulto-van Silfhout, A. T., Cappuyns, E., van der Werf, I. M., Mancini, G. M., Tzschach, A., et al. (2019). Clinical presentation of a complex neurodevelopmental disorder caused by mutations in ADNP. *Biol. Psych.* 85, 287–297.

Vandeweyer, G., Helsmoortel, C., Van Dijck, A., Vulto-van Silfhout, A. T., Coe, B. P., Bernier, R., et al. (2014). The transcriptional regulator ADNP links the BAF (SWI/SNF) complexes with Autism. *Am. J. Med. Genet. C Semin. Med. Genet.* 166, 315–326.

Wamaitha, S. E., del Valle, I., Cho, L. T., Wei, Y., Fogarty, N. M. E., Blakeley, B., et al. (2015). Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Gene. Dev.* 29, 1239–1255. doi: 10.1101/gad.257071.114

Wang, X., Lee, R. S., Alver, B. H., Haswell, J. R., Wang, S., Mieczkowski, J., et al. (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet.* 49, 289–295. doi: 10.1038/ng.3746

Yang, P. Y., Oldfield, A., Kim, T. Y., Yang, A., Yang, J. Y. H., Ho, J. W. K., et al. (2017). Integrative analysis identifies co-dependent gene expression regulation of BRG1 and CHD7 at distal regulatory sites in embryonic stem cells. *Bioinformatics* 33, 1916–1920. doi: 10.1093/bioinformatics/btx092

Zhang, X., Li, B., Li, W., Ma, L., Zheng, D., Li, L., et al. (2014). Transcriptional repression by the BRG1-SWI/SNF complex affects the pluripotency of human embryonic stem cells. *Stem. Cell Rep.* 3, 460–474. doi: 10.1016/j.stemcr.2014.07.004

Zhao, H., Han, Z., Liu, X., Gu, J., Tang, F., Wei, G., et al. (2017). Chromatin remodeler Chd4 Represses aberrant expression of Tbx3 and sustains self-renewal of embryonic stem cells. *J. Biol. Chem.* 292, 8507–8519.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 11 August 2020
doi: 10.3389/fgene.2020.00885

Neurodevelopmental Disorders Caused by Defective Chromatin Remodeling: Phenotypic Complexity Is Highlighted by a Review of ATRX Function

Sara Timpano[1] *and David J. Picketts*[1,2,3,4,5]*

[1] Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada, [2] Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, ON, Canada, [3] Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada, [4] Department of Medicine, University of Ottawa, Ottawa, ON, Canada, [5] University of Ottawa Brain and Mind Research Institute, Ottawa, ON, Canada

The ability to determine the genetic etiology of intellectual disability (ID) and neurodevelopmental disorders (NDD) has improved immensely over the last decade. One prevailing metric from these studies is the large percentage of genes encoding epigenetic regulators, including many members of the ATP-dependent chromatin remodeling enzyme family. Chromatin remodeling proteins can be subdivided into five classes that include SWI/SNF, ISWI, CHD, INO80, and ATRX. These proteins utilize the energy from ATP hydrolysis to alter nucleosome positioning and are implicated in many cellular processes. As such, defining their precise roles and contributions to brain development and disease pathogenesis has proven to be complex. In this review, we illustrate that complexity by reviewing the roles of ATRX on genome stability, replication, and transcriptional regulation and how these mechanisms provide key insight into the phenotype of ATR-X patients.

Keywords: intellectual disability, neurodevelopmental disorder, *ATRX*, ATR-X syndrome, chromatin remodeling, SWI/SNF

OPEN ACCESS

Edited by:
Mojgan Rastegar,
University of Manitoba, Canada

Reviewed by:
Craig Peterson,
University of Massachusetts Medical School, United States
Tom Moss,
Laval University, Canada

***Correspondence:**
David J. Picketts
dpicketts@ohri.ca

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *07 June 2020*
Accepted: *20 July 2020*
Published: *11 August 2020*

Citation:
Timpano S and Picketts DJ (2020) Neurodevelopmental Disorders Caused by Defective Chromatin Remodeling: Phenotypic Complexity Is Highlighted by a Review of ATRX Function. Front. Genet. 11:885. doi: 10.3389/fgene.2020.00885

INTRODUCTION

Neurodevelopmental disorders (NDD) are highly complex and heterogeneous conditions that have a global prevalence of approximately 2–3% of the population. Despite being aware of these conditions for over a century, it is only within the last decade that the development of exome and whole genome sequencing has dramatically enhanced the discovery of the underlying causes of these disorders. Indeed, the SysID database[1] list 1,334 genes (updated March 26, 2020) that contribute to intellectual disability (ID) (Kochinke et al., 2016), while approximately 100 genes are associated with autism spectrum disorder (ASD) (Satterstrom et al., 2020). Interestingly, a substantial proportion of NDD causing genes are involved in chromatin and/or transcriptional regulation including the broad family of ATP-dependent chromatin remodelers.

Chromatin remodelers utilize energy from ATP hydrolysis to alter nucleosome spacing/density or to facilitate histone variant exchange (Bowman and Poirier, 2015). There are four main families of ATP-dependent chromatin remodelers characterized by their conserved ATPase domain of the helicase II superfamily (**Figure 1**). These families are divided into the (1) SWI/SNF group,

[1] https://sysid.cmbi.umcn.nl/

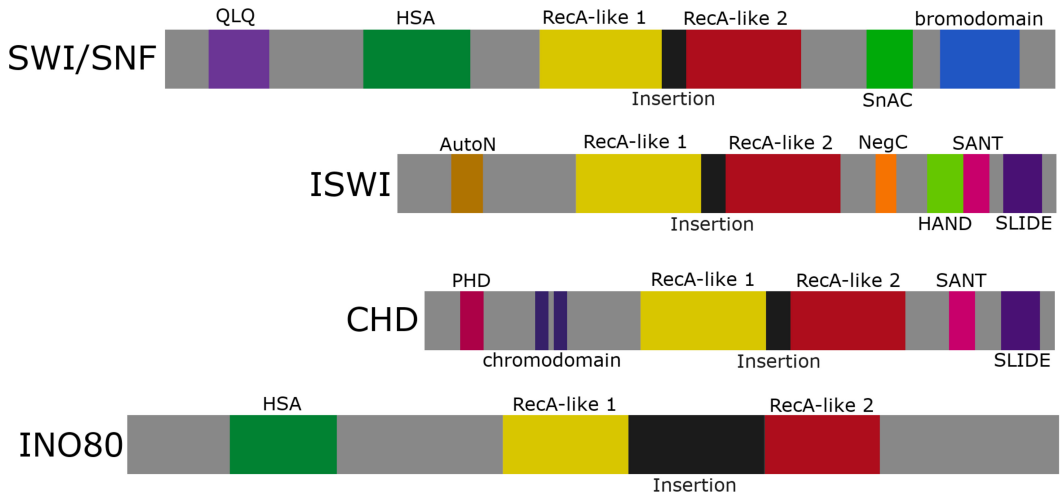

FIGURE 1 | The ATP-dependent chromatin remodeling family. Representation of the four chromatin remodeling groups: SWI/SNF, ISWI, CHD, and INO80. Each group contains an ATPase domain subdivided into RecA-like lobes 1 and 2 separated by a variable linker region (labeled insertion). SWI/SNF and INO80 share an HSA domain, while ISWI and CHD share a SANT and SLIDE domain.

large complexes made up of ∼15 subunits, (2) ISWI group, heterodimers and four subunit complexes, (3) CHD group, complexes that incorporate up to ∼10 subunits, and (4) INO80 group, ∼15 subunit complexes. In addition, the focus of this review is ATRX which represents one of several orphan families that have been less studied mechanistically. In addition to the ATPase domain that is subdivided into two RecA-like lobes, these chromatin remodeling enzymes are characterized by additional motifs that facilitate protein–protein interactions (e.g., HSA and QLQ domains), DNA interactions (e.g., HAND and SLIDE domains), and chromatin interactions (e.g., SANT, chromodomain, and bromodomain) (**Figure 1**). The SWI/SNF and INO80 family primarily promote transcription and DNA repair by sliding/ejecting nucleosomes (SWI/SNF/BRG1, BRM) or depositing histone variants (INO80/SRCAP). The ISWI and CHD family primarily mediate nucleosome maturation and spacing to promote chromatin formation post-replication, highly structured chromatin (ISWI), or transcriptional repression (CHD) (Clapier et al., 2017).

Mutations in these enzyme families results in aberrant gene expression that impinges on many cellular activities including DNA replication, DNA repair, as well as cell proliferation and differentiation. As indicated above, mutations in many of these family members lead to a wide range of NDD and symptoms (**Table 1**) with some of the more well-studied disorders being Coffin-Siris syndrome (CSS), Nicolaides-Baraitser syndrome (NCBS), CHARGE syndrome, and ATR-X syndrome. Moreover, it is becoming clear that mutations in multiple components of these remodeling complexes cause ID (**Table 1**) and can contribute to a spectrum of clinical phenotypes that is best illustrated by mutations in the SWI/SNF interacting partners (Bögershausen and Wollnik, 2018; van der Sluijs et al., 2019). The reader is referred to a number of recent reviews for detailed information of these different remodeler classes (Hota and Bruneau, 2016; Sokpor et al., 2017; Goodwin and Picketts, 2018; Alfert et al., 2019; Hoffmann and Spengler, 2019).

Here, we will review recent studies on ATRX to highlight the multiple biochemical functions chromatin remodeling proteins participate in, and the diverse set of mechanisms that, collectively, contribute to the complexity underlying the pathogenesis of NDD.

MOLECULAR GENETICS OF THE ATR-X SYNDROME

The ATR-X syndrome is a rare human congenital disorder with a wide range of symptoms that primarily affects males. Over 200 cases have been identified worldwide with and an estimated prevalence of <1–9/1,000,000 (Gibbons, 2006). Affected individuals display cognitive impairment typically described as severe ID, and many are non-verbal, capable of speaking or signing only a few words (Saugier-Veber et al., 1995; Guerrini et al., 2000). Originally, the presence of alpha thalassemia was used as a diagnostic tool to identify affected individuals, but there is variability in the hematological symptoms (Gibbons et al., 1995). The majority of patients are affected with microcephaly and skeletal malformations (Holmes and Gang, 1984; Carpenter et al., 1999). Muscle development is also impaired in most, leading to delayed motor development and hypotonia, while approximately one third of patients experience seizures (Lossi et al., 1999).

Although ATR-X syndrome patients present with a heterogeneous phenotype, the disease is caused by mutations in a single gene, the *ATRX* locus, which spans over 300 kbp on chromosome Xq13.3-21.1 (Gibbons et al., 1995, 2008; Picketts et al., 1996). The *ATRX* gene encodes two major trasncripts (**Figure 2**), one encoding the full length protein and a truncated isoform generated by an alternative splicing event that retains

TABLE 1 | ATP-dependent chromatin remodelers are a frequent cause of NDDs. List of NDD implicated genes which are incorporated into ATP-dependent chromatin remodeling complexes.

Family	Gene	Associated disease	References
SWI /SNF	ARID1A	ID, CSS	Tsurusaki et al. (2012); Kosho et al. (2013)
	ARID1B	ID, CSS, ASD	Hoyer et al. (2012); Santen et al. (2012); De Rubeis et al. (2014)
	ARID2	ID, CSS-like, NCBS-like	Shang et al. (2015); Bramswig et al. (2017)
	DPF2	ID, CSS-like	Vasileiou et al. (2018)
	PBRM	ASD	O'Roak et al. (2012b)
	SMARCA2	NCBS	Van Houdt et al. (2012); Tsurusaki et al. (2014b)
	SMARCB1	ID, CSS, Kleefstra	Kleefstra et al. (2012); Santen et al. (2013); Wieczorek et al. (2013)
	SMARCC1	ASD	Hormozdiari et al. (2015)
	SMARCC2	ID, ASD	Machol et al. (2019)
	SMARCE1	CSS-like	Tsurusaki et al. (2012); Kosho et al. (2013); Wieczorek et al. (2013)
	SMARCA4	CSS	Santen et al. (2013); Tsurusaki et al. (2014b)
	SOX11	ID, CSS-like	Tsurusaki et al. (2014a)
ISWI	BAZ1A	ID	Zaghlool et al. (2016)
	BAZ1B	Williams-Beuren syndrome	Lu et al. (1998); Peoples et al. (1998)
	BAZ2B	ID, ASD	Scott et al. (2020)
	BPTF	ID	Stankiewicz et al. (2017)
	SMARCA1	CSS-like, Rett syndrome-like, schizophrenia	Karaca et al. (2015); Homann et al. (2016); Lopes et al. (2016)
CHD	CHD2	Lennox-Gastaut syndrome, Doose Syndrome, Epileptic encephalopathy	Carvill et al. (2013)
	CHD3	Macrocephaly, ID, impaired speech/language	Snijders Blok et al. (2018)
	CHD4	Sifrim–Hitz–Weiss syndrome	Weiss et al. (2020)
	CHD5	ASD-like	Pisansky et al. (2017)
	CHD7	CHARGE syndrome, ASD	Vissers et al. (2004); Sanlaville et al. (2006); O'Roak et al. (2012b)
	CHD8	ASD	O'Roak et al. (2012a,b); De Rubeis et al. (2014); Neale et al. (2016)
INO80	INO80	Microcephaly, ID	Alazami et al. (2015)
	SRCAP	Floating-harbor syndrome	Hood et al. (2012)
	YY1AP1	ID	Guo et al. (2017)
ATRX	ATRX	ATR-X syndrome	Gibbons et al. (1995)

intron 11 and terminates translation prematurely (Garrick et al., 2004; Mitson et al., 2011). The full length transcript encodes a protein of 285 kDa in size while the shorter transcript generates a smaller truncated protein that is 180 kDa and lacks the ATP-dependent remodeling domain (Picketts et al., 1996; Garrick et al., 2004).

The *N*-terminus of the ATRX protein houses several motifs critical for its interaction with chromatin, including a heterochromatin protein 1 (HP1α) binding motif (PxVxL) (Lechner et al., 2005) and enhancer of zeste homolog 2 (EZH2) interaction domain (Cardoso et al., 1998), and the ATRX-DNMT3-DNMT3L (ADD) domain (Picketts et al., 1998; Xie et al., 1999). The ADD domain comprises a GATA-like zinc finger and a plant homeodomain (PHD)-like finger that targets the dual histone post translational modification (PTM), H3K9me3 and H3K4me0 (Argentaro et al., 2007; Dhayalan et al., 2011; Eustermann et al., 2011; Iwase et al., 2011). A region within the center of the polypeptide mediates death domain associated protein (DAXX) binding (Xue et al., 2003). Toward the *C*-terminus lies the highly conserved RecA-like lobes 1 and 2 that together are required for ATPase activity (Picketts et al., 1996), as well as mapped regions for interactions with the methyl-CpG-binding protein (MeCP2) (Nan et al., 2007) and the promyelocytic leukemia protein (PML) (Bérubé et al., 2008) (**Figure 2**).

The majority of ATR-X syndrome causing mutations are missense mutations mapping within the ADD (50%) and SNF2-like/helicase domains (30%) (Argentaro et al., 2007; Gibbons et al., 2008). To date, there has been a lack of genotype: phenotype correlations identified, although mutations within the ADD domain typically produce more severe psychomotor phenotypes compared to mutations in the SNF2-like/helicase domain (Badens et al., 2006).

It should also be noted that somatic mutations in the *ATRX* gene have been identified in a wide range of cancers that include pancreatic neuroendocrine tumors, gliomas, neuroblastomas, and sarcomas, which will not be discussed here but have been the focus of recent reviews (Watson et al., 2015; Dyer et al., 2017).

INTERACTING PARTNERS AND BIOCHEMICAL FUNCTIONS

All functional studies indicate that ATRX is a heterochromatin interacting protein. It localizes to pericentromeric heterochromatin, telomeres, PML nuclear bodies, and physically interacts with the HP1 family (McDowell et al., 1999; Berube et al., 2000; Tang et al., 2004). Later work demonstrated that ATRX could be recruited to the heterochromatin histone mark, H3K9me3, either indirectly by its interaction with HP1 or recruitment by MeCP2, and directly by binding of the ADD domain to H3K9me3 that lies adjacent to unmethylated H3K4 (Berube et al., 2000; Bannister et al., 2001; Nan et al., 2007; Eustermann et al., 2011; Iwase et al., 2011). ATRX and DAXX were identified as interacting partners by two separate groups, one using ATRX co-IP experiments and the other a Flag-DAXX pull-down approach (Xue et al., 2003; Tang et al., 2004). Further characterization showed that most of the endogenous ATRX protein is in a 1 MDa complex with DAXX, while DAXX also fractionates in a 700 kDa complex independent of ATRX. Deletion mutants were used to demonstrate that the ATRX/DAXX interaction was mediated through the PAH domain of DAXX and a region between the ADD and

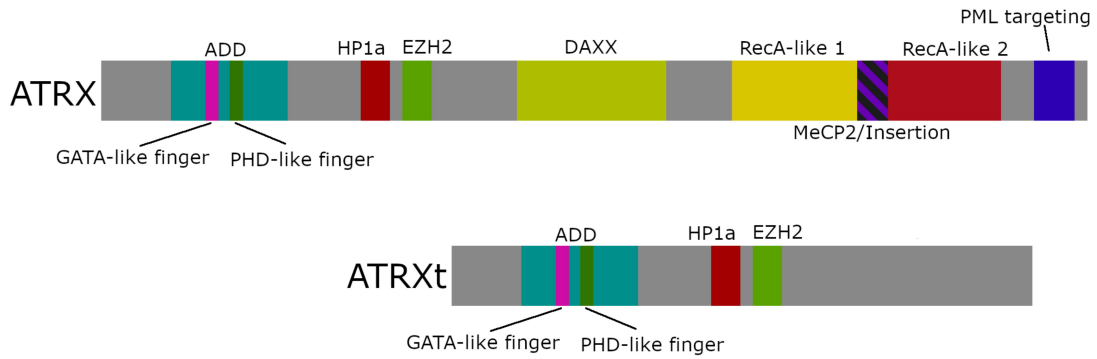

FIGURE 2 | ATRX domain structure. Schematic diagram of full-length ATRX (282 kDa), truncated ATRX (ATRXt; 180 kDa), and locations of the key protein interaction domains. The two isoforms share an ADD domain, a HP1α binding motif, an EZH2 binding motif, while the ADD domain is comprised of a GATA-like zinc finger and a PHD-like zinc finger. The full-length polypeptide also contains a DAXX binding motif, a SNF2-ATPase domain comprising RecA-like lobes 1 and 2 separated by a linker region containing a MeCP2 binding motif (MeCP2/Insertion), and a PML targeting motif.

SNF2 domains within ATRX (Tang et al., 2004). Both the ATRX/DAXX complex and recombinant ATRX protein had DNA or nucleosome stimulated ATPase activity which was impaired by patient mutations that localized to the ATPase domain (Xue et al., 2003; Tang et al., 2004). A mononucleosome disruption assay was used to demonstrate that the ATRX/DAXX complex could alter the DNAse I digestion pattern of the mononucleosome in the presence of ATP. The localization of the altered digestion pattern indicated that ATRX/DAXX disrupts DNA–histone interactions at the entry site of the nucleosome and does not alter nucleosome phasing. In addition, a triple-helix strand displacement assay was used to show that the ATRX/DAXX complex and ATRX alone had a DNA translocase property similar to the RSC and SWI/SNF complexes (Xue et al., 2003). More recent work has indicated that DAXX is an H3.3-specific histone chaperone that functions with ATRX to deposit the histone variant at pericentric and telomeric repeats, while DAXX functions independently of ATRX to repress retrotransposons (Lewis et al., 2010; Hoelper et al., 2017). In this regard, the ATRX/DAXX complex shows some similarities with the ISWI complex ACF and its interactions with the histone chaperone NAP1 (Gemmen et al., 2005; Torigoe et al., 2011). These properties could be used to reconstitute H3.3 containing nucleosomal arrays that might guide future *in vitro* biochemical studies to further define ATRX function during transcription or DNA replication (Peterson, 2009).

Indeed, in a series of papers ATRX, DAXX, and the histone variant H3.3 were shown to co-localize at telomeres where the ATRX/DAXX complex functions as a histone chaperone to deposit H3.3 into telomeric heterochromatin (Wong et al., 2009; Goldberg et al., 2010; Lewis et al., 2010). Further work showed that DAXX functions as the histone chaperone, that H3.3K9me3 deposition occurs in a replication-independent manner by the complex, and both H3.3 loading and heterochromatin organization by ATRX/DAXX is mediated by SUV39H1 and PML in PML-associated heterochromatin domains (Drané et al., 2010; Goldberg et al., 2010; Lewis et al., 2010; He et al., 2015; Udugama et al., 2015; Delbarre et al., 2017). Additionally, ATRX was shown to be critical for the formation of senescence-induced heterochromatin foci (SAHF) that help drive cancer cells into therapy-induced senescence (Kovatcheva et al., 2017). Finally, ATRX has been shown to bind to the *Xist* lncRNA to promote recruitment of the PRC2 repressive complex and facilitate stable heterochromatin formation of the silenced X-chromosome (Sarma et al., 2014). RNA binding remains an understudied role for ATRX, although several reports have shown a range of interactions with multiple lncRNAs including TERRA (telomeric repeat-containing RNA) (Chu et al., 2017; Nguyen et al., 2017), *ChRO1* in muscle (Park et al., 2018), and minor satellite RNAs at centromeric heterochromatin (Ren et al., 2020). These interactions are mediated through a unique *N*-terminal domain in ATRX to regulate differentiation, gene expression, DNA and histone methylation and chromatin compaction (Chu et al., 2017; Nguyen et al., 2017; Park et al., 2018; Ren et al., 2020).

A role for ATRX at heterochromatin was also strengthened by chromatin immunoprecipitation experiments that showed enriched ATRX binding at telomeres and centromeres. Interestingly, ATRX was also enriched at repetitive DNA elements while having a lower frequency of binding within gene bodies (Law et al., 2010). Further characterization showed that ATRX was prevalent at long terminal repeats of endogenous retrovirus sequences of family K (ERVK), at variable number tandem repeats (VNTRs) and at simple tandem repeats (Law et al., 2010). Many of the tandem repeats were GC-rich sequences that are predicted to form G-quadruplex secondary DNA structures (G4 DNA) including the telomeric repeats and some CpG islands. The formation of G4 DNA has been proposed to have important roles in the regulation of gene expression, as well as be prohibitive to DNA replication and transcription (Rhodes and Lipps, 2015; Valton and Prioleau, 2016; Varshney et al., 2020). *In vitro* studies confirmed that ATRX can bind to G4 DNA structures (Law et al., 2010). In addition, ATRX mutations have variable effects on α-globin expression including individuals with the same mutation. Law et al. (2010) demonstrate that one ATRX binding site lies within a GC-rich VNTR sequence 1 kb upstream of the HBM gene. The authors demonstrate a positive

correlation in ATR-X patients such that increasing VNTR repeat size increases the severity of the α-thalassemia as measured by the level of HbH inclusions in red blood cells. Since the sequence is a GC-rich VNTR that is predicted to form G4 quadruplexes, it was inferred that increasing repeat size increases the probability to form G4 DNA that subsequently alters HBM expression.

The ATRX protein was also shown to co-purify with the MRE11-RAD50-NBS1 (MRN) complex, an active player in the processing of DNA double strand breaks (DSB) that suggested ATRX was critical to maintain genome integrity (Leung et al., 2013). Consistent with this finding, ATRX knockdown studies in HeLa cells resulted in defects in mitotic progression and micronuclei formation from altered chromosome condensation and centromeric cohesion (Ritchie et al., 2008). Other studies indicated that ATRX loss impaired replication fork progression during S-phase resulting in telomere fragility, increased DSB, and mitotic catastrophe (Huh et al., 2012, 2016; Leung et al., 2013; Watson et al., 2013).

The ATRX-DAXX-H3.3 complex is critical for this heterochromatic formation and subsequent maintenance (Law et al., 2010; Eid et al., 2015; He et al., 2015; Udugama et al., 2015). H3.3 within telomeric regions is targeted for trimethylation on its K9 residue (He et al., 2015; Udugama et al., 2015). H3.3K9me3 recruits more ATRX-DAXX-H3.3 complexes, which in turn will deposit H3.3, creating a positive feedback loop required for maintaining telomere structure (Udugama et al., 2015). Failure to establish proper structure will reduce telomere integrity and result in an increase of non-coding telomeric transcript expression (He et al., 2015; Udugama et al., 2015).

The eclectic properties of the ATRX protein do not make it intuitively obvious how an aberration of these functions can result in a neurodevelopment disorder with cognitive deficits. In the remaining section, we discuss the characterization of mouse models and the insights they have provided into the pathophysiology of ATR-X patients and, more generally, the complex etiology of NDDs caused by defective epigenetic regulators.

DELINEATING PATHOPHYSIOLOGICAL MECHANISMS OF THE ATR-X SYNDROME

Functional Effects of Patient Mutations and Generation of Animal Models

One of the first questions addressed was do patient mutations affect protein stability and function? Immunoblots of extracts from patient-derived EBV-transformed B-lymphocytes showed significantly reduced levels of ATRX protein from all patients tested (McDowell et al., 1999; Cardoso et al., 2000). Interestingly, in patients with early premature stop codons (e.g., p.Arg37X), translation was initiated from an internal methionine that produced a smaller truncated protein at ∼30% levels leading to a milder phenotype (Howard et al., 2004; Abidi et al., 2005; Basehore et al., 2015). Utilizing recombinant proteins, other studies demonstrated that mutations within the ATPase domain attenuated ATPase activity but did not reduce it, while mutations in the ADD domain or the PML-targeting domain reduced localization to chromocenters and PML nuclear bodies, respectively (Cardoso et al., 2000; Bérubé et al., 2008). *Atrx*-null mutations in mice show defective extraembryonic trophoblast development and die embryonically at ∼E9.5 (Garrick et al., 2006). Collectively, these studies indicate that ATR-X syndrome causing mutations are functional hypomorphs, while more severe mutations are not found and are presumably non-viable.

Several different ATRX-deficient mouse lines have been generated and used for functional characterization. The most widely used model is a floxed allele in which loxP sites flanked exon 18 which encodes the ATP-binding pocket (Bérubé et al., 2005; Garrick et al., 2006). These animals have been crossed with several different tissue-specific Cre driver lines to inactivate ATRX in skeletal muscle progenitors (Huh et al., 2012), Sertoli cells (Bagheri-fam et al., 2011), osteobalsts (Solomon et al., 2013), chondrocytres (Solomon et al., 2009), the retina (Medina et al., 2009; Lagali et al., 2016), and the developing forebrain (Bérubé et al., 2005) among others. A second transgenic line ($Atrx^{\Delta E2}$) was developed by deleting exon 2 and replacing it with a SA-IRES-β-geo cassette (Nogami et al., 2011; Shioda et al., 2011). This mutation was meant to mimic the p.Arg37X mutation and make an *N*-terminally truncated ATRX protein by initiating translation from an internal methionine codon (Howard et al., 2004; Abidi et al., 2005). Both of these models will be discussed in more detail in the following sections. Finally, an overexpression transgenic line was created with the *ATRX* cDNA under control of a CMV enhancer/β-actin promoter which resulted in growth retardation, neural tube defects and a high incidence of embryonic lethality demonstrating the importance of ATRX dosage to normal development (Berube et al., 2002).

While each of these models has provided valuable insight into disease mechanisms (as highlighted below), the field still awaits a model whereby a single nucleotide variant is introduced into the *ATRX* gene to recreate a known patient mutation, such as the common p.Arg246Cys mutation within the ADD domain.

Replication Stress Impairs Progenitor Expansion Resulting in Microcephaly

Microcephaly is common to many NDDs and has also been observed in mouse models that deleted other genes encoding chromatin remodeling proteins (Ronan et al., 2013). Most ATR-X patients develop postnatal microcephaly and, in instances where CT or MRI scans have been performed, mild cerebral atrophy was detected. Similarly, three patient autopsy reports also described that the brains were smaller in size (Gibbons, 2006).

The first indication that ATRX may be critical for cell growth came from co-culture experiments of embryonic stem cells (ESC) from control or *Atrx*-null cells. This experiment demonstrated that the *Atrx*-null cells were underrepresented after 4-days of co-culture. Flow cytometry was used to examine cell cycle distribution but no differences were observed suggesting that cells may have transitioned to a slower cycling, differentiated cell type (Garrick et al., 2006). Given that ATRX has high expression in the developing forebrain, the $Atrx^{fl/fl}$ line was next crossed

with the forebrain-specific Foxg1-Cre line ($Atrx^{Foxg1Cre}$) that initiates Cre expression in the developing telencephalon at ∼E8.5 (Hébert and McConnell, 2000). Loss of ATRX caused a 25–30% reduction in cell number with a noticeably smaller neocortex and hippocampus including almost a complete absence of the dentate gyrus that likely contributed to early postnatal lethality (Bérubé et al., 2005). Similar to ESC co-culture experiments, BrdU-pulse labeling experiments suggested no differences in the proportion of cycling cells. However, there was a dramatic increase in the number of TUNEL+ cells leading to a reduction in the number of neurons that reached the cortical layers (Bérubé et al., 2005). Similarily, the $Atrx^{\Delta E2}$ mice were smaller and also showed brain hypocellularity, although to a milder extent (Nogami et al., 2011). *Atrx* inactivation in Sertoli and muscle cells, also showed a significant impact on the growth of the tissue (Bagheri-fam et al., 2011; Huh et al., 2012). However, a retina progenitor cell cKO only had a limited effect on the size of the mature tissue suggesting that defects in cell cycle progression lead to significant hypocellularity among tissues that require a rapid expansion over a narrow developmental timeframe (Medina et al., 2009).

Although not initially observed, delayed cell cycle progression through both S- and G2/M phases was later observed in other studies (Ritchie et al., 2008; Watson et al., 2013; Huh et al., 2016). For G2/M, the progression from prometaphase to metaphase was prolonged and associated with sister chromatid cohesion and congression defects that impaired proper separation at anaphase leading to DNA bridges and micronuclei (Ritchie et al., 2008). Evidence for DNA bridges and micronuclei in $Atrx^{Foxg1Cre}$ mice were also detected by high magnification microscopy at the apical surface on cortical sections of the neuroepithelium where cortical progenitors complete mitosis. Interestingly, a recent study has also demonstrated that ATRX promotes telomere cohesion between sister telomeres to mediate the repair of DNA DSB (Lovejoy et al., 2020).

Defects in S-phase were observed using BrdU-pulse chase flow cytometry experiments where a delay from G1 to S-phase and also from G2/M to the following G1 phase was identified (Huh et al., 2012). Co-labeling experiments demonstrated that ATRX associated with replicating chromatin during mid-late S-phase and cytological analysis showed a high prevalence of genomic instability that was enriched at telomeres and pericentromeric heterochromatin (Huh et al., 2012; Watson et al., 2013). Moreover, treatment with a compound that binds and stabilizes G4 DNA increased the number of telomere dysfunction induced foci (TIFs) and decreased cell viability suggesting that G4 DNA formation was the main cause of replicative stress (Watson et al., 2013). Other studies indicate that replication stress at telomeres may be mediated by increased TERRA transcription (Nguyen et al., 2017). TERRA levels are tightly regulated and critical for both telomere formation, replication and maintenance (Bettin et al., 2019). However, when TERRA levels increase, as shown for ATRX-null cells, it enhances R-loop (RNA-DNA hybrid) formation and G4 DNA stabilization, each of which increase replication fork stalling and collapse that then induces homology directed repair (HDR) and TIFs (Nguyen et al., 2017). The regulation of R-loops has also been proposed for other proteins that interact with G4 DNA during replication and/or at telomeres (Zhou et al., 2014; Ribeiro de Almeida et al., 2018; Toubiana and Selig, 2018; Maffia et al., 2020). It should also be mentioned here that somatic *ATRX* mutations, and to a lesser extent H3.3 and DAXX mutations, are prevalent in cancers characterized by ALT (alternative lengthening of telomeres), a HDR mechanism to maintain telomere length that is normally suppressed by ATRX (Heaphy et al., 2011; Lovejoy et al., 2012; Schwartzentruber et al., 2012; Pickett and Reddel, 2015; Verma and Greenberg, 2016).

Another indicator of replicative stress as a major impediment to growth of *Atrx*-null cells was demonstrated by studies showing an increased sensitivity to hydroxyurea, enhanced DSBs, and the use of DNA fiber analysis to show increased stalled replication forks and reduced origin firing (Leung et al., 2013; Clynes et al., 2014; Huh et al., 2016). Mechanistically, ATRX physically interacts with the MRN complex where it is thought to block HDR at stalled replication forks to allow for fork restart after the G4 DNA is resolved (Clynes et al., 2014). Indeed, one group demonstrated that fork protection could be restored by treatment with an Mre11 exonuclease inhibitor (Huh et al., 2016). This study also suggested that hyperactivation of poly (ADP-ribose) polymerase-1 (Parp-1) during neurogenesis may function as a compensatory mechanism to protect stalled replication forks from collapse and HDR, thus dampening the extent of cell loss during neurogenesis.

During mouse cortical development, the cortical layers are formed in sequential fashion from a pool of neural progenitor cells (NPC) that must continue to proliferate to maintain the pool size. Alterations in NPC proliferation depletes the pool often resulting in altered cell lamination typically observed as a reduction in upper layer neurons. For $Atrx^{Foxg1Cre}$ mice, the most proliferative NPCs that ultimately would become upper layer neurons are more susceptible to incur replication-induced DNA damage. Frequently the resulting genomic instability will occur at telomeres and pericentromeric heterochromatin, but it could also occur at other genomic regions that can form G4 DNA or similar secondary DNA structures that induce replication fork stalling and collapse. Accumulation of sufficient damage further leads to their demise and decreases neuron production and brain size. Indeed, the $Atrx^{Foxg1Cre}$ forebrain is reduced in size with a compromised production of upper layer neurons (Ritchie et al., 2014; Huh et al., 2016). Similar results have been observed in mice lacking CHD4 and SMARCA5 where the NPCs either fail to progress through the cell cycle or incur significant DNA damage, respectively, prior to undergoing apoptosis (Alvarez-Saavedra et al., 2014, 2019; Nitarska et al., 2016).

The SWI/SNF complex is also critical for brain development but utilizes different mechanisms than ATRX. The SWI/SNF complex is required during early neurogenesis for differentiation from radial glial progenitor cells into intermediate progenitor cells. This switch from a neural stem cell to an NPC is accompanied by the fundamental shift from the npBAF to nBAF complex, which involves the substitution of three subunits (BAF45, BAF53, and BAF55). Failure to switch leads to increased cell death, a small progenitor pool, and failure to further differentiate (Lessard et al., 2007; Wu et al., 2007; Bachmann et al., 2016). Interestingly, while SMARCA5 loss hampers NPC

proliferation, loss of its ISWI ortholog SMARCA1 fails to repress expression of proliferation genes resulting in delayed neuronal differentiation and a larger brain (Yip et al., 2016). Taken together, these examples highlight the importance of chromatin remodeling proteins to NPC homeostasis and provide insight into the multitude of mechanisms at work often resulting in a similar phenotype.

Transcriptional Deficits Associated With ATRX Mutations

Chromatin remodeling proteins were first identified as transcriptional coactivators and they continue to be implicated in the regulation of many genes. Since its identification, ATRX has also been presumed to regulate gene transcription. While there is a good level of understanding regarding how ATRX maintains genomic stability through the regulation of tandem repeats, telomeres and pericentromeric heterochromatin, the identification of direct transcriptional targets has proven more challenging. Initial ChIPseq experiments suggested that ATRX was bound at few promoters, gene bodies and regulatory elements (Law et al., 2010). Further work has suggested that ATRX binding may differ between tissues to ensure proper silencing of repetitive elements located near or within expressed genes in that particular tissue (Nguyen et al., 2017). Consistent with this idea, ATRX ChIPseq analysis of NPCs demonstrated a higher enrichment of binding sites at gene regulatory elements compared to what was observed in mouse ESCs suggesting that more genes may be under direct ATRX regulation within the brain (Law et al., 2010; He et al., 2015; Danussi et al., 2018). Another contributing factor to differential target gene expression is represented by ATRX effects on α-globin gene expression. Mutational analysis identified >15 ATR-X patients with the identical missense change (p.Arg246Cys), yet they showed a variable degree of hemoglobin H inclusions in blood samples, indicative of differing levels of α-globin expression (Gibbons et al., 1997). Repression of α-globin expression was dependent on the size of a GC-rich VNTR located within the globin gene cluster (Law et al., 2010). A second factor driving the tissue specificity and the variable effects was the formation of R-loops caused by the transcription of the GC-rich VNTR sequences (Nguyen et al., 2017). The larger sequences generate increased R-loops and G4 DNA structures that normally recruit ATRX to re-establish the normal chromatin structure. In the absence of ATRX the R-loop/G4 DNA is not resolved effectively which then impedes both replication and transcription processes (Nguyen et al., 2017). The slight stochastic nature of these effects likely also dampens readouts of differential expression from RNAseq experiments thereby raising the need for a scRNAseq approach in future studies.

Gene expression analysis of control and $Atrx^{Foxg1Cre}$ cortical samples at two timepoints (E13.5, P0.5) identified 202 and 304 differentially expressed genes (DEGs; ±1.5-fold change), respectively, with almost two-thirds of genes upregulated (Levy et al., 2008). Among these, 27 were common to both datasets including the downregulation of several ancestral pseudoautosomal region (aPAR) genes (Csfr2a, Dhrsxy, Cd99, and Asmtl) (Levy et al., 2008, 2014). In mouse, the aPAR genes are located in subtelomeric regions and contain potential G4 DNA sequences. Each gene analyzed had enriched histone H3.3 and ATRX binding within their gene body and showed reduced H3.3 levels when ATRX was absent (Levy et al., 2014). Interestingly, these intragene G4 DNA sequences also showed increased binding of RNA pol II in $Atrx^{Foxg1Cre}$ samples suggesting that transcription becomes impeded at these regions within the gene leading to reduced expression. The authors extended this finding to *Nlgn4*, a gene encoding a post-synaptic cell adhesion molecule implicated in ASDs (Jamain et al., 2003; Laumonnier et al., 2004). This result conflicted with the study on R-loop formation which found no differences in RNA pol II loading or histone modifications across genes containing the GC-rich repeats (Nguyen et al., 2017).

Other downregulated genes from this analysis include *Gbx2*, *NeuroD4*, *Wif1*, *Nxph1*, *Nxph2*, and *Mbp*, each of which could contribute to cognitive deficits observed in patients but require further analysis to asses their contribution to the phenotype (Levy et al., 2008, 2014). In a similar experiment in the retina, 173 DEGs were identified with two-thirds upregulated (109 genes) and one-third (64 genes) downregulated (Lagali et al., 2016). Most of these genes were involved in the regulation of glutamate activity, ion channel regulation or encoded neuroprotective peptides, with four shown to be also dysregulated in the cortex (*Csf2ra*, *Cbln4*, *Syt13*, and *Nlgn4*). Each of these studies showed that the mutant samples had only small numbers of genes with large changes in gene expression and, while some downregulated genes may impede transcriptional elongation, this mechanism may not be universal, particularly as it pertains to upregulated genes.

However, other indirect mechanisms have been explored to explain transcriptional dysregulation, particularly a loss of repression. The ATRX/DAXX complex is critical for loading H3.3 at telomeres and pericentromeric heterochromatin (Goldberg et al., 2010; Wong et al., 2010). Research over the last few years has expanded this regulation to include H3.3 deposition at endogenous retroviral elements, regions associated with imprinted genes, and some CpG islands (Elsässer et al., 2015; He et al., 2015; Sadic et al., 2015; Voon et al., 2015). At the telomere, the loss of ATRX affected the transcription of telomeric DNA and the non-coding RNA TERRA, although studies conflict on whether levels increase or decrease (Eid et al., 2015; Nguyen et al., 2017). Surprisingly, TERRA was also shown to bind to an additional ∼4,000 binding sites aside from the telomere where it co-localized with ATRX (Chu et al., 2017). Many of these sites were within introns and comprised GA repeats, however, depletion of TERRA usually resulted in downregulation while ATRX depletion increased expression (Chu et al., 2017). While it remains to be determined how this might affect neuronal gene expression, ATRX has been shown to interact with other lncRNAs including *Xist* to facilitate PRC2 silencing and *ChR01* that is required for heterochromatin reorganization in differentiating muscle cells (Sarma et al., 2014; Park et al., 2018). Indeed, ATRX binding to lncRNA or R-loops may be a key mechanism mediating transcriptional repression of specific target genes.

Histone H3.3 ChIPseq studies have also shown that it is enriched at the intracisternal A-particle endogenous retroviral elements (IAP/ERVs), which account for almost half of all mutation causing ERV insertions (Maksakova et al., 2006; Elsässer et al., 2015). Moreover, H3.3 deposition at these sites requires ATRX/DAXX to facilitate H3K9me3 and repression while depletion of ATRX, DAXX, or H3.3 results in reduction of the H3K9me3 mark and IAP/ERV derepression (Elsässer et al., 2015; He et al., 2015; Voon et al., 2015). In mouse ESCs, ERV derepression affected the expression of neighboring genes in a minority of cases with most genes neutral to ERV derepression. It raises the question of whether or not this type of derepression would affect many genes or occur rapidly within a post-mitotic neuron, and thus, function as a major effector in dysregulated gene expression in *Atrx*-null neurons. In this regard, a related study using cultured post-mitotic neurons demonstrated that the ADD domain can also bind the H3K9me3S10ph dual histone mark (Noh et al., 2015). This histone mark is rapidly induced by neuronal depolarization where it appears at centromeric and pericentromeric heterochromatin co-localized with ATRX to repress transcription of non-coding centromeric minor satellite sequences (Noh et al., 2015). While it is unclear what the impact of increased centromeric minor satellite transcription would have on disease pathology, it remains to be determined whether this dual mark affects activity-dependent transcription of genes mediating learning or memory.

It was also demonstrated that ATRX was bound to 56 CpG islands which was unexpected since they are often associated with active chromatin, typically promoters (Voon et al., 2015). However, these CpG islands were associated with H3K9me3, almost half were methylated and many corresponded to imprinted loci often residing in intragenic regions within a transcriptional unit (Voon et al., 2015). Indeed, in all cases examined ATRX was bound to the silenced imprinted allele which became reactivated in ATRX KO cells (Voon et al., 2015). This study contrasted somewhat with an independent report in which ATRX was recruited by MeCP2 to silence the active allele of several imprinted genes in the developing telencephalon (Levy et al., 2014). The difference in these studies may reflect differential regulation of imprinting loci in ESCs versus differentiating NPCs. Perhaps the most compelling example of derepression came from a study with the $Atrx^{\Delta E2}$ mice (Shioda et al., 2018). In this model, the authors identify a small list of 31 DEGs in the adult hippocampus but with most genes (23/31) downregulated (Shioda et al., 2018). Among the upregulated genes was an imprinted gene from the lymphocyte-regulated gene family, *Xlr3b*. Although *Xlr3b* had widespread expression across many tissues, it was only upregulated in the brain. Further work showed that ATRX bound to a G4 DNA sequence within the CpG island of the *Xlr3b* gene where it normally interacted with DAXX and H3.3 and recruited DNMT1 and DNMT3 to silence the gene. The subsequent overexpression of *Xlr3b* in the $Atrx^{\Delta E2}$ mice was shown to produce a protein that localized to dendritic RNA granules where it interacted with ribonucleoproteins, dynein proteins and the RNA-binding protein, TIA1, to regulate mRNA transport (Shioda et al., 2018). One of the targets identified was the mRNA for CAMK II-α which they had previously shown to be deregulated in these animals. Excitingly, they also showed that the G-quadruplex-binding ligand 5-aminolevulinic acid (5-ALA) was able to decrease RNApol II occupancy and *Xlr3b* expression in the $Atrx^{\Delta E2}$ mice, although methylation of the G4 DNA sequence within the CpG island was not affected. It seems that formation or stabilization of this G4 DNA sequence is required to activate the *Xlr3b* gene and that ATRX normally prevents this by facilitating heterochromatin formation. In this regard, mapping of G4 DNA sequences have shown an enriched number in gene regulatory elements where many function to increase transcription when stabilized (Hänsel-Hertsch et al., 2016). While G4 DNA stabilization occurs in the $Atrx^{\Delta E2}$ mice, further work is required to explain how 5-ALA represses Xlr3b transcription when it should stabilize the G4 DNA. Regardless, the derepression of G4 DNA within CpG islands and/or other regulatory elements is an exciting mechanism that can explain DEG upregulation, particularly when coupled with the finding that ATRX binding is increased at regulatory elements in NPCs compared to ESCs. Collectively, the derepression of tandem repeats, retrotransposable elements and G4 quadruplexes can all function to impinge on neuronal function.

Morphological, Behavioral, and Cell Non-autonomous Deficits

We have discussed global effects on DNA replication and transcription that occur in the absence of ATRX in the previous two sections. In this section, we review the morphological and functional repercussions of these deficits. Aside from being reduced in size, the $Atrx^{Foxg1\,Cre}$ mice had a normal cortical morphology with proper lamination although a reduction of upper layer neurons (Bérubé et al., 2005; Ritchie et al., 2014; Huh et al., 2016). The reduction in upper layer neurons may also contribute to the partial agenesis of the corpus callosum observed in some patients (Gibbons, 2006). The hippocampus was also reduced in size while the dentate gyrus consisted of only a few disorganized cells. Behavior analysis was not performed due to the early postnatal lethality, although female heterozygous mice showed impairment in spatial, contextual fear, and novel object recognition memory (Tamming et al., 2017).

The $Atrx^{\Delta E2}$ mice also had smaller brains but no differences in cell density within layers II/III of the prefrontal cortex (PFC) or hippocampus (Shioda et al., 2011). Examination of dendritic spines in the PFC showed that the $Atrx^{\Delta E2}$ mice had similar numbers but fewer mature spines and many more, thin, long immature spines (Shioda et al., 2011). Behavioral analysis indicated that the mice have impaired contextual fear memory (fear conditioning test), spatial memory (Y-maze), but not anxiety behaviors (Nogami et al., 2011; Shioda et al., 2011). Electrophysiology studies in hippocampal slices demonstrated reduced NMDAR-dependent long term potentiation (LTP) evoked by high stimulation frequency in hippocampal CA1 neurons which was mediated by increased CAMK2A and GluR1 phosphorylation (Nogami et al., 2011). This was in contrast to a later article by the same group that showed phosphorylated CAMK2A levels were reduced in the $Atrx^{\Delta E2}$ mice while 5-ALA

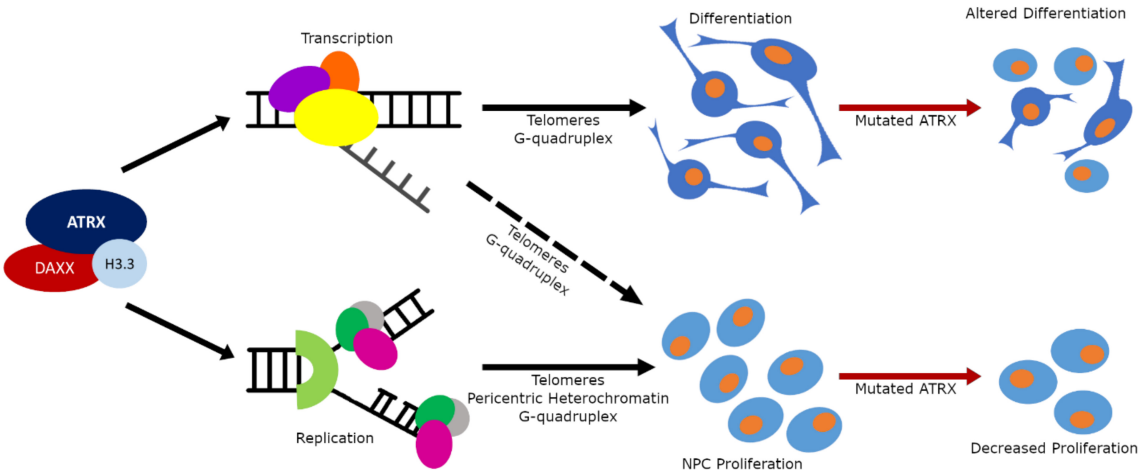

FIGURE 3 | The multiple functions of the ATRX protein. Schematic diagram of ATRX functional influence on brain development and its contribution to NDDs. Normally ATRX utilizes its chromatin remodeling activity to (1) influence transcription and DNA replication in heterochromatic regions to control the rate of proliferation in the neuronal progenitor cell population (bottom arm); and (2) to influence transcription in heterochromatic regions to control differentiation processes (top arm). When ATRX is mutated the cellular proliferation rates in progenitors is slowed resulting in a smaller progenitor population; and the differentiation processes are altered resulting in either dysfunctional cellular morphology or complete absence of specific cell types.

restored the levels at the synapse and the phosphorylation levels (Shioda et al., 2018). A recent article examining hippocampal function using CAMKII-Cre mice to inactivate *Atrx* in postnatal excitatory forebrain neurons demonstrated reduced paired-pulse facilitation and LTP in proximal and distal apical dendrites of CA1 synapses (Gugustea et al., 2019). This represented the first study of mice in which *Atrx* has been inactivated after neurogenesis and it will be interesting to ascertain the full characterization of these mice.

Studies of the retina have also provided useful information into ATRX function. Many ATR-X patients have visual problems although this has been an under-appreciated aspect of the phenotype (Medina et al., 2009). Inactivation of ATRX in retinal progenitors $Atrx^{Pax6Cre}$ resulted in a slight reduction in retina size and a specific reduction of interneurons, namely amacrine and horizontal cells (Medina et al., 2009). Surprisingly, $Atrx^{Pitf1aCre}$ mice that ablates ATRX in a bi-potential progenitor that generates amacrine or horizontal cells did not recapitulate the phenotype, while inactivation with a bipolar cell specific Cre driver ($Atrx^{Vsx2Cre}$) did not affect bipolar cell survival but did result in reduced amacrine and horizontal cells suggesting that interneuron survival was a cell non-autonomous effect (Lagali et al., 2016). Additional characterization of these mice showed that the bipolar axons were mislocalized within the inner plexiform layer and many had axonal swellings or tortuous paths to their targets. Gene expression analysis identified alterations in the glutamate pathway, ion channel regulation and altered expression of neuroprotective peptides. Altered axonal pathfinding was also observed in Drosophila XNP mutants, the homolog to the ATPase domain of ATRX (Sun et al., 2006). It will be important to further explore in greater detail whether axonal pathfinding is also altered within forebrain or hippocampal neurons.

PERSPECTIVES

Studies to date have indicated that ATRX has multiple roles during forebrain development that can contribute to the phenotype of ATR-X patients. It functions mainly as a heterochromatin interacting protein acting to ensure that repetitive DNA is properly packaged and organized into heterochromatin. We have highlighted how aberrations in heterochromatin maintenance leads to genomic instability and replication stress that impairs NPC expansion leading to a microcephalic brain (**Figure 3**). The loss of ATRX also affects gene expression typically resulting in increased gene derepression but also downregulation. It remains to be teased apart which targets are direct versus indirect, and when disrupted expression hampers neuronal function. It is likely that inactivation of ATRX in postmitotic neurons, following neurogenesis and lamination, will help define a role for ATRX target genes in altered synaptic activity and/or synaptic plasticity underlying cognitive impairment. Moreover, the contribution of other central nervous system cell types to the phenotype have not been explored. ATRX is expressed in glia and oligodendrocytes which are known to intimately communicate with neurons to mediate function, as shown recently in *Drosophila* glial ATRX dependent ensheathment of sensory neurons, for normal dendritic arborization and stimulus processing (Yadav et al., 2019). Intriguingly, MRI studies on ATR-X patients showed severe glial defects and white matter disruption, further stressing the need for research in this area (Wada et al., 2013; Lee et al., 2015). Importantly, a further understanding of ATRX function and its aberrant molecular pathways are required before potential treatments can be explored. In this regard, treatment with 5-ALA has shown promise in one animal model and it is being investigated in Japanese patients

(T. Wada, personal communication). ATRX is but one of many different chromatin remodeling proteins mutated in NDDs but it serves to demonstrate how complex these disorders are and how widely chromatin remodelers impact cellular activities.

AUTHOR CONTRIBUTIONS

ST and DP wrote and edited the manuscript together. ST generated the figures and table. Both authors contributed to the article and approved the submitted version.

FUNDING

The research supporting this work was funded by two operating grants (FRN133586 and FRN165994) from the Canadian Institute of Health Research awarded to DP.

ACKNOWLEDGMENTS

We are grateful to Dr. Alex Córdova and Valérie Cardin for their valued input and critical reading of the manuscript.

REFERENCES

Abidi, F. E., Cardoso, C., Lossi, A. M., Lowry, R. B., Depetris, D., Mattéi, M. G., et al. (2005). Mutation in the 5′ alternatively spliced region of the XNP/ATR-X gene causes Chudley-Lowry syndrome. *Eur. J. Hum. Genet.* 13, 176–183. doi: 10.1038/sj.ejhg.5201303

Alazami, A. M., Patel, N., Shamseldin, H. E., Anazi, S., Al-Dosari, M. S., Alzahrani, F., et al. (2015). Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* 10, 148–161. doi: 10.1016/j.celrep.2014.12.015

Alfert, A., Moreno, N., and Kerl, K. (2019). The BAF complex in development and disease. *Epigenet. Chromat.* 12, 1–15. doi: 10.1186/s13072-019-0264-y

Alvarez-Saavedra, M., De Repentigny, Y., Lagali, P. S., Raghu Ram, E. V. S., Yan, K., Hashem, E., et al. (2014). Snf2h-mediated chromatin organization and histone H1 dynamics govern cerebellar morphogenesis and neural maturation. *Nat. Commun.* 5:5181. doi: 10.1038/ncomms5181

Alvarez-Saavedra, M., Yan, K., De Repentigny, Y., Hashem, L. E., Chaudary, N., Sarwar, S., et al. (2019). Snf2h drives chromatin remodeling to prime upper layer cortical neuron development. *Front. Mol. Neurosci.* 12:243. doi: 10.3389/fnmol.2019.00243

Argentaro, A., Yang, J. C., Chapman, L., Kowalczyk, M. S., Gibbons, R. J., Higgs, D. R., et al. (2007). Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11939–11944. doi: 10.1073/pnas.0704057104

Bachmann, C., Nguyen, H., Rosenbusch, J., Pham, L., Rabe, T., Patwa, M., et al. (2016). mSWI/SNF (BAF) complexes are indispensable for the neurogenesis and development of embryonic olfactory epithelium. *PLoS Genet.* 12:6274. doi: 10.1371/journal.pgen.1006274

Badens, C., Lacoste, C., Philip, N., Martini, N., Courrier, S., Giuliano, F., et al. (2006). Mutations in PHD-like domain of the ATRX gene correlate with severe psychomotor impairment and severe urogenital abnormalities in patients with ATRX syndrome. *Clin. Genet.* 70, 57–62. doi: 10.1111/j.1399-0004.2006.00641.x

Bagheri-fam, S., Argentaro, A., Svingen, T., Combes, A. N., Sinclair, A. H., Koopman, P., et al. (2011). Defective survival of proliferating sertoli cells and androgen receptor function in a mouse model of the ATR-X syndrome. *Hum. Mol. Genet.* 20, 2213–2224. doi: 10.1093/hmg/ddr109

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., et al. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124. doi: 10.1038/35065138

Basehore, M. J., Michaelson-Cohen, R., Levy-Lahad, E., Sismani, C., Bird, L. M., Friez, M. J., et al. (2015). Alpha-thalassemia intellectual disability: variable phenotypic expression among males with a recurrent nonsense mutation - c.*109C*>T (p.R37X). *Clin. Genet.* 87, 461–466. doi: 10.1111/cge.12420

Bérubé, N. G., Healy, J., Medina, C. F., Wu, S., Hodgson, T., Jagla, M., et al. (2008). Patient mutations alter ATRX targeting to PML nuclear bodies. *Eur. J. Hum. Genet.* 16, 192–201. doi: 10.1038/sj.ejhg.5201943

Berube, N. G., Jagla, M., Smeenk, C., De Repentigny, Y., Kothary, R., and Picketts, D. J. (2002). Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Hum. Mol. Genet.* 11, 253–261. doi: 10.1093/hmg/11.3.253

Bérubé, N. G., Mangelsdorf, M., Jagla, M., Vanderluit, J., Garrick, D., Gibbons, R. J., et al. (2005). The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *J. Clin. Invest.* 115, 258–267. doi: 10.1172/JCI200522329

Berube, N. G., Smeenk, C. A., and Picketts, D. J. (2000). Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum. Mol. Genet.* 9, 539–547. doi: 10.1093/hmg/9.4.539

Bettin, N., Oss Pegorar, C., and Cusanelli, E. (2019). The emerging roles of TERRA in telomere maintenance and genome stability. *Cells* 8:246. doi: 10.3390/cells8030246

Bögershausen, N., and Wollnik, B. (2018). Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-related intellectual disability disorders. *Front. Mol. Neurosci.* 11:252. doi: 10.3389/fnmol.2018.00252

Bowman, G. D., and Poirier, M. G. (2015). Post-translational modifications of histones that influence nucleosome dynamics. *Chem. Rev.* 115, 2274–2295. doi: 10.1021/cr500350x

Bramswig, N. C., Caluseriu, O., Lüdecke, H. J., Bolduc, F. V., Noel, N. C. L., Wieland, T., et al. (2017). Heterozygosity for ARID2 loss-of-function mutations in individuals with a Coffin-Siris syndrome-like phenotype. *Hum. Genet.* 136, 297–305. doi: 10.1007/s00439-017-1757-z

Cardoso, C., Lutz, Y., Mignon, C., Compe, E., Depetris, D., Mattei, M. G., et al. (2000). ATR-X mutations cause impaired nuclear location and altered DNA binding properties of the XNP/ATR-X protein. *J. Med. Genet.* 37, 746–751. doi: 10.1136/jmg.37.10.746

Cardoso, C., Timsit, S., Villard, L., Khrestchatisky, M., Fontès, M., and Colleaux, L. (1998). Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. *Hum. Mol. Genet.* 7, 679–684. doi: 10.1093/hmg/7.4.679

Carpenter, N. J., Qu, Y., Curtis, M., and Patil, S. R. (1999). X-linked mental retardation syndrome with characteristic "coarse" facial appearance, brachydactyly, and short stature maps to proximal Xq. *Am. J. Med. Genet.* 85, 230–235. doi: 10.1002/(sici)1096-8628(19990730)85:3<230::aid-ajmg9>3.0.co;2-o

Carvill, G. L., Heavin, S. B., Yendle, S. C., McMahon, J. M., O'Roak, B. J., Cook, J., et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat. Genet.* 45, 825–830. doi: 10.1002/jmri.25711.PET/MRI

Chu, H., Cifuentes-rojas, C., Kesner, B., Aeby, E., Lee, H., Wei, C., et al. (2017). TERRA RNA antagonizes ATRX and protects telomeres. *Cell* 170, 86–101. doi: 10.1016/j.cell.2017.06.017.TERRA

Clapier, C. R., Iwasa, J., Cairns, B. R., and Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* 18, 407–422. doi: 10.1038/nrm.2017.26

Clynes, D., Jelinska, C., Xella, B., Ayyub, H., Taylor, S., Mitson, M., et al. (2014). ATRX dysfunction induces replication defects in primary mouse cells. *PLoS One* 9:e092915. doi: 10.1371/journal.pone.0092915

Danussi, C., Bose, P., Parthasarathy, P. T., Silberman, P. C., Van Arnam, J. S., Vitucci, M., et al. (2018). Atrx inactivation drives disease-defining phenotypes in glioma cells of origin through global epigenomic remodeling. *Nat. Commun.* 9, 1–15. doi: 10.1038/s41467-018-03476-3476

De Rubeis, S., He, X., Goldberg, A., Poultney, C., Samocha, K., Cicek, A., et al. (2014). Synaptic, transcriptional, and chromatin genes disrupted in autism. *Nature* 515, 209–215.

Delbarre, E., Ivanauskienne, K., Spirkoski, J., Shah, A., Vekterud, K., Oivind Moskaug, J., et al. (2017). PML protein organizes heterochromatin domains

where it regulates histone H3 . 3 loading by ATRX / DAXX. *Genome Res.* 27, 913–921. doi: 10.1101/gr.215830.116.6

Dhayalan, A., Tamas, R., Bock, I., Tattermusch, A., Dimitrova, E., Kudithipudi, S., et al. (2011). The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum. Mol. Genet.* 20, 2195–2203. doi: 10.1093/hmg/ddr107

Drané, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* 24, 1253–1265. doi: 10.1101/gad.566910

Dyer, M. A., Qadeer, Z. A., Valle-Garcia, D., and Bernstein, E. (2017). ATRX and DAXX: mechanisms and mutations. *Cold Spring Harb. Perspect. Med.* 7, 1–16. doi: 10.1101/cshperspect.a026567

Eid, R., Demattei, M.-V., Episkopou, H., Augé-Gouillou, C., Decottignies, A., Grandin, N., et al. (2015). Genetic inactivation of ATRX leads to a decrease in the amount of telomeric cohesin and level of telomere transcription in human glioma cells. *Mol. Cell. Biol.* 35, 2818–2830. doi: 10.1128/mcb.01317-1314

Elsässer, S. J., Noh, K.-M., Diaz, N., Allis, D., and Banaszynski, L. A. (2015). Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* 522, 240–244. doi: 10.1038/nature14345

Eustermann, S., Yang, J. C., Law, M. J., Amos, R., Chapman, L. M., Jelinska, C., et al. (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat. Struct. Mol. Biol.* 18, 777–782. doi: 10.1038/nsmb.2070

Garrick, D., Samara, V., McDowell, T. L., Smith, A. J. H., Dobbie, L., Higgs, D. R., et al. (2004). A conserved truncated isoform of the ATR-X syndrome protein lacking the SWI/SNF-homology domain. *Gene* 326, 23–34. doi: 10.1016/j.gene.2003.10.026

Garrick, D., Sharpe, J. A., Arkell, R., Dobbie, L., Smith, A. J. H., Wood, W. G., et al. (2006). Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues. *PLoS Genet.* 2:e58. doi: 10.1371/journal.pgen.0020058

Gemmen, G. J., Sim, R., Haushalter, K. A., Ke, P. C., Kadonaga, J. T., and Smith, D. E. (2005). Forced unraveling of nucleosomes assembled on heterogeneous DNA using core histones, NAP-1, and ACF. *J. Mol. Biol.* 351, 89–99. doi: 10.1016/j.jmb.2005.05.058

Gibbons, R. (2006). Alpha thalassaemia-mental retardation, X linked. *Orphanet J. Rare Dis.* 1, 1–9. doi: 10.1186/1750-1172-1-15

Gibbons, R. J., Bachoo, S., Picketts, D. J., Aftimos, S., Asenbauer, B., Bergoffen, J., et al. (1997). Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. *Nat. Genet.* 17, 146–148. doi: 10.1038/ng1097-146

Gibbons, R. J., Picketts, D. J., Villard, L., and Higgs, D. R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). *Cell* 80, 837–845. doi: 10.1016/0092-8674(95)90287-90282

Gibbons, R. J., Wada, T., Fisher, C. A., Malik, N., Mitson, M. J., Steensma, D. P., et al. (2008). Mutations in the chromatin-associated protein ATRX. *Hum. Mutat.* 29, 796–802. doi: 10.1002/humu.20734

Goldberg, A. D., Banaszynski, L. A., Noh, K., Lewis, P. W., Elsaesser, S. J., Stadler, S., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Science* 140, 678–691. doi: 10.1016/j.cell.2010.01.003.Distinct

Goodwin, L. R., and Picketts, D. J. (2018). The role of ISWI chromatin remodeling complexes in brain development and neurodevelopmental disorders. *Mol. Cell. Neurosci.* 87, 55–64. doi: 10.1016/j.mcn.2017.10.008

Guerrini, R., Shanahan, J. L., Carrozzo, R., Bonanni, P., Higgs, D. R., and Gibbons, R. J. (2000). Hemizygous mutation of the peripheral myelin protein 22 gene associated with Charcot-Marie-Tooth disease type 1. *Ann. Neurol.* 47, 117–121.

Gugustea, R., Tamming, R. J., Martin-Kenny, N., Bérubé, N. G., and Leung, L. S. (2019). Inactivation of ATRX in forebrain excitatory neurons affects hippocampal synaptic plasticity. *Hippocampus* 30, 565–581. doi: 10.1002/hipo.23174

Guo, D. C., Duan, X. Y., Regalado, E. S., Mellor-Crummey, L., Kwartler, C. S., Kim, D., et al. (2017). Loss-of-function mutations in YY1AP1 lead to grange syndrome and a fibromuscular dysplasia-like vascular disease. *Am. J. Hum. Genet.* 100, 21–30. doi: 10.1016/j.ajhg.2016.11.008

Hänsel-Hertsch, R., Beraldi, D., Lensing, S. V., Marsico, G., Zyner, K., Parry, A., et al. (2016). G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* 48, 1267–1272. doi: 10.1038/ng.3662

He, Q., Kim, H., Huang, R., Lu, W., Tang, M., Shi, F., et al. (2015). The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation. *Cell Stem Cell* 17, 273–286. doi: 10.1016/j.stem.2015.07.022

Heaphy, C. M., Wilde, R. F., De Jiao, Y., Klein, A. P., Edil, B. H., Shi, C., et al. (2011). Altered telomeres in tumors with ATRX and DAXX mutations. *Science* 333, 1–4. doi: 10.1126/science.1207313.Altered

Hébert, J. M., and McConnell, S. K. (2000). Targeting of cre to the Foxg1 (BF-1) locus mediates loxP recombination in the telencephalon and other developing head structures. *Dev. Biol.* 222, 296–306. doi: 10.1006/dbio.2000.9732

Hoelper, D., Huang, H., Jain, A. Y., Patel, D. J., and Lewis, P. W. (2017). Structural and mechanistic insights into ATRX-dependent and -independent functions of the histone chaperone DAXX. *Nat. Commun.* 8:1193. doi: 10.1038/s41467-017-01206-y

Hoffmann, A., and Spengler, D. (2019). Chromatin remodeling complex NuRD in neurodevelopment and neurodevelopmental disorders. *Front. Genet.* 10:682. doi: 10.3389/fgene.2019.00682

Holmes, L. B., and Gang, D. L. (1984). An X-linked mental retardation syndrome with craniofacial abnormalities, microcephaly and club foot. *Am. J. Med. Genet.* 17, 375–382. doi: 10.1002/ajmg.1320170131

Homann, O. R., Misura, K., Lamas, E., Sandrock, R. W., Nelson, P., Mcdonough, S. I., et al. (2016). Whole-genome sequencing in multiplex families with psychoses reveals mutations in the SHANK2 and SMARCA1 genes segregating with illness. *Mol. Psychiatry* 21, 1690–1695. doi: 10.1038/mp.2016.24.Whole-genome

Hood, R. L., Lines, M. A., Nikkel, S. M., Schwartzentruber, J., Beaulieu, C., Nowaczyk, M. J. M., et al. (2012). Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. *Am. J. Hum. Genet.* 90, 308–313. doi: 10.1016/j.ajhg.2011.12.001

Hormozdiari, F., Penn, O., Borenstein, E., and Eichler, E. E. (2015). The discovery of integrated gene networks for autism and related disorders. *Genome Res.* 25, 142–154. doi: 10.1101/gr.178855.114.142

Hota, S. K., and Bruneau, B. G. (2016). ATP-dependent chromatin remodeling during mammalian development. *Dev.* 143, 2882–2897. doi: 10.1242/dev.128892

Howard, M. T., Malik, N., Anderson, C. B., Voskuil, J. L. A., Atkins, J. F., and Gibbons, R. J. (2004). Attenuation of an amino-terminal premature stop codon mutation in the ATRX gene by an alternative mode of translational initiation. *J. Med. Genet.* 41, 951–956. doi: 10.1136/jmg.2004.020248

Hoyer, J., Ekici, A. B., Endele, S., Popp, B., Zweier, C., Wiesener, A., et al. (2012). Haploinsufficiency of ARID1B, a member of the SWI/SNF-A chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.* 90, 565–572. doi: 10.1016/j.ajhg.2012.02.007

Huh, M. S., Ivanochko, D., Hashem, L. E., Curtin, M., Delorme, M., Goodall, E., et al. (2016). Stalled replication forks within heterochromatin require ATRX for protection. *Cell Death Dis.* 7, 1–12. doi: 10.1038/cddis.2016.121

Huh, M. S., O'Dea, T. P., Ouazia, D., McKay, B. C., Parise, G., Parks, R. J., et al. (2012). Compromised genomic integrity impedes muscle growth after Atrx inactivation. *J. Clin. Invest.* 122, 4412–4423. doi: 10.1172/JCI63765

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P. W., Cochrane, J. C., et al. (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat. Struct. Mol. Biol.* 18, 769–776. doi: 10.1038/nsmb.2062

Jamain, S., Quach, H., Betancur, C., Råstam, M., Colineaux, C., Gillberg, C., et al. (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat. Genet.* 34, 27–29. doi: 10.1038/ng1136

Karaca, E., Harel, T., Pehlivan, D., Jhangiani, S. N., Gambin, T., Coban Akdemir, Z., et al. (2015). Genes that affect brain structure and function identified by rare variant analyses of mendelian neurologic Disease. *Neuron* 88, 499–513. doi: 10.1016/j.neuron.2015.09.048

Kleefstra, T., Kramer, J. M., Neveling, K., Willemsen, M. H., Koemans, T. S., Vissers, L. E. L. M., et al. (2012). Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* 91, 73–82. doi: 10.1016/j.ajhg.2012.05.003

Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., et al. (2016). Systematic phenomics analysis deconvolutes genes mutated in intellectual disability into biologically coherent modules. *Am. J. Hum. Genet.* 98, 149–164. doi: 10.1016/j.ajhg.2015.11.024

Kosho, T., Okamoto, N., Ohashi, H., Tsurusaki, Y., Imai, Y., Hibi-Ko, Y., et al. (2013). Clinical correlations of mutations affecting six components of the SWI/SNF complex: detailed description of 21 patients and a review of the literature. *Am. J. Med. Genet. Part A* 161, 1221–1237. doi: 10.1002/ajmg.a.35933

Kovatcheva, M., Liao, W., Klein, M. E., Robine, N., Geiger, H., Crago, A. M., et al. (2017). ATRX is a regulator of therapy induced senescence in human cells. *Nat. Commun.* 8, e540–e545. doi: 10.1038/s41467-017-00540-545

Lagali, P. S., Medina, C. F., Zhao, B. Y. H., Yan, K., Baker, A. N., Coupland, S. G., et al. (2016). Retinal interneuron survival requires non-cellautonomous Atrx activity. *Hum. Mol. Genet.* 25, 4787–4803. doi: 10.1093/hmg/ddw306

Laumonnier, F., Bonnet-Brilhault, F., Gomot, M., Blanc, R., David, A., Moizard, M. P., et al. (2004). X-Linked mental retardation and autism are associated with a mutation in the NLGN4 Gene, a Member of the neuroligin family. *Am. J. Hum. Genet.* 74, 552–557. doi: 10.1086/382137

Law, M. J., Lower, K. M., Voon, H. P. J., Hughes, J. R., Garrick, D., Viprakasit, V., et al. (2010). ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* 143, 367–378. doi: 10.1016/j.cell.2010.09.023

Lechner, M. S., Schultz, D. C., Negorev, D., Maul, G. G., and Rauscher, F. J. (2005). The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem. Biophys. Res. Commun.* 331, 929–937. doi: 10.1016/j.bbrc.2005.04.016

Lee, J. S., Lee, S., Lim, B. C., Kim, K. J., Hwang, Y. S., Choi, M., et al. (2015). Alpha-thalassemia X-linked intellectual disability syndrome identified by whole exome sequencing in two boys with white matter changes and developmental retardation. *Gene* 569, 318–322. doi: 10.1016/j.gene.2015.04.075

Lessard, J., Wu, J. I., Ranish, J. A., Wan, M., Winslow, M. M., Staahl, B. T., et al. (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* 55, 201–215. doi: 10.1016/j.neuron.2007.06.019.An

Leung, J. W. C., Ghosal, G., Wang, W., Shen, X., Wang, J., Li, L., et al. (2013). Alpha thalassemia/mental retardation syndrome X-linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress. *J. Biol. Chem.* 288, 6342–6350. doi: 10.1074/jbc.M112.411603

Levy, M. A., Fernandes, A. D., Tremblay, D. C., Seah, C., and Bérubé, N. G. (2008). The SWI/SNF protein ATRX co-regulates pseudoautosomal genes that have translocated to autosomes in the mouse genome. *BMC Genomics* 9:468. doi: 10.1186/1471-2164-9-468

Levy, M. A., Kernohan, K. D., Jiang, Y., and Bérubé, N. G. (2014). ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions. *Hum. Mol. Genet.* 24, 1824–1835. doi: 10.1093/hmg/ddu596

Lewis, P. W., Elsaesser, S. J., Noh, K. M., Stadler, S. C., and Allis, C. D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14075–14080. doi: 10.1073/pnas.1008850107

Lopes, F., Barbosa, M., Ameur, A., Soares, G., De Sá, J., Dias, A. I., et al. (2016). Identification of novel genetic causes of Rett syndrome-like phenotypes. *J. Med. Genet.* 53, 190–199. doi: 10.1136/jmedgenet-2015-103568

Lossi, A. M., Millan, J. M., Villard, L., Orellana, C., Cardoso, C., Prieto, F., et al. (1999). Genetic analysis of families with Parkinson disease that carry the Ala53Thr mutation in the gene encoding α-synudein [1]. *Am. J. Hum. Genet.* 65, 558–562. doi: 10.1086/302486

Lovejoy, C. A., Li, W., Reisenweber, S., Thongthip, S., Bruno, J., de Lange, T., et al. (2012). Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of Telomeres pathway. *PLoS Genet.* 8:2772. doi: 10.1371/journal.pgen.1002772

Lovejoy, C. A., Takai, K., Huh, M. S., Picketts, D. J., and de Lange, T. (2020). ATRX affects the repair of telomeric DSBs by promoting cohesion and a DAXX-dependent activity. *PLoS Biol.* 18:e3000594. doi: 10.1371/journal.pbio.3000594

Lu, X., Meng, X., Morris, C. A., and Keating, M. T. (1998). A novel human gene, WSTF, is deleted in Williams syndrome. *Genomics* 54, 241–249. doi: 10.1006/geno.1998.5578

Machol, K., Rousseau, J., Ehresmann, S., Garcia, T., Nguyen, T. T. M., Spillmann, R. C., et al. (2019). Expanding the spectrum of BAF-Related disorders: de novo variants in SMARCC2 Cause a syndrome with intellectual disability and developmental delay. *Am. J. Hum. Genet.* 104, 164–178. doi: 10.1016/j.ajhg.2018.11.007

Maffia, A., Ranise, C., and Sabbioneda, S. (2020). From R-loops to G-quadruplexes: emerging new threats for the replication fork. *Int. J. Mol. Sci.* 21:1506. doi: 10.3390/ijms21041506

Maksakova, I. A., Romanish, M. T., Gagnier, L., Dunn, C. A., Van De Lagemaat, L. N., and Mager, D. L. (2006). Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet.* 2:2. doi: 10.1371/journal.pgen.0020002

McDowell, T. L., Gibbons, R. J., Sutherland, H., O'Rourke, D. M., Bickmore, W. A., Pombo, A., et al. (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13983–13988. doi: 10.1073/pnas.96.24.13983

Medina, C. F., Mazerolle, C., Wang, Y., Bérubé, N. G., Coupland, S., Gibbons, R. J., et al. (2009). Altered visual function and interneuron survival in Atrx knockout mice: inference for the human syndrome. *Hum. Mol. Genet.* 18, 966–977. doi: 10.1093/hmg/ddn424

Mitson, M., Kelley, L. A., Sternberg, M. J. E., Higgs, D. R., and Gibbons, R. J. (2011). Functional significance of mutations in the Snf2 domain of ATRX. *Hum. Mol. Genet.* 20, 2603–2610. doi: 10.1093/hmg/ddr163

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M. J., Shu, X., et al. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2709–2714. doi: 10.1073/pnas.0608056104

Neale, B. M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K. E., Sabo, A., et al. (2016). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485:485. doi: 10.1016/j.physbeh.2017.03.040

Nguyen, D. T., Voon, H. P. J., Xella, B., Scott, C., Clynes, D., Babbs, C., et al. (2017). The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats. *EMBO Rep.* 18, 914–928. doi: 10.15252/embr.201643078

Nitarska, J., Smith, J. G., Sherlock, W. T., Hillege, M. M. G., Nott, A., Barshop, W. D., et al. (2016). A functional switch of NuRD chromatin remodeling complex subunits regulates mouse cortical development. *Cell Rep.* 17, 1683–1698. doi: 10.1016/j.celrep.2016.10.022

Nogami, T., Beppu, H., Tokoro, T., Moriguchi, S., Shioda, N., Fukunaga, K., et al. (2011). Reduced expression of the ATRX gene, a chromatin-remodeling factor, causes hippocampal dysfunction in mice. *Hippocampus* 21, 678–687. doi: 10.1002/hipo.20782

Noh, K. M., Maze, I., Zhao, D., Xiang, B., Wenderski, W., Lewis, P. W., et al. (2015). ATRX tolerates activity-dependent histone H3 methyl/phos switching to maintain repetitive element silencing in neurons. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6820–6827. doi: 10.1073/pnas.1411258112

O'Roak, B. J., Vives, L., Fu, W., Egertson, J. D., Stanaway, I. B., Phelps, I. G., et al. (2012a). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338, 1619–1622. doi: 10.1007/978-4-431-56050-0_20

O'Roak, B. J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B. P., et al. (2012b). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485, 246–250. doi: 10.1038/nature10989

Park, J., Lee, H., Han, N., Kwak, S., Lee, H. T., Kim, J. H., et al. (2018). Long non-coding RNA ChRO1 facilitates ATRX/DAXX-dependent H3.3 deposition for transcription-associated heterochromatin reorganization. *Nucleic Acids Res.* 46, 11759–11775. doi: 10.1093/nar/gky923

Peoples, R. J., Cisco, M. J., Kaplan, P., and Francke, U. (1998). Identification of the WBSCR9 gene, encoding a novel transcriptional regulator, in the Williams-Beuren syndrome deletion at 7q11.*23*. *Cytogenet. Cell Genet.* 82, 238–246. doi: 10.1159/000015110

Peterson, C. L. (2009). Reconstitution of nucleosomal arrays using recombinant Drosophila ACF and NAP1. *Cold Spring Harb. Protoc.* 4, 1–5. doi: 10.1101/pdb.prot5114

Pickett, H. A., and Reddel, R. R. (2015). Molecular mechanisms of activity and derepression of alternative lengthening of telomeres. *Nat. Struct. Mol. Biol.* 22, 875–880. doi: 10.1038/nsmb.3106

Picketts, D. J., Higgs, D. R., Bachoo, S., Blake, D. J., Quarrell, O. W. J., and Gibbons, R. J. (1996). ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum. Mol. Genet.* 5, 1899–1907. doi: 10.1093/hmg/5.12.1899

Picketts, D. J., Tastan, A. O., Higgs, D. R., and Gibbons, R. J. (1998). Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. *Mamm. Genome* 9, 400–403. doi: 10.1007/s003359900781

Pisansky, M. T., Young, A. E., O'Connor, M. B., Gottesman, I. I., Bagchi, A., and Gewirtz, J. C. (2017). Mice lacking the chromodomain helicase DNA-binding 5 chromatin remodeler display autism-like characteristics. *Transl. Psychiatry* 7:e1152. doi: 10.1038/tp.2017.111

Ren, W., Medeiros, N., Warneford-Thomson, R., Wulfridge, P., Yan, Q., Bian, J., et al. (2020). Disruption of ATRX-RNA interactions uncovers roles in ATRX localization and PRC2 function. *Nat. Commun.* 11:2219. doi: 10.1038/s41467-020-15902-15909

Rhodes, D., and Lipps, H. J. (2015). Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* 43, 8627–8637. doi: 10.1093/nar/gkv862

Ribeiro de Almeida, C., Dhir, S., Dhir, A., Moghaddam, A. E., Sattentau, Q., Meinhart, A., et al. (2018). RNA Helicase DDX1 converts RNA G-quadruplex structures into R-loops to promote IgH class switch recombination. *Mol. Cell* 70, 650–662. doi: 10.1016/j.molcel.2018.04.001

Ritchie, K., Seah, C., Moulin, J., Isaac, C., Dick, F., and Bérubé, N. G. (2008). Loss of ATRX leads to chromosome cohesion and congression defects. *J. Cell Biol.* 180, 315–324. doi: 10.1083/jcb.200706083

Ritchie, K., Watson, L. A., Davidson, B., Jiang, Y., and Bérubé, N. G. (2014). ATRX is required for maintenance of the neuroprogenitor cell pool in the embryonic mouse brain. *Biol. Open* 3, 1158–1163. doi: 10.1242/bio.20148730

Ronan, J. L., Wu, W., and Crabtree, G. R. (2013). From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet* 14, 347–359. doi: 10.1038/jid.2014.371

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., et al. (2015). Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep.* 16, 836–850. doi: 10.15252/embr.201439937

Sanlaville, D., Etchevers, H. C., Gonzales, M., Martinovic, J., Clément-Ziza, M., Delezoide, A. L., et al. (2006). Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J. Med. Genet.* 43, 211–217. doi: 10.1136/jmg.2005.036160

Santen, G. W. E., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* 44, 379–380. doi: 10.1038/ng.2217

Santen, G. W. E., Aten, E., Vulto-van Silfhout, A. T., Pottinger, C., van Bon, B. W. M., van Minderhout, I. J. H. M., et al. (2013). Coffin-siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum. Mutat.* 34, 1519–1528. doi: 10.1002/humu.22394

Sarma, K., Cifuentes-Rojas, C., Ergun, A., del Rosario, A., Jeon, Y., White, F., et al. (2014). ATRX directs binding of PRC2 to Xist RNA and polycomb targets. *Cell* 159, 869–883. doi: 10.1016/j.cell.2014.10.019

Satterstrom, F. K., Kosmicki, J. A., Wang, J., and Breen, M. S. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* 180, 568–584.

Saugier-Veber, P., Munnich, A., Lyonnet, S., Toutain, A., Moraine, C., Piussan, C., et al. (1995). Lumping Juberg-Marsidi syndrome and X-linked alpha-thalassemia/mental retardation syndrome? *Am. J. Med. Genet.* 55, 300–301. doi: 10.1002/ajmg.1320550310

Schwartzentruber, J., Korshunov, A., Liu, X. Y., Jones, D. T. W., Pfaff, E., Jacob, K., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482, 226–231. doi: 10.1038/nature10833

Scott, T., Guo, H., Eichler, E., Rosenfeld, J., Pang, K., Liu, Z., et al. (2020). BAZ2B haploinsufficiency as a cause of developmental delay, intellectual disability, and autism spectrum disorder. *Hum. Mutat.* 41, 921–925.

Shang, L., Cho, M. T., Retterer, K., Folk, L., Humberson, J., Rohena, L., et al. (2015). Mutations in ARID2 are associated with intellectual disabilities. *Neurogenetics* 16, 307–314. doi: 10.1007/s10048-015-0454-450

Shioda, N., Beppu, H., Fukuda, T., Li, E., Kitajima, I., and Fukunaga, K. (2011). Aberrant calcium/calmodulin-dependent protein kinase II (CaMKII) activity is associated with abnormal dendritic spine morphology in the ATRX mutant mouse brain. *J. Neurosci.* 31, 346–358. doi: 10.1523/JNEUROSCI.4816-10.2011

Shioda, N., Yabuki, Y., Yamaguchi, K., Onozato, M., Li, Y., Kurosawa, K., et al. (2018). Targeting G-quadruplex DNA as cognitive function therapy for ATR-X syndrome article. *Nat. Med.* 24, 802–813. doi: 10.1038/s41591-018-0018-6

Snijders Blok, L., Rousseau, J., Twist, J., Ehresmann, S., Takaku, M., Venselaar, H., et al. (2018). CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-018-06014-6016

Sokpor, G., Xie, Y., Rosenbusch, J., and Tuoc, T. (2017). Chromatin remodeling BAF (SWI/SNF) complexes in neural development and disorders. *Front. Mol. Neurosci.* 10:243. doi: 10.3389/fnmol.2017.00243

Solomon, L. A., Li, J. R., Bérubé, N. G., and Beier, F. (2009). Loss of ATRX in chondrocytes has minimal effects on skeletal development. *PLoS One* 4:7106. doi: 10.1371/journal.pone.0007106

Solomon, L. A., Russell, B. A., Makar, D., Bérubé, N. G., and Beier, F. (2013). Loss of ATRX does not confer susceptibility to osteoarthritis. *PLoS One* 8:85526. doi: 10.1371/journal.pone.0085526

Stankiewicz, P., Khan, T. N., Szafranski, P., Slattery, L., Streff, H., Vetrini, F., et al. (2017). Haploinsufficiency of the chromatin remodeler BPTF causes syndromic developmental and speech delay, postnatal microcephaly, and dysmorphic features. *Am. J. Hum. Genet.* 101, 503–515. doi: 10.1016/j.ajhg.2017.08.014

Sun, X., Morozova, T., and Sonnenfeld, M. (2006). Glial and neuronal functions of the Drosophila homolog of the human SWI/SNF gene ATR-X (DATR-X) and the jing zinc-finger gene specify the lateral positioning of longitudinal glia and axons. *Genetics* 173, 1397–1415. doi: 10.1534/genetics.106.057893

Tamming, R. J., Siu, J. R., Jiang, Y., Prado, M. A. M., Beier, F., and Bérubé, N. G. (2017). Mosaic expression of Atrx in the mouse central nervous system causes memory deficits. *DMM Dis. Model. Mech.* 10, 119–126. doi: 10.1242/dmm.027482

Tang, J., Wu, S., Liu, H., Stratt, R., Barak, O. G., Shiekhattar, R., et al. (2004). A novel transcription regulatory complex containing death domain-associated protein and the ATR-X Syndrome protein. *J. Biol. Chem.* 279, 20369–20377. doi: 10.1074/jbc.M401321200

Torigoe, S. E., Urwin, D. L., Ishii, H., Smith, D. E., and Kadonaga, J. T. (2011). Identification of a rapidly formed nonnucleosomal histone-DNA intermediate that is converted into chromatin by ACF. *Mol. Cell* 43, 638–648. doi: 10.1016/j.molcel.2011.07.017

Toubiana, S., and Selig, S. (2018). DNA:RNA hybrids at telomeres - when it is better to be out of the (R) loop. *FEBS J.* 285, 2552–2566. doi: 10.1111/febs.14464

Tsurusaki, Y., Koshimizu, E., Ohashi, H., Phadke, S., Kou, I., Shiina, M., et al. (2014a). De novo SOX11 mutations cause Coffin-Siris syndrome. *Nat. Commun.* 5, 1–7. doi: 10.1038/ncomms5011

Tsurusaki, Y., Okamoto, N., Ohashi, H., Mizuno, S., Matsumoto, N., Makita, Y., et al. (2014b). Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin. Genet.* 85, 548–554. doi: 10.1111/cge.12225

Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* 44, 376–378. doi: 10.1038/ng.2219

Udugama, M., Chang, F. T. M., Chan, F. L., Tang, M. C., Pickett, H. A., McGhie, J. D. R., et al. (2015). Histone variant H3.3 provides the heterochromatic *H3* lysine 9 tri-methylation mark at telomeres. *Nucleic Acids Res.* 43, 10227–10237. doi: 10.1093/nar/gkv847

Valton, A. L., and Prioleau, M. N. (2016). G-Quadruplexes in DNA replication: a problem or a necessity? *Trends Genet.* 32, 697–706. doi: 10.1016/j.tig.2016.09.004

van der Sluijs, P. J., Jansen, S., Vergano, S. A., Adachi-Fukuda, M., Alanay, Y., AlKindy, A., et al. (2019). The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome. *Genet. Med.* 21, 1295–1307. doi: 10.1038/s41436-018-0330-z

Van Houdt, J. K. J., Nowakowska, B. A., Sousa, S. B., Van Schaik, B. D. C., Seuntjens, E., Avonce, N., et al. (2012). Heterozygous missense mutations in SMARCA2 cause nicolaides-baraitser syndrome. *Nat. Genet.* 44, 445–449. doi: 10.1038/ng.1105

Varshney, D., Spiegel, J., Zyner, K., Tannahill, D., and Balasubramanian, S. (2020). The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* 21, 459–474. doi: 10.1038/s41580-020-0236-x

Vasileiou, G., Vergarajauregui, S., Endele, S., Popp, B., Büttner, C., Ekici, A. B., et al. (2018). Mutations in the BAF-complex subunit DPF2 are associated with coffin-siris syndrome. *Am. J. Hum. Genet.* 102, 468–479. doi: 10.1016/j.ajhg.2018.01.014

Verma, P., and Greenberg, R. A. (2016). Noncanonical views of homology-directed DNA repair. *Genes Dev.* 30, 1138–1154. doi: 10.1101/gad.280545.116

Vissers, L. E. L. M., Van Ravenswaaij, C. M. A., Admiraal, R., Hurst, J. A., De Vries, B. B. A., Janssen, I. M., et al. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.* 36, 955–957. doi: 10.1038/ng1407

Voon, H. P. J., Hughes, J. R., Rode, C., DeLaRosa-Velázquez, I. A., Jenuwein, T., Feil, R., et al. (2015). ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. *Cell Rep.* 11, 405–418. doi: 10.1016/j.celrep.2015.03.036

Wada, T., Ban, H., Matsufuji, M., Okamoto, N., Enomoto, K., Kurosawa, K., et al. (2013). Neuroradiologic features in α-linked α-thalassemia/mental retardation syndrome. *Am. J. Neuroradiol.* 34, 2034–2038. doi: 10.3174/ajnr.A3560

Watson, L. A., Goldberg, H., and Berube, N. G. (2015). Emerging roles of ATRX in cancer. *Epigenomics* 7, 1365–1378. doi: 10.2217/epi.15.82

Watson, L. A., Solomon, L. A., Li, J. R., Jiang, Y., Edwards, M., Shin-Ya, K., et al. (2013). Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *J. Clin. Invest.* 123, 2049–2063. doi: 10.1172/JCI65634

Weiss, K., Lazar, H. P., Kurolap, A., Martinez, A. F., Paperna, T., Cohen, L., et al. (2020). Correction: the CHD4-related syndrome: a comprehensive investigation of the clinical spectrum, genotype-phenotype correlations, and molecular basis (Genetics in Medicine, (2020), 22, 2, (389-397), 10.1038/s41436-019-0612-0). *Genet. Med.* 22:669. doi: 10.1038/s41436-019-0727-3

Wieczorek, D., Bögershausen, N., Beleggia, F., Steiner-Haldenstätt, S., Pohl, E., Li, Y., et al. (2013). A comprehensive molecular study on coffin-siris and nicolaides-baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.* 22, 5121–5135. doi: 10.1093/hmg/ddt366

Wong, L. H., McGhie, J. D., Sim, M., Anderson, M. A., Ahn, S., Hannan, R. D., et al. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res.* 20, 351–360. doi: 10.1101/gr.101477.109

Wong, L. H., Ren, H., Williams, E., McGhie, J., Ann, S., Sim, M., et al. (2009). Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome Res.* 19, 404–414. doi: 10.1101/gr.084947.108

Wu, Y., Chen, L., Scott, P. G., and Tredget, E. E. (2007). Mesenchymal stem cells enhance wound healing through differentiation and angiogenesis. *Stem Cells* 25, 2648–2659. doi: 10.1634/stemcells.2007-2226

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., et al. (1999). Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* 236, 87–95. doi: 10.1016/S0378-1119(99)00252-258

Xue, Y., Gibbons, R., Yan, Z., Yang, D., McDowell, T. L., Sechi, S., et al. (2003). The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10635–10640. doi: 10.1073/pnas.1937626100

Yadav, S., Younger, S. H., Zhang, L., Thompson-Peer, K. L., Li, T., Jan, L. Y., et al. (2019). Glial ensheathment of the somatodendritic compartment regulates sensory neuron structure and activity. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5126–5134. doi: 10.1073/pnas.1814456116

Yip, D. J., Corcoran, C. P., Alvarez-Saavedra, M., DeMaria, A., Rennick, S., Mears, A. J., et al. (2016). Snf2l regulates foxg1-dependent progenitor cell expansion in the developing brain. *Dev. Cell* 22, 871–878. doi: 10.1016/j.devcel.2012.01.020.Yip

Zaghlool, A., Halvardson, J., Zhao, J. J., Etemadikhah, M., Kalushkova, A., Konska, K., et al. (2016). A role for the chromatin-remodeling factor BAZ1A in neurodevelopment. *Hum. Mutat.* 37, 964–975. doi: 10.1002/humu.23034

Zhou, R., Zhang, J., Bochman, M. L., Zakian, V. A., and Ha, T. (2014). Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. *eLife* 3:e02190. doi: 10.7554/elife.02190

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 21 August 2020
doi: 10.3389/fcell.2020.00763

The MeCP2E1/E2-BDNF-*miR132* Homeostasis Regulatory Network Is Region-Dependent in the Human Brain and Is Impaired in Rett Syndrome Patients

*Shervin Pejhan[1], Marc R. Del Bigio[2] and Mojgan Rastegar[1]**

[1] Regenerative Medicine Program, Department of Biochemistry and Medical Genetics, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada, [2] Department of Pathology, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

OPEN ACCESS

Edited by:
Enrique Medina-Acosta, State University of the North Fluminense Darcy Ribeiro, Brazil

Reviewed by:
Maurizio D'Esposito, National Research Council (CNR), Italy
Zilong Qiu, Shanghai Institutes for Biological Sciences (CAS), China

***Correspondence:**
Mojgan Rastegar
mojgan.rastegar@umanitoba.ca

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Cell and Developmental Biology

Received: *06 April 2020*
Accepted: *21 July 2020*
Published: *21 August 2020*

Citation:
Pejhan S, Del Bigio MR and Rastegar M (2020) The MeCP2E1/E2-BDNF-miR132 Homeostasis Regulatory Network Is Region-Dependent in the Human Brain and Is Impaired in Rett Syndrome Patients. Front. Cell Dev. Biol. 8:763. doi: 10.3389/fcell.2020.00763

Rett Syndrome (RTT) is a rare and progressive neurodevelopmental disorder that is caused by *de novo* mutations in the X-linked Methyl CpG binding protein 2 (*MECP2*) gene and is subjected to X-chromosome inactivation. RTT is commonly associated with neurological regression, autistic features, motor control impairment, seizures, loss of speech and purposeful hand movements, mainly affecting females. Different animal and cellular model systems have tremendously contributed to our current knowledge about MeCP2 and RTT. However, the majority of these findings remain unexamined in the brain of RTT patients. Based on previous studies in rodent brains, the highly conserved neuronal microRNA "*miR132*" was suggested to be an inhibitor of MeCP2 expression. The neuronal *miR132* itself is induced by Brain Derived Neurotrophic Factor (BDNF), a neurotransmitter modulator, which in turn is controlled by MeCP2. This makes the basis of the *MECP2-BDNF-miR132* feedback regulatory loop in the brain. Here, we studied the components of this feedback regulatory network in humans, and its possible impairment in the brain of RTT patients. In this regard, we evaluated the transcript and protein levels of *MECP2*/MeCP2E1 and E2 isoforms, *BDNF*/BDNF, and *miR132* (both *3p* and *5p* strands) by real time RT-PCR, Western blot, and ELISA in four different regions of the human RTT brains and their age-, post-mortem delay-, and sex-matched controls. The transcript level of the studied elements was significantly compromised in RTT patients, even though the change was not identical in different parts of the brain. Our data indicates that MeCP2E1/E2-BDNF protein levels did not follow their corresponding transcript trends. Correlational studies suggested that the *MECP2E1/E2-BDNF-miR132* homeostasis regulation might not be similarly controlled in different parts of the human brain. Despite challenges in evaluating autopsy samples in rare diseases, our findings would help to shed some light on RTT pathobiology, and obscurities caused by limited studies on MeCP2 regulation in the human brain.

Keywords: Rett Syndrome, MeCP2 isoforms, MeCP2E1, MeCP2E2, BDNF, *miR132*, Epigenetics, DNA methylation

INTRODUCTION

The X-linked Methyl CpG binding protein 2 (*MECP2*) gene encodes for an important nuclear factor that binds to methylated DNA. MeCP2 protein is the prototype member of the methyl binding proteins, which was discovered almost three decades ago (Lewis et al., 1992). MeCP2 is a multi-functional epigenetic factor and its proper homeostasis regulation is critical for normal brain development (Delcuve et al., 2009; Zachariah and Rastegar, 2012; Rastegar, 2017b; Tillotson and Bird, 2019). Mutations in the *MECP2* gene located in *Xq28* are the underlying cause of over 90% of typical and a portion of atypical cases of Rett Syndrome (RTT) (Neul et al., 2008; Liyanage and Rastegar, 2014; Chahil and Bollu, 2020), a severe neurodevelopmental disorder that is primarily detected in females (Amir et al., 1999). The vast majority of these mutations happen *de novo* and mostly in the paternal germline (Girard et al., 2001; Trappe et al., 2001). There are also rare familial cases of RTT that may happen due to skewing of X chromosome inactivation (XCI) in the mother, who can pass *MECP2* mutation(s) on to male and female offspring as an asymptomatic carrier (Villard et al., 2000). The pattern of XCI has been studied in RTT, however, the results are not fully conclusive (Xinhua et al., 2008). Although the majority of classical cases of RTT have balanced XCI pattern in the brain tissue (Zoghbi et al., 1990; Shahbazian et al., 2002a,b), non-random XCI or skewed pattern, has also been reported in several cases (Renieri et al., 2003).

In mice and humans, the *Mecp2*/*MECP2* gene consists of 4 coding exons separated by 3 introns (Liyanage and Rastegar, 2014). Although several protein-coding and non-coding *MECP2* transcripts have been identified or predicted (Singh et al., 2008), the two most-studied splice variants of MeCP2 are known as MeCP2E1 and MeCP2E2 (Kriaucionis and Bird, 2004; Mnatzakanian et al., 2004). Accordingly, MeCP2E1 and E2 isoform-specific expression, regulation, function, and clinical relevance to RTT have been the focus of several research studies (Rastegar et al., 2009; Itoh et al., 2012; Kerr et al., 2012; Zachariah et al., 2012; Liyanage et al., 2013, 2019; Olson et al., 2014; Yasui et al., 2014; Sheikh et al., 2017; Vogel Ciernia et al., 2018; Martinez de Paz et al., 2019; Takeguchi et al., 2020). Information about other *MECP2* transcripts and their tissue- and cell-type specific expression could be obtained through RNA sequencing or data mining of related publically available data repositories. Among the known *MECP2* protein-coding transcripts, MeCP2E2 was the original protein variant that was discovered in 1992 (also referred to as MeCP2β or MeCP2A) (Lewis et al., 1992). In 2004, another splice variant of MeCP2 was identified, currently known as MeCP2E1 (also called MeCP2α or MeCP2B) (Kriaucionis and Bird, 2004; Mnatzakanian et al., 2004). MeCP2E1 is encoded by exons 1, 3, and 4; while MeCP2E2 is encoded by exons 2, 3, and 4. The two MeCP2E1 (498-aa) and E2 (486-aa) isoforms fully share MeCP2 protein domains and only differ in their short N-terminal regions that consist of 21-aa exclusive to the MeCP2E1, and 9-aa specific to MeCP2E2 (Kriaucionis and Bird, 2004; Mnatzakanian et al., 2004; Zachariah et al., 2012; Liyanage and Rastegar, 2014; Olson et al., 2014). The larger MeCP2E1 isoform is the dominant protein in the brain with relatively consistent expression throughout different brain regions of the adult mice, and earlier developmental onset of expression in the brain (Olson et al., 2014). Interestingly, E1-deficiency in mice mimics similar phenotype as mice lacking the whole *Mecp2* gene (Yasui et al., 2014), suggesting its importance in RTT pathobiology. In support of this conclusion, E2-deficient mice do not exhibit RTT-associated symptoms (Itoh et al., 2012), suggesting that E2 may not have direct relevance to RTT.

In the brain, compromised MeCP2 expression and/or function is associated with impaired brain function associated with autism, RTT, and *MECP2* duplication Syndrome (MDS) (Liyanage and Rastegar, 2014). Our previous studies have highlighted the cell type-specific role of DNA methylation, and murine strain in *Mecp2*/MeCP2E1/E2 regulation (Liyanage et al., 2013, 2015, 2019; Olson et al., 2014; Xu et al., 2019; Amiri et al., 2020). Studies have also shown that in rodent brain/brain cells MeCP2 homeostasis is controlled *via* a regulatory feedback loop consisting of MeCP2, brain-derived neurotrophic factor (BDNF), and a neuronal-specific microRNA named *miR132* (Klein et al., 2007). The monogenic character of RTT, mainly caused by mutations in the *MECP2* gene has prompted several RTT animal, and development of cellular (Rastegar et al., 2009; Marchetto et al., 2010; Yazdani et al., 2012; Ezeonwuka and Rastegar, 2014) model systems to study disease mechanism. However, the brain region-specific expression pattern of the suggested MeCP2-BDNF-*miR132* regulatory network in the human brains is unknown. Furthermore, it is not clear if the components of this homeostasis regulatory network are affected in the human RTT brain, mainly due to the limitation of access to human brain tissues for this rare disorder (1:10,000–15,000 females) (Amir et al., 1999).

Here, we investigated the expression of the suggested MeCP2 homeostasis regulatory network, while studying both MeCP2 isoforms in post-mortem human brain tissues of RTT patients and age-/sex-matched controls. We employed quantitative and semi-quantitative approaches including real time reverse transcriptase-PCR (RT-PCR), ELISA, and Western blot (WB) to study the components of MeCP2E1/E2-BDNF-*miR132* feedback loop at the transcript and protein levels. Our results indicate that *MECP2E1* transcripts are higher than *MECP2E2* in the human frontal cerebrum, hippocampus, amygdala, and cerebellum, and that the latter shows significantly higher levels of *MECP2E1* and *MECP2E2* compared to other tested brain regions. We further show that not only *MECP2E1* and *MECP2E2*, but also *BDNF* transcripts are significantly lower in RTT brains compared to controls. We observed that *miR132-3p* is the dominant strand in all studied human brain regions with significantly lower levels in the cerebellum. Combined correlational analysis of our transcript and protein studies suggests a potential regulatory feedback of *MECP2*/MeCP2-*BDNF*/BDNF in the brain, which might be independent of *miR132* in specific parts of the human brain.

Abbreviations: BDNF (protein)/*BDNF* (human gene), brain derived neurotrophic factor; ELISA, enzyme-linked immunosorbent assay; GAPDH (protein)/*GAPDH* (human gene), Glyceraldehyde 3-phosphate dehydrogenase; H&E, Hematoxylin and Eosin; IHC, immunohistochemistry; MeCP2 (protein)/*MECP2* (human gene)/*Mecp2* (mouse gene), Methyl CpG binding Protein 2; *miR132, microRNA132*; PMD, post-mortem delay; qRT-PCR, quantitative reverse transcription polymerase chain reaction; RTT, Rett syndrome; WB, Western blot.

MATERIALS AND METHODS

Ethics Statement

Our research studies on the human brain tissues presented in this manuscript were reviewed and approved by the "University of Manitoba Bannatyne Campus Research Ethics Board" and under an approved Health Research Board protocol #HS20095 (H2016:337).

Human Brain Tissues

Autopsy brain samples of three female RTT patients were analyzed in this study in parallel to their comparable age- and sex-matched controls. RTT brain samples were from: (A) a 13-year-old female Rett Syndrome patient with the most frequent MeCP2 mutation (T158M). The brain tissues for this patient was received by organ donation, and the related clinical data are already reported (Olson et al., 2018), (B) a 19-year-old female RTT patient (A201V). The brain tissue for this patient was also received through direct organ donation, and (C) the third patient was a 20-year-old female RTT patient (R255X), for whom the brain samples were received from NIH Neurobiobank (NIH # 4516). Tissues from three comparable female control (non-RTT) brain tissues were received from NIH Neurobiobank. **Table 1** provides the information about these brain samples. Two patients with T158M and R255X mutations are representatives of the first and third most common RTT-associated point mutations that cause the disease (RettBASE Database) (Krishnaraj et al., 2017). These two mutations (T158M and R255X) affect two of the main protein functional domains, namely the methyl binding domain (involved in DNA binding of MeCP2) and the transcriptional repression domain (involved in protein-protein interactions), respectively (Liyanage and Rastegar, 2014). The third patient with clinical diagnosis of RTT had a polymorphism (A201V) known to be associated with RTT. Based on potential link to disease symptoms such as cognitive, behavioral, and locomotor impairments in RTT patients, four brain regions were selected for our study that included the frontal cerebrum, hippocampus, amygdala, and cerebellum. Hematoxylin and Eosin (H&E) stained slides of the human brain tissue were prepared by paid services from the Health Sciences Center, Pathology Laboratory, Winnipeg, and Department of Human Anatomy and Cell Science, University of Manitoba. Imaging and analysis of the images were done by the authors.

TABLE 1 | The list of human brain samples used in our study.

NIH#	RTT/Control	Age (Years)	Sex	PMD
*	T158M	13	F	<6 h
*	A201V	19	F	24 h
4516	R255X	20	F	9 h
5401	Control	19	F	22 h
5646	Control	20	F	23 h
5287	Control	23	F	15 h
¤	Control	14	F	24 h
¤	Control	18	F	48 h

All patients and controls are Caucasian. Two RTT brain samples (T158M and A201V) were donated directly to our lab for research (*), and were coordinated through collaboration with Dr. Victoria Siu (Western University) who obtained proper consent for research, as reported previously (Olson et al., 2018; Pejhan et al., 2020). Two control brain samples for Hematoxylin and Eosin (H&E) staining were provided by the Health Sciences Centre Pathology Laboratory, Winnipeg, MB (¤), coordinated by Dr. Marc Del Bigio. Post-mortem delay (PMD) shows the interval between death and tissue preservation. Samples from NIH NeuroBioBank were obtained by Dr. Mojgan Rastegar and signed Material Transfer Agreement (MTA); F (Female).

Quantitative Real Time Reverse-Transcriptase Polymerase Chain Reaction (qRT-PCR)

Total RNA was extracted from the frontal cerebrum, hippocampus, amygdala, and cerebellum by using Trizole (Life Technologies Inc., 15596-062) as per manufacturer's guidelines, and as previously reported (Rastegar et al., 2004; Huang et al., 2005; Kobrossy et al., 2006). The extracted total RNA was treated with DNase I for eliminating traces of possible contamination with genomic DNA, by using Ambion TURBO DNA-free kit (Thermo Fisher Scientific). Total RNA was then processed for cDNA synthesis by Superscript III Reverse Transcriptase (Invitrogen), as described elsewhere (Pejhan et al., 2020). For quantitative RT-PCR, we used a Fast 7500 Real-Time PCR machine (Applied Biosystems) with SYBR Green-based RT-PCR Master Mix reagents (Applied Biosystems). Transcripts levels of *MECP2* isoforms and *BDNF* were assessed using specific primers that are listed in **Supplementary Table S1**, being normalized to *GAPDH* (*Glyceraldehyde 3-phosphate dehydrogenase)*, as previously described (Barber et al., 2013; Xu et al., 2019). We used Microsoft Excel 2.10 and GraphPad Prism 7 to analyze the results.

TaqMan Assay for *miR132* Detection

The TaqMan®MicroRNA assay kit from Thermo Fisher was used to analyze and quantify *miR132*. First cDNA was reverse transcribed from DNase-treated RNA by TaqMan®MicroRNA reverse transcription kit (Thermo Fisher Cat. No. 4366596), as per manufacturer's protocol and guidelines. The generated cDNA was used to perform RT-PCR reaction for *miR132-3p*, *miR132-5p*, and *U6* with the TaqMan primers from Thermo Fisher (hsa-*miR132-3p*, Cat. No. 000457, hsa-*miR132-5p*, Cat. No. 002132, and *U6* snRNA, Cat. No. 001973). We selected *U6* snRNA as an internal miRNA control, as suggested by previous studies (McDermott et al., 2013; Duan et al., 2018). The results for two *miR132* strands were normalized to *U6* snRNA (endogenous control), with the same analysis method as the other studied genes here.

Western Blot

Nuclear and cytoplasmic protein fractions were isolated from frozen frontal cerebrum, hippocampus, amygdala, and cerebellum samples by NE-PER Extraction Kit (Thermo Scientific, Cat. No. 78833), according to the manufacturer's instruction and as described previously (Olson et al., 2014, 2018). Western blot experiments were performed with the described protocols in previous reports (Rastegar et al., 2000; Wu et al., 2001; Gordon et al., 2003; Sheikholeslami et al., 2019). As housekeeping protein and loading control, GAPDH was checked on all membranes and the results for each antibody were

normalized to GAPDH signals. Validation of GAPDH in the nuclear and cytoplasmic extracts from human RTT and control brain tissues has been also reported recently (Olson et al., 2018). The list of primary and secondary antibodies is provided in **Supplementary Table S2**. MeCP2E1 and MeCP2E2 antibodies are custom-made and have been previously characterized and reported for detecting endogenous protein isoforms (Zachariah et al., 2012; Liyanage et al., 2013; Olson et al., 2014; Yasui et al., 2014). BDNF antibody was purchased from Abcam (108319).

Enzyme-Linked Immunosorbent Assay (ELISA)

We used ELISA to quantify BDNF protein in the cytoplasmic protein extracts of the human brain tissues. Human BDNF ELISA kit from Sigma-Aldrich (RAB0026) was used according to the manufacturer's instructions. For intra and inter-assay coefficients of variation, we used <10% and <12%, respectively, while the sensitivity was determined at 80 pg/ml. Absorbance was measured on SpectraMax M2e plate reader (Molecular Devices), and concentrations were calculated based on a standard curve that was generated by the Softmax Pro 5.3. BDNF levels were calculated as ng/mg of total protein.

Statistical Analysis

GraphPad Prism software was used for statistical analysis and generating the graphs, as reported (Barber et al., 2013; Liyanage et al., 2015; Sheikholeslami et al., 2019). Statistical significance was determined by one-way ANOVA with multiple pairwise comparisons among different brain regions for transcript assessments and protein quantifications by ELISA. We used Student's *t*-test to compare controls and RTT samples for these parametric data. Western blot non-parametric results were analyzed by Kruskal-Wallis, and Mann-Whitney tests. Pearson correlation study was used to measure the degree of the relationship between parametric variables, and Spearman's for non-parametric ones. The significance level was calculated at $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

RESULTS

MECP2E1 and MECP2E2 Transcripts Are Significantly Lower in RTT Patients Compared to Controls

In order to study *MECP2E1* and *MECP2E2* transcript levels, we performed qRT-PCR in isolated RNA samples from the frontal cerebrum, hippocampus, amygdala, and cerebellum of RTT patients and age-/sex-matched control brain tissues. Our results indicated that *MECP2E1* transcripts are significantly higher than *MECP2E2* transcripts in all the studied brain regions of the controls and RTT patients (**Figures 1A,B**). In control brain tissues, both isoforms showed higher transcript levels in the cerebellum compared to other brain regions, with a similar trend in RTT patients. Comparing *MECP2E1* and *MECP2E2* transcript levels in RTT patients versus controls, we observed significantly lower levels of *MECP2E1* and *MECP2E2* isoforms in RTT patients in all four tested brain regions (**Figure 1C**). The severity of such significantly reduced expression was brain region-dependent, ranging from 20% in the hippocampus to 60% in the frontal cerebrum (**Table 2**).

Transcript Detection of *miR132* and BDNF in the Brain of RTT Patients and Controls

Next, we aimed to investigate the transcript levels of *miR132* (both *3p* and *5p* strands) in the brain of RTT patients and control tissue samples. Isolated RNA from the selected four brain regions was subjected to TaqMan real time PCR. Our results indicated that between the two strands, *miR132-3p* is the dominant transcript in all tested brain regions of control and RTT brain tissue samples. Despite its significantly lower level, the *miR132-5p* strand still showed a similar pattern in comparison with the *miR132-3p* among different brain regions of RTT and controls (**Figures 1D–F**). Analysis of these results indicated that compared to controls, RTT brain regions had significantly lower levels of *miR132-3p* in the frontal cerebrum, hippocampus, and amygdala, but they were equal in the cerebellum (**Figure 1F**). The minimum ratio belonged to RTT frontal cerebrum with less than 18% of the control levels ($P < 0.001$). Amygdala with less than 43% ($P < 0.01$), and hippocampus with less than 66% of control level ($P < 0.05$) were the next two regions.

In order to study the components of MeCP2 homeostasis regulation, next we analyzed *BDNF* transcripts in the control and RTT brain tissues. In control brain samples, *BDNF* transcripts were relatively consistent in the frontal cerebrum, hippocampus, and amygdala. However, *BDNF* transcripts were found to be about 1.5 times higher in the cerebellum compared to the other tested brain regions ($P < 0.0001$) (**Figure 1G**). Similar analysis of RTT brain tissues showed significantly higher *BDNF* levels in the cerebellum compared to the frontal cerebrum ($P < 0.0001$) and amygdala ($P < 0.01$). Although, *BDNF* transcripts in the cerebellum were still higher compared to hippocampus, the detected change did not reach to any statistical significance (**Figure 1H**). Next, we compared BDNF transcript levels of control and RTT brain tissues. In all four tested brain regions of RTT samples; *BDNF* transcripts were significantly lower compared to controls (**Figure 1I**). Such significantly lower levels in RTT brains ranged from 30% reduced *BDNF* levels in the hippocampus ($P < 0.001$) and amygdala ($P < 0.01$) to 60% in the frontal cerebrum ($P < 0.0001$) (**Table 2**). Taken together, our results indicate that not only *miR132* (both *miR132-3p* and *miR132-5p)* but also *BDNF* transcripts are significantly lower in RTT patients, and that the *miR132-3p* is the dominant microRNA strand in the human brain, at least in the four tested brain regions in this study.

Expression Analysis of MeCP2 Isoforms in RTT and Control Human Brain Tissues

Others and us have reported that MeCP2 protein and transcript levels do not always correlate with each other (Balmer et al., 2003; Mullaney et al., 2004; Liyanage et al., 2013; McGowan and Pang, 2015). Therefore, we next analyzed the protein levels of MeCP2E1 and MeCP2E2 in the control and RTT patient brain tissues. As MeCP2 is a nuclear protein, we isolated the

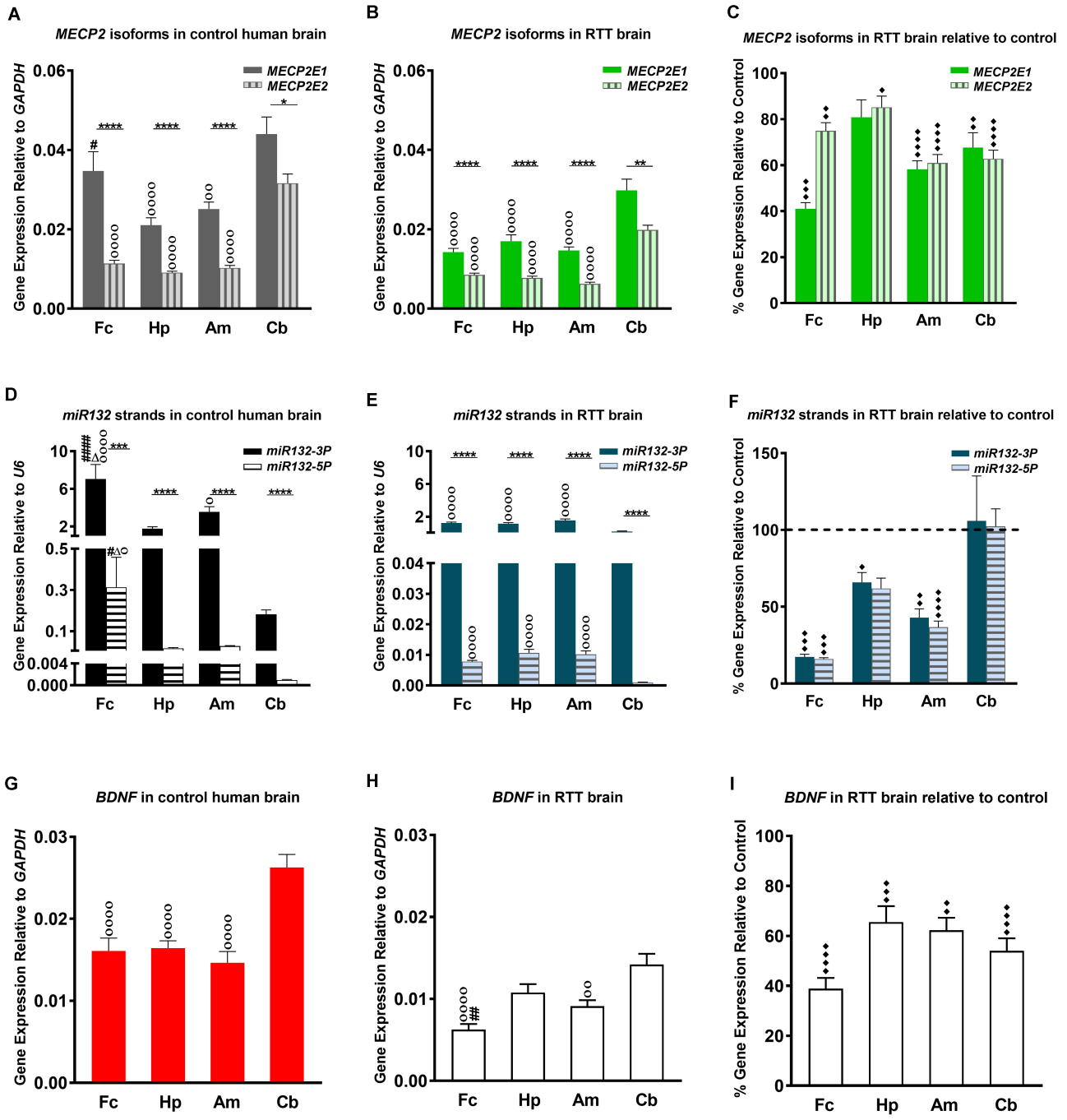

FIGURE 1 | Transcript levels of *MECP2E1*, *MECP2E2*, *BDNF*, and *miR132*. In all studied brain regions of three RTT **(B,E)** and controls **(A,D)**, *MECP2E1* and *miR132-3P* are significantly higher than *MECP2E2* and *miR132-5P*, respectively. Cerebellum in comparison with the other three brain regions shows the highest level for all the studied transcripts except for *microRNA132* **(A,B,D,E,G,H)**. Rett Syndrome (RTT) patients show lower transcript levels of *MECP2E1*, *MECP2E2*, *BDNF*, *miR132-3P*, and *miR132-5P* in four studied brain regions compared to their age-, and sex- matched controls **(C,F,I)**. Cerebellum has exceptionally similar level of *miRNAs* in RTT and controls **(F)**. Frontal cerebrum (Fc), Hippocampus (Hp), Amygdala (Am), Cerebellum (Cb); $N = 3 \pm$ Standard Error of the Mean (SEM); One-way ANOVA with multiple pairwise comparisons used among different brain regions. Student's *t*-test was used to compare controls and RTT samples for each parameter. The significance level was determined at single symbol $P < 0.05$, double symbols $P < 0.01$, triple symbols $P < 0.001$, and quadruple symbols $P < 0.0001$.

TABLE 2 | Summary of changes at transcript level in the studied element of four brain regions in RTT patients compared to controls.

	MECP2E1	*MECP2E2*	*BDNF*	*miR132-3p*
Frontal Cerebrum	~60%↓	~30%↓	~60%↓	~80%↓
Hippocampus	~20%↓ NS	~20%↓	~30%↓	~30%↓
Amygdala	~40%↓	~40%↓	~30%↓	~60%↓
Cerebellum	~30%↓	~40%↓	~40%↓	~10%↑ NS

NS, not statistically significant.

nuclear and cytoplasmic extracts of frozen brain tissues of RTT and control brain samples, including the frontal cerebrum, hippocampus, amygdala, and cerebellum. WB experiments on isolated nuclear extracts with our validated custom-made anti-MeCP2E1 and MeCP2E2 antibodies (Zachariah et al., 2012; Olson et al., 2014), indicated that both MeCP2E1 and MeCP2E2 proteins are detectable in RTT and control brain tissues at 72–75 kDa (**Figure 2**). Despite consistent detection of GAPDH in different samples (indicating comparable loading from each sample), MeCP2 isoforms were inconsistently detected in both the controls and RTT samples. In this regard, the 20 year-old control brain tissues showed lower level of both MeCP2 isoforms in the frontal cerebrum. Also, MeCP2E2 showed relatively low levels in two out of the three controls in the majority of tested brain regions (**Figure 2**). The post-mortem delay (PMD) of these control samples (**Figure 2**) from left to right are as following: 22 h (NIH Neurobiobank #5401, 19 years old female), 23 h (NIH Neurobiobank #5646, 20 years old female), and 15 h (NIH Neurobiobank #5287, 23 years old female). In case of any impact of PMD on protein detection of MeCP2E1 and MCP2E2, it is possible that our results on MeCP2 protein levels may not properly reflect the actual intact protein level, and that the protein may have undergone some degradation. We recently reported that the highest intact detection of MeCP2 in surgical samples is associated with the 1 h time-window after sample collection (Pejhan et al., 2020), which may not be feasible to achieve in all cases of post-mortem organ donations. In line with the detected MeCP2 levels with high variations, most RTT brain regions also showed inconsistent levels. Of note, the two RTT brain tissues (T158M and A201V) that we received through organ donation (with PMD of <6 and 24 h, respectively) showed higher MeCP2E1 and MeCP2E2 levels compared to almost all other tested samples. The fact that the A201V sample with 24 h PMD still shows much higher MeCP2

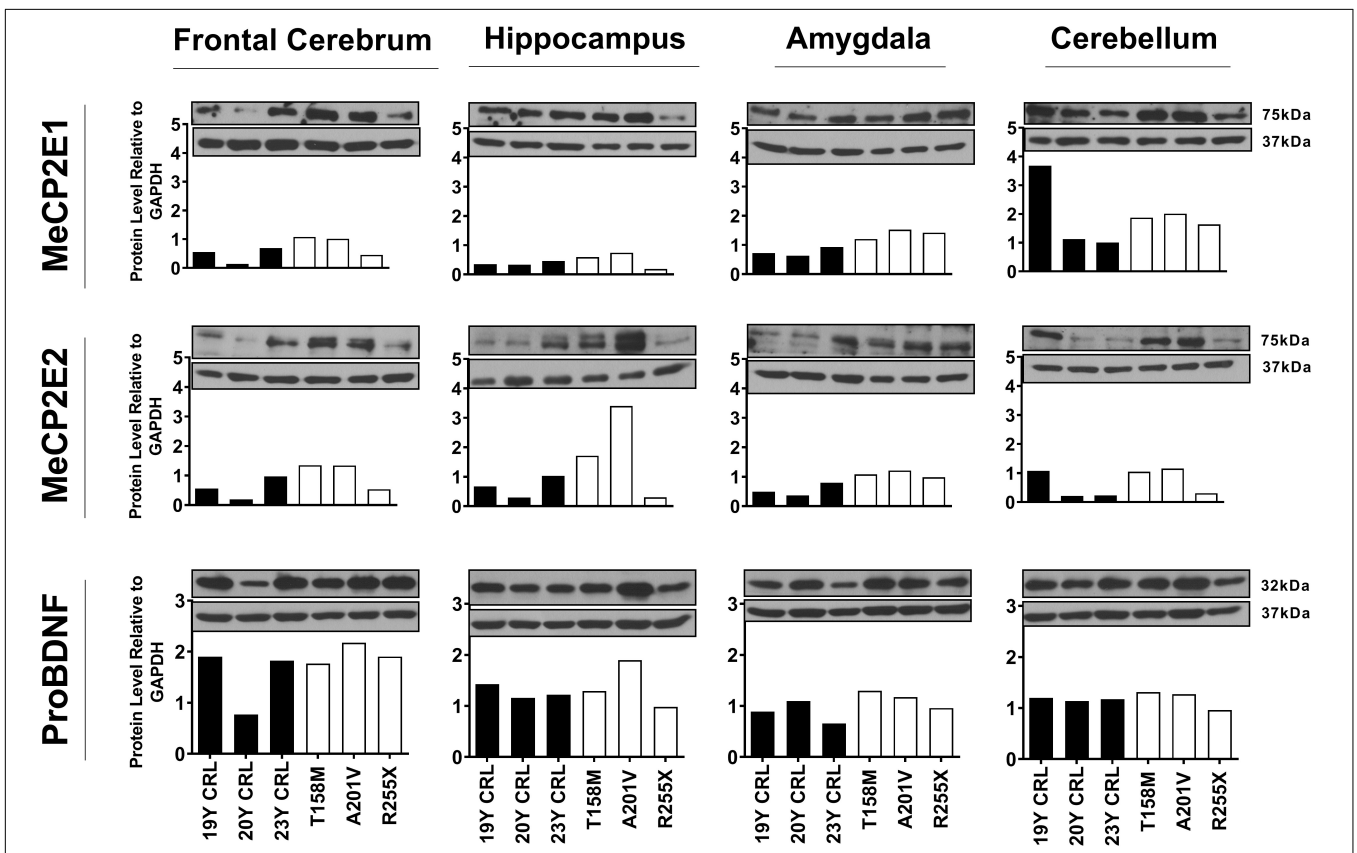

FIGURE 2 | Western blot results of MeCP2E1, MeCP2E2, and ProBDNF in human Rett Syndrome (RTT) and control brain tissues. Western Blot results for MeCP2 isoforms and ProBDNF normalized to GAPDH in three RTT patients and their age-, and sex-matched controls show inter-regional variations as well as variable amount of proteins among controls and RTTs. The full blots for ProBDNF (including mature BDNF) with GAPDH loading control are provided in **Supplementary Figure S1**. Control (CRL); Year Old (Y).

protein levels compared to even control samples, highlights that MeCP2 homeostasis at the protein level might be more complicated as previously thought. As expected, the R255X non-sense mutation showed relatively low level of both MeCP2 isoforms (at 72–75 kDa), originated from the non-mutated X-chromosome. Non-parametric comparisons between the controls and RTT samples did not show any significant differences of the MeCP2 isoforms (**Figures 3A–C**), which was not in concordance with the results of their corresponding transcript levels.

BDNF and ProBDNF Levels in the Brain of RTT Patients and Controls

Next, we investigated BDNF protein levels in these brain tissue samples. As BDNF is a cytoplasmic protein (Wetmore et al., 1991; Aliaga et al., 2009), we performed WB experiments with isolated cytoplasmic extracts. Our results showed the presence of the mature BDNF at around 14 kDa, and proBDNF at 32 kDa (**Supplementary Figure S1**). The detected bands were relatively consistent within the control group, with no obvious differences between the control and RTT brain tissues (**Figures 2**, **3D,E** and **Supplementary Figure S1**). Surprisingly, we detected very low levels of the mature BDNF in the cerebellum of both RTT and control tissues (**Supplementary Figure S1**). To confirm our findings, we further studied BDNF levels using ELISA kits (Sigma-Aldrich), as a parallel complementary approach. Based on ELISA results, control hippocampus samples showed statistically significant higher levels of BDNF compared to the frontal cerebrum and cerebellum. The average level of BDNF in control hippocampus was 1.81 and 2.02 times higher than the frontal cerebrum and cerebellum, respectively ($P < 0.05$). Also, the amygdala and hippocampus showed higher levels of the mature BDNF compared to the other two brain regions. However, inter-regional differences among RTT samples were not statistically significant (**Figure 3E**). As comparing Western blot results in **Figure 3D** with ELISA results in **Figure 3E** shows, regions with higher mature BDNF showed lower levels of ProBDNF precursor, although the detected difference among

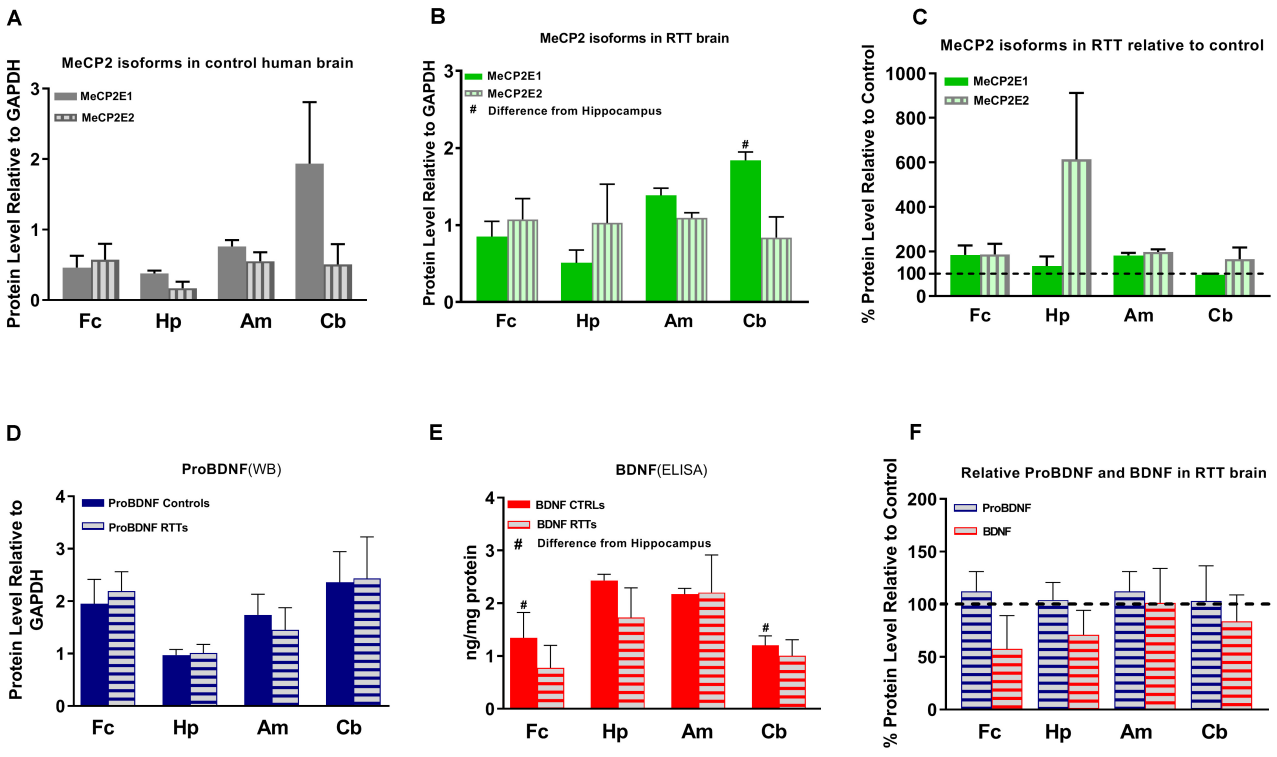

FIGURE 3 | Results of protein studies by Western blot and ELISA in different regions of Rett Syndrome (RTT) and control brain tissues. **(A)** Western blot results for MeCP2 isoforms quantified and analyzed in four brain regions (frontal cerebrum, hippocampus, amygdala, and cerebellum) for mean of controls show non-significant variations between the regions. **(B)** Western blot results for MeCP2 isoforms quantified and analyzed in four brain regions for mean of RTTs show higher level of MeCP2E1 isoform in the cerebellum compared to the frontal cerebrum or hippocampus. **(C)** Mean of MeCP2 isoforms in RTT and controls are not significantly different in relative RTT/control analysis. **(D)** Western blot results for ProBDNF quantified and analyzed in four brain regions for mean of controls and RTTs show non-significant inter-regional variations without significant difference between RTT and controls. **(E)** ELISA results for BDNF quantified and analyzed in four brain regions for mean of controls and RTTs show significantly higher levels in Hippocampus compared to frontal cerebrum and cerebellum of the controls. **(F)** Mean of ProBDNF and BDNF in RTT and controls are not significantly different in relative RTT/Control analysis. Frontal cerebrum (Fc); Hippocampus (Hp); Amygdala (Am); Cerebellum (Cb); $N = 3 \pm$ Standard Error of the Mean (SEM); One-way ANOVA (ELISA)/Kruskal-Wallis test (WB) with multiple pairwise comparisons used among different brain regions. Student's *t*-test (ELISA)/Mann-Whiney test (WB) was used to compare controls and RTT samples for each parameter. The significance level was determined at $P < 0.05$ (#).

these regions was not statistically significant. RTT samples had similar levels of proBDNF, with a trend of lower levels of mature BDNF in different regions of the brain, which was not statistically significant (**Figure 3F**).

Correlation Studies

Next, we performed correlation of transcript-transcript and transcript-protein analysis among the studied parameters of RTT and control brain tissues (**Figure 4**).

We observed that the two *MECP2* isoforms were positively correlated at the transcript level (red arrows in **Figure 4**, top panel), which was statistically significant for RTT and combined data from controls and RTT samples together ($P < 0.01$). However, at the protein level, the two MeCP2 isoforms showed statistically significant positive correlation only for the combined data, which is marked by red arrowhead ($P < 0.05$) (**Figure 4**, middle panel). The transcript-protein correlation was negative for MeCP2E2, which was statistically significant for the combined data ($P < 0.01$), marked with red arrowhead in **Figure 4** (bottom panel). The gray arrowhead in **Figure 4** (bottom panel) points to positive transcript-protein correlation for MeCP2E1 for the control data ($P < 0.05$). Our studies showed that although both *MECP2* isoforms have a statistically significant positive correlation with *BDNF* at the transcript level (green arrow in **Figure 4**, top panel); however, at the protein level, BDNF is negatively correlated only with MeCP2E1 in control samples (green arrowhead in **Figure 4**, middle panel). ProBDNF on the other hand, showed a positive correlation only with MeCP2E2 in combined data ($P < 0.01$), which is marked by purple arrowhead in **Figure 4**, middle panel. Furthermore, we observed a negative correlation between BDNF and its precursor ProBDNF, which was statistically significant ($P < 0.05$) for combined data (green arrowhead in **Figure 4**, middle panel, right graph). ProBDNF also showed a negative correlation with *BDNF* transcripts, which was statistically significant for both RTT and combined data (purple arrowheads in **Figure 4**, bottom panel). Finally, *miR132* showed a negative correlation with *MECP2E2*, which was statistically significant ($P < 0.01$) in RTT brains (blue arrow in **Figure 4**, top panel).

Histopathological Examination

In addition to molecular analysis of the RTT brain compared to controls, we then studied if these examined RTT brain tissues show any gross anatomical abnormalities at the pathological levels. Therefore, we performed histopathological investigation of the frontal cerebrum, hippocampus, and cerebellum in RTT patients and their sex- and age-matched controls by Hematoxylin and Eosin staining of the prepared slides (**Figure 5**). Formalin-fixed amygdala was not available for all samples, and was therefore removed from the H&E analysis. The microscopic evaluation was unremarkable without features of dead neurons or inflammation. Increased cell density and smaller pyramidal neurons in comparison with age- and sex-matched controls were observed in the hippocampus of RTT brain with the T158M mutation (**Figure 5A4** versus **A3**). In RTT brain with A201V mutation, the dentate gyrus of the hippocampal formation was slightly thin toward the end and focally hypocellular. These subtle findings of microscopic examination are shown in **Figure 5**.

DISCUSSION

Studies on Human Brain Tissues Support That Components of the MeCP2-BDNF-*miR132* Regulatory Loop Are Expressed in a Region-Dependent Manner in the Human Brain

Our finding to detect lower transcript levels of *MECP2* isoforms in the human RTT brain is in agreement with previous studies that reported decreased *MECP2* transcripts in Rett Syndrome (Chen et al., 2001; Petel-Galil et al., 2006). In our study, MeCP2E2, and ProBDNF showed a mainly negative correlation with their transcripts, which was opposite to what we saw for MeCP2E1. The correlation coefficient was not similar for RTT and controls in all cases (**Figure 4**). Complex regulation of protein expression in the human brain has been suggested as a possible reason behind the low predictive value of transcript expression for the corresponding protein(s) (Bauernfeind and Babbitt, 2017). Post-transcriptional control, and cell- or tissue-specific regulation rather than transcriptional-specific control are also among other possible explanations (Shahbazian et al., 2002a). In the human post-mortem RTT brain, *BDNF* transcripts showed a similar pattern to *MECP2E1/E2* with lower transcript expression, without any significant difference at the protein levels compared to controls. In this regard, our recent report on the formalin-fixed brain tissues added another layer to the complexity of MeCP2-BDNF expression. In this previous study, BDNF immunolabeling showed that different cells of the brain i.e., astrocytes, endothelial cells, and neurons could all be part of MeCP2 homeostasis (Pejhan et al., 2020). Therefore, it might be too simplistic to suggest a single regulatory mechanism without considering the cellular source of the proteins. Furthermore, the inter-regional variations that we showed in the elements of the suggested MeCP2 regulatory network in the human brain might be the result of variable number of cells and different cell types with different functions in each region, which makes finding a unique regulatory system in the whole brain of human less probable. Studying the third element of the MeCP2 homeostasis regulatory network proposed in rodents (*miR132*), a brain region-specific expression was identified specific to the cerebellum, where *MECP2E1-E2* and *BDNF* levels were the highest and *miR132* levels were found to be the lowest (**Figure 1**). Interestingly, cerebellum was the single brain region that both *miR132* strands were detected at similar levels between RTT patients and control samples. In general, lower transcript levels of *MECP2* isoforms in the human RTT brains were associated with lower level of *miR132* in all brain regions, except for the cerebellum. As indicated, cerebellum was different from the other three tested brain regions with similar level of *miR132* in RTT and control brains. Looking at an average of all brain regions in our correlation studies, we noticed a negative correlation between *miR132* and both *MECP2E1* and *MECP2E2* transcripts,

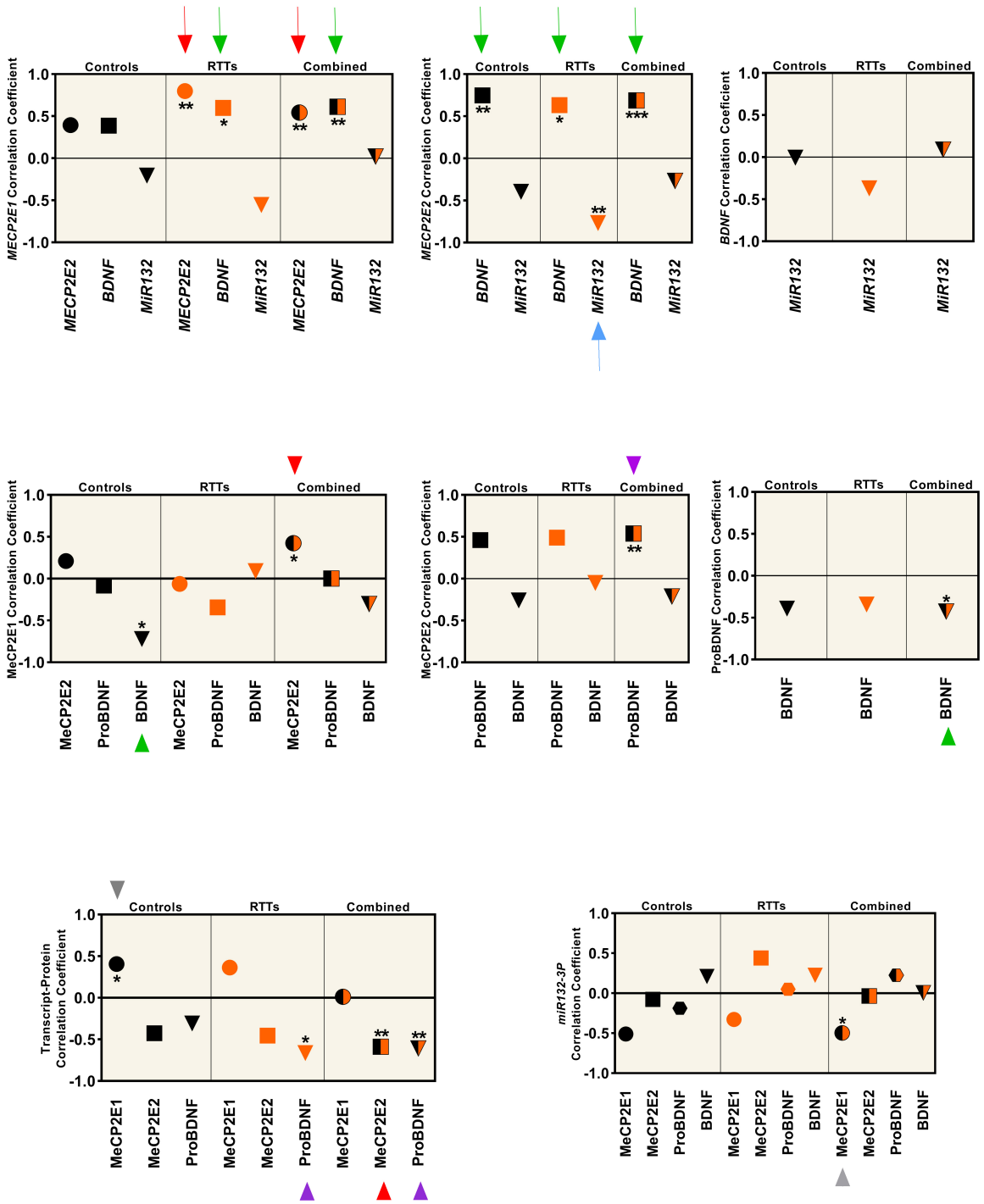

FIGURE 4 | Correlation analysis between all the studied parameters (MeCP2 isoforms, ProBDNF, BDNF, and *miR132*) from four brain regions (Frontal Cerebrum, Hippocampus, Amygdala, and Cerebellum) in transcript and protein level. All graphs represent the Pearson's (for parametric variables), or Spearman's (for non-parametric variables) correlation coefficient in controls (black), RTTs (orange), and controls and RTTs combined (black/orange). Statistical significance: ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; The charts are showing the data collected from four different brain regions (Frontal Cerebrum, Hippocampus, Amygdala and Cerebellum) of 3 patients and their age and sex matched controls ($N = 3$ controls + 3 RTTs). The two isoforms are positively correlated at the transcript level, which is significant for RTT and mixed data (red arrow), and *MECP2E2* is negatively correlated with its protein, which is significant in mixed data (red arrowhead). Both isoforms show positive correlation with BDNF at the transcript level. Green arrows show the significant ones. BDNF protein has a significant negative correlation with MeCP2E1 in control samples (green arrowhead). ProBDNF is positively correlated with MeCP2E2, which is significant in mixed data (purple arrowhead). There is a negative correlation between ProBDNF and BDNF transcript in both RTT and mixed data (purple arrowhead). The only significant correlation of *miR132* is a negative correlation with *MECP2E2* in RTTs (blue arrow). Note that all *miR132* transcripts in this Figure refer to *miR132-3P* strand.

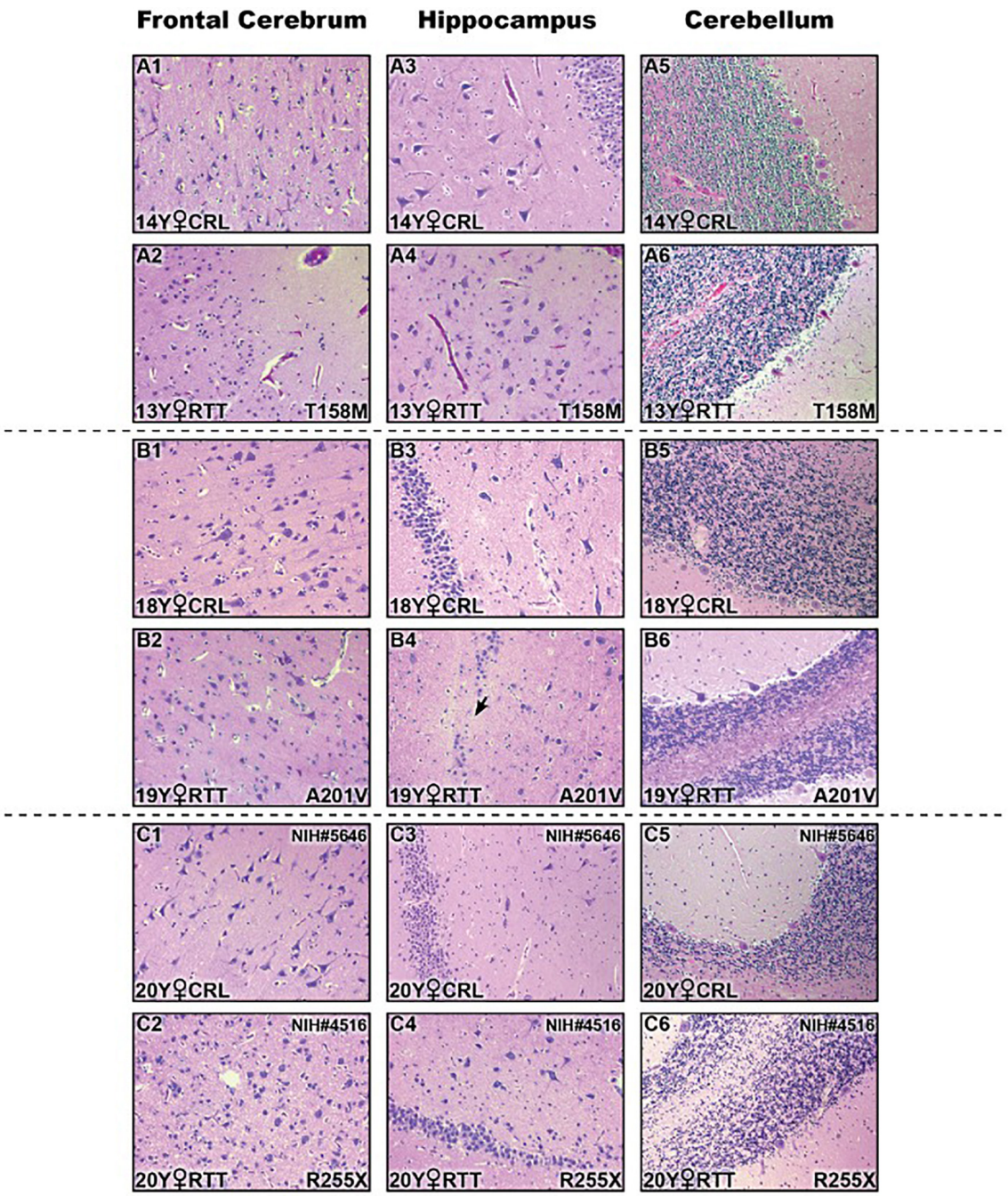

FIGURE 5 | Histopathologic studies of different brain regions in three Rett Syndrome (RTT) patients and their sex-, and age-matched controls stained by Hematoxylin and Eosin (H&E). There were not noticeable findings at the microscopic levels. Two subtle changes are shown here: small pyramidal neurons and increased cell density in hippocampal region of T158M brain **(A4)** compared to the control **(A3)**; thin and hypocellular dentate gyrus in the hippocampus of A201V brain (black arrow in **B4**) compared to the control **(B3)**; Control (CRL); Year old (Y); 200× magnification.

which was significant for the latter (**Figure 4**). Previous *in vivo* studies in rodents have shown that increased level of *miR132* is associated with decreased level of MeCP2 and BDNF in the rat hippocampus (Su et al., 2015), suggesting the potential existence of species-specific regulatory mechanisms between rodents and humans. In the future, we will be interested to study the detailed mechanism of MeCP2 homeostasis regulation in human brain cells using relevant *in vitro* studies. Densely packed neurons with smaller soma size and reduced dendritic complexity are important morphological features of RTT brains compared to controls. These characteristics have been extensively studied in *in vitro* differentiated human neurons (Marchetto et al., 2010;

Chin et al., 2016), and in rare studies on post-mortem brain samples of RTT patients (Armstrong et al., 1995). Although this was out of the scope of our current study; it would be important to investigate such changes in various type of neurons in different brain regions of RTT patients to establish if similar abnormalities are region, or cell type-specific in the human brain.

ProBDNF and Mature BDNF in RTT and Control Human Brain Samples

Most studies on MeCP2-BDNF interactions in RTT pathophysiology have focused on neurons while our previous study showed a clear astroglial/endothelial pattern of immunostaining for BDNF in human brain samples (Pejhan et al., 2020). Apart from the BDNF source of expression, its impairment in RTT patients is another controversy. Here, we showed lower *BDNF* mRNA expression in RTT brain. However, BDNF protein levels (detected by WB, ELISA, and previously by IHC (Pejhan et al., 2020) were not detected at lower levels in the human RTT patient brain tissues, suggesting RTT brain has comparable BDNF levels to controls. We detected equal to higher BDNF levels in the Purkinje cells of the RTT cerebellar tissues by immunohistological studies (Pejhan et al., 2020). Such weak predictive value of transcript expression for their corresponding proteins is not unusual and has been linked to complex regulation of protein expression in the human brain (Bauernfeind and Babbitt, 2017). Of note anti-BDNF antibody can detect both proBDNF and the mature protein in ELISA and IHC. Only WB experiment differentiates between proBDNF and the mature BDNF by molecular size, but it also showed no difference between control and RTT brain. Our WB results with lower level of mature BDNF and higher level of ProBDNF in the cerebellum are in line with our previous findings with IHC (Pejhan et al., 2020). Furthermore, BDNF immunoreactivity has been shown to increase with conditions such as hypoxia or neuroinflammation (Hartman et al., 2015; Satriotomo et al., 2016), which are part of the pathogenesis of different neurological diseases, not fully studied in the context of RTT brain. Our correlation studies at the transcript level of *MECP2* isoforms and *BDNF* were in line with increased level of MeCP2 and BDNF as part of a regulatory effect, suggested in previous studies on rat neurons and rat brain (Klein et al., 2007; Su et al., 2015). However, at the protein level, MeCP2E1 mainly showed a negative correlation with BDNF, and for MeCP2E2, the positive correlation was shown mainly with proBDNF, and not the mature protein (**Figure 4**). While our data provide important insight about the components of MeCP2E1/E2-BDNF-*miR132* homeostasis regulatory network in the human brain and its impairment in RTT patients, additional studies are required to investigate the conservation of this regulatory network from rodents to humans.

LIMITATIONS

Similar to previous studies that used post-mortem human brain tissues for Rett Syndrome, one limitation of our study was the absence of a large cohort of post-mortem RTT brain tissues. While RTT is a rare disease, we still obtained three independent biological replicates from different regions of the brain for human RTT patients and comparable controls, for statistical analysis of our results. Furthermore, our recent IHC examination of RTT human brain tissues (Pejhan et al., 2020) has shown the significance of cell-type variability for the elements studied in this project. Therefore, we acknowledge that studying the frozen brain lysate gives us an overall result for different cell types with potentially variable regulatory systems in a complex organ such as the human brain.

CONCLUSION

Our findings from post-mortem human brain tissues suggest the existence of a brain region-specific regulation of the MeCP2E1/E2-BDNF-*miR132* that might be differently regulated in the cerebellum. Although, MeCP2 homeostasis regulation is still poorly understood, our data suggest that it is compromised in the human RTT brain, at least at the transcript levels. Our findings suggest that similar to other model systems, *MECP2E1* and *MECP2E2* transcript and protein may not be fully correlated in the human brain. In addition, BDNF transcript and protein levels do not show the same pattern of altered expression in the human RTT brain, with potential regulatory mechanisms that control BDNF maturation from its precursor ProBDNF protein. Despite intensive efforts in studying MeCP2 function in neurodevelopment, immune system (Pecorelli et al., 2020), and human cancer (Neupane et al., 2016), our findings would help to shed some light on its still obscure regulatory network. Several more common neurodevelopmental conditions such as fetal alcohol spectrum disorders or autism spectrum disorders with the evidence for the role of MeCP2 in their molecular biology of disease (Liyanage et al., 2015, 2017; Rastegar, 2017a; Amiri et al., 2020) may also benefit from the results of our study. Further *in vitro* studies could prove to be important for addressing whether MeCP2 homeostasis regulation is conserved from rodents to humans.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available by authors, upon request.

ETHICS STATEMENT

Our research on human brain tissues was reviewed and approved by the University of Manitoba Bannatyne Campus Research Ethics Board with an approved Health Research Board protocol #HS20095 (H2016:337). Consent to perform research on donated human brain tissues was obtained from patient immediate family members.

AUTHOR CONTRIBUTIONS

SP and MR designed the experiments. SP performed the RT-PCR, WB, and ELISA. MD dissected the A201V and T158M human brain regions, and provided control tissues for H&E samples (**Figures 5A,B**). MR provided the conception and design, contributed reagents, materials, analysis tools, and research facilities. SP and MR wrote the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant 2016-06035 to MR, International Rett Syndrome Foundation (IRSF) Grant 3212 to MR, Ontario Rett Syndrome Association (ORSA) to MR, CIHR Tri-Council Bridge Funding to MR, and Rady Innovation Fund to MR and MD. SP is supported by IRSF, NSERC-DG, and Graduate Enhancement of Tri-Council Stipends (GETS) supplements to MR. MD held the Canada Research Chair in Developmental Neuropathology.

ACKNOWLEDGMENTS

Tissue samples and related data for R255X and frozen control tissues were obtained from the "University of Maryland Brain and Tissue Bank, which is a Brain and Tissue Repository of the NIH Biobank (at NIH NeuroBioBank Program: neurobiobank.nih.gov). The authors would like to sincerely thank the family members of the T158M and A201V patients for organ brain donation for research, as previously described in our publications (Olson et al., 2018; Pejhan et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.00763/full#supplementary-material

FIGURE S1 | Western blot analysis of ProBDNF and mature BDNF in Rett Syndrome (RTT) and control brain tissues. ProBDNF is detected around 32 kDa while the mature BDNF is detected around 14 kDa. The bands detected around 28 kDa have been observed by other groups and are suggested to be truncated protein (Mowla et al., 2001). Mature BDNF has a significantly lower level in the cerebellar tissues of controls and RTT tissues compared to other studied brain regions (frontal cerebrum, hippocampus, and amygdala). However, the level of ProBDNF in the cerebellum is comparable to other regions. Control (CRL); Year old (Y); Frontal cerebrum (Fc); Hippocampus (Hp); Amygdala (Am); Cerebellum (Cb); Kilodaton (kDa).

TABLE S1 | List of primers used for qRT-PCR in this project.

TABLE S2 | List of primary and secondary antibodies.

REFERENCES

Aliaga, E. E., Mendoza, I., and Tapia-Arancibia, L. (2009). Distinct subcellular localization of BDNF transcripts in cultured hypothalamic neurons and modification by neuronal activation. *J. Neural Transm.* 116, 23–32. doi: 10.1007/s00702-008-0159-8

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* 23, 185–188. doi: 10.1038/13810

Amiri, S., Davie, J. R., and Rastegar, M. (2020). Chronic ethanol exposure alters DNA methylation in neural stem cells: role of mouse strain and sex. *Mol. Neurobiol.* 57, 650–667. doi: 10.1007/s12035-019-01728-0

Armstrong, D., Dunn, J. K., Antalffy, B., and Trivedi, R. (1995). Selective dendritic alterations in the cortex of Rett syndrome. *J. Neuropathol. Exp. Neurol.* 54, 195–201. doi: 10.1097/00005072-199503000-00006

Balmer, D., Goldstine, J., Rao, Y. M., and LaSalle, J. M. (2003). Elevated methyl-CpG-binding protein 2 expression is acquired during postnatal human brain development and is correlated with alternative polyadenylation. *J. Mol. Med.* 81, 61–68. doi: 10.1007/s00109-002-0396-5

Barber, B. A., Liyanage, V. R., Zachariah, R. M., Olson, C. O., Bailey, M. A., and Rastegar, M. (2013). Dynamic expression of MEIS1 homeoprotein in E14.5 forebrain and differentiated forebrain-derived neural stem cells. *Ann. Anat.* 195, 431–440. doi: 10.1016/j.aanat.2013.04.005

Bauernfeind, A. L., and Babbitt, C. C. (2017). The predictive nature of transcript expression levels on protein expression in adult human brain. *BMC Genomics* 18:322. doi: 10.1186/s12864-017-3674-x

Chahil, G., and Bollu, P. C. (2020). *Rett Syndrome*. Treasure Island, FL: StatPearls.

Chen, R. Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* 27, 327–331. doi: 10.1038/85906

Chin, E. W., Marcy, G., Yoon, S. I., Ma, D., Rosales, F. J., Augustine, G. J., et al. (2016). Choline ameliorates disease phenotypes in human iPSC models of rett syndrome. *Neuromol. Med.* 18, 364–377. doi: 10.1007/s12017-016-8421-y

Delcuve, G. P., Rastegar, M., and Davie, J. R. (2009). Epigenetic control. *J. Cell Physiol.* 219, 243–250. doi: 10.1002/jcp.21678

Duan, Z. Y., Cai, G. Y., Li, J. J., Bu, R., Wang, N., Yin, P., et al. (2018). U6 can be used as a housekeeping gene for urinary sediment miRNA studies of IgA nephropathy. *Sci. Rep.* 8:10875. doi: 10.1038/s41598-018-29297-7

Ezeonwuka, C. D., and Rastegar, M. (2014). MeCP2-related diseases and animal models. *Diseases* 2, 45–70. doi: 10.3390/diseases2010045

Girard, M., Couvert, P., Carrie, A., Tardieu, M., Chelly, J., Beldjord, C., et al. (2001). Parental origin of de novo MECP2 mutations in Rett syndrome. *Eur. J. Hum. Genet.* 9, 231–236. doi: 10.1038/sj.ejhg.5200618

Gordon, J., Wu, C. H., Rastegar, M., and Safa, A. R. (2003). Beta2-microglobulin induces caspase-dependent apoptosis in the CCRF-HSB-2 human leukemia cell line independently of the caspase-3, -8 and -9 pathways but through increased reactive oxygen species. *Int. J. Cancer* 103, 316–327. doi: 10.1002/ijc.10828

Hartman, W., Helan, M., Smelter, D., Sathish, V., Thompson, M., Pabelick, C. M., et al. (2015). Role of hypoxia-induced brain derived neurotrophic factor in human pulmonary artery smooth muscle. *PLoS One* 10:e0129489. doi: 10.1371/journal.pone.0129489

Huang, H., Rastegar, M., Bodner, C., Goh, S. L., Rambaldi, I., and Featherstone, M. (2005). MEIS C termini harbor transcriptional activation domains that respond to cell signaling. *J. Biol. Chem.* 280, 10119–10127. doi: 10.1074/jbc.M413963200

Itoh, M., Tahimic, C. G., Ide, S., Otsuki, A., Sasaoka, T., Noguchi, S., et al. (2012). Methyl CpG-binding protein isoform MeCP2_e2 is dispensable for Rett syndrome phenotypes but essential for embryo viability and placenta development. *J. Biol. Chem.* 287, 13859–13867. doi: 10.1074/jbc.M111.309864

Kerr, B., Soto, C. J., Saez, M., Abrams, A., Walz, K., and Young, J. I. (2012). Transgenic complementation of MeCP2 deficiency: phenotypic rescue of

Mecp2-null mice by isoform-specific transgenes. *Eur. J. Hum. Genet.* 20, 69–76. doi: 10.1038/ejhg.2011.145

Klein, M. E., Lioy, D. T., Ma, L., Impey, S., Mandel, G., and Goodman, R. H. (2007). Homeostatic regulation of MeCP2 expression by a CREB-induced microRNA. *Nat. Neurosci.* 10, 1513–1514. doi: 10.1038/nn2010

Kobrossy, L., Rastegar, M., and Featherstone, M. (2006). Interplay between chromatin and trans-acting factors regulating the Hoxd4 promoter during neural differentiation. *J. Biol. Chem.* 281, 25926–25939. doi: 10.1074/jbc.M602555200

Kriaucionis, S., and Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res.* 32, 1818–1823. doi: 10.1093/nar/gkh349

Krishnaraj, R., Ho, G., and Christodoulou, J. (2017). RettBASE: Rett syndrome database update. *Hum. Mutat.* 38, 922–931. doi: 10.1002/humu.23263

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., et al. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905–914. doi: 10.1016/0092-8674(92)90610-o

Liyanage, V. R., Curtis, K., Zachariah, R. M., Chudley, A. E., and Rastegar, M. (2017). Overview of the genetic basis and epigenetic mechanisms that contribute to FASD pathobiology. *Curr. Top Med. Chem.* 17, 808–828. doi: 10.2174/1568026616666160414124816

Liyanage, V. R., and Rastegar, M. (2014). Rett syndrome and MeCP2. *Neuromol. Med.* 16, 231–264. doi: 10.1007/s12017-014-8295-9

Liyanage, V. R., Zachariah, R. M., Davie, J. R., and Rastegar, M. (2015). Ethanol deregulates Mecp2/MeCP2 in differentiating neural stem cells via interplay between 5-methylcytosine and 5-hydroxymethylcytosine at the Mecp2 regulatory elements. *Exp. Neurol.* 265, 102–117. doi: 10.1016/j.expneurol.2015.01.006

Liyanage, V. R., Zachariah, R. M., and Rastegar, M. (2013). Decitabine alters the expression of Mecp2 isoforms via dynamic DNA methylation at the Mecp2 regulatory elements in neural stem cells. *Mol. Autism* 4:46. doi: 10.1186/2040-2392-4-46

Liyanage, V. R. B., Olson, C. O., Zachariah, R. M., Davie, J. R., and Rastegar, M. (2019). DNA methylation contributes to the differential expression levels of Mecp2 in male mice neurons and astrocytes. *Int. J. Mol. Sci.* 20:1845. doi: 10.3390/ijms20081845

Marchetto, M. C., Carromeu, C., Acab, A., Yu, D., Yeo, G. W., Mu, Y., et al. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 143, 527–539. doi: 10.1016/j.cell.2010.10.016

Martinez de Paz, A., Khajavi, L., Martin, H., Claveria-Gimeno, R., Tom Dieck, S., Cheema, M. S., et al. (2019). MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. *Epigenet. Chromatin.* 12:63. doi: 10.1186/s13072-019-0298-1

McDermott, A. M., Kerin, M. J., and Miller, N. (2013). Identification and validation of miRNAs as endogenous controls for RQ-PCR in blood specimens for breast cancer studies. *PLoS One* 8:e83718. doi: 10.1371/journal.pone.0083718

McGowan, H., and Pang, Z. P. (2015). Regulatory functions and pathological relevance of the MECP2 3'UTR in the central nervous system. *Cell Regen.* 4:9. doi: 10.1186/s13619-015-0023-x

Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J., et al. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.* 36, 339–341. doi: 10.1038/ng1327

Mowla, S. J., Farhadi, H. F., Pareek, S., Atwal, J. K., Morris, S. J., Seidah, N. G., et al. (2001). Biosynthesis and post-translational processing of the precursor to brain-derived neurotrophic factor. *J. Biol. Chem.* 276, 12660–12666. doi: 10.1074/jbc.M008104200

Mullaney, B. C., Johnston, M. V., and Blue, M. E. (2004). Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. *Neuroscience* 123, 939–949. doi: 10.1016/j.neuroscience.2003.11.025

Neul, J. L., Fang, P., Barrish, J., Lane, J., Caeg, E. B., Smith, E. O., et al. (2008). Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* 70, 1313–1321. doi: 10.1212/01.wnl.0000291011.54508.aa

Neupane, M., Clark, A. P., Landini, S., Birkbak, N. J., Eklund, A. C., Lim, E., et al. (2016). MECP2 is a frequently amplified oncogene with a novel epigenetic mechanism that mimics the role of activated RAS in malignancy. *Cancer Discov.* 6, 45–58. doi: 10.1158/2159-8290.CD-15-0341

Olson, C. O., Pejhan, S., Kroft, D., Sheikholeslami, K., Fuss, D., Buist, M., et al. (2018). MECP2 Mutation interrupts nucleolin-mTOR-P70S6K signaling in Rett syndrome patients. *Front. Genet.* 9:635. doi: 10.3389/fgene.2018.00635

Olson, C. O., Zachariah, R. M., Ezeonwuka, C. D., Liyanage, V. R., and Rastegar, M. (2014). Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements. *PLoS One* 9:e90645. doi: 10.1371/journal.pone.0090645

Pecorelli, A., Cordone, V., Messano, N., Zhang, C., Falone, S., Amicarelli, F., et al. (2020). Altered inflammasome machinery as a key player in the perpetuation of Rett syndrome oxinflammation. *Redox Biol.* 28:101334. doi: 10.1016/j.redox.2019.101334

Pejhan, S., Mok Siu, V., Ang, L. C., Del Bigio, M. R., and Rastegar, M. (2020). Differential brain region-specific expression of MeCP2 and BDNF in Rett syndrome patients: a distinct grey-white matter variation. *Neuropathol. Appl. Neurobiol.* doi: 10.1111/nan.12619 [Epub ahead of print].

Petel-Galil, Y., Benteer, B., Galil, Y. P., Zeev, B. B., Greenbaum, I., Vecsler, M., et al. (2006). Comprehensive diagnosis of Rett's syndrome relying on genetic, epigenetic and expression evidence of deficiency of the methyl-CpG-binding protein 2 gene: study of a cohort of Israeli patients. *J. Med. Genet.* 43:e56. doi: 10.1136/jmg.2006.041285

Rastegar, M. (2017a). Editorial (Thematic issue: neuro epigenetics and neurodevelopmental disorders: from molecular mechanisms to cell fate commitments of the brain cells and human disease). *Curr. Top. Med. Chem.* 17, 769–770. doi: 10.2174/1568026616999160812144822

Rastegar, M. (2017b). "Epigenetics and cerebellar neurodevelopmental disorders," in *Development of the Cerebellum from Molecular Aspects to Diseases*, ed. H. Marzban (Cham: Springer International Publishing), doi: 10.1007/978-3-319-59749-2_10

Rastegar, M., Hotta, A., Pasceri, P., Makarem, M., Cheung, A. Y., Elliott, S., et al. (2009). MECP2 isoform-specific vectors with regulated expression for Rett syndrome gene therapy. *PLoS One* 4:e6810. doi: 10.1371/journal.pone.0006810

Rastegar, M., Kobrossy, L., Kovacs, E. N., Rambaldi, I., and Featherstone, M. (2004). Sequential histone modifications at Hoxd4 regulatory regions distinguish anterior from posterior embryonic compartments. *Mol. Cell Biol.* 24, 8090–8103. doi: 10.1128/MCB.24.18.8090-8103.2004

Rastegar, M., Rousseau, G. G., and Lemaigre, F. P. (2000). CCAAT/enhancer-binding protein-alpha is a component of the growth hormone-regulated network of liver transcription factors. *Endocrinology* 141, 1686–1692. doi: 10.1210/endo.141.5.7478

Renieri, A., Meloni, I., Longo, I., Ariani, F., Mari, F., Pescucci, C., et al. (2003). Rett syndrome: the complex nature of a monogenic disease. *J. Mol. Med.* 81, 346–354. doi: 10.1007/s00109-003-0444-9

Satriotomo, I., Nichols, N. L., Dale, E. A., Emery, A. T., Dahlberg, J. M., and Mitchell, G. S. (2016). Repetitive acute intermittent hypoxia increases growth/neurotrophic factor expression in non-respiratory motor neurons. *Neuroscience* 322, 479–488. doi: 10.1016/j.neuroscience.2016.02.060

Shahbazian, M. D., Antalffy, B., Armstrong, D. L., and Zoghbi, H. Y. (2002a). Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum. Mol. Genet.* 11, 115–124. doi: 10.1093/hmg/11.2.115

Shahbazian, M. D., Sun, Y., and Zoghbi, H. Y. (2002b). Balanced X chromosome inactivation patterns in the Rett syndrome brain. *Am. J. Med. Genet.* 111, 164–168. doi: 10.1002/ajmg.10557

Sheikh, I., de Paz, A. M., Akhtar, S., Ausio, J., and Vincent, J. B. (2017). MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation. *Hum. Mol. Genet.* 26, 4132–4141. doi: 10.1093/hmg/ddx300

Sheikholeslami, K., Ali Sher, A., Lockman, S., Kroft, D., Ganjibakhsh, M., Nejati-Koshki, K., et al. (2019). Simvastatin induces apoptosis in medulloblastoma brain tumor cells via mevalonate cascade prenylation substrates. *Cancers* 11:994. doi: 10.3390/cancers11070994

Singh, J., Saxena, A., Christodoulou, J., and Ravine, D. (2008). MECP2 genomic structure and function: insights from ENCODE. *Nucleic Acids Res.* 36, 6035–6047. doi: 10.1093/nar/gkn591

Su, M., Hong, J., Zhao, Y., Liu, S., and Xue, X. (2015). MeCP2 controls hippocampal brain-derived neurotrophic factor expression via homeostatic interactions with microRNA132 in rats with depression. *Mol. Med. Rep.* 12, 5399–5406. doi: 10.3892/mmr.2015.4104

Takeguchi, R., Takahashi, S., Kuroda, M., Tanaka, R., Suzuki, N., Tomonoh, Y., et al. (2020). MeCP2_e2 partially compensates for lack of MeCP2_e1: a male case of Rett syndrome. *Mol. Genet. Genomic Med.* 8:e1088. doi: 10.1002/mgg3.1088

Tillotson, R., and Bird, A. (2019). The Molecular Basis of MeCP2 Function in the Brain. *J. Mol. Biol.* 432, 1602–1623. doi: 10.1016/j.jmb.2019.10.004

Trappe, R., Laccone, F., Cobilanschi, J., Meins, M., Huppke, P., Hanefeld, F., et al. (2001). MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am. J. Hum. Genet.* 68, 1093–1101. doi: 10.1086/320109

Villard, L., Kpebe, A., Cardoso, C., Chelly, P. J., Tardieu, P. M., and Fontes, M. (2000). Two affected boys in a Rett syndrome family: clinical and molecular findings. *Neurology* 55, 1188–1193. doi: 10.1212/wnl.55.8.1188

Vogel Ciernia, A., Yasui, D. H., Pride, M. C., Durbin-Johnson, B., Noronha, A. B., Chang, A., et al. (2018). MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Hum. Mol. Genet.* 27, 4077–4093. doi: 10.1093/hmg/ddy301

Wetmore, C., Cao, Y. H., Pettersson, R. F., and Olson, L. (1991). Brain-derived neurotrophic factor: subcellular compartmentalization and interneuronal transfer as visualized with anti-peptide antibodies. *Proc. Natl. Acad. Sci. U.S.A.* 88, 9843–9847. doi: 10.1073/pnas.88.21.9843

Wu, C. H., Rastegar, M., Gordon, J., and Safa, A. R. (2001). beta(2)-microglobulin induces apoptosis in HL-60 human leukemia cell line and its multidrug resistant variants overexpressing MRP1 but lacking Bax or overexpressing P-glycoprotein. *Oncogene* 20, 7006–7020. doi: 10.1038/sj.onc.1204893

Xinhua, B., Shengling, J., Fuying, S., Hong, P., Meirong, L., and Wu, X. R. (2008). X chromosome inactivation in Rett Syndrome and its correlations with MECP2 mutations and phenotype. *J. Child Neurol.* 23, 22–25. doi: 10.1177/0883073807307077

Xu, W., Liyanage, V. R. B., MacAulay, A., Levy, R. D., and Curtis, K. (2019). Genome-Wide transcriptome landscape of embryonic brain-derived neural stem cells exposed to alcohol with strain-specific cross-examination in BL6 and CD1 Mice. *Sci. Rep.* 9:206. doi: 10.1038/s41598-018-36059-y

Yasui, D. H., Gonzales, M. L., Aflatooni, J. O., Crary, F. K., Hu, D. J., Gavino, B. J., et al. (2014). Mice with an isoform-ablating Mecp2 exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum. Mol. Genet.* 23, 2447–2458. doi: 10.1093/hmg/ddt640

Yazdani, M., Deogracias, R., Guy, J., Poot, R. A., Bird, A., and Barde, Y. A. (2012). Disease modeling using embryonic stem cells: MeCP2 regulates nuclear size and RNA synthesis in neurons. *Stem Cells* 30, 2128–2139. doi: 10.1002/stem.1180

Zachariah, R. M., Olson, C. O., Ezeonwuka, C., and Rastegar, M. (2012). Novel MeCP2 isoform-specific antibody reveals the endogenous MeCP2E1 expression in murine brain, primary neurons and astrocytes. *PLoS One* 7:e49763. doi: 10.1371/journal.pone.0049763

Zachariah, R. M., and Rastegar, M. (2012). Linking epigenetics to human disease and Rett syndrome: the emerging novel and challenging concepts in MeCP2 research. *Neural Plast.* 2012:415825. doi: 10.1155/2012/415825

Zoghbi, H. Y., Percy, A. K., Schultz, R. J., and Fill, C. (1990). Patterns of X chromosome inactivation in the Rett syndrome. *Brain Dev.* 12, 131–135. doi: 10.1016/s0387-7604(12)80194-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 15 September 2020
doi: 10.3389/fgene.2020.563718

Preclinical and Clinical Epigenetic-Based Reconsideration of Beckwith-Wiedemann Syndrome

Chiara Papulino†, Ugo Chianese†, Maria Maddalena Nicoletti, Rosaria Benedetti‡ and Lucia Altucci*‡*

Department of Precision Medicine, Università degli Studi della Campania "Luigi Vanvitelli", Naples, Italy

Epigenetics has achieved a profound impact in the biomedical field, providing new experimental opportunities and innovative therapeutic strategies to face a plethora of diseases. In the rare diseases *scenario,* Beckwith-Wiedemann syndrome (BWS) is a pediatric pathological condition characterized by a complex molecular basis, showing alterations in the expression of different growth-regulating genes. The molecular origin of BWS is associated with impairments in the genomic imprinting of two domains at the 11p15.5 chromosomal region. The first domain contains three different regions: insulin growth like factor gene (*IGF2*), *H19,* and abnormally methylated DMR1 region. The second domain consists of cell proliferation and regulating-genes such as *CDKN1C* gene encoding for cyclin kinase inhibitor its role is to block cell proliferation. Although most cases are sporadic, about 5–10% of BWS patients have inheritance characteristics. In the 11p15.5 region, some of the patients have maternal chromosomal rearrangements while others have Uniparental Paternal Disomy UPD(11)pat. Defects in DNA methylation cause alteration of genes and the genomic structure equilibrium leading uncontrolled cell proliferation, which is a typical tumorigenesis event. Indeed, in BWS patients an increased childhood tumor predisposition is observed. Here, we summarize the latest knowledge on BWS and focus on the impact of epigenetic alterations to an increased cancer risk development and to metabolic disorders. Moreover, we highlight the correlation between assisted reproductive technologies and this rare disease. We also discuss intriguing aspects of BWS in twinning. Epigenetic therapies in clinical trials have already demonstrated effectiveness in oncological and non-oncological diseases. In this review, we propose a potential "epigenetic-based" approaches may unveil new therapeutic options for BWS patients. Although the complexity of the syndrome is high, patients can be able to lead a normal life but tumor predispositions might impair life expectancy. In this sense epigenetic therapies should have a supporting role in order to guarantee a good prognosis.

Keywords: Beckwith-Wiedemann syndrome, rare diseases, cancer predisposition, epigenetics, metabolic disorders, DNA methylation, monozygotic twins

OPEN ACCESS

Edited by:
Mojgan Rastegar,
University of Manitoba, Canada

Reviewed by:
Alessandro Mussa,
University of Turin, Italy
Silvia Russo,
Istituto Auxologico Italiano (IRCCS), Italy
Lidia Larizza,
Italian Auxological Institute (IRCCS), Italy

***Correspondence:**
Rosaria Benedetti
rosaria.benedetti@unicampania.it
Lucia Altucci
lucia.altucci@unicampania.it

†These authors have contributed equally to this work

‡These authors share last authorship

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *19 May 2020*
Accepted: *26 August 2020*
Published: *15 September 2020*

Citation:
Papulino C, Chianese U, Nicoletti MM, Benedetti R and Altucci L (2020) Preclinical and Clinical Epigenetic-Based Reconsideration of Beckwith-Wiedemann Syndrome. Front. Genet. 11:563718. doi: 10.3389/fgene.2020.563718

Abbreviations: ART, assisted reproductive technologies; BWS, Beckwith-Wiedemann syndrome; BWSp, Beckwith-Wiedemann syndrome spectrum; CNVs, copy number variations; CTCF, CCCTC-binding factor; DMR, differentially methylated region; DZ, dizygotic twins; GOM, gain of methylation; IG-DMR, intergenic differentially methylated region; ICR1, imprinting control region 1; ICR2, imprinting control region 2; LLD, leg length discrepancy; LOF, large offspring syndrome; LOM, loss of methylation; MS-MLPA, methylation-specific multiplex ligation-dependent probe amplification; MZ, monozygotic twins; ncRNA, non-coding RNA; TSS-DMR, transcription start site differentially methylated region; UPD, uniparental disomy.

INTRODUCTION

Epigenetic alterations play a crucial role in both cancer and non-oncological diseases (Mau and Yung, 2014) regulating DNA methylation, histone modifications (Nowacka-Zawisza and Wisnik, 2017) and micro-RNA expression that ultimately determines gene expression (Abi Khalil, 2014).

Some rare congenital diseases present alterations in epigenetically regulated genes (Niculescu and Lupu, 2011), providing the rationale to interfere via an epigenetic rebalance to mitigate/overcome these conditions (Nguyen, 2019). BWS is an example of a complex disease characterized by genetic and epigenetic aberrations on chromosomal region 11p15.5 comprising a telomeric and a centromeric domain, both regulated by genomic imprinting (Gomes et al., 2009). BWS is characterized by juvenile abnormal overgrowth (Choufani et al., 2010; Eggermann et al., 2015), affecting 1 child in 10,340–13,700 live births worldwide (Mussa et al., 2013). The associated phenotype may have a range of manifestations (Maas et al., 2016; Dawkins et al., 2018), although the most common clinical features are macroglossia, present in half of the patients with a molecular defect at 11p15.5 (Gaston et al., 2001; Brioude et al., 2018), visceromegalia and abdominal wall defects. Minor features include ear pits, hypoglycemia, nephromegaly, and isolated lateralized overgrowth (Mussa et al., 2016b). BWS is equally incident in males and females, but in monozygotic twins, there is observed an excess in females (Weksberg et al., 2010), and the discordant disease presentation suggests an important epigenetic role (Bliek et al., 2009b).

BWS is considered an imprinting disorder (IDs) affecting growth, development, and metabolism (Netchine et al., 2013) often caused by alterations in imprinting control regions (ICRs) in the parental 11p15.5 region (Abramowitz and Bartolomei, 2012; Begemann et al., 2012; Abi Khalil, 2014). ICRs impairment leads to abnormal methylation state deregulating genes as *CDNK1C, H19, IGF2,* and *KCNQ1OT1* involved in growth, so provoking the onset of BWS features (Wang et al., 2020). The expression/inactivation of a gene is related to differentially methylated regions (DMRs).

The majority of IDs such as BWS present the same four classes of molecular alterations lead to imbalanced gene expression: uniparental disomy (UPD), epimutation (aberrant methylation marks), copy number variations (CNVs) and point mutations in imprinted genes. Since DNA methylation marks are transmitted through generations, an interesting approach for BWS resides in the epigenetic-regulated inheritance study, which is still in an early phase. In that case, the major limitation for it to be carried on is due to the low patients' number (Nguyen, 2019). For this reason, multi-institutional collaborations are required in order to reach a statistically significant number of patients for the constitution of an observational cohort or a clinical trial (Griggs et al., 2009). Moreover, the low BWS prevalence requires special combined efforts to improve diagnosis, care, and prevention (Ayme and Schmidtke, 2007). In order to overcome this limitation and trying to boost patient care, the Coordination of Rare Diseases at Sanford (CoRDS) started in 2013 a program (NCT01793168) providing a centralized-international patients registry for all rare diseases. This program helps in the identification, advance treatments, and therapies on a large cohort of patients.

ROLE OF EPIGENETIC ALTERATIONS IN BWS

Epigenetic alterations associated with BWS are present on two different domains, independently methylated in the 11p15.5 gene cluster (Krzyzewska et al., 2019): *H19/IGF2*:IG-DMR (intergenic differentially methylated region) and *KCNQ1OT1*:TSS-DMR (transcription start site differentially methylated region) (**Figure 1**). The first region is also known as ICR1 consisting of two main imprinted genes, *IGF2* and *H19*. *IGF2* is paternally expressed and it encodes s for a fetal growth factor implicated in BWS pathogenesis. *H19* is a maternally expressed allele, non-coding RNA, and its function is still unclear, although it has been postulated a tumor suppressor role (Cai and Cullen, 2007; Raveh et al., 2015). *IGF2* deregulation determines overgrowth and region-specific tumor development in BWS (Schofield et al., 2001). *IGF2* and *H19* genes are oppositely imprinted with an enhanced competition (Maher and Reik, 2000). The *H19/IGF2*:IG-DMR domain works as a chromatin insulator regulating the expression of *IGF2* and *H19* genes. The chromatin insulator is located 2 Kb upstream of the *H19* gene and displays CTCF binding sites (Hark et al., 2000). Once CTCF binding is prevented, the enhancers can interact with *IGF2* promoters. Oppositely the unmethylated maternal allele allows CTCF binding to the insulator elements blocking the downstream enhancers of *H19* to access the *IGF2* promoter (Soejima and Higashimoto, 2013) (**Figure 2A**). H19/IGF2:IG-DMR-specific histone marks have also been reported: the methylated region presents H3K9me3 and H4K20me3, whereas the unmethylated area carries H3K4me2/3 and H3/H4 acetylated (Nativio et al., 2011). In 5-10% of BWS patients observed with a gain of methylation (GOM) of the H19/IGF2: IG-DMR on the unmethylated maternal allele (**Figure 2B**) (Eggermann et al., 2014; Brioude et al., 2018). The aberrant methylation is also characterized by variations in histone signatures. Accessible modification of H3K9ac and bivalent H3K4me2/H3K27me3 are converted to the repressive H3K9me3 and H4K20me3 (Nativio et al., 2011). These chromatin alterations prevent CTCF binding to the maternal *H19/IGF2*:IG-DMR, thus enabling the enhancers to access the *IGF2* promoter and leading both to a biallelic expression of *IGF2* and reduced expression of *H19*. In detail, *H19/IGF2*:IG-DMR is characterized by 59-bp elements hD1 and hD2, composed mainly of two Sox-Oct motif-like sequences (SO motifs) and a single Oct motif (rO) and are accompanied by triple-repeat sequences containing CBS1-3 or CBS4-6 (also known as CBSs). Recently, by using transplantation in mice, it was reported that H19/IGF2:IG-DMR methylation status is regulated by the CBSs and SO motifs (Hori et al., 2012). Interestingly in BWS patients, SO motifs are mutated or deleted. Three Oct motifs variants found in BWS patients disrupt hD1-dependent DNA demethylation and cause the stack of methylation (Kubo et al., 2020). Single nucleotide substitutions in hD1 Oct motifs present in BWS

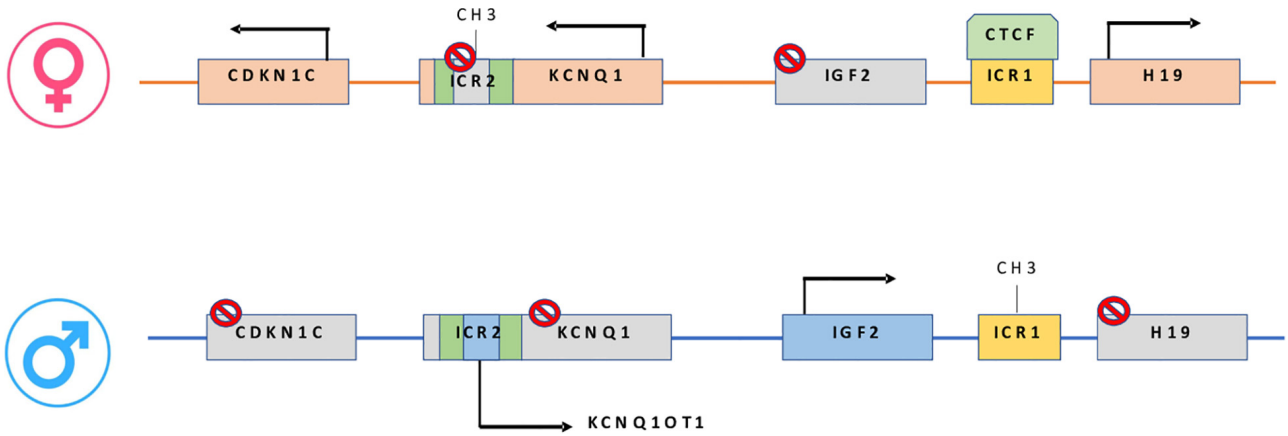

FIGURE 1 | On the maternal allele: CTCF binding H19/IGF2:IG-DMR, prevents IGF2 from accessing the shared enhancers located on the H19 side. On the paternal allele: H19/IGF2:IG-DMR methylation prevents CTCF from binding, allowing IGF2 to engage the enhancers and silencing the H19 promoter. In maternal KCNQ1OT1:TSS-DMR, methylation induces KCNQ1-CDKN1C expression while paternally unmethylated state represses KCNQ1-CDKN1C expression and increase ncKCNQ1OT1 transcription.

patients show hypermethylation in *H19/IGF2*:IG-DMR coupled with histone modifications: maternal aberrant DNA methylation is connected with reduction of H3K4me2 and H3K9ac and increase of H3K9me3 and H3K27me3 (Kubo et al., 2020). *KCNQ1OT1*:TSS-DMR contains six different imprinted genes: *KCNQ1*, *KCNQ1OT1*, *CDKN1C*, *SLC22A18*, *TSSC3*, and *PHLDA2*. *CDKN1C* is a maternally expressed gene encoding for a cell cycle regulator, a cyclin-dependent kinase inhibitor. It regulates in a negative manner cell proliferation. *KCNQ1OT1* is a paternally expressed long non-coding RNA capable to inhibit the expression of genes in the domain *in cis*. The KCNQ1OT1:TSS-DMR of this domain is localized in the intron 10 of the *KCNQ1* gene, and it is methylated on the maternal but not on the paternal allele; thus the unmethylated paternal one allows the transcription of *KCNQ1OT1*:TSS-DMR, preventing the *CDKN1C* and *KCNQ1* gene expression. In mice, this DMR interacts with G9a and the PRC2 complex leading to repressive histone modifications such as H3K9me3 and H3K27me3. The methylated *KCNQ1OT1*:TSS-DMR form present on the maternal allele prevents the transcription of the long non-coding RNA, *KCNQ1OT1,* allowing the expression of several genes present in the domain such as *CDKN1C*. Loss of methylation (LOM) on the maternal allele occurs in 50% of BWS patients (Eggermann et al., 2014; Brioude et al., 2018). This molecular defect is accompanied by changes in histone modifications, such as the loss of H3K9me2 (Soejima and Higashimoto, 2013). LOM is responsible for the biallelic expression of the *KCNQ1OT1* transcript, which reduces the *CDNK1C* expression provoking BWS phenotype. *CDNK1C* gene acquires biallelic methylation in some tumors and in 5% of BWS cases, this gene presents point mutations (Du et al., 2003). Another molecular alteration is UPD(11)p present in 10-20% of BWS patients (Eggermann et al., 2014; Brioude et al., 2018) (**Figure 2C**). UPD condition results when both chromosomes -or part of them- is/are inherited from one parent. In BWS, the UPD is paternal and involves the 11p15 region implying no maternal contribution for it. UPD is related to mitotic recombination during embryonic development. BWS patients with UPD(11)pat show mosaicism suggesting to be a post-zygotic event. The mosaicism level of UPD in different tissues is strongly associated with the pathological phenotype. The size of the UPD region or the level of mosaicism are not correlated with the severity of the disorder (Cooper et al., 2007). Patients with a mosaic UPD(11)pat for the entire chromosome 11 present clinical features comparable to UPD(11)pat cases restricted to a small part of 11p. BWS patients present UPD(11)pat isodisomic, thus determining increased *IGF2* expression and reduced *H19* expression (Wang et al., 2020). In some BWS patients, it is reported that β*2SP* gene, a TGF-β/Smad3/4 adaptor protein, a potent tumor suppressor is epigenetically silenced, thus observing a loss or markedly decreased expression of β*2SP* related to aberrant DNA methylation of CpG islands around its promoter region (Yao et al., 2010).

BWS AS MULTI-LOCUS IMPRINTING DISORDER

Some BWS patients but also other imprinting disorder cases present additional methylation defects at different imprinted loci determining a pathological condition known as multi-locus imprinting disturbance (MLID) that potentially alters the expression of multiple imprinted gene clusters (Bens et al., 2016; Sanchez-Delgado et al., 2016). The MLID frequency is related to the type of imprinting disorder and appears to vary depending on the technology sensitivity used to analyze and on Imprinted Differentially Methylated Regions (IDMRs) investigated (Poole et al., 2013). Currently, BWS shows at frequencies around 50% in cases with LOM at KCNQ1OT1:TSS-DMR (Court et al., 2013), while MLID condition is less observed among patients with GOM at H19/IGF2:IG-DMR (Maeda et al., 2014). Some of MLID BWS patients present LOM or GOM at both maternal and paternal IDMRs (Court et al., 2013; Poole et al., 2013)

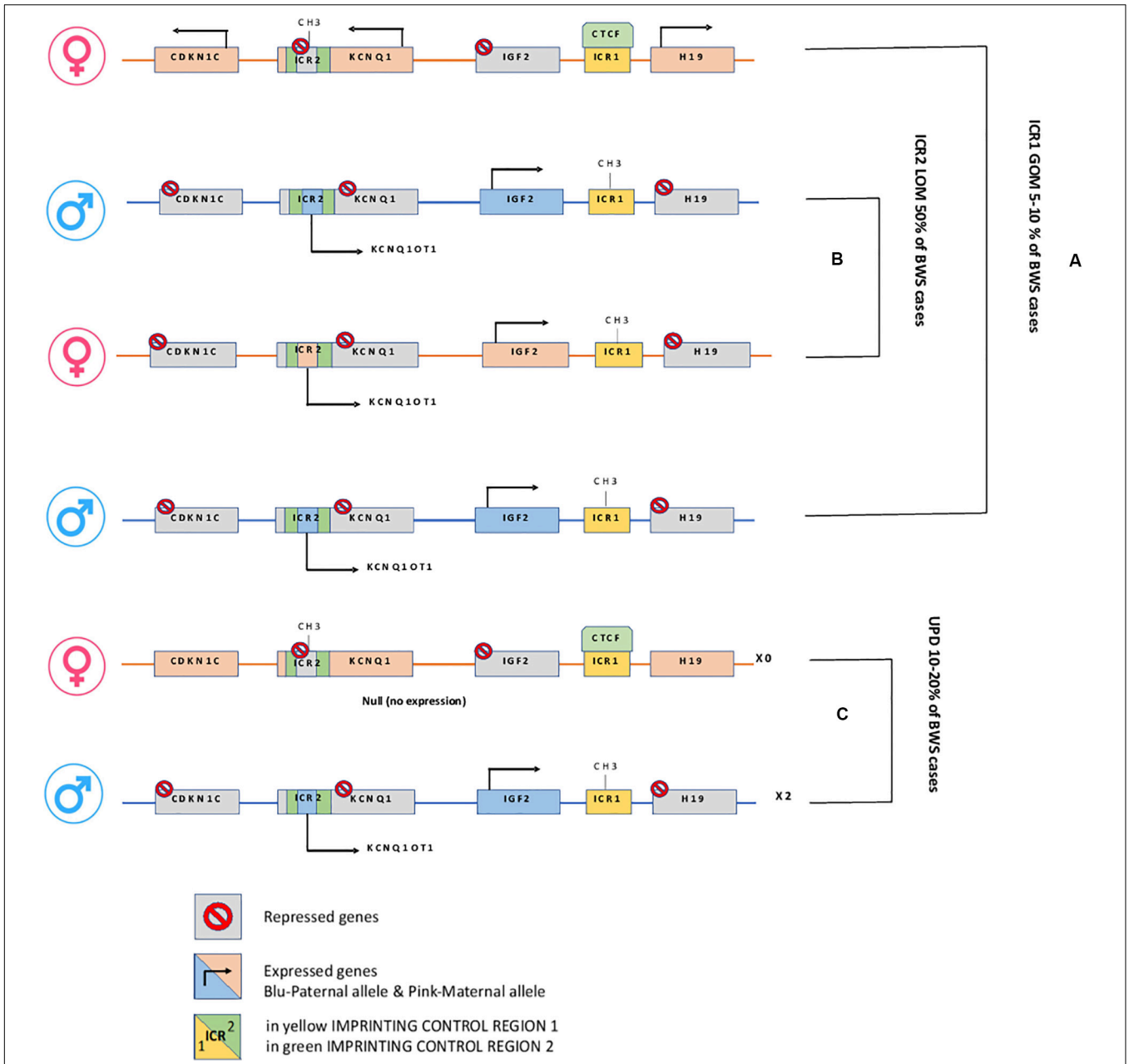

FIGURE 2 | H19/IGF2:IG-DMR and KCNQ1OT1:TSS-DMR. **(A)** In Maternal allele H19/IGF2:IG-DMR GOM leads to overexpression of the growth factor IGF2. **(B)** In Maternal KCNQ1OT1:TSS-DMR LOM, the ncRNA *KCNQ1OT1* is expressed while *cis*-linked genes are repressed. **(C)** Paternal uniparental disomy.

while others have a hypomethylation syndrome restricted to maternally imprinted genes (Boonen et al., 2008; Baple et al., 2011). MLID condition usually presents BWS clinical features but in some cases in patients shows complex or atypical phenotypes, conceivably reflecting the loci and tissues mosaicism (Docherty et al., 2014; Begemann et al., 2018) and probably as a consequence of the dominance of one locus on other (Zhang et al., 1997; Azzi et al., 2009). In a cohort study, MLID was identified only in BWS patients with KCNQ1OT1:TSS-DMR LOM and hypomethylation was found only at maternally iDMRs, a condition described as multiple maternal hypomethylation syndrome (MMHS) (Bliek et al., 2009b; Sano et al., 2016; Fontana et al., 2018). A small number of manuscripts on MLID reported as GOM at paternally methylated iDMRs in BWS patients (Court et al., 2013; Maeda et al., 2014). The most frequently altered iDMRs found in BWS patients with MLID are *PLAGL1*, *GRB10*, *MEST*, *GNAS*, *IGF2R*, and *ZNF331*. From a pathological point of view, usually, MLID BWS patients have a decreased level of body weight compared to the one who characterized by a single molecular defect in the 11p15.5 region. Specifically, MLID patients show features not

typical of BWS, such as speech retardation, apnea, and feeding difficulties (Bliek et al., 2009b). MLID-associated clinical signs may also only manifest as patients grow up; therefore MLID analysis after molecular confirmation of a specific ID, could guide a patient-tailored follow-up to track subclinical signs before their manifestation (Bakker et al., 2016). In 2005 for the first time there were reported two patients with MLID presented transient neonatal diabetes mellitus (TNDM) and hypomethylation of both iDMRs of *KCNQ1OT1* gene (Arima et al., 2005) and of *PLAGL1*, an antiapoptotic (Shuman et al., 2006) gene located at chromosome 6q24 (Bliek et al., 2009b). A further broad analysis revealed that the MLID prevalence is higher in BWS cases than in other imprinting disorders (Brioude et al., 2018). However, despite the growing number of studies, the etiology of MLID is still unclear as well as mechanisms underlying the co-regulation of imprinting marks across the genome. However, MLID causative mutations have been identified in members of the NLRP and zinc-finger protein families in few BWS patients (Docherty et al., 2014). Their role in the imprinting process and the pattern of inheritance has yet to be fully elucidated. Molecular characterization of MLID is fundamental not only to define the clinical diagnosis of IDs better but also to evidence common functional networks at the basis of the imprinting genome-wide deregulation.

GENES CANDIDATES IN OVERGROWTH FEATURE IN BWS

BWS is a childhood cancer predisposition disorder with increased risk of embryonic tumors, predominately Wilms tumor, and hepatoblastoma (Duffy et al., 2018). The chromosome 11p15.5 region contains imprinted genes that are fetal growth regulators (Brioude et al., 2018). The molecular analysis of domain 1 and 2 has identified two candidates related to tumor development in BWS: *CDKN1C* (*p57Kip2*) and *IGF2*. Their aberrations are associated with growth and development disturbances. *IGF2* is a cell cycle regulator gene encoding for a growth factor with different functions in promoting cell growth and proliferation during fetal development (Park et al., 2017). *IGF2* binds three receptors, IGF1 receptor (IGF1R), insulin receptor isoform A (IR-A), and the IGF1R–IR-A hybrid receptor, resulting in a cascade of intracellular events and promoting cell survival and mitogenesis (Alvino et al., 2011). *IGF2* loss of imprinting from DMR dysregulation on the maternal allele increases its signals, thus promoting growth and anti-apoptosis processes (Gallagher and LeRoith, 2010). In addition, the dysregulation of *IGF2* expression has been recently observed also in several tumor onsets such as for breast, ovarian, esophageal, and colorectal cancer, and its presence has been reported and associated with poor prognosis (Livingstone, 2013). *IGF2* is involved in prenatal skeletal muscle growth and in muscle regeneration in adults. It has been studied using *in vitro* differentiation models and *in vivo* loss of function in mouse models (Park et al., 2017). In mice, the *IGF2* overexpression determines BWS clinical features leading to overgrowth, polyhydramnios, fetal and neonatal lethality, disproportionate organ overgrowth, and macroglossia (Murrell et al., 2004). The *CDKN1C* encodes for a negative regulator of the cell cycle, and its function is to inhibit several Cyclin/CdK complexes. This gene maps in the centromeric region of 11p15, and it is paternally imprinted, thus implying its expression on the maternal allele. *CDKN1C* is considered a tumor suppressor gene. It blocks cell proliferation by inhibiting cell cycle progression, tissue invasion, metastasis, angiogenesis, and promotes apoptosis and cell differentiation (Kavanagh and Joseph, 2011). CDKN1C protein is composed of 316 amino acids assembled in three different domains: the N-terminal domain is fundamental for CdK inhibition, a central repeating sequence, and a C-terminal domain working as the PCNA binding domain homolog to p27Kip1 it interacts with PCNA protein (Stampone et al., 2018).

In BWS patients, the missense variants of *CDNK1C* involve highly conserved amino acids of the N-terminal domain with the CdK inhibitory role, impairing its function or its cellular localization. Until 2014 there were described 33 different point mutations in coding and non-coding regions of the gene (Eggermann et al., 2014). Most of them were missense mutations in frame-shift, leading to the production of a truncated form of the protein (Brioude et al., 2015). 5–10% of BWS patients show pathogenic variants of *CDKN1C* gene on the maternal allele, thus determining its loss of function (Eggermann et al., 2014; Brioude et al., 2018). Several *CDKN1C* gene point mutations have been identified, 40% with a BWS family history (Romanelli et al., 2010; Eggermann et al., 2014; Brioude et al., 2018). In the Human Gene Mutation Database, 65 different variants are reported in BWS patients exhibiting different clinical features including polydactyly, extra nipple, genital anomalies, and cleft palate, strongly dependent on epigenetic/chromosomal abnormalities at chromosome 11p15.5. BWS patients carrying *CDKN1C* mutations show a higher frequency of abdominal wall defects and omphalocele (Brioude et al., 2015), whereas mice lacking of a maternal copy of the *CDKN1C* gene present a phenotype close to BWS with gastrointestinal tract abnormalities or exomphalos (Yan et al., 1997; Zhang et al., 1997). *CDKN1C* mutations have also been correlated to several types of cancer, such as colorectal, lymph-hematologic, breast cancer (Larson et al., 2005), and it emphasizes the presence of commonalities features between BWS and cancer predisposition.

THE CLINICAL SPECTRUM OF BWS

Genetic and epigenetic changes frequently lead to different clinical phenotypes reported in the clinical criteria applied for the BWS definition. In **Figure 3** are reported genetic and epigenetic abnormalities in association with BWS patients' recurrence.

An abnormal epigenetic variability may cause the onset of pathological mosaic states, one with a normal genotype (Zhang et al., 1997) and the other one carrying modified epigenetic information (Sazhenova and Lebedev, 2008). Thus in 2018, due to this high complexity, the Consensus Group introduced the concept of BWS spectrum (BWSp) including patients with a clinical diagnosis of BWS with or without an epigenetic alteration at the 11p15 locus, patients with "atypical BWS" (a condition defined by fewer cardinal and suggestive features than

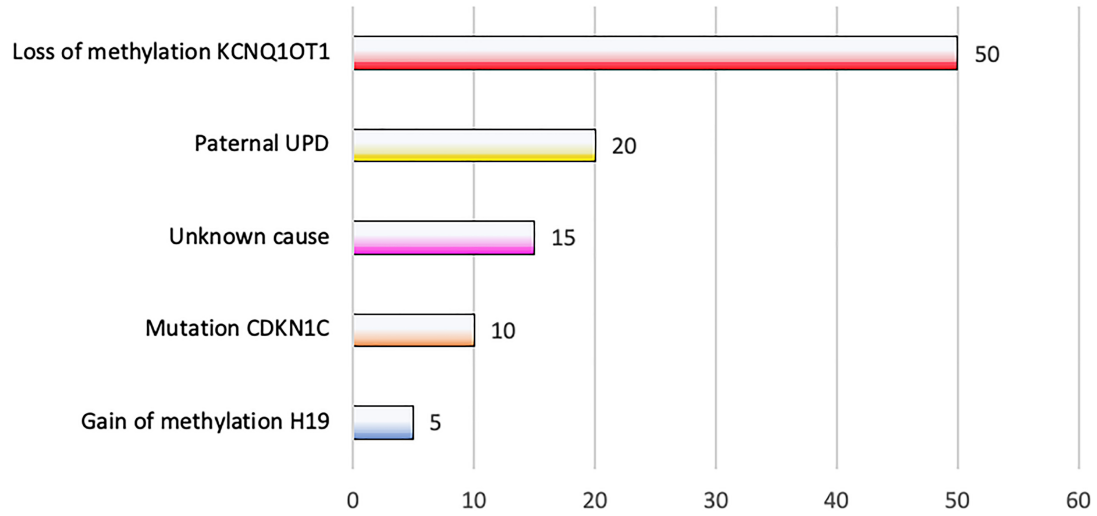

FIGURE 3 | Beckwith-Wiedemann syndrome frequency of molecular defects.

those needed for a BWS clinical diagnosis) and an epigenetic change at the BWS locus and patients with "isolated lateralized overgrowth." The accurate identification of clinical aspects is crucial for the diagnosis: cardinal features include macroglossia, exomphalos, lateralized overgrowth, multifocal Wilms tumor, prolonged hyperinsulinism, and distinct pathological findings are unique to BWS. Suggestive features are less specific but may help the clinical diagnosis and the indication for molecular testing. These features are birth weight (>2SDS above the mean), facial nevus simplex, polyhydramnios and/or placentomegaly, ear creases and/or pits, transient hypoglycemia (lasting < 1 week), typical BWSp tumors (neuroblastoma, rhabdomyosarcoma, unilateral Wilms tumor, hepatoblastoma, adrenocortical carcinoma or pheochromocytoma), nephromegaly and/or hepatomegaly, umbilical hernia and/or diastasis recti (Brioude et al., 2018). The consensus criteria for clinical diagnosis apply a scoring system: each cardinal feature gets two points while each suggestive feature gets one point. For the clinical diagnosis, a score $\geq$ 4 is required, and the molecular confirmation of 11p15.5 anomalies needs to be applied. A genetic test is also required with a score $\geq$ 2 with further evaluation of a BWS expert (Brioude et al., 2018).

BWS DEDICATED MOLECULAR INVESTIGATIONS AND PRENATAL TEST

BWS has genetic and epigenetic abnormalities, and it is phenotypically associated with different clinical aspects; therefore, the analysis of the molecular features of patients is essential for their management and treatment. First molecular testing procedure is usually applied evaluating mosaicism using DNA from blood leukocytes; then, the analysis may be implemented on additional samples such as buccal swabs, skin fibroblasts, or cells of mesenchymal origins. The procedure evaluates H19/IGF2:IG-DMR and KCNQ1OT1:TSS-DMR methylation levels as well as DMR copy number variations (CNVs) (Brioude et al., 2018). The most common diagnostic test for this purpose is Methylation-Specific Multiplex Ligation-dependent Probe Amplification (MS-MLPA) able to detect the percent of methylation and DMR copy number status simultaneously (Priolo et al., 2008; Scott et al., 2008; Wang et al., 2020). Low-level mosaicism patients require more sensitive methylation-specific techniques such as MS-PCR, MS quantitative PCR and in particular chromosomal microarray analysis (CMA) can determine the length of the UPD(11)pat region (Coffee et al., 2006; Azzi et al., 2011; Russo et al., 2016). Single gene sequencing of *CDKN1C* mutations can clarify the recurrence risk in family members (Brioude et al., 2018). Chromosome microarray might be required to define the size and the nature of duplication or deletion in the case of CNV (Baskin et al., 2014; Liu et al., 2015), whereas fluorescence *in situ* hybridization (FISH) or karyotyping may also provide the identification of chromosomal translocations (Bi et al., 2013). Prenatal molecular tests may allow handling the disease before the birth despite results reliability, and the ethical issues need to be taken into account (Brioude et al., 2018). This type of analysis requires samples such as chorionic villus (CVS) cells, amniotic fluid (AF) cells or fetal blood cells (native and cultured) (Brioude et al., 2018) but it is essential to consider that CVS cells could have different methylation patterns of the 11p15.5 region from embryonic tissues (Paganini et al., 2015) showing false-positive results due to tissue mosaicism (Brioude et al., 2018). For this reason, it is highly recommended that

a multicenter audit of cases in order to implement different methods and get a correct diagnosis (Brioude et al., 2018). Different methods and diagnostic rates might be applied (Eggermann et al., 2016; Brioude et al., 2018). Quantitative PCR from amniocytes and cord blood leucocytes in 15 weeks pregnant women (see also NCT01842659) might result useful to evaluate the imprinting region 11p15.5, the Methylation Index (MI), the Interclass Correlation Coefficient (ICC) and the agreement between these two samples. Biochemical screening results such as elevated levels of free β-human chorionic gonadotropin (hCG) in the first trimester and/or increased α-fetoprotein (αFP) levels in the second trimester (associated with exomphalos) can be associated with BWS in the fetus (Gocmen et al., 2005; Kagan et al., 2015). BWS diagnosis can also result from an ultrasonographic detection of an anterior abdominal wall defect, macroglossia, or, less accurately, from macrosomia, visceromegaly, polyhydramnios, placentomegaly, or pancreatic overgrowth. BWS diagnosis can also be proposed by prenatal ultrasound scan (USS) identifying placental mesenchymal dysplasia, urinary tract abnormalities, cardiac defects, adrenal cysts and masses (Kagan et al., 2015; Brioude et al., 2018). Williams et al., 2005 reported diagnostic scheme based on the identification by ultrasound examination of different findings: at least two main (i.e., an abdominal wall defect, macroglossia, or macrosomia), or one main and two minor findings are required to lead BWS diagnosis (i.e., nephromegaly/dysgenesis, adrenal cytomegaly, aneuploidy/abnormal loci, or polyhydramnios) (Williams et al., 2005).

PREDISPOSITION IN TUMOR DEVELOPMENT: ESTIMATED CANCER RISK IN BWS

Epigenetic aberrations effects in BWS phenotype may determine different tumor predisposition (Segers et al., 2012). Numerous pieces of evidence suggest that cancer is strongly related to the patient's age (Kalish et al., 2016; MacFarland et al., 2018). There is a high risk during the first 4 years of life (11%) reduced to 3% in the 4–10 years-old patients (DeBaun and Tucker, 1998). Screening and diagnostic procedures should be applied for early diagnosis. The screening phase consists of repeated abdominal ultrasound (every 3–4 months during the first 8-10 years of life) and serum α-fetoprotein (aFP) measurement (every 3 months during the first 30 months of life) (Mussa et al., 2019). These procedures are applied to all patients regardless of the molecular diagnosis and the genotype (Zhang et al., 1997).

BWS molecular subtypes are related to tumor predisposition, and each subtype is associated with the development of one (or more) type of cancer (Mussa et al., 2016a). In **Table 1** are reported the molecular subgroups and cancer types occurrence.

The four main molecular subtypes of BWS (KCNQ1OT1:TSS-DMR-LOM, H19/IGF2:IG-DMR -GOM, UPD, and *CDKN1C* mutations) are characterized by specific genotype-phenotype correlation to tumor development risk (Ibrahim et al., 2014; Mussa et al., 2016a). Patients with 11p15 telomeric domain defects (H19/IGF2:IG-DMR -GOM and UPD) have a higher risk of developing cancer compared to cases presenting centromeric aberrations (KCNQ1OT1:TSS-DMR-LOM mutation and *CDKN1C*) (Soejima and Higashimoto, 2013; Maas et al., 2016). The molecular subtype characterization of each BWS case increases those patients at the highest cancer risk focusing on a specific cancer. Subtypes analysis and embryonic tumor incidence of BWS cases may guarantee a more efficient surveillance optimizing tumor screening. As previously reported, a meta-analysis collecting results from seven studies, including 1370 genotyped BWS patients identified 102 cases with BWS-related malignancies (Zhang et al., 1997; Mussa et al., 2016a). It is interesting to note how different was the prevalence among the molecular subtypes: 2.5% (21/836) in KCNQ1OT1:TSS-DMR-LOM, 13.8% (47/341) in UPD, 22.8% (28/123) in H19/IGF2:IG-DMR -GOM and 8.6% (6/70) in patients with *CDKN1C* mutations. Wilms tumor represents the most common cancer in BWS patients (Rump et al., 2005) with a high prevalence in combination with telomeric defects in H19/IGF2:IG-DMR -GOM and UPD subgroups (21.1% vs. 6.2%, respectively, $P < 0.001$); subjects with centromeric defects display a lower rate. The tumor development in H19/IGF2:IG-DMR -GOM cases is significantly higher compared to UPD cases. Adrenal carcinoma has only been observed in UPD (1.5%, $P < 0.001$). Hepatoblastoma development has been associated with UPD (4.7%, $P < 0.001$) (Mussa et al., 2016a) although it is also observed in patients with KCNQ1OT1:TSS-DMR-LOM (0.7%) and H19/IGF2:IG-DMR -GOM (0.8%, $P < 0.001$). Neuroblastic

TABLE 1 | Association of BWS molecular defects in subgroups and tumor risk frequency.

Molecular defect	Alteration	Frequency of molecular defects	Tumor risk compared with other molecular subgroups
H19/IGF2:IG-DMR HYPERmethylation	Hypermethylation	5–10%	High risk of Wilms tumor
UPD(11)pat	Paternal UPD	20%	High risk of Wilms tumor and hepatoblastoma
KCNQ1OT1:TSS-DMR HYPOmethylation	Hypomethylation	50%	Tumor incidence is lower than the other molecular subgroups and is very variable
CDKN1C mutations	Loss of function mutations	5–10% (Sporadic cases 5%; Familial cases 40%)	Low risk of Wilms tumor

Adapted from Cooper et al. (2005), Eggermann et al. (2014), Maas et al. (2016), Mussa et al. (2016a), and Brioude et al. (2018).

tumors have been correlated with *CDKN1C* mutations (4.3%, $P = 0.003$) despite also observed in KCNQ1OT1:TSS-DMR-LOM (0.5%) and UPD cases (0.9%) with lower prevalence (Brioude et al., 2018). Additional studies are required to implement the available data since some associations seem to be clinically relevant, although not statistically significant (Mussa et al., 2016a). In general, this meta-analysis confirmed previous studies reporting the most common histotypes associated with BWS, such as Wilms tumor, hepatoblastoma, neuroblastic tumors, adrenal carcinoma, and rhabdomyosarcoma (Lapunzina, 2005; Shuman et al., 2006). The overall tumor risk of H19/IGF2:IG-DMR GOM is ∼23%, specifically with a 21% risk of developing Wilms tumor (Maas et al., 2016). The lowest tumor risk regards KCNQ1OT1:TSS-DMR hypomethylated subgroup, although there is a remarkable tumor variability. This subgroup shows different cancer types: hepatoblastoma, rhabdomyosarcoma, and gonadoblastoma but not in Wilms tumor. Patients with H19/IGF2:IG-DMR hypermethylated present Wilms tumor and hepatoblastoma their recurrence is related to *IGF2* overexpression during cancer development (Akmal et al., 1995; Rump et al., 2005; Maas et al., 2016; Mussa et al., 2017; Brioude et al., 2019). Several studies observing different cohorts of BWS patients confirmed higher tumor risk associated with the H19/IGF2:IG-DMR hypermethylated and UPD(11)pat subgroup and high frequency for Wilms tumor and hepatoblastoma (Maas et al., 2016; Ounap, 2016; Brioude et al., 2018; Kamien et al., 2018; MacFarland et al., 2018; Wang et al., 2020). Wilms tumor rate is more frequent in the H19/IGF2:IG-DMR subgroup than in the cases observed for the UPD (Wang et al., 2020). The UPD subgroup is associated with a high prevalence of hemihyperplasia and hepatoblastoma (Maas et al., 2016; Mussa et al., 2017). Only four cases of BWS children belonging to KCNQ1OT1:TSS-DMR hypomethylated subgroup had Wilms tumor or nephrogenic remnants. Adrenocortical tumors have a percentage of 3% in BWS cases, and few of them are associated with the LOM *KCNQ1OT1* gene (Alsultan et al., 2008; Mama et al., 2014). Two cases of *KCNQ1OT1* LOM with adrenocortical tumors also recently observed (Wijnen et al., 2012), and an additional one recently reported, although neither of these patients presented typical phenotypic features of BWS (Wijnen et al., 2012; Eltan et al., 2020). Guidelines are discordant because the European consensus does not recommend this screening (Brioude et al., 2018) that is required in the USA consensus (Kalish et al., 2017).

METABOLIC IMBALANCE

Metabolic disorders are one of the major clinical conditions in BWS (Schiff et al., 1973). Among metabolic imbalances, hyperinsulinemia/hypoglycemia are pathological states distressing 50% of BWS patients (Martinez y Martinez et al., 1992). Congenital hyperinsulinism (HI) is associated with a dysregulation of insulin secretion from pancreatic β-cells, and the molecular etiology of HI is due to mutations in ABCC8 and KCNJ11 genes located at the 11p15 region encoding for two subunits of the pancreatic β-cell ATP-sensitive potassium channel (KATP channel), SUR1 and Kir6.2 (Tung et al., 2020). Although only a few cases of BWS show HI, 50% of BWS neonates present transient HI while 5% have persistent HI requiring medical and/or surgical management (Senniappan et al., 2015; Kalish et al., 2016). Hypoglycemia can start in the neonatal period during the first days of life (Sweet et al., 2013). Glycemic disorders are related to aberrations of tumor suppressor genes (*IGF2*, *H19,* and *p57KIP2)* located in the 11p15 region and associated with BWS (Lee et al., 1999). Anomalies in the type 1 sulfonylurea receptor (SUR1) gene on chromosome 11p15 have been reported (de Lonlay-Debeney et al., 1999; Glaser et al., 2000). *IGF2* is overexpressed in 20% of BWS individuals (Weksberg et al., 1993), and its loss of imprinting is responsible for hypoglycemia (Lee et al., 1999). Loss of imprinting and UPD on the paternal allele in the 11p15 region (Slavotinek et al., 1997) leads to *IGF2* gene overexpression. IGF2 protein binding to the insulin receptor sustains the hyperinsulinemia condition. It is therefore not surprising that hyperinsulinemic hypoglycemia is often associated with BWS diagnosis. Hypoglycemia BWS cases reported hyperinsulinism and inappropriate insulin secretion (Shepherd et al., 2000). Pancreatic β cell dysregulation has been linked to the cause of hyperinsulinism in BWS (Stanley, 1997). In many histological analyses, hypertrophy and hyperplasia (Lteif and Schwenk, 1999) were observed, strengthening the correlation between incorrect pancreatic activity and BWS (Laje et al., 2013). However, the progression toward type 1 diabetes has not been documented (Leibowitz et al., 1995). The high levels of insulin in the blood cause a prolonged lowering of glucose concentrations, although the mechanism of insulin release in the different secretagogues occurs in different ways (Munns and Batch, 2001). In most BWS patients, life expectancy is good (Weng et al., 1995), and the metabolic imbalance tends to improve within time. However, in some cases, prolonged drug treatment is required to control hypoglycemia (Shilyansky et al., 1997). To obtain better results for glycemic control, it depends upon timeliness (Aynsley-Green et al., 2000). Unfortunately, in 20% of BWS cases, hypoglycemia is difficult to control and may cause severe decompensations leading to neurological alterations with severe repercussions in cognitive function development as intellectual impairment (Cresto et al., 1998). In severe cases, partial pancreatectomy may be required (Martinez y Martinez et al., 1992; Laje et al., 2013).

MONOZYGOTIC TWINS DISCORDANCE

In twins, it has been observed a discordant monozygotic phenomenon (Weksberg et al., 2002), whereas BWS afflicts one subject, although the other twin may have some characteristics of the syndrome. Theoretically, monozygotic twins (MZ) resulting from a single zygote should have identical genomes. However, several examples of genetic differences have been reported among MZ, thus suggesting that somatic changes may occur after conception (Bliek et al., 2009a). The mosaicism is leading to a discordance between MZ in BWS, and it is linked to an epigenetic event triggered by the twinning process. Cells involved in this event spread among embryos in a multiple pregnancy creating a mosaic distribution being responsible for the variable

phenotypic spectrum observed in BWS cases (Cohen et al., 2019). The timing of epigenetic aberrations influences the twinning, the degree of severity of BWS, and the degree of mosaicism. The theory of "diffuse mosaicism" proposed by Cohen et al. (2019) outlines the time points when the epigenetic event occurs in relation to twinning and the determination of chorionicity. In singleton gestations, the epigenetic aberration has been reported to occur first in non-mosaic patients during embryogenesis and subsequently in mosaic patients. In dichorionic gestations, the zygosity determines the timing of the event. It occurs earlier in the dizygotic dichorionic pregnancies than in the monozygotic dichorionic pregnancies (Cohen et al., 2019). Previous research has shown that an epigenetic event before twinning leads to the formation of two different clonal cell populations (Bell and Spector, 2011). These different cell clones repel each other and trigger the twinning event leading to the formation of separate cell masses (Hall, 1996; Machin, 1996; Weksberg et al., 2002). Therefore it is possible that a cell group carrying an imprinting alteration of *KCNQ1OT1* (LOM) could preferably increase its growth rate compared to normal cells, thus generating asymmetry of the entire cell mass and increasing the possibility of cell clones separation genotypically distinct (Weksberg et al., 2002). Most BWS MZ exhibit *KCNQ1OT1* imprinting defects (in KCNQ1OT1:TSS-DMR), indicating that monozygotic twinning is mechanically linked to the imprinting error or, conversely, that epigenetic alterations in KvDMR1 (LOM) can increase the possibility of monozygotic twinning. KCNQ1OT1:TSS-DMR hypomethylation is related to a failure in methylation, and it coincides or occurs shortly after the twinning event (Bliek et al., 2009a; Castillo-Fernandez et al., 2014).

Differences in genotype and phenotype can be attributed to various causes, including non-random inactivation X (Machin, 1996). These pieces of evidence agree with the higher incidence observed in monozygotic discordant female twins, suggesting that monozygotic twinning, genomic imprinting, and X inactivation may be mechanically and temporally related events (Lubinsky and Hall, 1991; Weksberg et al., 2002). In most reports, there is a high prevalence of MZ female and very few cases of discordant male twins (Weksberg et al., 2002; Smith et al., 2006; Bliek et al., 2009a; Tierling et al., 2011). In 250 BWS patients, 20 sets of monozygotic and 2 sets of dizygotic twins with high prevalence (16 out of 20) for female MZ were identified (Weksberg et al., 2002). Later in 2009, another study showed a high MZ twinning rate of 2.5% as compared with 0.3–0.4% among normal twins, while dizygotic (DZ) cases reported 0.75% as rate with 0.7–1.1% of prevalence. A female excess among BWS multiple births was observed in this study. In 10 MZ, 9 were females, and all 3 cases of DZ were females (Bliek et al., 2009a). More recently, in 2019, the high incidence of monozygotic female in a cohort of 26 BWS twins was confirmed (Cohen et al., 2019). The significant female preponderance of the MZ discordant for BWS could be associated with a variety of sex-related factors. For example, the developmental error can occur equally in male and female embryos, demonstrating a lethal effect on male MZ or in alternative the delay of early development in female embryos compared to men (Hall, 1996). It is subordinated to the X inactivation process, and it can increase the susceptibility in female MZ embryos to certain errors. The double discrepancy is due to the failure of Dnmt1o (DNMT1oocyte) to maintain methylation in phase S of a cell cycle occurring before or during the twinning event (Bestor, 2003). It has been hypothesized that the excess of twins in BWS patients is secondary to X inactivation with delayed embryogenesis allowing the acquisition of errors such as failure of maintenance methylation (Lubinsky and Hall, 1991; Orstavik et al., 1995; Hall, 1996; Hall and Lopez-Rangel, 1996; Weksberg et al., 2005). This evidence supports the high incidence of female BWS MZ. In two more extensive studies (Gaston et al., 2001; Weksberg et al., 2002) KCNQ1OT1:TSS-DMR hypomethylation from DNA of blood samples has been reported in affected and unaffected twins of discordant couples. It has been proposed that the aberrant methylation in the blood of the healthy twin is caused by vascular connections in the placenta shared by both MZ (monochorionic, diamniotic). This failure occurs in the eight-cell blastocyst stage preceding the moment when MZ (monochorionic, diamniotic) twinning is established. However, not all methylation defects involve twinning, and most BWS patients are singletons (Cohen et al., 2019). An explanation of these assumptions is related to the theory of endangered twins in which twinning occurs in all cases, but the second fetus is reabsorbed in early pregnancy (Landy and Keith, 1998). The twin discordance of BWS patients is observed not only in females but also in males (Smith et al., 2006). A pair of male MZ discordant for BWS was reported, and among these, the affected twin had paternal UPD for chromosome 11p15. The second male twin pair was concordant, and both demonstrated H19/IGF2:IG-DMR hypermethylation, thus suggesting BWS-related molecular heterogeneity in male MZ (Smith et al., 2006).

ASSISTED REPRODUCTION TECHNIQUES (ART) AND BWS EPIGENOME

The association between ART and syndromes related to epigenetic defects has been reported in several cases such as BWS (Dhont et al., 1999; Ferraretti et al., 2012, 2013; Kupka et al., 2014), large offspring syndrome in ruminants (LOF) (de Mouzon et al., 2012) and Angelman syndrome (MIM 105830) (Talaulikar and Arulkumaran, 2012) (MRC Working Party on Children Conceived by *In Vitro* Fertilization, 1990). In 1995 a BWS patient conceived through ART (Sutcliffe et al., 1995) was reported, Young et al. (1998) described the LOF, etiologically correlated with *in vitro* fertilization (Kupka et al., 2014). The LOF shows a significant effect on the phenotype connected with ART (Li et al., 2019), and this syndrome is a model for BWS (Chen et al., 2013) presenting similar phenotypic abnormalities (Young et al., 2001). In LOF affected bovines was reported an association between multiple loci imprinted defects and ART. This evidence underlines how ART may cause imprinting disturbances (Mussa et al., 2017). Most of the patients suffering from imprinting disturbances conceived via *in vitro* fertilization (Kupka et al., 2014) and intracytoplasmic sperm injection (ICSI) showed aberrant imprinted DNA methylation (Hattori et al., 2019). In the paper of DeBaun et al. (2003) were identified seven

sporadic cases conceived by ART and epigenetic alterations were present in six of them generally associated with BWS. Several studies have further explored this association (Maher and Reik, 2000; Gicquel et al., 2003; Halliday et al., 2004; Chang et al., 2005; Rossignol et al., 2006; Sutcliffe et al., 2006; Doornbos et al., 2007; Lim et al., 2009; Hiura et al., 2012; Tee et al., 2013). The most recent literature also has corroborated the hypothesis that BWS is related to ART (Vermeiden and Bernardus, 2013; Hattori et al., 2019). Indeed, LOM of KCNQ1OT1:TSS-DMR represents the molecular defect found in BWS patients conceived through ART (Gomes et al., 2009; Mussa et al., 2017). In 2015, the clinical study NCT00773825 on the association between BWS and ART was completed, although the results are not yet reported. The authors investigated the methylation status at nine different loci and other epigenetic marks using Southern blot and methyl-specific quantitative PCR in three groups of patients: children naturally conceived, children conceived after ovarian stimulation, but *in vivo* fertilization, and a group of children conceived after ovarian stimulation and *in vitro* fertilization. The aim of this trial was to determine if children born following ART exhibit an increased risk of imprinting defects. Moreover, previous results showed that ART might favor imprinting alterations at the centromeric imprinted 11p15 locus and, consequently, the incidence of BWS. Some of BWS patients reported DNA methylation defects abnormal methylation at loci other than the 11p15 region. This condition was present in both BWS patients naturally, and ART conceived (Gicquel et al., 2008), suggesting that ART procedure could be not specifically involved in the loss of methylation at various imprinting loci (Rossignol et al., 2006).

TREATMENT APPROACHES AND PERSPECTIVES FOR EPI-BASED THERAPIES

Currently, there is no dedicated therapy for BWS, and all available treatments are mainly addressing clinical features for ameliorating the quality of life. The current on-going clinical trial NCT01916148 aims to determine the ability F-DOPA PET, a PET radiotracer, to detect focal lesions prior to surgical intervention in BWS patients with hyperinsulinemic hypoglycemia. This study is useful to guarantee an early diagnosis of the pathology and to manage available treatments for BWS patients. Macroglossia occurs in 90% of BWS patients and may regress spontaneously in some children, but 40% of them undergo surgery to reduce tongue size (Brioude et al., 2018). The regional overgrowth in BWS can be progressive or non-progressive (Burkardt et al., 2019). It occurs in 43-60% of patients, and the management is related to the affected limbs (Wang et al., 2020). Leg length discrepancy (LLD) can influence negative life quality and may require shoe lifts or, in some cases, surgical correction (Ghanem et al., 2011; Brioude et al., 2018). The asymmetric overgrowth of the upper limbs generally does not require surgery (Brioude et al., 2018). New strategies are necessary to improve the possible and available treatment for BWS patients (Swinney and Xia, 2014). Many studies have demonstrated that by reprogramming the epigenetic landscape, it is possible to modulate the defects present in the genome leading to the treatment of different diseases (Miranda Furtado et al., 2019). Epigenetic markers can be targeted by activators or inhibitors of epigenetic-modifying proteins (Lauschke et al., 2018), the so-called "epidrugs" currently used mainly in tumor treatments, such as hematological malignancies (Woods et al., 2015; Mazzone et al., 2017). The effectiveness of these treatments is slowly widening toward new fields of application. Epidrugs are demonstrating their potential in other pathologies, such as infectious diseases, metabolic and cardiovascular disorders (Das et al., 2009; Dunn and Rao, 2017). There are promising clinical advances in epigenetics toward new drug discovery (Crea et al., 2011; Glasgow et al., 2015; Lundstrom, 2017) and biomarkers (Shao et al., 2018) in order to limit epigenetic mutation effects (Vitiello et al., 2015; Inamura, 2017). Epigenetic (i.e., hypermethylation of the tumor suppressor gene promoters) and genetic mutations of epigenetic enzymes (loss or gain of function) can be used as predictors of therapy response in different types of diseases (Welch and Clegg, 2010). Prolonged re-expression of epigenetically silenced genes has been demonstrated for various genes, including tumor suppressor genes (Gaur et al., 2015). BWS might represent one of the many new challenges for epigenetic treatment-based applications. For example, it is possible to mitigate CpG island methylation on maternal and/or paternal allele to restore the normal transcriptional activity in the imprinting control regions (Bartolomei and Ferguson-Smith, 2011). Epigenetic treatments might become a valid opportunity for aberrations in 11p15.5 imprinted region (Smith et al., 2012). The Food and Drug Administration (FDA) in the United States has approved several DNA methylation inhibitors, including cytidine analogs 5-azacitidine and zebularine and nucleoside analogs. Histone deacetylase inhibitors (Cheng et al., 2019) such as suberoylanilide hydroxamic acid (SAHA, trade name Vorinostat), romidepsin (trade name Istodax), Valproic acid (VPA) and trichostatin A (TSA) (Kelly et al., 2010; Heerboth et al., 2014) disrupt deacetylation process. New therapeutic programs and technique advancements might be applied to reprogram the epigenetic circuit and to counteract chronic symptoms. Since the main BWS targets are determined by deregulation on epigenetic processes, it should be interesting to evaluate potential targets by using drugs against the activity of DNA methyltransferases and or histone deacetylases and histone acetylation. These drugs should be potential treatments against BWS. Ideally, the first application may be a combination therapy to control disease progression. For example, surgical techniques for phenotypic abnormalities control (Wang et al., 2020) might be supported by the use of epi-based treatments for the metabolic imbalances in young BWS patients. The epigenetic sensitization to radiotherapy might provide promising results in BWS affected by Wilms tumor. Moreover, given the potential therapeutic role of epigenetic modulating agents in metabolic disorders (Crispo et al., 2019), it is tempting to hypothesize that in the near future, by targeting epi-modifiers and remodelers might prove beneficial also in BWS. Clearly, given the genome and epigenome heterogeneity and complexity of this disease, patient's stratification may represent a '*conditio sine qua non*' for future epi-based applications.

CONCLUSION

The progress in epigenetic drugs discovery (Morera et al., 2016; Esteller, 2017; Velasco and Francastel, 2019), the involvement of epigenetic mutations in a wide range of diseases (Dirks et al., 2016; Graca et al., 2016; Jones et al., 2016; Berdasco and Esteller, 2019) and the newly chromatin-based identified disease biomarkers (Dirks et al., 2016) have suggested epi-based approaches as a promising tool for clinical applications. Evidently, potential new therapeutic options require better clinical knowledge. Indeed, the target identification and characterization are at the basis for a correct therapy. In the case of BWS there are some levels of complexity to be decrypted. Being a rare disease, the generally low number of patients is a bottleneck and a hindrance to the development of dedicated therapeutic approaches nor the causal identification of "druggable" targets proven beneficial for the restoration of the health status or, at least, for symptoms defeat. The lesson learned from BWS homozygotic twins not only suggests the existence of a phenotypic link between epigenome deregulation and the complexity of BWS disease, but it also indicates the potential of targeting the epigenome pharmacologically in BWS to at least obtain beneficial effects ameliorating the quality of life. Indeed, on one side, the different disease phenotypes in homozygotic BWS twins have consolidated the idea that the disease penetrance is epigenome-regulated, despite somatic changes may occur after conception. On the other side, this proves in humans the potential of a therapeutic approach targeting BWS epigenome. Recently we have just started to understand the pivotal role of epigenome deregulation in BWS (and other rare diseases) supporting the development of diagnostic, prognostic, and therapeutic approaches also based on this notion. In this perspective, therapeutic approaches might also be applied to epimutations related to tumor predisposition in BWS, which might gain benefit from the use of treatment schemes, including or based on chromatin acting drugs.

AUTHOR CONTRIBUTIONS

CP, UC, RB, and LA: conceptualization. LA and RB: funding acquisition. CP, UC, MN, and RB: writing – original draft preparation. RB, LA, and MN: writing – review and editing. MN: writing and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by "Epigenetic Hallmarks of Multiple Sclerosis" (acronym Epi-MS) (id:415, Merit Ranking Area ERC LS) in VALERE 2019 Program; V:ALERE 2020 – Progetto competitivo "CIRCE" in risposta al bando D.R. n. 138 del 17/02/2020 Program; Blueprint 282510; MIUR20152TE5PK; EPICHEMBIO CM1406; EPIGEN-MIUR-CNR; AIRC-17217; VALERE: Vanvitelli per la Ricerca; Campania Regional Government Technology Platform Lotta alle Patologie Oncologiche: iCURE; Campania Regional Government FASE2: IDEAL. MIUR, Proof of Concept POC01_00043. POR Campania FSE 2014-2020 ASSE III.

REFERENCES

Abi Khalil, C. (2014). The emerging role of epigenetics in cardiovascular disease. *Ther. Adv. Chronic. Dis.* 5, 178–187. doi: 10.1177/2040622314529325

Abramowitz, L. K., and Bartolomei, M. S. (2012). Genomic imprinting: recognition and marking of imprinted loci. *Curr. Opin. Genet. Dev.* 22, 72–78. doi: 10.1016/j.gde.2011.12.001

Alsultan, A., Lovell, M. A., Hayes, K. L., Allshouse, M. J., and Garrington, T. P. (2008). Simultaneous occurrence of right adrenocortical tumor and left adrenal neuroblastoma in an infant with Beckwith-Wiedemann syndrome. *Pediatr. Blood Cancer* 51, 695–698. doi: 10.1002/pbc.21694

Akmal, S. N., Yun, K., MacLay, J., Higami, Y., and Ikeda, T. (1995). Insulin-like growth factor 2 and insulin-like growth factor binding protein 2 expression in hepatoblastoma. *Hum. Pathol.* 26, 846–851. doi: 10.1016/0046-8177(95)90005-5

Alvino, C. L., Ong, S. C., McNeil, K. A., Delaine, C., Booker, G. W., Wallace, J. C., et al. (2011). Understanding the mechanism of insulin and insulin-like growth factor (IGF) receptor activation by IGF-II. *PLoS One* 6:e27488. doi: 10.1371/journal.pone.0027488

Arima, T., Kamikihara, T., Hayashida, T., Kato, K., Inoue, T., Shirayoshi, Y., et al. (2005). ZAC, LIT1 (KCNQ1OT1) and p57KIP2 (CDKN1C) are in an imprinted gene network that may play a role in Beckwith-Wiedemann syndrome. *Nucleic Acids Res.* 33, 2650–2660. doi: 10.1093/nar/gki555

Ayme, S., and Schmidtke, J. (2007). Networking for rare diseases: a necessity for Europe. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 50, 1477–1483. doi: 10.1007/s00103-007-0381-9

Aynsley-Green, A., Hussain, K., Hall, J., Saudubray, J. M., Nihoul-Fekete, C., De Lonlay-Debeney, P., et al. (2000). Practical management of hyperinsulinism in infancy. *Arch. Dis. Child Fetal Neonatal Ed.* 82, F98–F107. doi: 10.1136/fn.82.2.f98

Azzi, S., Rossignol, S., Steunou, V., Sas, T., Thibaud, N., Danton, F., et al. (2009). Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. *Hum. Mol. Genet.* 18, 4724–4733. doi: 10.1093/hmg/ddp435

Azzi, S., Steunou, V., Rousseau, A., Rossignol, S., Thibaud, N., Danton, F., et al. (2011). Allele-specific methylated multiplex real-time quantitative PCR (ASMM RTQ-PCR), a powerful method for diagnosing loss of imprinting of the 11p15 region in Russell Silver and Beckwith Wiedemann syndromes. *Hum. Mutat.* 32, 249–258. doi: 10.1002/humu.21403

Bakker, J. P., Wang, R., Weng, J., Aloia, M. S., Toth, C., Morrical, M. G., et al. (2016). Motivational Enhancement for Increasing Adherence to CPAP: a randomized controlled trial. *Chest* 150, 337–345. doi: 10.1016/j.chest.2016.03.019

Baple, E. L., Poole, R. L., Mansour, S., Willoughby, C., Temple, I. K., Docherty, L. E., et al. (2011). An atypical case of hypomethylation at multiple imprinted loci. *Eur. J. Hum. Genet.* 19, 360–362. doi: 10.1038/ejhg.2010.218

Bartolomei, M. S., and Ferguson-Smith, A. C. (2011). Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* 3:a002592. doi: 10.1101/cshperspect.a002592

Baskin, B., Choufani, S., Chen, Y. A., Shuman, C., Parkinson, N., Lemyre, E., et al. (2014). High frequency of copy number variations (CNVs) in the chromosome 11p15 region in patients with Beckwith-Wiedemann syndrome. *Hum. Genet.* 133, 321–330. doi: 10.1007/s00439-013-1379-z

Begemann, M., Rezwan, F. I., Beygo, J., Docherty, L. E., Kolarova, J., Schroeder, C., et al. (2018). Maternal variants in NLRP and other maternal effect proteins are associated with multilocus imprinting disturbance in offspring. *J. Med. Genet.* 55, 497–504. doi: 10.1136/jmedgenet-2017-105190

Begemann, M., Spengler, S., Gogiel, M., Grasshoff, U., Bonin, M., Betz, R. C., et al. (2012). Clinical significance of copy number variations in the 11p15.5

imprinting control regions: new cases and review of the literature. *J. Med. Genet.* 49, 547–553. doi: 10.1136/jmedgenet-2012-100967

Bell, J. T., and Spector, T. D. (2011). A twin approach to unraveling epigenetics. *Trends Genet.* 27, 116–125. doi: 10.1016/j.tig.2010.12.005

Bens, S., Kolarova, J., Beygo, J., Buiting, K., Caliebe, A., Eggermann, T., et al. (2016). Phenotypic spectrum and extent of DNA methylation defects associated with multilocus imprinting disturbances. *Epigenomics* 8, 801–816. doi: 10.2217/epi-2016-0007

Berdasco, M., and Esteller, M. (2019). Clinical epigenetics: seizing opportunities for translation. *Nat. Rev. Genet.* 20, 109–127. doi: 10.1038/s41576-018-0074-2

Bestor, T. H. (2003). Imprinting errors and developmental asymmetry. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 1411–1415. doi: 10.1098/rstb.2003.1323

Bi, W., Borgan, C., Pursley, A. N., Hixson, P., Shaw, C. A., Bacino, C. A., et al. (2013). Comparison of chromosome analysis and chromosomal microarray analysis: What is the value of chromosome analysis in today's genomic array era? *Genet. Med.* 15, 450–457. doi: 10.1038/gim.2012.152

Bliek, J., Alders, M., Maas, S. M., Oostra, R. J., Mackay, D. M., van der Lip, K., et al. (2009a). Lessons from BWS twins: complex maternal and paternal hypomethylation and a common source of haematopoietic stem cells. *Eur. J. Hum. Genet.* 17, 1625–1634. doi: 10.1038/ejhg.2009.77

Bliek, J., Verde, G., Callaway, J., Maas, S. M., De Crescenzo, A., Sparago, A., et al. (2009b). Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* 17, 611–619. doi: 10.1038/ejhg.2008.233

Boonen, S. E., Porksen, S., Mackay, D. J., Oestergaard, E., Olsen, B., Brondum-Nielsen, K., et al. (2008). Clinical characterisation of the multiple maternal hypomethylation syndrome in siblings. *Eur. J. Hum. Genet.* 16, 453–461. doi: 10.1038/sj.ejhg.5201993

Brioude, F., Kalish, J. M., Mussa, A., Foster, A. C., Bliek, J., Ferrero, G. B., et al. (2018). Expert consensus document: clinical and molecular diagnosis, screening and management of Beckwith-Wiedemann syndrome: an international consensus statement. *Nat. Rev. Endocrinol.* 14, 229–249. doi: 10.1038/nrendo.2017.166

Brioude, F., Netchine, I., Praz, F., Le Jule, M., Calmel, C., Lacombe, D., et al. (2015). Mutations of the Imprinted CDKN1C Gene as a Cause of the Overgrowth Beckwith-Wiedemann Syndrome: clinical Spectrum and Functional Characterization. *Hum. Mutat.* 36, 894–902. doi: 10.1002/humu.22824

Brioude, F., Toutain, A., Giabicani, E., Cottereau, E., Cormier-Daire, V., and Netchine, I. (2019). Overgrowth syndromes - clinical and molecular aspects and tumour risk. *Nat. Rev. Endocrinol.* 15, 299–311. doi: 10.1038/s41574-019-0180-z

Burkardt, D. D., Tatton-Brown, K., Dobyns, W., and Graham, J. M. Jr. (2019). Approach to overgrowth syndromes in the genome era. *Am J. Med. Genet. C Semin. Med. Genet.* 181, 483–490. doi: 10.1002/ajmg.c.31757

Cai, X., and Cullen, B. R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 13, 313–316. doi: 10.1261/rna.351707

Castillo-Fernandez, J. E., Spector, T. D., and Bell, J. T. (2014). Epigenetics of discordant monozygotic twins: implications for disease. *Genome Med.* 6:60. doi: 10.1186/s13073-014-0060-z

Chang, A. S., Moley, K. H., Wangler, M., Feinberg, A. P., and Debaun, M. R. (2005). Association between Beckwith-Wiedemann syndrome and assisted reproductive technology: a case series of 19 patients. *Fertil. Steril.* 83, 349–354. doi: 10.1016/j.fertnstert.2004.07.964

Chen, Z., Robbins, K. M., Wells, K. D., and Rivera, R. M. (2013). Large offspring syndrome: a bovine model for the human loss-of-imprinting overgrowth syndrome Beckwith-Wiedemann. *Epigenetics* 8, 591–601. doi: 10.4161/epi.24655

Cheng, Y., He, C., Wang, M., Ma, X., Mo, F., Yang, S., et al. (2019). Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials. *Signal Transduct. Target. Ther.* 4:62. doi: 10.1038/s41392-019-0095-0

Choufani, S., Shuman, C., and Weksberg, R. (2010). Beckwith-Wiedemann syndrome. *Am J. Med. Genet. C Semin. Med. Genet.* 154C, 343–354. doi: 10.1002/ajmg.c.30267

Coffee, B., Muralidharan, K., Highsmith, W. E. Jr., Lapunzina, P., and Warren, S. T. (2006). Molecular diagnosis of Beckwith-Wiedemann syndrome using quantitative methylation-sensitive polymerase chain reaction. *Genet. Med.* 8, 628–634. doi: 10.1097/01.gim.0000237770.42442.cc

Cohen, J. L., Duffy, K. A., Sajorda, B. J., Hathaway, E. R., Gonzalez-Gandolfi, C. X., Richards-Yutz, J., et al. (2019). Diagnosis and management of the phenotypic spectrum of twins with Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. A* 179, 1139–1147. doi: 10.1002/ajmg.a.61164

Cooper, W. N., Curley, R., Macdonald, F., and Maher, E. R. (2007). Mitotic recombination and uniparental disomy in Beckwith-Wiedemann syndrome. *Genomics* 89, 613–617. doi: 10.1016/j.ygeno.2007.01.005

Cooper, W. N., Luharia, A., Evans, G. A., Raza, H., Haire, A. C., Grundy, R., et al. (2005). Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* 13, 1025–1032. doi: 10.1038/sj.ejhg.5201463

Court, F., Martin-Trujillo, A., Romanelli, V., Garin, I., Iglesias-Platas, I., Salafsky, I., et al. (2013). Genome-wide allelic methylation analysis reveals disease-specific susceptibility to multiple methylation defects in imprinting syndromes. *Hum. Mutat.* 34, 595–602. doi: 10.1002/humu.22276

Crea, F., Nobili, S., Paolicchi, E., Perrone, G., Napoli, C., Landini, I., et al. (2011). Epigenetics and chemoresistance in colorectal cancer: an opportunity for treatment tailoring and novel therapeutic strategies. *Drug Resist. Updat.* 14, 280–296. doi: 10.1016/j.drup.2011.08.001

Cresto, J. C., Abdenur, J. P., Bergada, I., and Martino, R. (1998). Long-term follow up of persistent hyperinsulinaemic hypoglycaemia of infancy. *Arch. Dis. Child* 79, 440–444. doi: 10.1136/adc.79.5.440

Crispo, F., Condelli, V., Lepore, S., Notarangelo, T., Sgambato, A., Esposito, F., et al. (2019). Metabolic dysregulations and epigenetics: a bidirectional interplay that drives tumor progression. *Cells* 8:798. doi: 10.3390/cells8080798

Das, R., Hampton, D. D., and Jirtle, R. L. (2009). Imprinting evolution and human health. *Mamm. Genome* 20, 563–572. doi: 10.1007/s00335-009-9229-y

Dawkins, H. J. S., Draghia-Akli, R., Lasko, P., Lau, L. P. L., Jonker, A. H., Cutillo, C. M., et al. (2018). Progress in rare diseases research 2010-2016: an IRDiRC perspective. *Clin. Transl. Sci.* 11, 11–20. doi: 10.1111/cts.12501

de Lonlay-Debeney, P., Poggi-Travert, F., Fournet, J. C., Sempoux, C., Dionisi Vici, C., Brunelle, F., et al. (1999). Clinical features of 52 neonates with hyperinsulinism. *N. Engl. J. Med.* 340, 1169–1175. doi: 10.1056/NEJM199904153401505

de Mouzon, J., Goossens, V., Bhattacharya, S., Castilla, J. A., Ferraretti, A. P., Korsak, V., et al. (2012). Assisted reproductive technology in Europe, 2007: results generated from European registers by ESHRE. *Hum. Reprod.* 27, 954–966. doi: 10.1093/humrep/des023

DeBaun, M. R., Niemitz, E. L., and Feinberg, A. P. (2003). Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *Am. J. Hum. Genet.* 72, 156–160. doi: 10.1086/346031

DeBaun, M. R., and Tucker, M. A. (1998). Risk of cancer during the first four years of life in children from The Beckwith-Wiedemann Syndrome Registry. *J. Pediatr.* 132(3 Pt 1), 398–400. doi: 10.1016/s0022-3476(98)70008-3

Dhont, M., De Sutter, P., Ruyssinck, G., Martens, G., and Bekaert, A. (1999). Perinatal outcome of pregnancies after assisted reproduction: a case-control study. *Am. J. Obstet. Gynecol.* 181, 688–695. doi: 10.1016/s0002-9378(99)70514-4

Dirks, R. A., Stunnenberg, H. G., and Marks, H. (2016). Genome-wide epigenomic profiling for biomarker discovery. *Clin. Epigenetics* 8:122. doi: 10.1186/s13148-016-0284-4

Docherty, L. E., Rezwan, F. I., Poole, R. L., Jagoe, H., Lake, H., Lockett, G. A., et al. (2014). Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes. *J. Med. Genet.* 51, 229–238. doi: 10.1136/jmedgenet-2013-102116

Doornbos, M. E., Maas, S. M., McDonnell, J., Vermeiden, J. P., and Hennekam, R. C. (2007). Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study. *Hum. Reprod.* 22, 2476–2480. doi: 10.1093/humrep/dem172

Du, M., Beatty, L. G., Zhou, W., Lew, J., Schoenherr, C., Weksberg, R., et al. (2003). Insulator and silencer sequences in the imprinted region of human chromosome 11p15.5. *Hum. Mol. Genet.* 12, 1927–1939. doi: 10.1093/hmg/ddg194

Duffy, K. A., Grand, K. L., Zelley, K., and Kalish, J. M. (2018). Tumor Screening in Beckwith-Wiedemann syndrome: parental perspectives. *J. Genet. Couns.* 27, 844–853. doi: 10.1007/s10897-017-0182-8

Dunn, J., and Rao, S. (2017). Epigenetics and immunotherapy: the current state of play. *Mol. Immunol.* 87, 227–239. doi: 10.1016/j.molimm.2017.04.012

Eggermann, T., Binder, G., Brioude, F., Maher, E. R., Lapunzina, P., Cubellis, M. V., et al. (2014). CDKN1C mutations: two sides of the same coin. *Trends Mol. Med.* 20, 614–622. doi: 10.1016/j.molmed.2014.09.001

Eggermann, T., Brioude, F., Russo, S., Lombardi, M. P., Bliek, J., Maher, E. R., et al. (2016). Prenatal molecular testing for Beckwith-Wiedemann and Silver-Russell syndromes: a challenge for molecular analysis and genetic counseling. *Eur. J. Hum. Genet.* 24, 784–793. doi: 10.1038/ejhg.2015.224

Eggermann, T., Perez de Nanclares, G., Maher, E. R., Temple, I. K., Tumer, Z., Monk, D., et al. (2015). Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. *Clin. Epigenetics* 7:123. doi: 10.1186/s13148-015-0143-8

Eltan, M., Arslan Ates, E., Cerit, K., Menevse, T. S., Kaygusuz, S. B., Eker, N., et al. (2020). Adrenocortical carcinoma in atypical Beckwith-Wiedemann syndrome due to loss of methylation at imprinting control region 2. *Pediatr. Blood Cancer* 67:e28042. doi: 10.1002/pbc.28042

Esteller, M. (2017). Epigenetic drugs: more than meets the eye. *Epigenetics* 12:307. doi: 10.1080/15592294.2017.1322881

Ferraretti, A. P., Goossens, V., de Mouzon, J., Bhattacharya, S., Castilla, J. A., Korsak, V., et al. (2012). Assisted reproductive technology in Europe, 2008: results generated from European registers by ESHRE. *Hum. Reprod.* 27, 2571–2584. doi: 10.1093/humrep/des255

Ferraretti, A. P., Goossens, V., Kupka, M., Bhattacharya, S., de Mouzon, J., Castilla, J. A., et al. (2013). Assisted reproductive technology in Europe, 2009: results generated from European registers by ESHRE. *Hum. Reprod.* 28, 2318–2331. doi: 10.1093/humrep/det278

Fontana, L., Bedeschi, M. F., Maitz, S., Cereda, A., Fare, C., Motta, S., et al. (2018). Characterization of multi-locus imprinting disturbances and underlying genetic defects in patients with chromosome 11p15.5 related imprinting disorders. *Epigenetics* 13, 897–909. doi: 10.1080/15592294.2018.1514230

Gallagher, E. J., and LeRoith, D. (2010). The proliferating role of insulin and insulin-like growth factors in cancer. *Trends Endocrinol. Metab.* 21, 610–618. doi: 10.1016/j.tem.2010.06.007

Gaston, V., Le Bouc, Y., Soupre, V., Burglen, L., Donadieu, J., Oro, H., et al. (2001). Analysis of the methylation status of the KCNQ1OT and H19 genes in leukocyte DNA for the diagnosis and prognosis of Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* 9, 409–418. doi: 10.1038/sj.ejhg.5200649

Gaur, S., Wen, Y., Song, J. H., Parikh, N. U., Mangala, L. S., Blessing, A. M., et al. (2015). Chitosan nanoparticle-mediated delivery of miRNA-34a decreases prostate tumor growth in the bone and its expression induces non-canonical autophagy. *Oncotarget* 6, 29161–29177. doi: 10.18632/oncotarget.4971

Ghanem, I., Karam, J. A., and Widmann, R. F. (2011). Surgical epiphysiodesis indications and techniques: update. *Curr. Opin. Pediatr.* 23, 53–59. doi: 10.1097/MOP.0b013e32834231b3

Gicquel, C., El-Osta, A., and Le Bouc, Y. (2008). Epigenetic regulation and fetal programming. *Best Pract. Res. Clin. Endocrinol. Metab.* 22, 1–16. doi: 10.1016/j.beem.2007.07.009

Gicquel, C., Gaston, V., Mandelbaum, J., Siffroi, J. P., Flahault, A., and Le Bouc, Y. (2003). In vitro fertilization may increase the risk of Beckwith-Wiedemann syndrome related to the abnormal imprinting of the KCN1OT gene. *Am. J. Hum. Genet.* 72, 1338–1341. doi: 10.1086/374824

Glaser, B., Thornton, P., Otonkoski, T., and Junien, C. (2000). Genetics of neonatal hyperinsulinism. *Arch. Dis. Child Fetal Neonatal Ed.* 82, F79–F86. doi: 10.1136/fn.82.2.f79

Glasgow, M. D., and Chougule, M. B. (2015). Recent developments in active tumor targeted multifunctional nanoparticles for combination chemotherapy in cancer treatment and imaging. *J. Biomed. Nanotechnol.* 11, 1859–1898. doi: 10.1166/jbn.2015.2145

Gocmen, R., Basaran, C., Karcaaltincaba, M., Cinar, A., Yurdakok, M., Akata, D., et al. (2005). Bilateral hemorrhagic adrenal cysts in an incomplete form of Beckwith-Wiedemann syndrome: MRI and prenatal US findings. *Abdom. Imaging* 30, 786–789. doi: 10.1007/s00261-005-0337-1

Gomes, M. V., Huber, J., Ferriani, R. A., Amaral Neto, A. M., and Ramos, E. S. (2009). Abnormal methylation at the KvDMR1 imprinting control region in clinically normal children conceived by assisted reproductive technologies. *Mol. Hum. Reprod.* 15, 471–477. doi: 10.1093/molehr/gap038

Graca, I., Pereira-Silva, E., Henrique, R., Packham, G., Crabb, S. J., and Jeronimo, C. (2016). Epigenetic modulators as therapeutic targets in prostate cancer. *Clin. Epigenetics* 8:98. doi: 10.1186/s13148-016-0264-8

Griggs, R. C., Batshaw, M., Dunkle, M., Gopal-Srivastava, R., Kaye, E., Krischer, J., et al. (2009). Clinical research for rare disease: opportunities, challenges, and solutions. *Mol. Genet. Metab.* 96, 20–26. doi: 10.1016/j.ymgme.2008.10.003

Hall, J. G. (1996). Twins and twinning. *Am. J. Med. Genet.* 61, 202–204.

Hall, J. G., and Lopez-Rangel, E. (1996). Embryologic development and monozygotic twinning. *Acta Genet. Med. Gemellol.* 45, 53–57. doi: 10.1017/s0001566000001094

Halliday, J., Oke, K., Breheny, S., Algar, E., and David, J. A. (2004). Beckwith-Wiedemann syndrome and IVF: a case-control study. *Am. J. Hum. Genet.* 75, 526–528. doi: 10.1086/423902

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486–489. doi: 10.1038/35013106

Hattori, H., Hiura, H., Kitamura, A., Miyauchi, N., Kobayashi, N., Takahashi, S., et al. (2019). Association of four imprinting disorders and ART. *Clin. Epigenetics* 11:21. doi: 10.1186/s13148-019-0623-3

Heerboth, S., Lapinska, K., Snyder, N., Leary, M., Rollinson, S., and Sarkar, S. (2014). Use of epigenetic drugs in disease: an overview. *Genet. Epigenet.* 6, 9–19. doi: 10.4137/GEG.S12270

Hiura, H., Okae, H., Miyauchi, N., Sato, F., Sato, A., Van De Pette, M., et al. (2012). Characterization of DNA methylation errors in patients with imprinting disorders conceived by assisted reproduction technologies. *Hum. Reprod.* 27, 2541–2548. doi: 10.1093/humrep/des197

Hori, N., Yamane, M., Kouno, K., and Sato, K. (2012). Induction of DNA demethylation depending on two sets of Sox2 and adjacent Oct3/4 binding sites (Sox-Oct motifs) within the mouse H19/insulin-like growth factor 2 (Igf2) imprinted control region. *J. Biol. Chem.* 287, 44006–44016. doi: 10.1074/jbc.M112.424580

Ibrahim, A., Kirby, G., Hardy, C., Dias, R. P., Tee, L., Lim, D., et al. (2014). Methylation analysis and diagnostics of Beckwith-Wiedemann syndrome in 1,000 subjects. *Clin. Epigenetics* 6:11. doi: 10.1186/1868-7083-6-11

Inamura, K. (2017). Major tumor suppressor and oncogenic non-coding RNAs: clinical relevance in lung cancer. *Cells* 6:12. doi: 10.3390/cells6020012

Jones, P. A., Issa, J. P., and Baylin, S. (2016). Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.* 17, 630–641. doi: 10.1038/nrg.2016.93

Kagan, K. O., Berg, C., Dufke, A., Geipel, A., Hoopmann, M., and Abele, H. (2015). Novel fetal and maternal sonographic findings in confirmed cases of Beckwith-Wiedemann syndrome. *Prenat. Diagn.* 35, 394–399. doi: 10.1002/pd.4555

Kalish, J. M., Boodhansingh, K. E., Bhatti, T. R., Ganguly, A., Conlin, L. K., Becker, S. A., et al. (2016). Congenital hyperinsulinism in children with paternal 11p uniparental isodisomy and Beckwith-Wiedemann syndrome. *J. Med. Genet.* 53, 53–61. doi: 10.1136/jmedgenet-2015-103394

Kalish, J. M., Doros, L., Helman, L. J., Hennekam, R. C., Kuiper, R. P., Maas, S. M., et al. (2017). Surveillance Recommendations for Children with Overgrowth Syndromes and Predisposition to Wilms Tumors and Hepatoblastoma. *Clin. Cancer Res.* 23, e115–e122. doi: 10.1158/1078-0432.CCR-17-0710

Kamien, B., Ronan, A., Poke, G., Sinnerbrink, I., Baynam, G., Ward, M., et al. (2018). A clinical review of generalized overgrowth syndromes in the era of massively parallel sequencing. *Mol. Syndromol.* 9, 70–82. doi: 10.1159/000484532

Kavanagh, E., and Joseph, B. (2011). The hallmarks of CDKN1C (p57. KIP2) in cancer. *Biochim. Biophys. Acta* 1816, 50–56. doi: 10.1016/j.bbcan.2011.03.002

Kelly, T. K., De Carvalho, D. D., and Jones, P. A. (2010). Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.* 28, 1069–1078. doi: 10.1038/nbt.1678

Krzyzewska, I. M., Alders, M., Maas, S. M., Bliek, J., Venema, A., Henneman, P., et al. (2019). Genome-wide methylation profiling of Beckwith-Wiedemann syndrome patients without molecular confirmation after routine diagnostics. *Clin. Epigenetics* 11:53. doi: 10.1186/s13148-019-0649-6

Kubo, S., Murata, C., Okamura, H., Sakasegawa, T., Sakurai, C., Hatsuzawa, K., et al. (2020). Oct motif variants in Beckwith-Wiedemann syndrome patients disrupt maintenance of the hypomethylated state of the H19/IGF2 imprinting control region. *FEBS Lett.* 594, 1517–1531. doi: 10.1002/1873-3468.13750

Kupka, M. S., Ferraretti, A. P., de Mouzon, J., Erb, K., D'Hooghe, T., Castilla, J. A., et al. (2014). Assisted reproductive technology in Europe, 2010: results generated from European registers by ESHREdagger. *Hum. Reprod.* 29, 2099–2113. doi: 10.1093/humrep/deu175

Laje, P., Palladino, A. A., Bhatti, T. R., States, L. J., Stanley, C. A., and Adzick, N. S. (2013). Pancreatic surgery in infants with Beckwith-Wiedemann syndrome and hyperinsulinism. *J. Pediatr. Surg.* 48, 2511–2516. doi: 10.1016/j.jpedsurg.2013.05.016

Landy, H. J., and Keith, L. G. (1998). The vanishing twin: a review. *Hum. Reprod. Update* 4, 177–183. doi: 10.1093/humupd/4.2.177

Lapunzina, P. (2005). Risk of tumorigenesis in overgrowth syndromes: a comprehensive review. *Am. J. Med. Genet. C Semin. Med. Genet.* 137C, 53–71. doi: 10.1002/ajmg.c.30064

Larson, P. S., Schlechter, B. L., de las Morenas, A., Garber, J. E., Cupples, L. A., and Rosenberg, C. L. (2005). Allele imbalance, or loss of heterozygosity, in normal breast epithelium of sporadic breast cancer cases and BRCA1 gene mutation carriers is increased compared with reduction mammoplasty tissues. *J. Clin. Oncol.* 23, 8613–8619. doi: 10.1200/JCO.2005.02.1451

Lauschke, V. M., Barragan, I., and Ingelman-Sundberg, M. (2018). Pharmacoepigenetics and toxicoepigenetics: novel mechanistic insights and therapeutic opportunities. *Annu. Rev. Pharmacol. Toxicol.* 58, 161–185. doi: 10.1146/annurev-pharmtox-010617-053021

Lee, M. P., DeBaun, M. R., Mitsuya, K., Galonek, H. L., Brandenburg, S., Oshimura, M., et al. (1999). Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5203–5208. doi: 10.1073/pnas.96.9.5203

Leibowitz, G., Glaser, B., Higazi, A. A., Salameh, M., Cerasi, E., and Landau, H. (1995). Hyperinsulinemic hypoglycemia of infancy (nesidioblastosis) in clinical remission: high incidence of diabetes mellitus and persistent beta-cell dysfunction at long-term follow-up. *J. Clin. Endocrinol. Metab.* 80, 386–392. doi: 10.1210/jcem.80.2.7852494

Li, Y., Hagen, D. E., Ji, T., Bakhtiarizadeh, M. R., Frederic, W. M., Traxler, E. M., et al. (2019). Altered microRNA expression profiles in large offspring syndrome and Beckwith-Wiedemann syndrome. *Epigenetics* 14, 850–876. doi: 10.1080/15592294.2019.1615357

Lim, D., Bowdin, S. C., Tee, L., Kirby, G. A., Blair, E., Fryer, A., et al. (2009). Clinical and molecular genetic features of Beckwith-Wiedemann syndrome associated with assisted reproductive technologies. *Hum. Reprod.* 24, 741–747. doi: 10.1093/humrep/den406

Liu, W., Zhang, R., Wei, J., Zhang, H., Yu, G., Li, Z., et al. (2015). Rapid diagnosis of imprinting disorders involving copy number variation and uniparental disomy using genome-wide SNP microarrays. *Cytogenet. Genome Res.* 146, 9–18. doi: 10.1159/000435847

Livingstone, C. (2013). IGF2 and cancer. *Endocr. Relat. Cancer* 20, R321–R339. doi: 10.1530/ERC-13-0231

Lteif, A. N., and Schwenk, W. F. (1999). Hypoglycemia in infants and children. *Endocrinol. Metab. Clin. North Am.* 28, 619–646, vii. doi: 10.1016/s0889-8529(05)70091-8

Lubinsky, M. S., and Hall, J. G. (1991). Genomic imprinting, monozygous twinning, and X inactivation. *Lancet* 337:1288. doi: 10.1016/0140-6736(91)92956-3

Lundstrom, K. (2017). Cell-impedance-based label-free technology for the identification of new drugs. *Expert Opin. Drug Discov.* 12, 335–343. doi: 10.1080/17460441.2017.1297419

Maas, S. M., Vansenne, F., Kadouch, D. J., Ibrahim, A., Bliek, J., Hopman, S., et al. (2016). Phenotype, cancer risk, and surveillance in Beckwith-Wiedemann syndrome depending on molecular genetic subgroups. *Am J. Med. Genet. A* 170, 2248–2260. doi: 10.1002/ajmg.a.37801

MacFarland, S. P., Duffy, K. A., Bhatti, T. R., Bagatell, R., Balamuth, N. J., Brodeur, G. M., et al. (2018). Diagnosis of Beckwith-Wiedemann syndrome in children presenting with Wilms tumor. *Pediatr. Blood Cancer* 65:e27296. doi: 10.1002/pbc.27296

Machin, G. A. (1996). Some causes of genotypic and phenotypic discordance in monozygotic twin pairs. *Am. J. Med. Genet.* 61, 216–228. doi: 10.1002/(sici)1096-8628(19960122)61:3<216::aid-ajmg5>3.0.co;2-s

Maeda, T., Higashimoto, K., Jozaki, K., Yatsuki, H., Nakabayashi, K., Makita, Y., et al. (2014). Comprehensive and quantitative multilocus methylation analysis reveals the susceptibility of specific imprinted differentially methylated regions to aberrant methylation in Beckwith-Wiedemann syndrome with epimutations. *Genet. Med.* 16, 903–912. doi: 10.1038/gim.2014.46

Maher, E. R., and Reik, W. (2000). Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *J. Clin. Invest.* 105, 247–252. doi: 10.1172/JCI9340

Mama, N., H'Mida, D., Lahmar, I., Yacoubi, M. T., and Tlili-Graiess, K. (2014). PHACES syndrome associated with carcinoid endobronchial tumor. *Pediatr. Radiol.* 44, 621–624. doi: 10.1007/s00247-013-2820-0

Martinez y Martinez, R., Martinez-Carboney, R., Ocampo-Campos, R., Rivera, H., Gomez Plascencia y Castillo, J., Cuevas, A., et al. (1992). Wiedemann-Beckwith syndrome: clinical, cytogenetical and radiological observations in 39 new cases. *Genet. Couns.* 3, 67–76.

Mau, T., and Yung, R. (2014). Potential of epigenetic therapies in non-cancerous conditions. *Front. Genet.* 5:438. doi: 10.3389/fgene.2014.00438

Mazzone, R., Zwergel, C., Mai, A., and Valente, S. (2017). Epi-drugs in combination with immunotherapy: a new avenue to improve anticancer efficacy. *Clin. Epigenetics* 9:59. doi: 10.1186/s13148-017-0358-y

Miranda Furtado, C. L., Dos Santos Luciano, M. C., Silva Santos, R. D., Furtado, G. P., Moraes, M. O., and Pessoa, C. (2019). Epidrugs: targeting epigenetic marks in cancer treatment. *Epigenetics* 14, 1164–1176. doi: 10.1080/15592294.2019.1640546

Morera, L., Lubbert, M., and Jung, M. (2016). Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin. Epigenetics* 8, 57. doi: 10.1186/s13148-016-0223-4

Munns, C. F., and Batch, J. A. (2001). Hyperinsulinism and Beckwith-Wiedemann syndrome. *Arch. Dis. Child Fetal Neonatal Ed.* 84, F67–F69. doi: 10.1136/fn.84.1.f67

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat. Genet.* 36, 889–893. doi: 10.1038/ng1402

Mussa, A., Ciuffreda, V. P., Sauro, P., Pagliardini, V., Pagliardini, S., Carli, D., et al. (2019). Longitudinal Monitoring of Alpha-Fetoprotein by Dried Blood Spot for Hepatoblastoma Screening in Beckwith(-)Wiedemann Syndrome. *Cancers* 11:86. doi: 10.3390/cancers11010086

Mussa, A., Molinatto, C., Baldassarre, G., Riberi, E., Russo, S., Larizza, L., et al. (2016a). Cancer Risk in Beckwith-Wiedemann syndrome: a systematic review and meta-analysis outlining a novel (Epi)Genotype specific histotype targeted screening protocol. *J. Pediatr.* 176, 142–149.e1. doi: 10.1016/j.jpeds.2016.05.038

Mussa, A., Molinatto, C., Cerrato, F., Palumbo, O., Carella, M., Baldassarre, G., et al. (2017). Assisted reproductive techniques and risk of beckwith-wiedemann syndrome. *Pediatrics* 140:e20164311. doi: 10.1542/peds.2016-4311

Mussa, A., Russo, S., De Crescenzo, A., Chiesa, N., Molinatto, C., Selicorni, A., et al. (2013). Prevalence of Beckwith-Wiedemann syndrome in North West of Italy. *Am. J. Med. Genet. A* 161A, 2481–2486. doi: 10.1002/ajmg.a.36080

Mussa, A., Russo, S., de Crescenzo, A., Freschi, A., Calzari, L., Maitz, S., et al. (2016b). Fetal growth patterns in Beckwith-Wiedemann syndrome. *Clin. Genet.* 90, 21–27. doi: 10.1111/cge.12759

Nativio, R., Sparago, A., Ito, Y., Weksberg, R., Riccio, A., and Murrell, A. (2011). Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith-Wiedemann syndrome and Silver-Russell syndrome. *Hum. Mol. Genet.* 20, 1363–1374. doi: 10.1093/hmg/ddr018

Netchine, I., Rossignol, S., Azzi, S., and Le Bouc, Y. (2013). Epigenetic anomalies in childhood growth disorders. *Nestle Nutr. Inst. Workshop Ser.* 71, 65–73. doi: 10.1159/000342568

Nguyen, K. V. (2019). Potential epigenomic co-management in rare diseases and epigenetic therapy. *Nucleosides Nucleotides Nucleic Acids* 38, 752–780. doi: 10.1080/15257770.2019.1594893

Niculescu, M. D., and Lupu, D. S. (2011). Nutritional influence on epigenetics and effects on longevity. *Curr. Opin. Clin. Nutr. Metab. Care* 14, 35–40. doi: 10.1097/MCO.0b013e328340ff7c

Nowacka-Zawisza, M., and Wisnik, E. (2017). DNA methylation and histone modifications as epigenetic regulation in prostate cancer (Review). *Oncol. Rep.* 38, 2587–2596. doi: 10.3892/or.2017.5972

Orstavik, R. E., Tommerup, N., Eiklid, K., and Orstavik, K. H. (1995). Non-random X chromosome inactivation in an affected twin in a monozygotic twin pair discordant for Wiedemann-Beckwith syndrome. *Am. J. Med. Genet.* 56, 210–214. doi: 10.1002/ajmg.1320560219

Ounap, K. (2016). Silver-russell syndrome and Beckwith-Wiedemann syndrome: opposite phenotypes with heterogeneous molecular etiology. *Mol. Syndromol.* 7, 110–121. doi: 10.1159/000447413

Paganini, L., Carlessi, N., Fontana, L., Silipigni, R., Motta, S., Fiori, S., et al. (2015). Beckwith-Wiedemann syndrome prenatal diagnosis by methylation analysis in chorionic villi. *Epigenetics* 10, 643–649. doi: 10.1080/15592294.2015.1057383

Park, K. S., Mitra, A., Rahat, B., Kim, K., and Pfeifer, K. (2017). Loss of imprinting mutations define both distinct and overlapping roles for misexpression of IGF2 and of H19 lncRNA. *Nucleic Acids Res.* 45, 12766–12779. doi: 10.1093/nar/gkx896

Poole, R. L., Docherty, L. E., Al Sayegh, A., Caliebe, A., Turner, C., Baple, E., et al. (2013). Targeted methylation testing of a patient cohort broadens the epigenetic and clinical description of imprinting disorders. *Am. J. Med. Genet. A* 161A, 2174–2182. doi: 10.1002/ajmg.a.36049

Priolo, M., Sparago, A., Mammi, C., Cerrato, F., Lagana, C., and Riccio, A. (2008). MS-MLPA is a specific and sensitive technique for detecting all chromosome 11p15.5 imprinting defects of BWS and SRS in a single-tube experiment. *Eur. J. Hum. Genet.* 16, 565–571. doi: 10.1038/sj.ejhg.5202001

Raveh, E., Matouk, I. J., Gilon, M., and Hochberg, A. (2015). The H19 Long non-coding RNA in cancer initiation, progression and metastasis - a proposed unifying theory. *Mol. Cancer* 14:184. doi: 10.1186/s12943-015-0458-2

Romanelli, V., Belinchon, A., Benito-Sanz, S., Martinez-Glez, V., Gracia-Bouthelier, R., Heath, K. E., et al. (2010). CDKN1C (p57(Kip2)) analysis in Beckwith-Wiedemann syndrome (BWS) patients: genotype-phenotype correlations, novel mutations, and polymorphisms. *Am. J. Med. Genet. A* 152A, 1390–1397. doi: 10.1002/ajmg.a.33453

Rossignol, S., Steunou, V., Chalas, C., Kerjean, A., Rigolet, M., Viegas-Pequignot, E., et al. (2006). The epigenetic imprinting defect of patients with Beckwith-Wiedemann syndrome born after assisted reproductive technology is not restricted to the 11p15 region. *J. Med. Genet.* 43, 902–907. doi: 10.1136/jmg.2006.042135

Rump, P., Zeegers, M. P., and van Essen, A. J. (2005). Tumor risk in Beckwith-Wiedemann syndrome: a review and meta-analysis. *Am. J. Med. Genet. A* 136, 95–104. doi: 10.1002/ajmg.a.30729

Russo, S., Calzari, L., Mussa, A., Mainini, E., Cassina, M., Di Candia, S., et al. (2016). A multi-method approach to the molecular diagnosis of overt and borderline 11p15.5 defects underlying Silver-Russell and Beckwith-Wiedemann syndromes. *Clin. Epigenetics* 8:23. doi: 10.1186/s13148-016-0183-8

Sanchez-Delgado, M., Riccio, A., Eggermann, T., Maher, E. R., Lapunzina, P., Mackay, D., et al. (2016). Causes and consequences of multi-locus imprinting disturbances in humans. *Trends Genet.* 32, 444–455. doi: 10.1016/j.tig.2016.05.001

Sano, S., Matsubara, K., Nagasaki, K., Kikuchi, T., Nakabayashi, K., Hata, K., et al. (2016). Beckwith-Wiedemann syndrome and pseudohypoparathyroidism type Ib in a patient with multilocus imprinting disturbance: a female-dominant phenomenon? *J. Hum. Genet.* 61, 765–769. doi: 10.1038/jhg.2016.45

Sazhenova, E. A., and Lebedev, I. N. (2008). [Epimutations of the KCNQ1OT1 imprinting center of chromosome 11 in early human embryo lethality]. *Genetika* 44, 1609–1616.

Schiff, D., Colle, E., Wells, D., and Stern, L. (1973). Metabolic aspects of the Beckwith-Wiedemann syndrome. *J. Pediatr.* 82, 258–262. doi: 10.1016/s0022-3476(73)80163-5

Schofield, P. N., Joyce, J. A., Lam, W. K., Grandjean, V., Ferguson-Smith, A., Reik, W., et al. (2001). Genomic imprinting and cancer; new paradigms in the genetics of neoplasia. *Toxicol. Lett.* 120, 151–160. doi: 10.1016/s0378-4274(01)00294-6

Scott, R. H., Douglas, J., Baskcomb, L., Nygren, A. O., Birch, J. M., Cole, T. R., et al. (2008). Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) robustly detects and distinguishes 11p15 abnormalities associated with overgrowth and growth retardation. *J. Med. Genet.* 45, 106–113. doi: 10.1136/jmg.2007.053207

Segers, H., Kersseboom, R., Alders, M., Pieters, R., Wagner, A., and van den Heuvel-Eibrink, M. M. (2012). Frequency of WT1 and 11p15 constitutional aberrations and phenotypic correlation in childhood Wilms tumour patients. *Eur. J. Cancer* 48, 3249–3256. doi: 10.1016/j.ejca.2012.06.008

Senniappan, S., Ismail, D., Shipster, C., Beesley, C., and Hussain, K. (2015). The heterogeneity of hyperinsulinaemic hypoglycaemia in 19 patients with Beckwith-Wiedemann syndrome due to KvDMR1 hypomethylation. *J. Pediatr. Endocrinol. Metab* 28, 83–86. doi: 10.1515/jpem-2013-0390

Shao, Q., Xu, J., Deng, R., Wei, W., Zhou, B., Yue, C., et al. (2018). Long non-coding RNA-422 acts as a tumor suppressor in colorectal cancer. *Biochem. Biophys. Res. Commun.* 495, 539–545. doi: 10.1016/j.bbrc.2017.10.076

Shepherd, R. M., Cosgrove, K. E., O'Brien, R. E., Barnes, P. D., Ammala, C., and Dunne, M. J. (2000). Hyperinsulinism of infancy: towards an understanding of unregulated insulin release. European Network for Research into Hyperinsulinism in Infancy. *Arch. Dis. Child Fetal Neonatal Ed.* 82, F87–F97. doi: 10.1136/fn.82.2.f87

Shilyansky, J., Cutz, E., and Filler, R. M. (1997). Endogenous hyperinsulinism: diagnosis, management, and long-term follow-up. *Semin. Pediatr. Surg.* 6, 115–120.

Shuman, C., Smith, A. C., Steele, L., Ray, P. N., Clericuzio, C., Zackai, E., et al. (2006). Constitutional UPD for chromosome 11p15 in individuals with isolated hemihyperplasia is associated with high tumor risk and occurs following assisted reproductive technologies. *Am. J. Med. Genet. A* 140, 1497–1503. doi: 10.1002/ajmg.a.31323

Slavotinek, A., Gaunt, L., and Donnai, D. (1997). Paternally inherited duplications of 11p15.5 and Beckwith-Wiedemann syndrome. *J. Med. Genet.* 34, 819–826. doi: 10.1136/jmg.34.10.819

Smith, A. C., Rubin, T., Shuman, C., Estabrooks, L., Aylsworth, A. S., McDonald, M. T., et al. (2006). New chromosome 11p15 epigenotypes identified in male monozygotic twins with Beckwith-Wiedemann syndrome. *Cytogenet. Genome Res.* 113, 313–317. doi: 10.1159/000090847

Smith, A. C., Suzuki, M., Thompson, R., Choufani, S., Higgins, M. J., Chiu, I. W., et al. (2012). Maternal gametic transmission of translocations or inversions of human chromosome 11p15.5 results in regional DNA hypermethylation and downregulation of CDKN1C expression. *Genomics* 99, 25–35. doi: 10.1016/j.ygeno.2011.10.007

Soejima, H., and Higashimoto, K. (2013). Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *J. Hum. Genet.* 58, 402–409. doi: 10.1038/jhg.2013.51

Stampone, E., Caldarelli, I., Zullo, A., Bencivenga, D., Mancini, F. P., Della Ragione, F., et al. (2018). Genetic and epigenetic control of CDKN1C expression: importance in cell commitment and differentiation, tissue homeostasis and human diseases. *Int. J. Mol. Sci.* 19:1055. doi: 10.3390/ijms19041055

Stanley, C. A. (1997). Hyperinsulinism in infants and children. *Pediatr. Clin. North Am.* 44, 363–374. doi: 10.1016/s0031-3955(05)70481-8

Sutcliffe, A. G., D'Souza, S. W., Cadman, J., Richards, B., McKinlay, I. A., and Lieberman, B. (1995). Minor congenital anomalies, major congenital malformations and development in children conceived from cryopreserved embryos. *Hum. Reprod.* 10, 3332–3337. doi: 10.1093/oxfordjournals.humrep.a135915

Sutcliffe, A. G., Peters, C. J., Bowdin, S., Temple, K., Reardon, W., Wilson, L., et al. (2006). Assisted reproductive therapies and imprinting disorders–a preliminary British survey. *Hum. Reprod.* 21, 1009–1011. doi: 10.1093/humrep/dei405

Sweet, C. B., Grayson, S., and Polak, M. (2013). Management strategies for neonatal hypoglycemia. *J. Pediatr. Pharmacol. Ther.* 18, 199–208. doi: 10.5863/1551-6776-18.3.199

Swinney, D. C., and Xia, S. (2014). The discovery of medicines for rare diseases. *Future Med. Chem.* 6, 987–1002. doi: 10.4155/fmc.14.65

Talaulikar, V. S., and Arulkumaran, S. (2012). Reproductive outcomes after assisted conception. *Obstet. Gynecol. Surv.* 67, 566–583. doi: 10.1097/OGX.0b013e31826a5d4a

Tee, L., Lim, D. H., Dias, R. P., Baudement, M. O., Slater, A. A., Kirby, G., et al. (2013). Epimutation profiling in Beckwith-Wiedemann syndrome: relationship with assisted reproductive technology. *Clin. Epigenetics* 5:23. doi: 10.1186/1868-7083-5-23

Tierling, S., Souren, N. Y., Reither, S., Zang, K. D., Meng-Hentschel, J., Leitner, D., et al. (2011). DNA methylation studies on imprinted loci in a male monozygotic twin pair discordant for Beckwith-Wiedemann syndrome. *Clin. Genet.* 79, 546–553. doi: 10.1111/j.1399-0004.2010.01482.x

Tung, J. Y., Lai, S. H. Y., Au, S. L. K., Yeung, K. S., Kan, A. S. Y., Loong, F., et al. (2020). Coexistence of paternally-inherited ABCC8 mutation and mosaic paternal uniparental disomy 11p hyperinsulinism. *Int. J. Pediatr. Endocrinol.* 2020:13. doi: 10.1186/s13633-020-00083-5

Velasco, G., and Francastel, C. (2019). Genetics meets DNA methylation in rare diseases. *Clin. Genet.* 95, 210–220. doi: 10.1111/cge.13480

Vermeiden, J. P., and Bernardus, R. E. (2013). Are imprinting disorders more prevalent after human in vitro fertilization or intracytoplasmic sperm injection? *Fertil. Steril.* 99, 642–651. doi: 10.1016/j.fertnstert.2013.01.125

Vitiello, M., Tuccoli, A., and Poliseno, L. (2015). Long non-coding RNAs in cancer: implications for personalized therapy. *Cell. Oncol.* 38, 17–28. doi: 10.1007/s13402-014-0180-x

Wang, R., Xiao, Y., Li, D., Hu, H., Li, X., Ge, T., et al. (2020). Clinical and molecular features of children with Beckwith-Wiedemann syndrome in China: a single-center retrospective cohort study. *Ital. J. Pediatr.* 46:55. doi: 10.1186/s13052-020-0819-3

Weksberg, R., Shen, D. R., Fei, Y. L., Song, Q. L., and Squire, J. (1993). Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. *Nat. Genet.* 5, 143–150. doi: 10.1038/ng1093-143

Weksberg, R., Shuman, C., and Beckwith, J. B. (2010). Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* 18, 8–14. doi: 10.1038/ejhg.2009.106

Weksberg, R., Shuman, C., Caluseriu, O., Smith, A. C., Fei, Y. L., Nishikawa, J., et al. (2002). Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* 11, 1317–1325. doi: 10.1093/hmg/11.11.1317

Weksberg, R., Shuman, C., and Smith, A. C. (2005). Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 137C, 12–23. doi: 10.1002/ajmg.c.30058

Welch, G. R., and Clegg, J. S. (2010). From protoplasmic theory to cellular systems biology: a 150-year reflection. *Am. J. Physiol. Cell Physiol.* 298, C1280–C1290. doi: 10.1152/ajpcell.00016.2010

Weng, E. Y., Mortier, G. R., and Graham, J. M. Jr. (1995). Beckwith-Wiedemann syndrome. An update and review for the primary pediatrician. *Clin. Pediatr.* 34, 317–326. doi: 10.1177/000992289503400605

Wijnen, M., Alders, M., Zwaan, C. M., Wagner, A., and van den Heuvel-Eibrink, M. M. (2012). KCNQ1OT1 hypomethylation: a novel disguised genetic predisposition in sporadic pediatric adrenocortical tumors? *Pediatr. Blood Cancer* 59, 565–566. doi: 10.1002/pbc.23398

Williams, D. H., Gauthier, D. W., and Maizels, M. (2005). Prenatal diagnosis of Beckwith-Wiedemann syndrome. *Prenat. Diagn.* 25, 879–884. doi: 10.1002/pd.1155

Woods, D. M., Sodre, A. L., Villagra, A., Sarnaik, A., Sotomayor, E. M., and Weber, J. (2015). HDAC Inhibition Upregulates PD-1 Ligands in Melanoma and Augments Immunotherapy with PD-1 Blockade. *Cancer Immunol. Res.* 3, 1375–1385. doi: 10.1158/2326-6066.CIR-15-0077-T

Yan, Y., Frisen, J., Lee, M. H., Massague, J., and Barbacid, M. (1997). Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development. *Genes Dev.* 11, 973–983. doi: 10.1101/gad.11.8.973

Yao, Z. X., Jogunoori, W., Choufani, S., Rashid, A., Blake, T., Yao, W., et al. (2010). Epigenetic silencing of beta-spectrin, a TGF-beta signaling/scaffolding protein in a human cancer stem cell disorder: Beckwith-Wiedemann syndrome. *J. Biol. Chem.* 285, 36112–36120. doi: 10.1074/jbc.M110.162347

Young, L. E., Fernandes, K., McEvoy, T. G., Butterwith, S. C., Gutierrez, C. G., Carolan, C., et al. (2001). Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. *Nat. Genet.* 27, 153–154. doi: 10.1038/84769

Young, L. E., Sinclair, K. D., and Wilmut, I. (1998). Large offspring syndrome in cattle and sheep. *Rev. Reprod.* 3, 155–163. doi: 10.1530/ror.0.0030155

Zhang, P., Liegeois, N. J., Wong, C., Finegold, M., Hou, H., Thompson, J. C., et al. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. *Nature* 387, 151–158. doi: 10.1038/387151a0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 23 November 2020
doi: 10.3389/fgene.2020.605263

O-GlcNAc: Regulator of Signaling and Epigenetics Linked to X-linked Intellectual Disability

*Daniel Konzman, Lara K. Abramowitz, Agata Steenackers, Mana Mohan Mukherjee, Hyun-Jin Na and John A. Hanover**

Laboratory of Cellular and Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, United States

Cellular identity in multicellular organisms is maintained by characteristic transcriptional networks, nutrient consumption, energy production and metabolite utilization. Integrating these cell-specific programs are epigenetic modifiers, whose activity is often dependent on nutrients and their metabolites to function as substrates and co-factors. Emerging data has highlighted the role of the nutrient-sensing enzyme *O*-GlcNAc transferase (OGT) as an epigenetic modifier essential in coordinating cellular transcriptional programs and metabolic homeostasis. OGT utilizes the end-product of the hexosamine biosynthetic pathway to modify proteins with *O*-linked β-D-*N*-acetylglucosamine (*O*-GlcNAc). The levels of the modification are held in check by the *O*-GlcNAcase (OGA). Studies from model organisms and human disease underscore the conserved function these two enzymes of *O*-GlcNAc cycling play in transcriptional regulation, cellular plasticity and mitochondrial reprogramming. Here, we review these findings and present an integrated view of how *O*-GlcNAc cycling may contribute to cellular memory and transgenerational inheritance of responses to parental stress. We focus on a rare human genetic disorder where mutant forms of OGT are inherited or acquired *de novo*. Ongoing analysis of this disorder, OGT- X-linked intellectual disability (OGT-XLID), provides a window into how epigenetic factors linked to *O*-GlcNAc cycling may influence neurodevelopment.

Keywords: O-linked β-D-*N*-acetylglucosamine (*O*-GlcNAc), X-linked intellectual disability (XLID), epigenetics, histone modification, DNA methylation, nutrient-sensing

OPEN ACCESS

Edited by:
Dag H. Yasui, University of California, Davis, United States

Reviewed by:
Richard Alan Katz, Fox Chase Cancer Center, United States
Beisi Xu, St. Jude Children's Research Hospital, United States

***Correspondence:**
John A. Hanover
john.hanover@nih.gov; jah@helix.nih.gov

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *11 September 2020*
Accepted: *20 October 2020*
Published: *23 November 2020*

Citation:
Konzman D, Abramowitz LK, Steenackers A, Mukherjee MM, Na H-J and Hanover JA (2020) O-GlcNAc: Regulator of Signaling and Epigenetics Linked to X-linked Intellectual Disability. Front. Genet. 11:605263. doi: 10.3389/fgene.2020.605263

INTRODUCTION

Throughout the lifetime of an organism there is a requirement to be able to adapt to environmental changes, whether that be development, stress, or nutritional state. This adaptation requires changes in transcriptional programs that allow an appropriate gene regulatory network response. One way this is achieved is through epigenetic regulation: heritable modifications of DNA and histones that influence complex networks impacting transcription, DNA replication, and DNA repair (Janke et al., 2015). Importantly, the activity of all epigenetic modifying enzymes relies on the availability of specific metabolites. This sets up signaling pathways in which cells have evolved the ability to detect nutrient alterations and respond with epigenetic changes to coordinate appropriate transcriptional programs. Some of the most well defined metabolites and their influence on

epigenetic modifications include: acetyl-CoA and histone acetylation; sirtuins, NAD^+, and histone deacetylation; S-adenosylmethionine and DNA/histone methylation; FAD, α-ketoglutarate, and DNA/histone demethylation (Janke et al., 2015; Etchegaray and Mostoslavsky, 2016; Wong et al., 2017; Schvartzman et al., 2018). Here, we focus on the conserved epigenetic role that the nutrient-sensitive post-translational modification (PTM) *O*-GlcNAc, and the enzymes involved in *O*-GlcNAc cycling, the *O*-GlcNAc transferase (OGT) and the *O*-GlcNAcase (MGEA5 or OGA) have in coordinating transcriptional responses to environmental changes.

Neurodevelopment is one such process that is heavily coordinated by the crosstalk between the genome, environment and metabolic flux. Alterations during this process could lead to neurodevelopmental disorders, which constitute a broad range of disorders that originate during development of the central nervous system. As more and more sequences become available from neurodevelopmental disorder patients, it is becoming increasingly clear the essential role chromatin modifiers and epigenetic regulators play in these diseases. Of the hundreds of genes with mutations or copy number variations that are causative of neurodevelopmental disorders, chromatin regulation is the second most common association behind synaptic function (Gabriele et al., 2018). Interestingly, recent data has identified placental OGT levels as a biomarker for neurodevelopmental disorders (Howerton et al., 2013; Howerton and Bale, 2014) and mutations in the OGT gene as being causative to a subset of X-linked intellectual disability (XLID) patients (Pravata et al., 2020b).

O-GlcNAc transferase uses the end-product of the hexosamine biosynthetic pathway (HBP), UDP-GlcNAc, to add a single GlcNAc monosaccharide onto serines and threonines of target intracellular proteins. To form UDP-GlcNAc, the HBP incorporates intermediate metabolites derived from carbohydrates, amino acids, fat, and nucleotides (**Figure 1**). Thus, OGT is uniquely positioned to sense environmental changes and respond through modifying key targets. A diverse array of more than 4,000 proteins have been identified to be *O*-GlcNAc modified (Ma and Hart, 2014), including transcription factors, epigenetic modulators, enzymes, kinases, mitochondrial proteins, structural proteins, nuclear porins, and components of vesicular trafficking pathways. How OGT recognizes specific target proteins remains relatively unclear, although it has preference for intrinsically disordered domains (Nishikawa et al., 2010). *O*-GlcNAcylation can influence modified proteins in various ways such as through crosstalk or competition with other PTMs like phosphorylation, altering enzyme activity, impacting protein stability, influencing subcellular localization, or altering binding partners (Bond and Hanover, 2013, 2015). Through this broad range of targets, *O*-GlcNAc influences many basic molecular processes including transcription, translation, proteostasis, and signaling [for more extensive reviews on the broad roles of *O*-GlcNAc and its targets see Bond and Hanover (2015) and Yang and Qian (2017)]. Regulation of *O*-GlcNAcylation has proven to be essential as too little or too much *O*-GlcNAc is associated with a number of diseases such as cancer, metabolic syndromes, and neurodegenerative diseases. Recently, OGT has been linked to a rare neurodevelopmental disorder known as OGT-XLID, which we hypothesize could be caused by dysfunction of the essential role of *O*-GlcNAc as an epigenetic regulator, the focus of this review.

OGT, OGA, and *O*-GlcNAc itself have all emerged as critical epigenetic regulators (**Figure 1**) essential for stem cell maintenance, development, and in the nervous system. Studies in model organisms have highlighted the conserved role OGT plays in gene regulation and cellular identity. Since the initial findings in *Drosophila* which defined the gene encoding OGT as a Polycomb group member critical for Hox gene silencing (Gambetta et al., 2009), *O*-GlcNAcylation has been recognized to play multifaceted roles in epigenetic regulation. Recent studies in *C. elegans* have demonstrated that OGT plays a critical role in preventing transitions between terminal cell fates. Advances in mass spectrometry technology have uncovered *O*-GlcNAcylation of histones, and is now widely believed to be part of the "histone code." Studies have also indicated that *O*-GlcNAc plays a role in histone exchange and is essential for the DNA damage response (Na et al., 2020). Further, OGT has been found in complex with TET proteins, signifying a potential role in DNA demethylation (Chen et al., 2013; Deplus et al., 2013; Vella et al., 2013). *O*-GlcNac has also been defined to regulate transcription through modification of key transcription factors as well as the RNA polymerase II C-terminal repeat domain (Lewis and Hanover, 2014).

In this review we dissect the conserved and wide-ranging roles *O*-GlcNAc cycling has in regulating transcriptional networks which contribute toward maintaining cellular identity. We discuss how placental OGT could be a marker for placental stress contributing toward neurodevelopmental disorders. Lastly, we examine how OGT's role in epigenetic regulation contributes toward mutations manifesting as an XLID in human patients.

O-GlcNAc CYCLING, EPIGENETICS AND CELL FATE DETERMINATION: LESSONS FROM MODEL ORGANISMS

Super Sex Combs: OGT and Hox Gene Repression

Across species, *O*-GlcNAc has a key role to play in development. Though its functions are numerous, OGT is particularly important in the spatiotemporal control of gene expression through its regulation of Hox genes. Conserved across bilaterians, the Hox genes encode critical transcription factors which specify the body plan along the anterior-posterior axis. Inappropriate regulation of these genes results in homeotic transformations, developmental errors in which one region of an animal inappropriately adopts the characteristics of another region. Two major groups of genes control the spatial expression of Hox genes: transcriptional repression by the Polycomb group (PcG) and activation by the trithorax group (trxG). In *Drosophila*, the gene encoding OGT was first described genetically by mutations that caused homeotic phenotypes including antenna-to-leg and wing-to-haltere transformations (Butler et al., 2019). Without

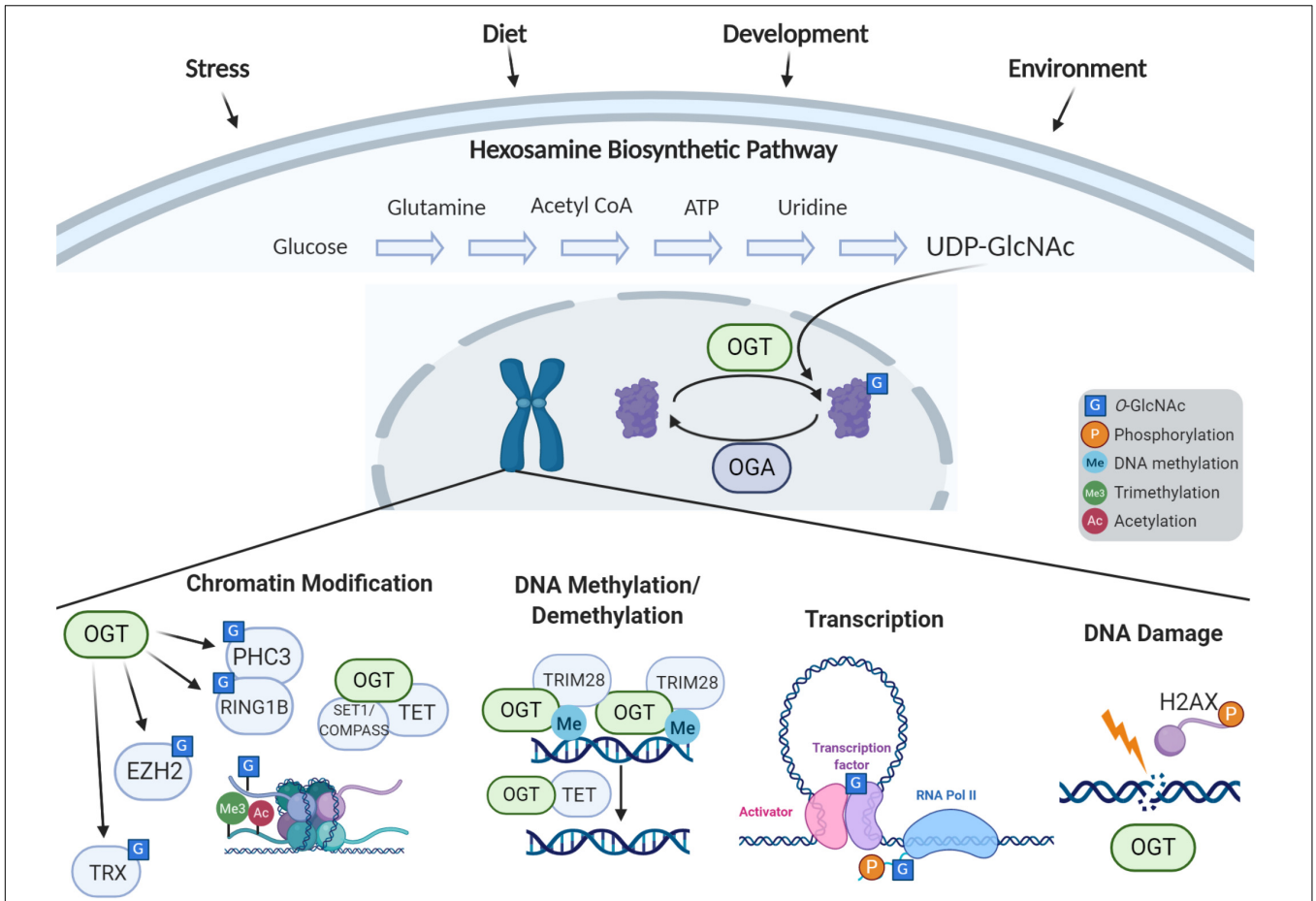

FIGURE 1 | Cellular response to stress via *O*-GlcNAc-dependent epigenetic regulation. The hexosamine biosynthetic pathway incorporates nutrients and intermediate metabolites to form the nucleotide sugar UDP-GlcNAc. The *O*-GlcNAc Transferase (OGT) utilizes this nucleotide sugar as the sugar donor to modify diverse nucleocytoplasmic proteins including transcription factors, chromatin modifiers, and histones. This modification can be removed by *O*-GlcNAcase (OGA), whose levels are also implicated in regulation of gene expression. The process of *O*-GlcNAc cycling enables the cell to sense and integrate nutrient and signaling information and respond to extracellular stressors. *O*-GlcNAc influences histone modifications through the interactions of OGT with the Polycomb complex and SET1/COMPASS, OGA's interactions with the trithorax complex, and direct *O*-GlcNAcylation of histones. OGT associates with TRIM28 to induce DNA methylation-dependent repression, and OGT is also involved in demethylation of DNA through the TET proteins. *O*-GlcNAc influences transcription by modification of transcription factors and the C-terminal domain of RNA Polymerase II. *O*-GlcNAc is also involved in the DNA damage response.

knowing the enzymatic function of the protein it encodes, the gene was named *super sex combs (sxc)* for its loss of function phenotype. Like PcG genes *Polycomb (Pc)* and *extra sex combs (esc)*, *sxc* mutations affect the development of the sex combs, a structure on the forelimbs of *Drosophila* males. Similar to other components of the PcG, *sxc* was required for the repression of Hox genes in tissues where they should not be expressed. Derepression of Hox genes could explain the homeotic transformations associated with *sxc* mutations. Epistasis experiments confirmed this, as *sxc* flies lacking the Hox gene *Ubx* had more normal wing development (Ingham, 1984). Later, immunostaining of embryos showed aberrant expression of at least six PcG target genes including *Ubx* and *Abd-B* (Gambetta et al., 2009; Gambetta and Müller, 2014).

Experiments on the PcG gene *polyhomeotic (ph)* demonstrated genetic interactions with *sxc*. Flies carrying *sxc* and *ph* mutations had enhanced *sxc* phenotypes, with greater numbers of sex combs, stronger developmental defects, and lethality even earlier in development (Cheng et al., 1994). These findings uncovered further genetic interactions and demonstrated that *sxc* functions with the PcG to prevent ectopic expression of Hox genes.

In 2009, the protein product of *sxc* was determined to be the *O*-GlcNAc transferase, with expression of an *Ogt* transgene rescuing the pupal lethality of *sxc* mutant animals (Sinclair et al., 2009). Further studies better defined the interaction between *sxc/Ogt* and *ph*. In fact, co-IP experiments indicated that PH was *O*-GlcNAc modified (Gambetta et al., 2009). In extracts of larvae with homozygous *sxc/Ogt* mutations and no maternal *Ogt* contribution, PH was shown to aggregate into high molecular weight assemblies, which likely impaired its function (Gambetta and Müller, 2014). Loss of *O*-GlcNAcylation of PH resulted in *sxc* phenotypes, demonstrating modification was required to prevent aggregation, allowing PH to function with the PcG (Gambetta and Müller, 2014). Later, mass spectrometry of mouse

embryonic stem cell (ESC) lysates demonstrated PHC3, the mammalian homolog of PH, is *O*-GlcNAc modified, suggesting the regulation of the PcG by *O*-GlcNAc is conserved (Myers et al., 2011). In human ESCs, the core subunit of PRC1, RING1B, is also modified and plays a role in targeting the complex (Maury et al., 2015). These findings clearly demonstrate the necessity of *O*-GlcNAc for the proper repression of Hox genes through the PcG component PH.

In mammals, the role of *O*-GlcNAc in PcG repression has diverged in certain ways. While PRC2 protein E(z) is not *O*-GlcNAc modified in *Drosophila* (Gambetta et al., 2009), several studies have identified its mammalian homolog EZH2 as modified, resulting in increased protein stability (Chu et al., 2014; Jiang et al., 2019). The effect of *O*-GlcNAc modification on PRC2-mediated H3K27me3 is controversial, with several publications describing a decrease associated with OGT depletion (Chu et al., 2014; Butler et al., 2019) and others noting no change (Myers et al., 2011; Forma et al., 2018; Jiang et al., 2019). These discrepancies likely arise from different cell lines being used, though one study found OGT knockdown prevented H3K27me3 changes associated with learning at specific gene promoters in mouse hippocampal tissue (Butler et al., 2019). This suggests the effects of *O*-GlcNAc may play roles in targeting PcG repression which are only apparent in certain developmental and genomic contexts.

O-GlcNAc transferase additionally interacts directly with the mammalian HOXA1 protein, as was found in a yeast-two-hybrid experiment (Lambert et al., 2012) and confirmed by co-IP (Draime et al., 2018). Though the authors did not find evidence for *O*-GlcNAc affecting localization, stability, or transcription factor activity (Draime et al., 2018), they only tested OGT and HOXA1 overexpression, so further experimentation may reveal additional layers of Hox gene regulation by *O*-GlcNAc. Although the phenotypes of *sxc/Ogt* flies are mostly consistent with those of other PcG genes, several findings suggest *sxc/Ogt* is also involved in Hox gene activation through trxG. Alone, both *sxc/Ogt* and *Asx* mutations produce anterior-to-posterior transformations associated with PcG mutations, but *sxc/Ogt*;*Asx* double heterozygotes display posterior-to-anterior transformations, which are typically associated with trxG mutations (Milne et al., 1999). This suggests *sxc/Ogt* acts not only as a Hox gene repressor through PcG but also as an activator through trxG. In fact, TRX and related transcriptional activators SET1/COMPASS and ASH1 have been shown to be *O*-GlcNAc modifiable, and staining for these proteins overlaps with *O*-GlcNAc on polytene chromosomes (Akan et al., 2016). The interaction between *O*-GlcNAc and trxG also appears in mammals, with *O*-GlcNAc modification of the H3K4 methyltransferase MLL5 working to stabilize this trxG enzyme (Ding et al., 2015). Thus, in addition to the necessary role of *O*-GlcNAc in the Polycomb repressive complex, this modification is also involved in transcriptional activation, highlighting the diverse routes in which *O*-GlcNAc regulates gene expression.

The appropriate spatiotemporal expression of Hox genes is crucial to the proper development of an organism. These regulatory systems which repress and activate Hox genes can cause severe developmental disorders when out of balance (Quinonez and Innis, 2014). Common clinical presentation for PcG mutations include intellectual disability and growth defects (Deevy and Bracken, 2019), which mirror those of OGT-XLID. Thus, the role of OGT in developmental regulation may contribute to symptoms of this disorder.

O-GlcNAc Transferase in Cell Fate Plasticity

Developmental plasticity is necessary for the development of multicellular organisms, but this plasticity must be restricted when cells reach their terminal fates. Several labs have worked to define which factors are involved in the epigenetic processes that impose barriers between cell fates using the nematode *C. elegans*. The worm is a model system perfectly suited to study the genetics of development, as *C. elegans* are relatively simple multicellular animals, with a rigidly regulated series of cell divisions and fate decisions. Several studies have employed genetic screens to determine which genes are required to maintain proper cell fates. Two independent studies have found the sole nemotode *OGT* ortholog *ogt-1* is involved in the restriction of cellular plasticity.

Both experiments are based around genetic screens using strains sensitized to cell fate transformation by ectopic overexpression of the CHE-1 transcription factor. CHE-1 is a master regulator which is necessary to establish the fate of a specific pair of neurons called the ASE neurons, which can be monitored with a cell-specific reporter: *gcy-5::gfp* (Tursun et al., 2011). Due to mechanisms that impose barriers between cell fates, ectopic expression of *che-1* later in development does not have any effects on gene expression or cell fate. Even when overexpressed in tissues that would not normally transcribe *che-1*, neither activation of CHE-1 target gene expression nor induction of ASE fate are observed in otherwise wild-type animals (Tursun et al., 2011). However, they found that mutation of the histone chaperone *lin-53* allowed *che-1* overexpression to implement neural fate in germ cells (Tursun et al., 2011), suggesting a role for histones and potentially histone modifications in maintaining cellular identity. Using similar genetic systems, more recent studies have uncovered additional factors, including *ogt-1* (Hajduskova et al., 2019; Rahe and Hobert, 2019), that are required for maintenance of cellular identity.

Hajduskova et al. (2019) performed an RNAi screen of 730 candidate genes with known or predicted roles in chromatin modifications and remodeling. For this screen, expression of *che-1* was induced with a heat-shock promoter and reporter GFP signal was monitored. Ectopic GFP expression was detected with knockdown of 10 of the 730 tested genes. Different gene knockdowns allowed GFP expression in different tissues, suggesting tissue-specific regulatory mechanisms. This study focused on one hit in particular: *mrg-1*, an ortholog of the mammalian gene *MRG15*. Knockdown of *mrg-1* enabled *che-1* overexpression to reprogram germline cells to neuron-like cells. To further study the process by which MRG-1 works as a barrier between cell fates, the authors performed IP-MS to identify MRG-1 binding partners, and uncovered OGT-1. The IP-MS experiments also identified other known chromatin

interactors such as *set-26* (H3K9 methyltransferase) and *sin-3* (histone deacetylase), which is itself a predicted interactor with OGT-1 (Yang et al., 2002).

A similar screen was performed in an independent lab which looked for genes involved in protecting epidermal cells from transdifferentiation (Rahe and Hobert, 2019). In this study, they overexpressed *che-1*, but restricted its expression to the epidermis starting at the end of the final larval stage (L4) and continuing into adulthood. Without mutagenesis, *gcy-5::gfp* expression was limited to ASE neurons, but mutations in seven genes caused ectopic expression of the reporter. Three independent missense mutations were isolated in the *ogt-1* gene. These three mutations were at highly conserved residues within the catalytic domains of OGT-1, in close proximity to sites previously reported to be necessary for enzymatic activity in human OGT (Lazarus et al., 2011). Other genes this screen identified included *dot-1.1* (an H3K79 methyltransferase) and *pmk-1*/p38-alpha MAPK. In mammalian neuroblastoma cells, OGT and p38 MAPK have been shown to physically interact to drive *O*-GlcNAcylation of specific targets (Cheung and Hart, 2008).

Though these studies were performed in nematodes in the context of ectopic gene expression, these results point toward an important role for *O*-GlcNAcylation in the control of cell fate plasticity. *ogt-1* came out of these two independent unbiased methods in the context of two different tissue types, suggesting it plays this role broadly. Though it has not been studied in as much depth in mammalian systems, some studies implicate *O*-GlcNAc in related processes. Much evidence points toward *O*-GlcNAc homeostasis being critical in stem and progenitor cells, which we will discuss further through its role in DNA damage and transcription factor regulation. In addition to its role restricting plasticity late in development, *O*-GlcNAc is critical in the networks of transcription factors which enable pluripotency in early development and adult stem cells which we will discuss in detail in its own section below. The developmental phenotypes of OGT-XLID patients and related transcriptomic data demonstrate the importance of *O*-GlcNAc in regulating sensitive cell fate specification events (Selvan et al., 2018). The importance of *O*-GlcNAc in cell fate may be intimately related to its roles in histone dynamics.

O-GlcNAc CYCLING AND HISTONES

Direct Modification of Histones by *O*-GlcNAc

O-GlcNAc affects chromatin and histone dynamics in a number of different ways. As described above, major complexes that modify histones are regulated by *O*-GlcNAc cycling. *O*-GlcNAc is also involved in the histone exchange required for DNA repair, and can modify histones directly.

Sakabe et al. (2010) first demonstrated direct evidence linking *O*-GlcNAcylation to the histone code, the holy grail of modern epigenetics (Jenuwein and Allis, 2001). Using an arduous combination of techniques they found *O*-GlcNAc on histones H2A, H2B, and H4. Acetylated histones were modified by *O*-GlcNAc and *O*-GlcNAcylation increased upon heat stress. Heat shock and OGT overexpression also modestly increased chromatin condensation. Using a chemical enrichment procedure, they mapped three *O*-GlcNAc sites on histones: H2AT101, H2BS36, and H4S47 (Sakabe et al., 2010). Additionally, they also provided evidence that other sites, including probable sites for modification of the remaining histone H3 must exist in the histone preparations. The identification of *O*-GlcNAc sites near known DNA interaction sites lead the authors to speculate that *O*-GlcNAcylation could induce major changes in chromatin structure not only by regulating peptide backbone conformation but also due to it being considerably larger than other common PTMs. One particularly intriguing *O*-GlcNAcylated site discussed was H4S47 (Sakabe et al., 2010). In yeast, mutation of this site to a cysteine induced activation of SWI/SNF targets independent of the SWI/SNF chromatin remodeling complex.

Since then, a total 17 different histone *O*-GlcNAcylation sites on H2A, H2AX, H2B, H3, H3.3, and H4 have been reported (see **Table 1** for the full list with references). Evidence for these sites primarily comes from identification techniques such as chemoenzymatic detection, immunoblotting, selective enzymatic labeling, and lectin staining, in combinations with various mutation experiments. Due to the indirect nature of these methods, some skepticism has been raised about the existence of histone *O*-GlcNAcylation (Fujiki et al., 2011; Schouppe et al., 2011; Zhang et al., 2011; Hahne et al., 2012; Deplus et al., 2013; Lercher et al., 2015; Ronningen et al., 2015; Chen and Yu, 2016; Hirosawa et al., 2016; Hayakawa et al., 2017). However, the presence of *O*-GlcNAc at the two sites, H2AS40 and H3.3T32, have been confirmed based on the recognition of endogenous *O*-GlcNAc by tandem mass spectrometry (MS) analysis of histones isolated from mammalian cells (Sakabe et al., 2010; Fong et al., 2012), providing the most robust evidence for the *O*-GlcNAcylation of histones. Direct identification of peptidyl *O*-GlcNAcylated Ser/Thr by MS is burdensome due to its unstable nature, limiting attempts to confirm additional modification sites. Some studies have disputed the existence of histone *O*-GlcNAcylation, including one which reported an inability to detect modified histones in cultured mammalian cells (Gagnon et al., 2015; Gambetta and Muller, 2015). Still, indirect evidence suggests that *O*-GlcNAcylation is an important form of histone PTM.

The 17 identified histone *O*-GlcNAcylation sites play various important roles in different biological functions. *O*-GlcNAc is identified at higher levels on H3 during interphase than mitosis, inversely related with phosphorylation, suggesting PTM crosstalk. Also, an increase in *O*-GlcNAcylation was observed to reduce mitosis specific phosphorylation at Ser10, Ser28, and Thr32. Inhibition of OGA hindered the transition from G2 to M phase of the cell cycle, showing a phenotype similar to hindering mitosis-specific phosphorylation on H3 delivering a mechanistic switch that orchestrates the G2-M transition of the cell cycle (Fong et al., 2012).

The combination of the extensive diversity of histone modifications allows for the complexity and flexibility of epigenetic regulation (Chen and Yu, 2016). It is now possible to map the genome- wide distribution and colocalization

TABLE 1 | Reported sites of direct *O*-GlcNAc modification of histones.

Sites	Sample	Enrichment	Detection	Functions	References
H2AS40	mES cells	RP-HPLC and mAb	*O*-GlcNAc site by HCD tandem MS	Tightly relates with the differentiation in mouse trophoblast stem cells	Hirosawa et al., 2016; Hayakawa et al., 2017
H2AT101	HeLa cells	GalNAz labeling and DTT tagging	DTT tag by CID tandem MS	May be a part of the histone code	Sakabe et al., 2010
	Recombinant histone	*In vitro* reaction with OGT	*O*-GlcNAc site by ETD tandem MS	Facilitates H2B K120 monoubiquitination, for transcriptional activation	Fujiki et al., 2011
H2AXT101	HeLa cells	Laser micro-irradiation, immunofluorescence (IF) staining and microscope image acquisition	Abolished *O*-GlcNAc signal by CTD110.6 Ab for FLAG-tagged H2AXT139A mutant	The *O*-GlcNAcylation negatively regulates DNA double-strand break-induced phosphorylation of H2AX and MDC1 by restraining the expansion of these phosphorylation events from the sites of DNA damage.	Chen and Yu, 2016
H2AXS139				Co-localizes with DNA damage foci, may function in DNA damage repair	
H2BS36	HeLa cells	GalNAz labeling and DTT tagging	DTT tag by CID tandem MS	May be a part of the histone code	Sakabe et al., 2010
H2BS52	Various cell lines	Proteome-wide studies without any specific enrichment	Large scale CID tandem MS using the Oscore software, which assesses presence of *O*-GlcNAcylation	Suggests that *O*-GlcNAc and phosphorylation are not necessarily mutually exclusive but can occur simultaneously at adjacent sites.	Hahne et al., 2012
H2BS55					
H2BS56					
H2BS64	Calf thymus	Lectin-pulldown and butylamine tagging	Butylamine tagging by CID tandem MS	Suggest the presence of *O*-GlcNAc-modified proteins among the lectin-binding partners	Schouppe et al., 2011
H2BS112	Various cell lines	Immunofluorescence, Chromatin immunoprecipitation, Immunoblotting	Immunofluorescence, Chromatin immunoprecipitation, Immunoblotting	Preserves a stable chromatin and represses gene transcription at the early stage of adipocyte differentiation	Ronningen et al., 2015
	HEK293T and *Tet2* knockout mouse	HT-pulldown and Chromatin IP	IP	Direct physical link between OGT and TET2/3 proteins provide new insight into the regulation and function of OGT in the cell.	Deplus et al., 2013
	HeLa cell	*In vitro* reaction with OGT	ETD–MS/MS mapping	Facilitates H2BK120 monoubiquitination, for transcriptional activation	Fujiki et al., 2011
H3S91	Recombinant histone	*In vitro* reaction with OGT	*O*-GlcNAc site by ETD tandem MS	Preserves stable chromatin in the early stages of cell differentiation and may repress gene transcription in adipocytes	Fujiki et al., 2011
H3S112					
H3S123					
H3S10	HEK239 cell	Overexpression and IP by tag Ab	Abolished *O*-GlcNAc signal by lectin staining of FLAG-tagged H3S10A mutant	Competitively reduces the levels of H3S10 phosphorylation, therefore regulates the pathway that H3S10P involved in, such as passing the G2-M phase check point, regulating the H4K16ac	Zhang et al., 2011
H3.3T32	HeLa cell	IP by anti-H3 Ab	*O*-GlcNAc site by ETD tandem MS	Increases the phosphorylation of Thr32, Ser28, and Ser10, which are the specific mark of mitosis	Fong et al., 2012
H3.3T80	Calf thymus	Lectin-pulldown and butylamine tagging	Butylamine tag by CID tandem MS	Suggest the presence of *O*-GlcNAc-modified proteins among the lectin-binding partners	Schouppe et al., 2011
H4S47	HeLa cell	GalNAz labeling and DTT tagging	DTT tag by CID tandem MS	May be a part of the histone code	Sakabe et al., 2010

mES cells, mouse Embryonic Stem cells; RP-HPLC, reversed-phase high performance liquid chromatography; Ab, antibody; mAb, monoclonal antibody; GalNAz, azide-modified galactose; DTT, dithiothreitol; IP, immunoprecipitation; MS, mass spectrometry analysis; HCD, higher-energy collisional dissociation; CID, collision-induced dissociation; ETD, electron-transfer dissociation.

of histone modifications at high resolutions using ChIP-seq, revealing the many amalgamations of histone modification crosstalk, such as precondition, mutual exclusion, or coexistence (Pick et al., 2014). H2BS112 *O*-GlcNAcylation functions as a precondition for H2BK120 monoubiquitination, with GlcNAc acting as an anchor for ubiquitin ligase, ultimately resulting in transcriptional activation via H3K4me3 (Fujiki et al., 2011). An increase in the intracellular level of UDP-GlcNAc induces an increase in histone *O*-GlcNAcylation and a partial suppression in H3S10ph, suggesting these modifications are mutually exclusive. Further examples of competition between *O*-GlcNAcylation and phosphorylation have been reported for the H3T32, H3S10, and H2AXS139 sites (Zhang et al., 2011; Fong et al., 2012; Chen and Yu, 2016). Future studies should, therefore, investigate phosphorylation of other residues for which *O*-GlcNAc modification has been reported. The crosstalk between *O*-GlcNAcylation and acetylation should also be validated considering the presence of a HAT-like domain in OGA, although the enzymatic activity of this domain is a point of controversy (Torres and Hart, 1984; Toleman et al., 2006; Kim et al., 2007; Butkinaree et al., 2008; Hayakawa et al., 2013; Rao et al., 2013).

O-GlcNAcylation of site-specific adapter proteins directly regulate the stability of H2A/H2B dimers in the nucleosome in synthetic *O*-GlcNAcylated histones (Lercher et al., 2015). To generate homogenous *O*-GlcNAc modified nucleosomes, one study generated H2AT101 *O*-GlcNAc mimics by replacing the threonine with cysteine and using a series of chemical reactions *in vitro* to stably link GlcNAc to the thiol. The authors showed that H2AT101 GlcNAcylation destabilized the H3/H4 tetramer-H2A/B dimer interface reducing nucleosome stability. Thus, regulation of nucleosome stability by OGT-dependent GlcNAc transfer may contribute to transcriptional regulation. *O*-GlcNAc's role in histone dynamics and in modifying variant histones make up an additional layer of histone regulation.

H2AX and the DNA Damage Response

Histone modifications play a crucial role in chromatin organization through processes including DNA metabolism, replication, transcription, and repair. Modification and exchange of histones can also reorganize chromatin to allow DNA repair machinery to access damaged chromosomal DNA (Downs et al., 2004). H2AX is a histone variant that differs from H2A at various amino acid residues along the entire protein and in its C-terminal extensions (Bonisch and Hake, 2012). The importance of this histone variant is highlighted by the phenotypes of knockout mice, which show radiation sensitivity, developmental delay, and male infertility (Celeste et al., 2002). H2AX is a central player in the DNA damage response (DDR) when phosphorylated at serine 139 (γH2AX) (Wahl and Carr, 2001), and as mentioned previously, can also be *O*-GlcNAc modified at this site (Liu and Li, 2018). γH2AX is incorporated into nucleosomes at double strand break sites, where it promotes accumulation of DNA repair proteins (Wahl and Carr, 2001). γ-phosphorylation is an early event in the DSB damage response induced by the ATM and ATR kinases, which additionally activate kinases Chk1 and Chk2 (Wahl and Carr, 2001).

In the growth and development of an organism, DNA damage poses a serious risk, as mutations will propogate from progenitors to their daughter cells. To maintain genome integrity, the cell cycle is regulated by the DDR pathway following DNA damage stress, with *O*-GlcNAc involved by modifying the arrangement of histones and kinases (Hanover et al., 2018; Liu and Li, 2018). Blocking *O*-GlcNAc transferase activity leads to delayed DSB repair, reduced cell proliferation, and increased cell senescence *in vivo*, while increased *O*-GlcNAc promotes DSB repair and hyper-proliferation *in vivo* and *in vitro* (Efimova et al., 2019). These findings suggest *O*-GlcNAc is necessary to protect the genome and for proper cell cycle progression. These effects are likely related to OGT's recruitment to sites of DNA damage, where it modifies H2A and mediator of DNA damage checkpoint 1 (MDC1) (Chen and Yu, 2016). In addition, one report has suggested OGT transfers GlcNAc onto H2AXS139, the same site as γ-phosphorylation (Chen and Yu, 2016). The authors suggest *O*-GlcNAc inhibits the DDR, though other studies suggest *O*-GlcNAc activates the DDR pathway (Efimova et al., 2019; Na et al., 2020).

The development of a multicellular organism is an energy-intensive and error-prone process. Stem and progenitor cells require high levels of glucose to grow and proliferate. A consequence of this energy use is the generation of reactive oxygen species (ROS), which cause damage, cellular stress, and DNA breaks. As such, DDR-related factors are prominent among proteins that accumulate *O*-GlcNAc when cells are stressed by ROS (Katai et al., 2016). Reciprocally, the DDR pathway has been shown to increase ROS levels (Rowe et al., 2008), pointing to a complex interplay between these processes. For example, high glucose has been demonstrated to elevate levels of *O*-GlcNAc, ROS, and DNA damage (Hu et al., 2019). The elevation of ROS in high glucose conditions may be linked to *O*-GlcNAc, as inhibition of OGT decreases ROS levels in a dose-dependent manner, and ultimately reduces neural tube defects in embryos of diabetic mice (Kim et al., 2017). Recently, we have reported that stress induces hyper-proliferation, *O*-GlcNAcylation, and DDR in *Drosophila* intestinal stem cells (Na et al., 2020). Likewise, genetic elevation of *O*-GlcNAc by deletion of *Oga* induced proliferation and DDR in fly intestinal stem cells, mouse embryonic fibroblasts, and mouse ESCs (Na et al., 2020). Previous work had shown that Chk1 phosphorylates OGT, which stabilizes the protein (Li et al., 2017). Through this interaction, we demonstrate that *O*-GlcNAc participates in an autoregulatory feedback loop where CHK1/CHK2 stabilizes OGT, allowing further *O*-GlcNAcylation that continues to activate the DDR pathway (Na et al., 2020) (**Figure 2**).

Thus, *O*-GlcNAcylation is a key regulator of the DDR pathway, which is crucial in supporting the development of a healthy organism. This further establishes *O*-GlcNAc's role in cell identity discussed above, which will be explored below in our discussion of transcription factor modification. Observations of abnormal neural proliferation and developmental delay in mice harboring an *Oga* deletion in the brain (Olivier-Van Stichelen et al., 2017) suggest *O*-GlcNAc homeostasis is needed for the proper development of sensitive tissues such as the brain. Thus, the pathways discussed above may contribute to the developmental

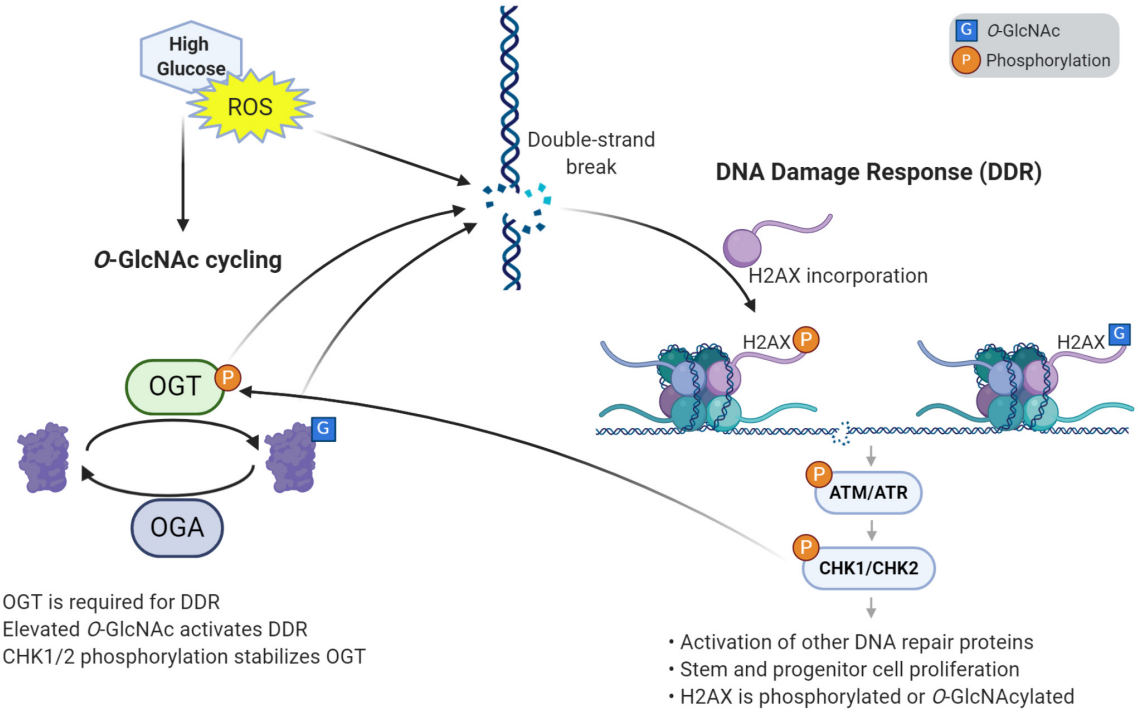

FIGURE 2 | *O*-GlcNAcylation induces the DNA damage response and proliferation in stem cells. High glucose and reactive oxygen species (ROS) cause damage to DNA such as double strand breaks. *O*-GlcNAcylation is required to respond to DNA damage, through its accumulation at damage sites and activation of the DNA damage response (DDR) pathway. Response to DNA damage also requires the incorporation of variant histone H2AX, which is phosphorylated and can also be *O*-GlcNAc modified. Elevated *O*-GlcNAc induces stem and progenitor cell proliferation and the DNA damage response through γH2AX and a phosphorylation cascade of effector kinases. One of the downstream kinases, CHK1/2, phosphorylates OGT, which stabilizes the enzyme. This autoregulatory loop helps further promote the DNA damage response and proliferation.

defects associated with OGT-XLID. Beyond the effects *O*-GlcNAc has on chromatin dynamics, *O*-GlcNAc interacts with DNA methylation pathways to regulate gene expression.

OGT AND DNA METHYLATION

DNA methylation is a critical epigenetic modification in mammals which occurs predominantly at the 5-position carbon on cytosine residues (5mC) followed by guanines. This epigenetic mark is involved in a variety of functions in the mammalian genome, including X-chromosome inactivation, gene silencing, genomic stability, cellular identity, and genomic imprinting (SanMiguel and Bartolomei, 2018). Two models for DNA methylation-dependent repression have been described. The first is a direct mechanism in which the presence of 5mC inhibits binding of transcription factors to DNA, thereby silencing gene expression. The second model is an indirect mechanism that involves recruitment of proteins that bind methylated DNA and associate with chromatin modifiers. These models are not mutually exclusive and can work in concert (Klose and Bird, 2006). Despite being stable and heritable, DNA methylation is also highly dynamic, particularly during development. Active DNA demethylation is mediated by the TET family proteins TET1, TET2, and TET3. These proteins iteratively oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009; Ito et al., 2010, 2011; He et al., 2011). These modifications can be transient intermediates in the demethylation process which are ultimately removed by base excision repair (Cortellino et al., 2011; He et al., 2011), or can act on their own as a stable modifications (Bachman et al., 2014, 2015). Importantly, OGT is a well established binding partner of the TET proteins and has recently been described to play a critical role in DNA methylation-dependent repression.

Recent studies have revealed that OGT is found in complex with the three TET proteins. OGT was found to be associated with TET1 at gene promoters of transcriptionally active genes in mouse ESCs. This interaction was required for OGT binding to chromatin and enhanced TET1 activity (Vella et al., 2013). Further, *O*-GlcNAcylation of TET1 was shown to regulate TET1 stability (Shi et al., 2013a), and OGT enhanced TET1 activity *in vitro* (Hrit et al., 2018). Disrupting the OGT-TET1 interaction in mouse ESCs resulted in increased 5mC and compensatory increases in TET2, accompanied by transcriptional changes (Hrit et al., 2018). These data indicate that the TET1-OGT complex is critical for proper pluripotency gene regulatory networks which maintain stem cell identity.

TET2 and TET3 also form complexes with OGT. Interaction of TET2 with OGT associates at transcriptional start sites

and facilitates histone *O*-GlcNAcylation. HCF-1 as part of the SET1/COMPASS complexes is a specific target of the TET2/3-OGT complex and promotes H3K4me3. A closer look at the TET3-OGT interaction indicated that this interaction stabilized OGT and enhanced chromatin association (Ito et al., 2014). However, another study found that *O*-GlcNAcylation of TET3 promoted its export from the nucleus, thereby inhibiting TET3 function (Zhang et al., 2014). Further complicating the regulation of TET proteins by OGT is the fact that they are also highly phosphorylated (Bauer et al., 2015) setting up the possibility of PTM crosstalk. While the functional relationship between OGT and TET proteins still remains controversial, there are some common themes in these papers. TET, OGT, and H2BS112 *O*-GlcNAc are colocalized in the genome, largely at CpG islands containing promoters of actively transcribed genes (Chen et al., 2013; Deplus et al., 2013; Vella et al., 2013). Thus this interaction likely reinforces active transcription to maintain cell intrinsic transcriptional program, possibly impacting development and gene expression in OGT-XLID.

In addition to its effects on the process of DNA demethylation, OGT has been found to be involved in gene silencing mediated by DNA methylation. OGT selectively associates with the scaffolding protein TRIM28 only in the presence of methylated DNA (Boulard et al., 2020). It has been proposed that *O*-GlcNAcylation of chromatin modifiers that interact with TRIM28 is required at the sites of retrotransposon promoters to repress their transcription (Boulard et al., 2020). This suggests disruption of *O*-GlcNAc cycling may lead to increased genome instability in addition to the contribution of impaired DNA damage response described above.

TRANSCRIPTION FACTOR *O*-GlcNAcylation

Gene expression is largely regulated by transcription factors, which themselves are heavily regulated by PTMs. *O*-GlcNAcylation is one of the major modifications that affects transcription factor functions, modulating their localization, stability, interacting partners, resulting in gene activation or silencing. While many studies have defined the role of *O*-GlcNAc modification of key transcription factors regulating processes like immune activation (Chang et al., 2020), here we focus on the role O-GlcNAc has in pluripotency and differentiation as well as in the nervous system and glucose and lipid metabolism in the liver as it relates more directly to neurodevelopmental disorders that will be described (**Figure 3**).

O-GlcNAcylation in Development and Cell Differentiation

Throughout development, the process of differentiation allows a multicellular organism to form a complex system of multiple tissues and cell types. In adulthood, stem cells retain the capacity to differentiate allowing for tissue repair. *O*-GlcNAcylation during development is essential, as deletion of *Ogt* in mouse ESCs is lethal (Shafi et al., 2000). Further, reduced expression of *Ogt* disrupts stem cell self-renewal through deregulation of pluripotency transcriptional networks (Jang et al., 2012). This is likely due to OGT's role in chromatin remodeling as described above, as well as through direct modification of transcription factors involved in pluripotency. In fact, the Yamanaka factors OCT4, SOX2, and C-MYC are all *O*-GlcNAc modified (Chou et al., 1995; Jang et al., 2012), and increasing *O*-GlcNAc in ESCs hampers normal differentiation (Jang et al., 2012). In fact, deleting *Oga* in the mouse is largely perinatal lethal (Keembiyehetty et al., 2015). Mice in which *Oga* was conditionally deleted in the brain exhibited microcephaly, high body fat percentage, hypotonia, and delayed development of the brain associated with abnormal cell proliferation and migration. This phenotype is related to transcriptional changes of pluripotency markers, including *Sox2*, *Nanog*, and *Otx2* (Olivier-Van Stichelen et al., 2017). The transcription factor *Sox2* and other transcription factors involved in the maintenance of mouse ESC pluripotency and stem cell self-renewal all require *O*-GlcNAcylation (Myers et al., 2011; Jang et al., 2012).

Cellular differentiation is regulated by *O*-GlcNAcylation, as lowered UDP-GlcNAc levels by HBP inhibition blocks the differentiation of adipocytes *in vitro* (Ishihara et al., 2010). Normal adipocyte differentiation in cell culture requires the transcription factors C/EBPα and C/EBPβ. It has been reported that C/EBPβ is modified by *O*-GlcNAc, and diminished *O*-GlcNAcylation reduces C/EBPα protein levels, suggesting the modification stabilizes the protein (Ishihara et al., 2010). Mass spectrometry analysis indicated two *O*-GlcNAc sites on Ser180 and Ser181, which are very close to C/EBPβ phosphorylation sites at Thr188, Ser184, and Thr179 (Li et al., 2009). The sequential phosphorylation of C/EBPβ at Thr188 then Ser184 by MAPK or CDK2, and Thr179 by GSK3β is required for DNA binding and transcriptional activity (Kim et al., 2007). Increased *O*-GlcNAcylation during adipocyte differentiation prevents C/EBPβ phosphorylation and subsequently delays adipocyte differentiation (Li et al., 2009). In addition, mutations of Ser180/181 rescued the phenotype induced by *O*-GlcNAcylation which suggests that the transcriptional activity of C/EBPβ is regulated by phosphorylation and *O*-GlcNAcylation in a competitive manner by alternative occupancy at adjacent sites (Li et al., 2009).

Hematopoietic stem cell (HSC) maintenance also requires balanced *O*-GlcNAc cycling. In fact, when *Oga* has been deleted in HSCs in mice, these mice exhibit diminished HSC pools as well as reduced intermediate progenitor populations. The elevated *O*-GlcNAcylation in progenitor cells of these mice were correlated with transcriptional changes in factors involved in adult stem cell maintenance, lineage specification and nutrient uptake (Abramowitz et al., 2019).

Transcription Factor *O*-GlcNAcylation in the Nervous System

O-GlcNAcylation has been extensively studied in the brain due to its critical role during development described above, and in neurodegenerative diseases like Alzheimer's (Griffith and Schmitz, 1995; Olivier-Van Stichelen et al., 2017). Emerging data has highlighted the role of transcription factor *O*-GlcNAcylation

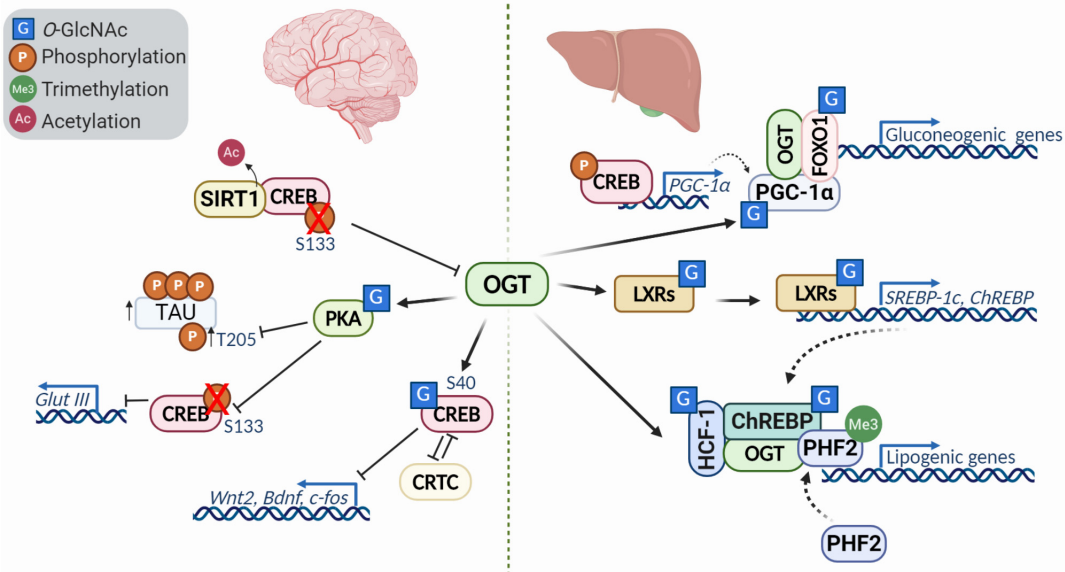

FIGURE 3 | Diversity of transcription factor *O*-GlcNAcylation in the brain and liver. *O*-GlcNacylation of transcription factors can act to modulate their activity, localization, stability, binding partners and other post translational modifications. In the brain **(left)**, *O*-GlcNAcylation of CREB represses its activity and downregulates transcriptional targets like *Wnt2* and *Bdnf* in a CRTC (cAMP-regulated transcriptional co-activators)-dependent manner. Reduced *O*-GlcNAc of protein kinase A (PKA) is characteristic of Alzheimer's disease, and leads to decreased phosphorylation of CREB at S133 and increased phosphorylation of tau at T205. Further, deacetylation of CREB by SIRT1, causes downregulation of OGT. In the liver **(right)**, modification of PGC-1a and association with OGT promotes FOXO1 *O*-GlcNAcylation, enhancing activity and promoting transcription gluconeogenic genes. The liver X receptors (LXRs) are nutrient sensors with critical roles in lipid metabolism, glucose homeostasis, and metabolism. *O*-GlcNAc modification of LXRs induces expression of SREBP-1 and ChREBP, key factors involved in hepatic lipogenesis and energy metabolism. HCF-1 *O*-GlcNAcylation stimulates OGT recruitment and modification of ChREBP, enhancing ChREBP activity and promoting recruitment of the activator PHF2 which binds the activating histone mark H3K4me3.

during neuronal development, synaptic plasticity and memory. CREB, a key transcription factor involved in learning and memory, is *O*-GlcNAc modified in the TAFII130 binding domain, a component of the TFIID transcriptional complex (Lamarre-Vincent and Hsieh-Wilson, 2003). *O*-GlcNAcylation of CREB impairs its interaction with TAFII130 and represses CREB transcriptional activity *in vitro* (Lamarre-Vincent and Hsieh-Wilson, 2003). Further studies demonstrated that *O*-GlcNAcylation of CREB at Ser40 modulates dendrite and axonal elongation with downregulation of Wnt2 and BDNF signaling (Rexach et al., 2012). It has also been suggested that glycosylation of CREB has a significant impact on long-term memory consolidation (Rexach et al., 2012). In accordance with this hypothesis, an independent research group has provided a link between *O*-GlcNAcylation, Protein Kinase A (PKA)-CREB signaling, and memory loss in Alzheimer's disease. Alzheimer's disease is associated with a decrease in *O*-GlcNAcylation, which has been shown to influence aggregating proteins such as tau (Akan et al., 2018). PKAs can be *O*-GlcNAc modified, which influences their localization, activity, and phosphorylation (Xie et al., 2016). The inhibition of PKA-CREB signaling by *O*-GlcNAcylation was associated with impaired learning and memory in mice.

While CREB itself is regulated by *O*-GlcNAc, it can also regulate OGT expression. CREB can be deactivated by the deacetylase SIRT1, thereby reducing OGT expression and promoting tau phosphorylation, one of the major events in the course of Alzheimer's disease (Lu et al., 2020).

With important roles in regulating the transcriptional networks required for proper development, differentiation and within the nervous system, it is unsurprising that deregulation in *O*-GlcNAc homeostasis is associated with neurodevelopmental diseases.

Transcription Factor *O*-GlcNAcylation in Carbohydrate and Lipid Metabolism

Beyond the brain, CREB also plays an important role in the liver where it regulates hepatic glucose and lipid metabolism (Herzig et al., 2001; Dentin et al., 2008). During prolonged fasting, CREB stimulates the gluconeogenic program with the coactivator PGC-1 (Herzig et al., 2001). PGC-1α (peroxisome proliferator-activated receptor gamma, co-activator 1 alpha) is a transcriptional co-activator that controls energy and nutrient homeostasis by coordinating gene expression. PGC-1α has been shown to form a complex with OGT and be *O*-GlcNAcylated at Ser333. Moreover, increased glucose levels lead to FOXO1 (Forkhead box other 1) *O*-GlcNAcylation via the PGC-1α/OGT complex, enhancing transcriptional activity (Housley et al., 2008). In addition, increased *O*-GlcNAc, either by addition of glucosamine or an OGA inhibitor, enhanced FOXO1 target gene expression in HepG2 cells (Kuo et al., 2008).

Insulin has two main functions within the liver: (1) downregulation of gluconeogenesis genes by initiating inhibitory phosphorylation of FOXO1, and (2) promotion of lipogenic pathways through activation of SHREBP-1c

(Brown and Goldstein, 2008). *O*-GlcNAcylation regulates both pathways by attenuating insulin signaling and activating lipogenic pathways. The liver X receptors (LXRs) are described as nutritional sensors for lipid metabolism, glucose homeostasis and inflammation, and are posttranslationally modified by phosphorylation, acetylation, and *O*-GlcNAcylation (Anthonisen et al., 2010). Increased glucose levels leads to LXR *O*-GlcNAcylation, inducing SREBP-1c (sterol regulatory element binding protein 1c) expression. SREBP-1c is a major player of gene expression in hepatic lipogenesis (Anthonisen et al., 2010). LXR has been shown to interact and co-localize with OGT *in vitro* and *in vivo*. Additionally, LXR enhanced the expression of *SREBP-1c* and *ChREBP*α/β under hyperglycemic conditions (Bindesbøll et al., 2015). The effects of *O*-GlcNAc on the LXR pathway may extend beyond the liver, as this was the most dysregulated pathway found by transcriptomics of OGT-XLID mutant cells (Selvan et al., 2018). The LXR pathway has been shown to be critical in the development of dopaminergic neurons from stem cells (Ma et al., 2009), providing a possible link between this pathway and the developmental disorder.

ChREBP (carbohydrate responsive element binding protein) is a key factor of energy metabolism in the liver and is regulated by *O*-GlcNAcylation. High glucose or OGT expression increased ChREBP *O*-GlcNAcylation, stabilizing the protein and increasing expression of its target genes. *Oga* overexpression in mouse livers markedly reduced ChREBP *O*-GlcNAcylation and decreased abundance of ChREBP targets (Guinez et al., 2011). These results clearly demonstrated that *O*-GlcNAcylation of ChREBP increases its stability and activity. In addition, a recent study identified HCF-1 as a modulator of ChREBP activity and the lipogenic program, which is glucose dependent. Elevated glucose induced HCF-1 *O*-GlcNAcylation and HCF-1/ChREBP complex formation, where HCF-1 recruits OGT to further promote *O*-GlcNAcylation of ChREBP (Lane et al., 2019). Moreover, HCF-1/ChREBP complex formation was associated with the recruitment of epigenetic activator PHF2, which binds to H3K4me3 and enhances transcription (Lane et al., 2019). These data demonstrated that lipogenic gene expression is under the control of epigenetic modulations, ChREBP *O*-GlcNAcylation, and activation by HCF-1.

Through regulation of transcription factors, *O*-GlcNAc is able to alter patterns of gene expression in response to nutrient status and stress. Mutations in OGT may impair its targeting of specific pathways and disrupt their normal function in responding to the environment and coordinating development.

PLACENTAL OGT AS A BIOMARKER FOR NEURODEVELOPMENTAL DISEASE: A MODEL OF TRANSGENERATIONAL INHERITANCE

Prenatal development is a particularly vulnerable time, when tight regulation of transcriptional networks is required to transform a fertilized egg into a multicellular organism requiring complex tissue development. In mammals, energy flow from the mother to the fetus is mediated by the placenta, which transfers macronutrients, gases, and metabolites into fetal circulation. Thus, the placenta transmits nutritional and stress information from the mother to the developing fetus. Glucose is the primary fuel for the fetus, which is provided by maternal transfer through the placenta. The fetal brain is a particularly nutritionally demanding tissue during fetal development. Brain regions are more sensitive to nutrient availability at particular gestational stages. Interestingly, a recent hypothesis to explain the male-biased presentation of neurodevelopmental diseases focuses on sex differences in the placenta in relaying signals to the developing brain regarding maternal perturbations (Charil et al., 2010; Howerton and Bale, 2012, 2014; Gabory et al., 2013; Howerton et al., 2013; Davis and Pfaff, 2014; Nugent and Bale, 2015; Bale, 2016).

As an X-linked gene and having the ability to transmit nutritional information to the nucleus, OGT is a unique candidate to signal maternal stress through the placenta to the developing fetus in a sex-dependent manner. Interestingly, *O*-GlcNAcylation of placental proteins correlates with maternal glycemic index (Dela Justina et al., 2018). In fact, a genome-wide screen looking for sex-specific changes in placental transcription after exposure to early prenatal stress identified *Ogt* as a top candidate for exhibiting sexually dimorphic expression in the placenta and changes in expression as a response to maternal stress (Howerton et al., 2013). This study found that male mouse placentas had about half the amount of OGT protein, corresponding to decreased total *O*-GlcNAc levels and even lower levels of OGT and *O*-GlcNAc upon prenatal stress, as compared to their female counterparts (Howerton et al., 2013). Hemizygous and homozygous mice that had *Ogt* deleted specifically in the placenta recapitulated models of early prenatal stress presenting with hypothylamic mitochondrial dysfunction characterized by transcriptional changes and altered cytochrome c oxidase activity (Howerton and Bale, 2014).

As discussed previously, OGT plays a critical role in regulating gene expression, particularly as a key regulator of Polycomb repression. Analysis into how placental expression of *Ogt* could causally contribute to neurodevelopmental disorders in the offspring have highlighted OGT's role in regulating the H3K27me3 repressive histone mark. Interestingly, Nugent et al. (2018), found that OGT establishes sex differences in placental H3K27me3, with an enrichment found in females. This sex difference was *Ogt* dependent, as genetic reduction of *Ogt* in mice masculinized female placental H3K27me3 levels (Nugent et al., 2018). It was hypothesized that the higher levels of H3K27me3 allow for resiliency and less transcriptional response to maternal stress in the female than the male placenta. One deregulated gene of particular interest was *Hsd17b3* (17-β-hydroxysteroid dehydrogenase-3). Through a ChIP-seq experiment on placental tissue, the researchers identified a correlation between *O*-GlcNAc occupancy and *Hsd17b3* expression and that *O*-GlcNAc is significantly reduced by exposure to early prenatal stress (Howerton and Bale, 2014). This gene codes a key enzyme in testosterone biosynthesis. Examination of testosterone in a model of early prenatal stress male placentas showed a significant reduction in testosterone levels, potentially contributing to hypothalamic changes in offspring (Howerton and Bale, 2014). Taken together, these observations supports the possibility that acting as an epigenetic modifier in the placenta, OGT is able

to pass transgenerational stress signals from the mother to the offspring in a sex-specific manner.

OGT AND X-LINKED INTELLECTUAL DISABILITY (OGT-XLID): A DISEASE OF *O*-GlcNAc IMBALANCE

X-linked intellectual disabilities are a group of a neurodevelopmental disorders representing about 5-10% of all cases of intellectual disability (Pravata et al., 2020b). Over 200 genes have been linked to XLID, although some of the candidates remain controversial. Over the past few years, a syndromic form of XLID affecting multiple families has been described which co-segregates with variants in the human OGT gene (Vaidyanathan et al., 2017; Willems et al., 2017; Selvan et al., 2018; Pravata et al., 2019, 2020a,b). At present, some 14 patients from 8 families have been analyzed which show non-synonymous variants in the OGT gene (Pravata et al., 2020a). All the patients carrying OGT variants were found to have decreased intellectual ability with IQ scores well below 70. In addition, the patients exhibit both mental and physical developmental delay, intrauterine growth retardation, low birth weight, short stature, restricted language skills, and drooling. A summary of XLID-causative variants and their position in the proposed OGT structure is shown in **Figure 4A**. It is notable that all of the inherited variants are in the TPR repeats (Bouazzi et al., 2015; Vaidyanathan et al., 2017; Willems et al., 2017; Selvan et al., 2018), which are implicated in both substrate binding and complex formation with other epigenetic regulators. Two mutations appeared *de novo* and are present in the catalytic domain, such as the N567L variant found in two female twins (Pravata et al., 2019), and the N648Y variant found in a male patient (Pravata et al., 2020a). On the basis of the common features of these patients and the fact that the clinical features co-segregate with the OGT gene it has been proposed that this syndrome be classified as a congenital disorder of glycosylation (CDG) termed OGT-CDG (Pravata et al., 2020b). Identification of this disorder strongly reinforces a growing body of evidence that *O*-GlcNAcylation plays a key role in development, particularly neurodevelopment. The involvement of OGT in numerous epigenetic pathways, with the propensity of epigenetic disorders to manifest in neurodevelopmental disorders (Gabriele et al., 2018) suggests possible mechanisms causing this form of XLID. Thus, the features of this disorder allows a dissection and discussion of the role of OGT and *O*-GlcNAcylation in many aspects of human physiology (**Figure 4B**).

OGT Activity: Enzyme Stability, Target Recognition and Responsiveness to Hexosamine Flux

The previous examination of XLID patients have revealed changes in OGT activity and stability (Vaidyanathan et al., 2017; Willems et al., 2017; Selvan et al., 2018). Interestingly, several reports have suggested little to no change in *O*-GlcNAc levels (Vaidyanathan et al., 2017; Willems et al., 2017). One exception is the N648Y mutation which shows reduced enzymatic activity (Pravata et al., 2019). However, the tools we have available may be too insensitive to detect small changes or changes in subsets of substrates. In particular, the mutations associated with TPR repeats have the potential to affect substrate or complex recognition without direct effect on the catalytic domain. Our previous structural work on the TPRs allowed us to propose a role for an asparagine ladder motif in substrate recognition (Jinek et al., 2004). Recently this hypothesis was confirmed by additional structural studies (Levine et al., 2018). Thus substrate recognition may be affected by the XLID mutations.

The mutations in OGT may also destabilize the OGT protein. This was observed with the OGT L254F mutation in particular (Pravata et al., 2020a,b). However, current data reveal that while OGT-XLID variants may be destabilized, OGT protein levels in the majority of cell lines carrying the XLID variants were minimally altered (Pravata et al., 2020b). It is yet to be determined what effect these variants might have on levels of UDP-GlcNAc derived from hexosamine flux. Since OGT is a critical regulator of glucose utilization and metabolic reprogramming, changes in both glucose utilization and synthesis of UDP-GlcNAc may be altered in OGT-XLID (Saeed et al., 2016; Hanover et al., 2018).

X-inactivation and XLID

The X chromosome has a unique pattern of inheritance compared to autosomal chromosomes. The X chromosome in males is inherited from the mother making recessive X-linked mutations predominant in males. The significance of the X-chromosomal location OGT has been previously discussed (Shafi et al., 2000; Hanover et al., 2003, 2012; Love et al., 2003, 2010; O'Donnell et al., 2004; Abramowitz et al., 2014; Olivier-Van Stichelen and Hanover, 2014, 2015; Olivier-Van Stichelen et al., 2014). In females, X-inactivation is primarily random, with individual cells inactivating one of the X chromosome employing a long non-coding RNA (*Xist*) and chromatin modifiers such as the Polycomb complexes (Love et al., 2010; Brockdorff, 2013; Froberg et al., 2013; Shi et al., 2013b; Monfort and Wutz, 2020). Interestingly, female XLID patients with *de novo* mutations at OGT N567K have been shown to exhibit extreme skewing in X-inactivation (98%) although it is unclear which of the two X chromosome are inactivated (Pravata et al., 2019). In these patients, the levels of *O*-GlcNAc are affected and this reduction can be recapitulated *in vitro* and in other model systems (Pravata et al., 2019). Given these findings, it is unclear why such skewing occurs in the context of this *de novo* mutation arising in female probands. OGT could contribute to DNA methylation/demethylation associated with the X-inactivation process. In addition, the role of OGT in Polycomb repression previously discussed raises the possibility that alterations in *O*-GlcNAc cycling could contribute to the skewing observed.

The OGA Gene and Compensation for OGT Perturbations

In many instances, variants in OGT associated with XLID resulted in a reduction in levels of OGA (Vaidyanathan et al., 2017; Willems et al., 2017; Selvan et al., 2018; Pravata et al.,

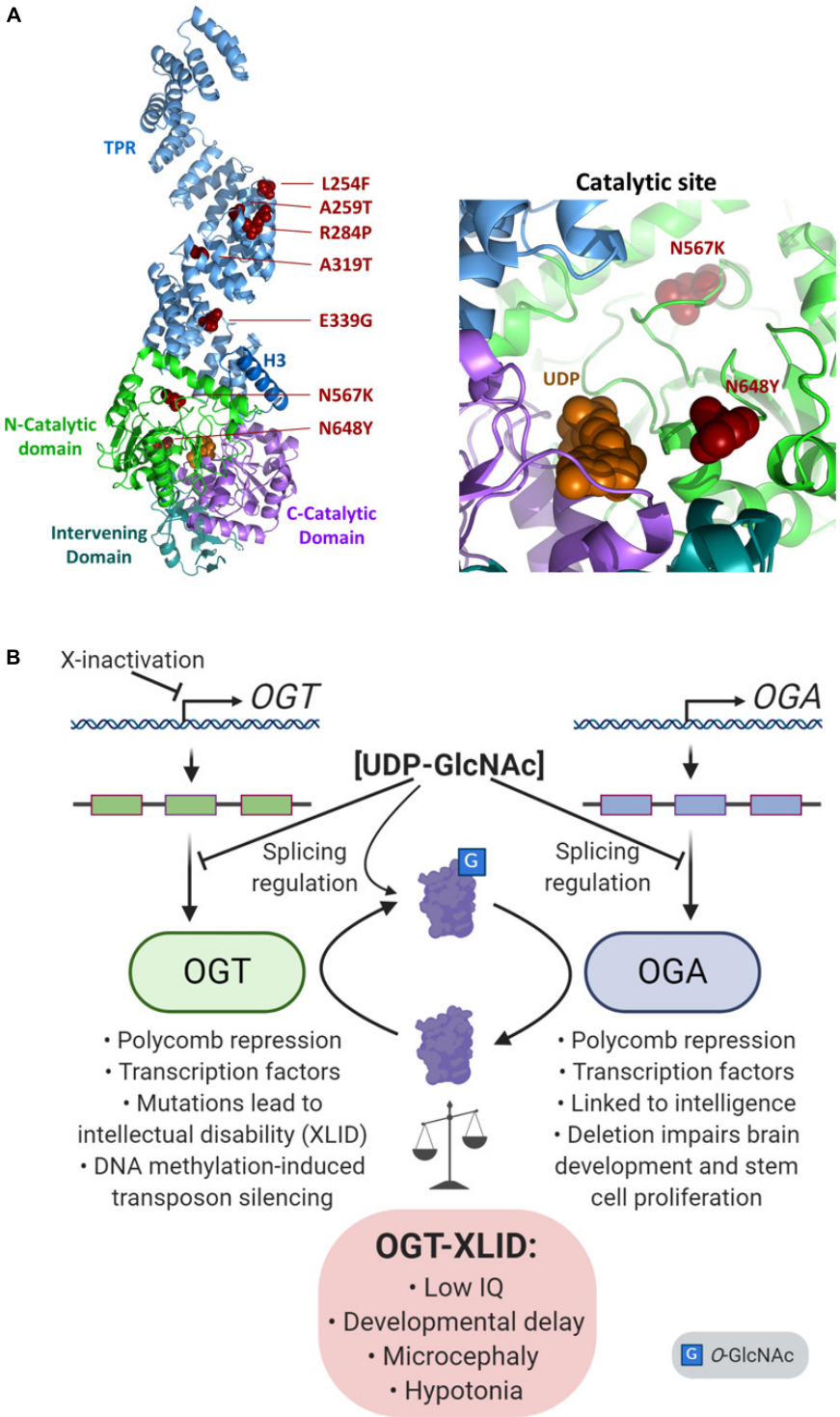

FIGURE 4 | Imbalance of *O*-GlcNAc cycling leads to epigenetic changes and disease. **(A)** The presumed structure of OGT (Lazarus et al., 2011) shown as a cartoon in complex with UDP (orange spheres), with sites of OGT-XLID causative variants highlighted (red spheres). Each domain is colored-coded and labeled with its name. To better show the variants in the catalytic domain, a zoomed panel rotated 180° about the *y*-axis from the full structure is shown on the right. **(B)** Expression of both *OGT* and *OGA* are regulated by cellular concentrations of UDP-GlcNAc through splicing mechanisms. As an X-chromosome gene, *OGT* is also regulated by X-inactivation. OGT and OGA proteins dynamically cycle *O*-GlcNAc to regulate many epigenetic mechanisms such as the Polycomb repressive complexes and regulation of transcription factors. In addition, both genes are linked to intelligence: *OGT* mutations cause intellectual disability, and a GWAS meta-analysis found *OGA* is associated with intelligence (Savage et al., 2018). Deletion of *Oga* disrupts proper neural development and the proliferation of mouse embryonic stem cells (Olivier-Van Stichelen et al., 2017).

2020a,b). OGA can upregulate gene expression of *OGT* through activation of the transcription factor C/EBP-β (Vaidyanathan et al., 2017), in addition to this protein's role in differentiation discussed above. Similarly, loss of *OGA* leads to an elevation in OGT protein levels (Keembiyehetty et al., 2015). The genes encoding OGA and OGT exhibit complex splicing patterns and recent findings suggest that mechanisms exist which serve to limit the translation of alternatively spliced species of the enzymes of *O*-GlcNAc cycling (**Figure 4B**) (Hanover et al., 2003; Park et al., 2017; Parra et al., 2020; Willems et al., 2017; Tan et al., 2020). *O*-GlcNAc regulates this process, suggesting a distinct feedback mechanism limiting the production of certain isoforms in response to *O*-GlcNAc elevation. In one study, an internal silencing site has been identified which appears to alter splicing of *OGT* mRNA, limiting its translation (Park et al., 2017). A similar mechanism is at play with OGA, where abundance of the enzyme is regulated by alternative splicing in response to *O*-GlcNAc levels (Tan et al., 2020). Thus, homeostatic systems are in place to limit the deregulation of total *O*-GlcNAc. However, lowering OGA levels can have a dramatic effect on neurodevelopment. We recently reported that mice in which *Oga* was deleted in the brain showed numerous phenotypes including microcephaly, enlarged ventricles, hypotonia, and developmental delay, strongly suggesting a possible link between OGT-XLID variants and perturbations of OGA levels (Olivier-Van Stichelen et al., 2017). In addition, a genome-wide association meta-analysis in 269,867 individuals identified *MGEA5/OGA* as one of the genes closely associated with intelligence (Savage et al., 2018). Studies on several OGT-XLID variants report a decrease in OGA levels (Vaidyanathan et al., 2017; Willems et al., 2017), although others have shown no evidence for a change in OGA levels (Selvan et al., 2018; Pravata et al., 2020a). So clear is this linkage between OGT variants and OGA that lowered OGA levels have been suggested as a diagnostic for XLID (Pravata et al., 2020b).

XLID and Other OGT Functions: HCF-1 Cleavage

O-GlcNAc transferase strongly associates with HCF-1 and heavily modifies it. In combination, OGT-HCF-1 forms interactions with numerous epigenetic complexes (Lazarus et al., 2013; Janetzko et al., 2016; Kapuria et al., 2016, 2018; Levine and Walker, 2016; Leturcq et al., 2017; Vaidyanathan et al., 2017; Gao et al., 2018). In addition, OGT cleaves HCF-1 using a catalytic mechanism which has been recently studied (Lazarus et al., 2013; Bhuiyan et al., 2015; Janetzko et al., 2016; Kapuria et al., 2016, 2018). Some of the OGT variants show changes in HCF-1 cleavage, but this does not seem to be universal to the disorder (Vaidyanathan et al., 2017; Willems et al., 2017; Selvan et al., 2018; Pravata et al., 2020a,b). Mutations in the gene encoding HCF-1 have been found to cause another form of XLID (Koufaris et al., 2016), suggesting a possible overlap in the mechanism of these disorders. Unlike OGT-XLID, this disorder does not include obvious developmental and morphological abnormalities, but HCF-1 dysregulation may still be an important contributor to key features of OGT-XLID.

OGT Interactions With Epigenetic Complexes

O-GlcNAc transferase interacts with numerous protein complexes associated with epigenetic regulation including the Polycomb repressive complexes, the pluripotency network associated with Oct4, Sin3A-HDAC complexes, and many others (Love et al., 2010; Hanover et al., 2012; Lewis and Hanover, 2014; Levine and Walker, 2016; Leturcq et al., 2017; Gao et al., 2018). The mutations seen in OGT-XLID could disrupt subsets of these interactions with accompanying changes in *O*-GlcNAc modification leading to a more pleiotropic deregulation of development. Improper regulation of these epigenetic complexes is known to impact development. For example, mutations in PRC2 proteins EZH2, SUZ12, and EED can cause Weaver Syndrome, an autism spectrum disorder that presents with developmental delay (Imagawa et al., 2017). These overlapping phenotypes may suggest common molecular pathways are at play with OGT-XLID.

At the present, it is difficult to examine these interactions quantitatively, but sensitive methods to examine both protein-protein interactions and *O*-GlcNAc turnover in those complexes are under continuous development.

OGT AND NEURODEVELOPMENTAL DISEASE: SUMMARY AND IMPLICATIONS

The identification of OGT-XLID as a rare human disorder suggests that non-synonymous variants of OGT can be tolerated to a limited degree. These OGT variants are hypomorphic, leading to only modest changes in *O*-GlcNAc levels due to compensatory changes in the *O*-GlcNAcase expression. Patients with OGT-XLID show numerous developmental and neurodevelopmental deficits resulting in a form of intellectual disability. This intellectual disability phenotype results from changes in neurodevelopment which strongly suggests that *O*-GlcNAc addition and removal may play a particularly important role in the functioning and development of the brain. This is perhaps not surprising given the importance of glucose and its metabolites in the physiology of the central nervous system. In addition, the complexity of the human brain originates from a finely tuned developmental process influenced by both genome and environment (Gabriele et al., 2018). A particularly sensitive time of brain development occurs *in utero* when glucose is provided primarily by maternal transfer through the placenta. This underscores the importance of maternal nutritional and environmental state and proper placental glucose metabolism for fetal brain development.

Development of the human cortex is incredibly sensitive, in part due to the requirement for two waves of rapid proliferation of progenitor cells (Florio and Huttner, 2014). During these developmental periods, cells are particularly prone to the accumulation of genetic lesions. Errors in DNA replication and repair can induce single nucleotide variants and insertions-deletions (Ernst, 2016). Transposable elements

are another major source of genome instability which have been linked to neurological disorders (Doyle et al., 2017). Mutations in OGT may derepress transposable elements considering the recent findings of OGT working with TRIM28 to silence retrotransposons in a DNA methylation-dependant manner (Boulard et al., 2020). TRIM28 silencing is active in neural progenitor cells, and heterozygous mice present behavioral phenotypes which suggest an important role in brain development (Whitelaw et al., 2010; Grassi et al., 2019). The roles OGT plays both in activating the DNA damage response and in the silencing of retrotransposons help maintain genome integrity in the developing brain. The contribution of DNA damage and transposable element activity to OGT-XLID phenotypes should be investigated further.

Thus, the functional and morphological features of the human brain render it highly vulnerable to both genetically and environmentally induced alterations. In addition, there is an increasing awareness that chromatin regulation may be central to understanding neurodevelopmental disorders (Gabriele et al., 2018). The insights gained from analysis of both OGT-XLID patients and deregulated placental OGT, are likely to provide a platform for understanding how the *O*-GlcNAc pathway is integrated in human physiology. We have highlighted the role of *O*-GlcNAc cycling in numerous epigenetic complexes regulating development and differentiation. We have also examined the role of *O*-GlcNAc cycling in stem cell differentiation and the regulation of DNA damage response signaling. Finally, we have argued that compensatory mechanisms may be in place to limit the impact of the OGT mutations including X-inactivation of OGT, intron retention of OGT transcripts, OGA down regulation by Polycomb repression, intron retention, and transcription. These highly varied modes of regulation serve to buffer the effects of the OGT mutations, but may themselves have phenotypic consequences.

AUTHOR CONTRIBUTIONS

DK, LA, and JH edited the manuscript. All authors contributed to the literature search, wrote specific sections of the manuscript, and agreed on the final version.

FUNDING

This review was supported by the Intramural NIDDK.

ACKNOWLEDGMENTS

Figures 1–3 and **4B** were created with BioRender.com. **Figure 4A** was created using PyMol.

REFERENCES

Abramowitz, L., Olivier-Van Stichelen, S., and Hanover, J. (2014). Chromosome imbalance as a driver of sex disparity in disease. *J. Genomics* 2, 77–88. doi: 10.7150/jgen.8123

Abramowitz, L. K., Harly, C., Das, A., Bhandoola, A., and Hanover, J. A. (2019). Blocked O-GlcNAc cycling disrupts mouse hematopoeitic stem cell maintenance and early T cell development. *Sci. Rep.* 9:12569. doi: 10.1038/s41598-019-48991-8

Akan, I., Love, D. C., Harwood, K. R., Bond, M. R., and Hanover, J. A. (2016). Drosophila O-GlcNAcase Deletion Globally Perturbs Chromatin O-GlcNAcylation. *J. Biol. Chem.* 291, 9906–9919. doi: 10.1074/jbc.M115.704783

Akan, I., Olivier-Van Stichelen, S., Bond, M. R., and Hanover, J. A. (2018). Nutrient-driven O-GlcNAc in proteostasis and neurodegeneration. *J. Neurochem.* 144, 7–34. doi: 10.1111/jnc.14242

Anthonisen, E. H., Berven, L., Holm, S., Nygard, M., Nebb, H. I., and Gronning-Wang, L. M. (2010). Nuclear receptor liver X receptor is O-GlcNAc-modified in response to glucose. *J. Biol. Chem.* 285, 1607–1615. doi: 10.1074/jbc.M109.082685

Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H. E., Iurlaro, M., Reik, W., et al. (2015). 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.* 11, 555–557. doi: 10.1038/nchembio.1848

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* 6, 1049–1055. doi: 10.1038/nchem.2064

Bale, T. L. (2016). The placenta and neurodevelopment: sex differences in prenatal vulnerability. *Dialogues Clin. Neurosci.* 18, 459–464. doi: 10.31887/dcns.2016.18.4/tbale

Bauer, C., Göbel, K., Nagaraj, N., Colantuoni, C., Wang, M., Müller, U., et al. (2015). Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* 290, 4801–4812. doi: 10.1074/jbc.M114.605881

Bhuiyan, T., Waridel, P., Kapuria, V., Zoete, V., and Herr, W. (2015). Distinct OGT-Binding Sites Promote HCF-1 Cleavage. *PLoS One* 10:e0136636. doi: 10.1371/journal.pone.0136636

Bindesbøll, C., Fan, Q., Nørgaard, R. C., MacPherson, L., Ruan, H. B., Wu, J., et al. (2015). Liver X receptor regulates hepatic nuclear O-GlcNAc signaling and carbohydrate responsive element-binding protein activity. *J. Lipid Res.* 56, 771–785. doi: 10.1194/jlr.M049130

Bond, M. R., and Hanover, J. A. (2013). O-GlcNAc cycling: a link between metabolism and chronic disease. *Annu. Rev. Nutr.* 33, 205–229. doi: 10.1146/annurev-nutr-071812-161240

Bond, M. R., and Hanover, J. A. (2015). A little sugar goes a long way: the cell biology of O-GlcNAc. *J. Cell Biol.* 208, 869–880. doi: 10.1083/jcb.201501101

Bonisch, C., and Hake, S. B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* 40, 10719–10741. doi: 10.1093/nar/gks865

Bouazzi, H., Lesca, G., Trujillo, C., Alwasiyah, M. K., and Munnich, A. (2015). Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. *Clin. Case Rep.* 3, 604–609. doi: 10.1002/ccr3.301

Boulard, M., Rucli, S., Edwards, J. R., and Bestor, T. H. (2020). Methylation-directed glycosylation of chromatin factors represses retrotransposon promoters. *Proc. Natl. Acad. Sci. U.S.A.* 117, 14292–14298. doi: 10.1073/pnas.1912074117

Brockdorff, N. (2013). Noncoding RNA and Polycomb recruitment. *RNA* 19, 429–442. doi: 10.1261/rna.037598.112

Brown, M. S., and Goldstein, J. L. (2008). Selective versus total insulin resistance: a pathogenic paradox. *Cell Metab.* 7, 95–96. doi: 10.1016/j.cmet.2007.12.009

Butkinaree, C., Cheung, W. D., Park, S., Park, K., Barber, M., and Hart, G. W. (2008). Characterization of beta-N-acetylglucosaminidase cleavage by caspase-3 during apoptosis. *J. Biol. Chem.* 283, 23557–23566. doi: 10.1074/jbc.M804116200

Butler, A. A., Sanchez, R. G., Jarome, T. J., Webb, W. M., and Lubin, F. D. (2019). O-GlcNAc and EZH2-mediated epigenetic regulation of gene expression during consolidation of fear memories. *Learn. Mem.* 26, 373–379. doi: 10.1101/lm.049023.118

Celeste, A., Petersen, S., Romanienko, P. J., Fernandez-Capetillo, O., Chen, H. T., Sedelnikova, O. A., et al. (2002). Genomic instability in mice lacking histone H2AX. *Science* 296, 922–927. doi: 10.1126/science.1069398

Chang, Y. H., Weng, C. L., and Lin, K. I. (2020). O-GlcNAcylation and its role in the immune system. *J. Biomed. Sci.* 27:57. doi: 10.1186/s12929-020-00648-9

Charil, A., Laplante, D. P., Vaillancourt, C., and King, S. (2010). Prenatal stress and brain development. *Brain Res. Rev.* 65, 56–79. doi: 10.1016/j.brainresrev.2010.06.002

Chen, Q., Chen, Y., Bian, C., Fujiki, R., and Yu, X. (2013). TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* 493, 561–564. doi: 10.1038/nature11742

Chen, Q., and Yu, X. (2016). OGT restrains the expansion of DNA damage signaling. *Nucleic Acids Res.* 44, 9266–9278. doi: 10.1093/nar/gkw663

Cheng, N. N., Sinclair, D. A., Campbell, R. B., and Brock, H. W. (1994). Interactions of polyhomeotic with Polycomb group genes of Drosophila melanogaster. *Genetics* 138, 1151–1162.

Cheung, W. D., and Hart, G. W. (2008). AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. *J. Biol. Chem.* 283, 13009–13020. doi: 10.1074/jbc.M8012222200

Chou, T. Y., Dang, C. V., and Hart, G. W. (1995). Glycosylation of the c-Myc transactivation domain. *Proc. Natl. Acad. Sci. U.S.A.* 92, 4417–4421. doi: 10.1073/pnas.92.10.4417

Chu, C. S., Lo, P. W., Yeh, Y. H., Hsu, P. H., Peng, S. H., Teng, Y. C., et al. (2014). O-GlcNAcylation regulates EZH2 protein stability and function. *Proc. Natl. Acad. Sci. U.S.A.* 111, 1355–1360. doi: 10.1073/pnas.1323226111

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146, 67–79. doi: 10.1016/j.cell.2011.06.020

Davis, E. P., and Pfaff, D. (2014). Sexually dimorphic responses to early adversity: implications for affective problems and autism spectrum disorder. *Psychoneuroendocrinology* 49, 11–25. doi: 10.1016/j.psyneuen.2014.06.014

Deevy, O., and Bracken, A. P. (2019). PRC2 functions in development and congenital disorders. *Development* 146:dev181354. doi: 10.1242/dev.181354

Dela Justina, V., Dos Passos Junior, R. R., Bressan, A. F., Tostes, R. C., Carneiro, F. S., Soares, T. S., et al. (2018). O-linked N-acetyl-glucosamine deposition in placental proteins varies according to maternal glycemic levels. *Life Sci.* 205, 18–25. doi: 10.1016/j.lfs.2018.05.013

Dentin, R., Hedrick, S., Xie, J., Yates, J. III, and Montminy, M. (2008). Hepatic glucose sensing via the CREB coactivator CRTC2. *Science* 319, 1402–1405. doi: 10.1126/science.1151363

Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Mendez, J., Murphy, N., et al. (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* 32, 645–655. doi: 10.1038/emboj.2012.357

Ding, X., Jiang, W., Zhou, P., Liu, L., Wan, X., Yuan, X., et al. (2015). Mixed Lineage Leukemia 5 (MLL5) Protein Stability Is Cooperatively Regulated by O-GlcNac Transferase (OGT) and Ubiquitin Specific Protease 7 (USP7). *PLoS One* 10:e0145023. doi: 10.1371/journal.pone.0145023

Downs, J. A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., et al. (2004). Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol. Cell* 16, 979–990. doi: 10.1016/j.molcel.2004.12.003

Doyle, G. A., Crist, R. C., Karatas, E. T., Hammond, M. J., Ewing, A. D., Ferraro, T. N., et al. (2017). Analysis of LINE-1 Elements in DNA from Postmortem Brains of Individuals with Schizophrenia. *Neuropsychopharmacology* 42, 2602–2611. doi: 10.1038/npp.2017.115

Draime, A., Bridoux, L., Belpaire, M., Pringels, T., Degand, H., Morsomme, P., et al. (2018). The O-GlcNAc transferase OGT interacts with and post-translationally modifies the transcription factor HOXA1. *FEBS Lett.* 592, 1185–1201. doi: 10.1002/1873-3468.13015

Efimova, E. V., Appelbe, O. K., Ricco, N., Lee, S. S., Liu, Y., Wolfgeher, D. J., et al. (2019). O-GlcNAcylation enhances double-strand break repair, promotes cancer cell proliferation, and prevents therapy-induced senescence in irradiated tumors. *Mol. Cancer Res.* 17, 1338–1350. doi: 10.1158/1541-7786.mcr-18-1025

Ernst, C. (2016). Proliferation and differentiation deficits are a major convergence point for neurodevelopmental disorders. *Trends Neurosci.* 39, 290–299. doi: 10.1016/j.tins.2016.03.001

Etchegaray, J. P., and Mostoslavsky, R. (2016). Interplay between metabolism and epigenetics: a nuclear adaptation to environmental changes. *Mol. Cell* 62, 695–711. doi: 10.1016/j.molcel.2016.05.029

Florio, M., and Huttner, W. B. (2014). Neural progenitors, neurogenesis and the evolution of the neocortex. *Development* 141, 2182–2194. doi: 10.1242/dev.090571

Fong, J. J., Nguyen, B. L., Bridger, R., Medrano, E. E., Wells, L., Pan, S., et al. (2012). beta-N-Acetylglucosamine (O-GlcNAc) is a novel regulator of mitosis-specific phosphorylations on histone H3. *J. Biol. Chem.* 287, 12195–12203. doi: 10.1074/jbc.M111.315804

Forma, E., Jóźwiak, P., Ciesielski, P., Zaczek, A., Starska, K., Bryś, M., et al. (2018). Impact of OGT deregulation on EZH2 target genes FOXA1 and FOXC1 expression in breast cancer cells. *PLoS One* 13:e0198351. doi: 10.1371/journal.pone.0198351

Froberg, J., Yang, L., and Lee, J. (2013). Guided by RNAs: X-inactivation as a model for lncRNA function. *J. Mol. Biol.* 425, 3698–3706. doi: 10.1016/j.jmb.2013.06.031

Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S., et al. (2011). GlcNAcylation of histone H2B facilitates its monoubiquitination. *Nature* 480, 557–560. doi: 10.1038/nature10656

Gabory, A., Roseboom, T. J., Moore, T., Moore, L. G., and Junien, C. (2013). Placental contribution to the origins of sexual dimorphism in health and diseases: sex chromosomes and epigenetics. *Biol. Sex Differ.* 4:5. doi: 10.1186/2042-6410-4-5

Gabriele, M., Lopez Tobon, A., D'Agostino, G., and Testa, G. (2018). The chromatin basis of neurodevelopmental disorders: rethinking dysfunction along the molecular and temporal axes. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 84(Pt B), 306–327. doi: 10.1016/j.pnpbp.2017.12.013

Gagnon, J., Daou, S., Zamorano, N., Iannantuono, N. V., Hammond-Martel, I., Mashtalir, N., et al. (2015). Undetectable histone O-GlcNAcylation in mammalian cells. *Epigenetics* 10, 677–691. doi: 10.1080/15592294.2015.1060387

Gambetta, M. C., and Müller, J. (2014). O-GlcNAcylation prevents aggregation of the Polycomb group repressor polyhomeotic. *Dev. Cell* 31, 629–639. doi: 10.1016/j.devcel.2014.10.020

Gambetta, M. C., and Muller, J. (2015). A critical perspective of the diverse roles of O-GlcNAc transferase in chromatin. *Chromosoma* 124, 429–442. doi: 10.1007/s00412-015-0513-1

Gambetta, M. C., Oktaba, K., and Muller, J. (2009). Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. *Science* 325, 93–96. doi: 10.1126/science.1169727

Gao, J., Yang, Y., Qiu, R., Zhang, K., Teng, X., Liu, R., et al. (2018). Proteomic analysis of the OGT interactome: novel links to epithelial-mesenchymal transition and metastasis of cervical cancer. *Carcinogenesis* 39, 1222–1234. doi: 10.1093/carcin/bgy097

Grassi, D. A., Jönsson, M. E., Brattås, P. L., and Jakobsson, J. (2019). TRIM28 and the control of transposable elements in the brain. *Brain Res.* 1705, 43–47. doi: 10.1016/j.brainres.2018.02.043

Griffith, L. S., and Schmitz, B. (1995). O-linked N-acetylglucosamine is upregulated in Alzheimer brains. *Biochem. Biophys. Res. Commun.* 213, 424–431. doi: 10.1006/bbrc.1995.2149

Guinez, C., Filhoulaud, G., Rayah-Benhamed, F., Marmier, S., Dubuquoy, C., Dentin, R., et al. (2011). O-GlcNAcylation increases ChREBP protein content and transcriptional activity in the liver. *Diabetes* 60, 1399–1413. doi: 10.2337/db10-0452

Hahne, H., Gholami, A. M., and Kuster, B. (2012). Discovery of O-GlcNAc-modified proteins in published large-scale proteome data. *Mol. Cell. Proteomics* 11, 843–850. doi: 10.1074/mcp.M112.019463

Hajduskova, M., Baytek, G., Kolundzic, E., Gosdschan, A., Kazmierczak, M., Ofenbauer, A., et al. (2019). MRG-1/MRG15 Is a Barrier for Germ Cell to Neuron Reprogramming in Caenorhabditis elegans. *Genetics* 211, 121–139. doi: 10.1534/genetics.118.301674

Hanover, J., Chen, W., and Bond, M. (2018). O-GlcNAc in cancer: an Oncometabolism-fueled vicious cycle. *J. Bioenerg. Biomembr.* 50, 155–173. doi: 10.1007/s10863-018-9751-2

Hanover, J., Krause, M., and Love, D. (2012). Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation. *Nat. Rev. Mol. Cell Biol.* 13, 312–321. doi: 10.1038/nrm3334

Hanover, J., Yu, S., Lubas, W., Shin, S., Ragano-Caracciola, M., Kochran, J., et al. (2003). Mitochondrial and nucleocytoplasmic isoforms of O-linked GlcNAc transferase encoded by a single mammalian gene. *Arch. Biochem. Biophys.* 409, 287–297. doi: 10.1016/s0003-9861(02)00578-7

Hayakawa, K., Hirosawa, M., Tabei, Y., Arai, D., Tanaka, S., Murakami, N., et al. (2013). Epigenetic switching by the metabolism-sensing factors in the generation of orexin neurons from mouse embryonic stem cells. *J. Biol. Chem.* 288, 17099–17110. doi: 10.1074/jbc.M113.455899

Hayakawa, K., Hirosawa, M., Tani, R., Yoneda, C., Tanaka, S., and Shiota, K. (2017). H2A O-GlcNAcylation at serine 40 functions genomic protection in association with acetylated H2AZ or gammaH2AX. *Epigenetics Chromatin* 10:51. doi: 10.1186/s13072-017-0157-x

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307. doi: 10.1126/science.1210944

Herzig, S., Long, F., Jhala, U. S., Hedrick, S., Quinn, R., Bauer, A., et al. (2001). CREB regulates hepatic gluconeogenesis through the coactivator PGC-1. *Nature* 413, 179–183. doi: 10.1038/35093131

Hirosawa, M., Hayakawa, K., Yoneda, C., Arai, D., Shiota, H., Suzuki, T., et al. (2016). Novel O-GlcNAcylation on Ser(40) of canonical H2A isoforms specific to viviparity. *Sci. Rep.* 6:31785. doi: 10.1038/srep31785

Housley, M. P., Rodgers, J. T., Udeshi, N. D., Kelly, T. J., Shabanowitz, J., Hunt, D. F., et al. (2008). O-GlcNAc regulates FoxO activation in response to glucose. *J. Biol. Chem.* 283, 16283–16292. doi: 10.1074/jbc.M802240200

Howerton, C. L., and Bale, T. L. (2012). Prenatal programing: at the intersection of maternal stress and immune activation. *Horm. Behav.* 62, 237–242. doi: 10.1016/j.yhbeh.2012.03.007

Howerton, C. L., and Bale, T. L. (2014). Targeted placental deletion of OGT recapitulates the prenatal stress phenotype including hypothalamic mitochondrial dysfunction. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9639–9644. doi: 10.1073/pnas.1401203111

Howerton, C. L., Morgan, C. P., Fischer, D. B., and Bale, T. L. (2013). O-GlcNAc transferase (OGT) as a placental biomarker of maternal stress and reprogramming of CNS gene transcription in development. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5169–5174. doi: 10.1073/pnas.1300065110

Hrit, J., Goodrich, L., Li, C., Wang, B. A., Nie, J., Cui, X., et al. (2018). OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *eLife* 7:e34870. doi: 10.7554/eLife.34870

Hu, C. M., Tien, S. C., Hsieh, P. K., Jeng, Y. M., Chang, M. C., Chang, Y. T., et al. (2019). High Glucose Triggers Nucleotide Imbalance through O-GlcNAcylation of Key Enzymes and Induces KRAS Mutation in Pancreatic Cells. *Cell Metab.* 29, 1334–1349.e10. doi: 10.1016/j.cmet.2019.02.005

Imagawa, E., Higashimoto, K., Sakai, Y., Numakura, C., Okamoto, N., Matsunaga, S., et al. (2017). Mutations in genes encoding polycomb repressive complex 2 subunits cause Weaver syndrome. *Hum. Mutat.* 38, 637–648. doi: 10.1002/humu.23200

Ingham, P. W. (1984). A gene that regulates the bithorax complex differentially in larval and adult cells of Drosophila. *Cell* 37, 815–823. doi: 10.1016/0092-8674(84)90416-1

Ishihara, K., Takahashi, I., Tsuchiya, Y., Hasegawa, M., and Kamemura, K. (2010). Characteristic increase in nucleocytoplasmic protein glycosylation by O-GlcNAc in 3T3-L1 adipocyte differentiation. *Biochem. Biophys. Res. Commun.* 398, 489–494. doi: 10.1016/j.bbrc.2010.06.105

Ito, R., Katsura, S., Shimada, H., Tsuchiya, H., Hada, M., Okumura, T., et al. (2014). TET3-OGT interaction increases the stability and the presence of OGT in chromatin. *Genes Cells* 19, 52–65. doi: 10.1111/gtc.12107

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129–1133. doi: 10.1038/nature09303

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., et al. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333, 1300–1303. doi: 10.1126/science.1210597

Janetzko, J., Trauger, S., Lazarus, M., and Walker, S. (2016). How the glycosyltransferase OGT catalyzes amide bond cleavage. *Nat. Chem. Biol.* 12, 899–901. doi: 10.1038/nchembio.2173

Jang, H., Kim, T. W., Yoon, S., Choi, S. Y., Kang, T. W., Kim, S. Y., et al. (2012). O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* 11, 62–74. doi: 10.1016/j.stem.2012.03.001

Janke, R., Dodson, A. E., and Rine, J. (2015). Metabolism and epigenetics. *Annu. Rev. Cell Dev. Biol.* 31, 473–496. doi: 10.1146/annurev-cellbio-100814-125544

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074–1080. doi: 10.1126/science.1063127

Jiang, M., Xu, B., Li, X., Shang, Y., Chu, Y., Wang, W., et al. (2019). O-GlcNAcylation promotes colorectal cancer metastasis via the miR-101-O-GlcNAc/EZH2 regulatory feedback circuit. *Oncogene* 38, 301–316. doi: 10.1038/s41388-018-0435-5

Jinek, M., Rehwinkel, J., Lazarus, B. D., Izaurralde, E., Hanover, J. A., and Conti, E. (2004). The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat. Struct. Mol. Biol.* 11, 1001–1007. doi: 10.1038/nsmb833

Kapuria, V., Röhrig, U., Bhuiyan, T., Borodkin, V., van Aalten, D., Zoete, V., et al. (2016). Proteolysis of HCF-1 by Ser/Thr glycosylation-incompetent O-GlcNAc transferase:UDP-GlcNAc complexes. *Genes Dev.* 30, 960–972. doi: 10.1101/gad.275925.115

Kapuria, V., Röhrig, U., Waridel, P., Lammers, F., Borodkin, V., van Aalten, D., et al. (2018). The conserved threonine-rich region of the HCF-1PRO repeat activates promiscuous OGT:UDP-GlcNAc glycosylation and proteolysis activities. *J. Biol. Chem.* 293, 17754–17768. doi: 10.1074/jbc.RA118.004185

Katai, E., Pal, J., Poor, V. S., Purewal, R., Miseta, A., and Nagy, T. (2016). Oxidative stress induces transient O-GlcNAc elevation and tau dephosphorylation in SH-SY5Y cells. *J. Cell. Mol. Med.* 20, 2269–2277. doi: 10.1111/jcmm.12910

Keembiyehetty, C., Love, D., Harwood, K., Gavrilova, O., Comly, M., and Hanover, J. (2015). Conditional knock-out reveals a requirement for O-linked N-Acetylglucosaminase (O-GlcNAcase) in metabolic homeostasis. *J. Biol. Chem.* 290, 7097–7113. doi: 10.1074/jbc.M114.617779

Kim, E. J., Amorelli, B., Abdo, M., Thomas, C. J., Love, D. C., Knapp, S., et al. (2007). Distinctive Inhibition of O-GlcNAcase Isoforms by an α-GlcNAc Thiolsulfonate. *J. Am. Chem. Soc.* 129, 14854–14855. doi: 10.1021/ja076038u

Kim, G., Cao, L., Reece, E. A., and Zhao, Z. (2017). Impact of protein O-GlcNAcylation on neural tube malformation in diabetic embryopathy. *Sci. Rep.* 7:11107. doi: 10.1038/s41598-017-11655-6

Klose, R. J., and Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* 31, 89–97. doi: 10.1016/j.tibs.2005.12.008

Koufaris, C., Alexandrou, A., Tanteles, G. A., Anastasiadou, V., and Sismani, C. (2016). A novel HCFC1 variant in male siblings with intellectual disability and microcephaly in the absence of cobalamin disorder. *Biomed. Rep.* 4, 215–218. doi: 10.3892/br.2015.559

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324, 929–930. doi: 10.1126/science.1169786

Kuo, M., Zilberfarb, V., Gangneux, N., Christeff, N., and Issad, T. (2008). O-glycosylation of FoxO1 increases its transcriptional activity towards the glucose 6-phosphatase gene. *FEBS Lett.* 582, 829–834. doi: 10.1016/j.febslet.2008.02.010

Lamarre-Vincent, N., and Hsieh-Wilson, L. C. (2003). Dynamic glycosylation of the transcription factor CREB: a potential role in gene regulation. *J. Am. Chem. Soc.* 125, 6612–6613. doi: 10.1021/ja028200t

Lambert, B., Vandeputte, J., Remacle, S., Bergiers, I., Simonis, N., Twizere, J. C., et al. (2012). Protein interactions of the transcription factor Hoxa1. *BMC Dev. Biol.* 12:29. doi: 10.1186/1471-213x-12-29

Lane, E. A., Choi, D. W., Garcia-Haro, L., Levine, Z. G., Tedoldi, M., Walker, S., et al. (2019). HCF-1 Regulates De Novo Lipogenesis through a Nutrient-Sensitive Complex with ChREBP. *Mol. Cell* 75, 357–371.e7. doi: 10.1016/j.molcel.2019.05.019

Lazarus, M., Jiang, J., Kapuria, V., Bhuiyan, T., Janetzko, J., Zandberg, W., et al. (2013). HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science* 342, 1235–1239. doi: 10.1126/science.1243990

Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011). Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* 469, 564–567. doi: 10.1038/nature09638

Lercher, L., Raj, R., Patel, N. A., Price, J., Mohammed, S., Robinson, C. V., et al. (2015). Generation of a synthetic GlcNAcylated nucleosome reveals regulation of stability by H2A-Thr101 GlcNAcylation. *Nat. Commun.* 6:7978. doi: 10.1038/ncomms8978

Leturcq, M., Lefebvre, T., and Vercoutter-Edouart, A. (2017). O-GlcNAcylation and chromatin remodeling in mammals: an up-to-date overview. *Biochem. Soc. Trans.* 45, 323–338. doi: 10.1042/BST20160388

Levine, Z., Fan, C., Melicher, M., Orman, M., Benjamin, T., and Walker, S. (2018). O-GlcNAc Transferase Recognizes Protein Substrates Using an Asparagine Ladder in the Tetratricopeptide Repeat (TPR) Superhelix. *J. Am. Chem. Soc.* 140, 3510–3513. doi: 10.1021/jacs.7b13546

Levine, Z., and Walker, S. (2016). The Biochemistry of O-GlcNAc Transferase: Which Functions Make It Essential in Mammalian Cells. *Annu. Rev. Biochem.* 85, 631–657. doi: 10.1146/annurev-biochem-060713-035344

Lewis, B., and Hanover, J. (2014). O-GlcNAc and the epigenetic regulation of gene expression. *J. Biol. Chem.* 289, 34440–34448. doi: 10.1074/jbc.R114.595439

Li, X., Molina, H., Huang, H., Zhang, Y. Y., Liu, M., Qian, S. W., et al. (2009). O-linked N-acetylglucosamine modification on CCAAT enhancer-binding protein beta: role during adipocyte differentiation. *J. Biol. Chem.* 284, 19248–19254. doi: 10.1074/jbc.M109.005678

Li, Z., Li, X., Nai, S., Geng, Q., Liao, J., Xu, X., et al. (2017). Checkpoint kinase 1-induced phosphorylation of O-linked beta-N-acetylglucosamine transferase regulates the intermediate filament network during cytokinesis. *J. Biol. Chem.* 292, 19548–19555. doi: 10.1074/jbc.M117.811646

Liu, C., and Li, J. (2018). O-GlcNAc: a sweetheart of the cell cycle and DNA Damage Response. *Front. Endocrinol.* 9:415. doi: 10.3389/fendo.2018.00415

Love, D., Kochan, J., Cathey, R., Shin, S., Hanover, J., and Kochran, J. (2003). Mitochondrial and nucleocytoplasmic targeting of O-linked GlcNAc transferase. *J. Cell Sci.* 116(Pt 4), 647–654. doi: 10.1242/jcs.00246

Love, D., Krause, M., and Hanover, J. (2010). O-GlcNAc cycling: emerging roles in development and epigenetics. *Semin. Cell Dev. Biol.* 21, 646–654. doi: 10.1016/j.semcdb.2010.05.001

Lu, S., Yin, X., Wang, J., Gu, Q., Huang, Q., Jin, N., et al. (2020). SIRT1 regulates O-GlcNAcylation of tau through OGT. *Aging* 12, 7042–7055. doi: 10.18632/aging.103062

Ma, D. K., Ming, G. L., and Song, H. (2009). Oxysterols drive dopaminergic neurogenesis from stem cells. *Cell Stem Cell* 5, 343–344. doi: 10.1016/j.stem.2009.09.001

Ma, J., and Hart, G. W. (2014). O-GlcNAc profiling: from proteins to proteomes. *Clin. Proteomics* 11:8. doi: 10.1186/1559-0275-11-8

Maury, J. J., El Farran, C. A., Ng, D., Loh, Y. H., Bi, X., Bardor, M., et al. (2015). RING1B O-GlcNAcylation regulates gene targeting of polycomb repressive complex 1 in human embryonic stem cells. *Stem Cell Res.* 15, 182–189. doi: 10.1016/j.scr.2015.06.007

Milne, T. A., Sinclair, D. A., and Brock, H. W. (1999). The Additional sex combs gene of Drosophila is required for activation and repression of homeotic loci, and interacts specifically with Polycomb and super sex combs. *Mol. Gen. Genet.* 261, 753–761. doi: 10.1007/s004380050018

Monfort, A., and Wutz, A. (2020). The B-side of Xist. *F1000Res* 9:F1000 Faculty Rev-55. doi: 10.12688/f1000research.21362.1

Myers, S. A., Panning, B., and Burlingame, A. L. (2011). Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9490–9495. doi: 10.1073/pnas.1019289108

Na, H.-J., Akan, I., Abramowitz, L. K., and Hanover, J. A. (2020). Nutrient-Driven O-GlcNAcylation Controls DNA Damage Repair Signaling and Stem/Progenitor Cell Homeostasis. *Cell Rep.* 31:107632. doi: 10.1016/j.celrep.2020.107632

Nishikawa, I., Nakajima, Y., Ito, M., Fukuchi, S., Homma, K., and Nishikawa, K. (2010). Computational prediction of O-linked glycosylation sites that preferentially map on intrinsically disordered regions of extracellular proteins. *Int. J. Mol. Sci.* 11, 4991–5008. doi: 10.3390/ijms11124991

Nugent, B. M., and Bale, T. L. (2015). The omniscient placenta: metabolic and epigenetic regulation of fetal programming. *Front. Neuroendocrinol.* 39:28–37. doi: 10.1016/j.yfrne.2015.09.001

Nugent, B. M., O'Donnell, C. M., Epperson, C. N., and Bale, T. L. (2018). Placental H3K27me3 establishes female resilience to prenatal insults. *Nat. Commun.* 9:2555. doi: 10.1038/s41467-018-04992-1

O'Donnell, N., Zachara, N., Hart, G., and Marth, J. (2004). Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability. *Mol. Cell. Biol.* 24, 1680–1690. doi: 10.1128/mcb.24.4.1680-1690.2004

Olivier-Van Stichelen, S., Abramowitz, L., and Hanover, J. (2014). X marks the spot: does it matter that O-GlcNAc transferase is an X-linked gene. *Biochem. Biophys. Res. Commun.* 453, 201–207. doi: 10.1016/j.bbrc.2014.06.068

Olivier-Van Stichelen, S., and Hanover, J. (2014). X-inactivation normalizes O-GlcNAc transferase levels and generates an O-GlcNAc-depleted Barr body. *Front. Genet.* 5:256. doi: 10.3389/fgene.2014.00256

Olivier-Van Stichelen, S., and Hanover, J. (2015). You are what you eat: O-linked N-acetylglucosamine in disease, development and epigenetics. *Curr. Opin. Clin. Nutr. Metab. Care* 18, 339–345. doi: 10.1097/MCO.0000000000000188

Olivier-Van Stichelen, S., Wang, P., Comly, M., Love, D., and Hanover, J. (2017). Nutrient-driven O-linked N-acetylglucosamine (O-GlcNAc) cycling impacts neurodevelopmental timing and metabolism. *J. Biol. Chem.* 292, 6076–6085. doi: 10.1074/jbc.M116.774042

Park, S., Zhou, X., Pendleton, K., Hunter, O., Kohler, J., O'Donnell, K., et al. (2017). A Conserved Splicing Silencer Dynamically Regulates O-GlcNAc Transferase Intron Retention and O-GlcNAc Homeostasis. *Cell Rep.* 20, 1088–1099. doi: 10.1016/j.celrep.2017.07.017

Parra, M., Zhang, W., Vu, J., DeWitt, M., and Conboy, J. (2020). Antisense targeting of decoy exons can reduce intron retention and increase protein expression in human erythroblasts. *RNA* 26, 996–1005. doi: 10.1261/rna.075028.120

Pick, H., Kilic, S., and Fierz, B. (2014). Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function. *Biochim. Biophys. Acta* 1839, 644–656. doi: 10.1016/j.bbagrm.2014.04.016

Pravata, V., Gundogdu, M., Bartual, S., Ferenbach, A., Stavridis, M., Õunap, K., et al. (2020a). A missense mutation in the catalytic domain of O-GlcNAc transferase links perturbations in protein O-GlcNAcylation to X-linked intellectual disability. *FEBS Lett.* 594, 717–727. doi: 10.1002/1873-3468.13640

Pravata, V., Muha, V., Gundogdu, M., Ferenbach, A., Kakade, P., Vandadi, V., et al. (2019). Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14961–14970. doi: 10.1073/pnas.1900065116

Pravata, V., Omelková, M., Stavridis, M., Desbiens, C., Stephen, H., Lefeber, D., et al. (2020b). An intellectual disability syndrome with single-nucleotide variants in O-GlcNAc transferase. *Eur. J. Hum. Genet.* 28, 706–714. doi: 10.1038/s41431-020-0589-9

Quinonez, S. C., and Innis, J. W. (2014). Human HOX gene disorders. *Mol. Genet. Metab.* 111, 4–15. doi: 10.1016/j.ymgme.2013.10.012

Rahe, D. P., and Hobert, O. (2019). Restriction of Cellular Plasticity of Differentiated Cells Mediated by Chromatin Modifiers, Transcription Factors and Protein Kinases. *G3* 9, 2287–2302. doi: 10.1534/g3.119.400328

Rao, F. V., Schuttelkopf, A. W., Dorfmueller, H. C., Ferenbach, A. T., Navratilova, I., and van Aalten, D. M. (2013). Structure of a bacterial putative acetyltransferase defines the fold of the human O-GlcNAcase C-terminal domain. *Open Biol.* 3:130021. doi: 10.1098/rsob.130021

Rexach, J. E., Clark, P. M., Mason, D. E., Neve, R. L., Peters, E. C., and Hsieh-Wilson, L. C. (2012). Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat. Chem. Biol.* 8, 253–261. doi: 10.1038/nchembio.770

Ronningen, T., Shah, A., Oldenburg, A. R., Vekterud, K., Delbarre, E., Moskaug, J. O., et al. (2015). Prepatterning of differentiation-driven nuclear lamin A/C-associated chromatin domains by GlcNAcylated histone H2B. *Genome Res.* 25, 1825–1835. doi: 10.1101/gr.193748.115

Rowe, L. A., Degtyareva, N., and Doetsch, P. W. (2008). DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. *Free Radic. Biol. Med.* 45, 1167–1177. doi: 10.1016/j.freeradbiomed.2008.07.018

Saeed, M., Ahmad, J., Kanwal, S., Holowatyj, A., Sheikh, I., Zafar Paracha, R., et al. (2016). Formal modeling and analysis of the hexosamine biosynthetic pathway: role of O-linked N-acetylglucosamine transferase in oncogenesis and cancer progression. *PeerJ* 4:e2348. doi: 10.7717/peerj.2348

Sakabe, K., Wang, Z., and Hart, G. W. (2010). Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19915–19920. doi: 10.1073/pnas.1009023107

SanMiguel, J. M., and Bartolomei, M. S. (2018). DNA methylation dynamics of genomic imprinting in mouse development. *Biol. Reprod.* 99, 252–262. doi: 10.1093/biolre/ioy036

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* 50, 912–919. doi: 10.1038/s41588-018-0152-6

Schouppe, D., Ghesquiere, B., Menschaert, G., De Vos, W. H., Bourque, S., Trooskens, G., et al. (2011). Interaction of the tobacco lectin with histone proteins. *Plant Physiol.* 155, 1091–1102. doi: 10.1104/pp.110.170134

Schvartzman, J. M., Thompson, C. B., and Finley, L. W. S. (2018). Metabolic regulation of chromatin modifications and gene expression. *J. Cell Biol.* 217, 2247–2259. doi: 10.1083/jcb.201803061

Selvan, N., George, S., Serajee, F., Shaw, M., Hobson, L., Kalscheuer, V., et al. (2018). *O*-GlcNAc transferase missense mutations linked to X-linked intellectual disability deregulate genes involved in cell fate determination and signaling. *J. Biol. Chem.* 293, 10810–10824. doi: 10.1074/jbc.RA118.002583

Shafi, R., Iyer, S., Ellies, L., O'Donnell, N., Marek, K., Chui, D., et al. (2000). The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5735–5739. doi: 10.1073/pnas.100471497

Shi, F. T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M. A., et al. (2013a). Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J. Biol. Chem.* 288, 20776–20784. doi: 10.1074/jbc.M113.460386

Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013b). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159–166. doi: 10.1016/j.canlet.2013.06.013

Sinclair, D. A., Syrzycka, M., Macauley, M. S., Rastgardani, T., Komljenovic, I., Vocadlo, D. J., et al. (2009). Drosophila O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc). *Proc. Natl. Acad. Sci. U.S.A.* 106, 13427–13432. doi: 10.1073/pnas.0904638106

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935. doi: 10.1126/science.1170116

Tan, Z., Fei, G., Paulo, J., Bellaousov, S., Martin, S., Duveau, D., et al. (2020). O-GlcNAc regulates gene expression by controlling detained intron splicing. *Nucleic Acids Res.* 48, 5656–5669. doi: 10.1093/nar/gkaa263

Toleman, C. A., Paterson, A. J., and Kudlow, J. E. (2006). The Histone Acetyltransferase NCOAT Contains a Zinc Finger-like Motif Involved in Substrate Recognition. *J. Biol. Chem.* 281, 3918–3925. doi: 10.1074/jbc.m510485200

Torres, C. R., and Hart, G. W. (1984). Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem.* 259, 3308–3317.

Tursun, B., Patel, T., Kratsios, P., and Hobert, O. (2011). Direct conversion of C. elegans germ cells into specific neuron types. *Science* 331, 304–308. doi: 10.1126/science.1199082

Vaidyanathan, K., Niranjan, T., Selvan, N., Teo, C., May, M., Patel, S., et al. (2017). Identification and characterization of a missense mutation in the O-linked β-N-acetylglucosamine (O-GlcNAc) transferase gene that segregates with X-linked intellectual disability. *J. Biol. Chem.* 292, 8948–8963. doi: 10.1074/jbc.M116.771030

Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A., et al. (2013). Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol. Cell* 49, 645–656. doi: 10.1016/j.molcel.2012.12.019

Wahl, G. M., and Carr, A. M. (2001). The evolution of diverse biological responses to DNA damage: insights from yeast and p53. *Nat. Cell Biol.* 3, E277–E286. doi: 10.1038/ncb1201-e277

Whitelaw, N. C., Chong, S., Morgan, D. K., Nestor, C., Bruxner, T. J., Ashe, A., et al. (2010). Reduced levels of two modifiers of epigenetic gene silencing, Dnmt3a and Trim28, cause increased phenotypic noise. *Genome Biol.* 11:R111. doi: 10.1186/gb-2010-11-11-r111

Willems, A., Gundogdu, M., Kempers, M., Giltay, J., Pfundt, R., Elferink, M., et al. (2017). Mutations in N-acetylglucosamine (O-GlcNAc) transferase in patients with X-linked intellectual disability. *J. Biol. Chem.* 292, 12621–12631. doi: 10.1074/jbc.M117.790097

Wong, C. C., Qian, Y., and Yu, J. (2017). Interplay between epigenetics and metabolism in oncogenesis: mechanisms and therapeutic approaches. *Oncogene* 36, 3359–3374. doi: 10.1038/onc.2016.485

Xie, S., Jin, N., Gu, J., Shi, J., Sun, J., Chu, D., et al. (2016). O-GlcNAcylation of protein kinase A catalytic subunits enhances its activity: a mechanism linked to learning and memory deficits in Alzheimer's disease. *Aging Cell* 15, 455–464. doi: 10.1111/acel.12449

Yang, X., and Qian, K. (2017). Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* 18, 452–465. doi: 10.1038/nrm.2017.22

Yang, X., Zhang, F., and Kudlow, J. E. (2002). Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell* 110, 69–80. doi: 10.1016/s0092-8674(02)00810-3

Zhang, Q., Liu, X., Gao, W., Li, P., Hou, J., Li, J., et al. (2014). Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked beta-N-acetylglucosamine transferase (OGT). *J. Biol. Chem.* 289, 5986–5996. doi: 10.1074/jbc.M113.524140

Zhang, S., Roche, K., Nasheuer, H. P., and Lowndes, N. F. (2011). Modification of histones by sugar beta-N-acetylglucosamine (GlcNAc) occurs on multiple residues, including histone H3 serine 10, and is cell cycle-regulated. *J. Biol. Chem.* 286, 37483–37495. doi: 10.1074/jbc.M111.284885

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 08 January 2021
doi: 10.3389/fgene.2020.613098

Impaired Regulation of Histone Methylation and Acetylation Underlies Specific Neurodevelopmental Disorders

Merrick S. Fallah[1,2†], *Dora Szarics*[1,2†], *Clara M. Robson*[1,2] *and James H. Eubanks*[1,2,3,4,5*]

[1] Division of Experimental and Translational Neuroscience, Krembil Research Institute, University Health Network, Toronto, ON, Canada, [2] Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada, [3] Department of Surgery (Neurosurgery), University of Toronto, Toronto, ON, Canada, [4] Institute of Medical Science, University of Toronto, Toronto, ON, Canada, [5] Department of Physiology, University of Toronto, Toronto, ON, Canada

Epigenetic processes are critical for governing the complex spatiotemporal patterns of gene expression in neurodevelopment. One such mechanism is the dynamic network of post-translational histone modifications that facilitate recruitment of transcription factors or even directly alter chromatin structure to modulate gene expression. This is a tightly regulated system, and mutations affecting the function of a single histone-modifying enzyme can shift the normal epigenetic balance and cause detrimental developmental consequences. In this review, we will examine select neurodevelopmental conditions that arise from mutations in genes encoding enzymes that regulate histone methylation and acetylation. The methylation-related conditions discussed include Wiedemann-Steiner, Kabuki, and Sotos syndromes, and the acetylation-related conditions include Rubinstein-Taybi, KAT6A, genitopatellar/Say-Barber-Biesecker-Young-Simpson, and brachydactyly mental retardation syndromes. In particular, we will discuss the clinical/phenotypic and genetic basis of these conditions and the model systems that have been developed to better elucidate cellular and systemic pathological mechanisms.

Keywords: epigenetics, histone, acetylation, methylation, neurodevelopment

OPEN ACCESS

Edited by:
Dag H. Yasui,
University of California, Davis, United States

Reviewed by:
Abhijit Shukla,
Cornell University, United States
Nathalie Berube,
Western University, Canada

***Correspondence:**
James H. Eubanks
jeubanks@uhnres.utoronto.ca

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *01 October 2020*
Accepted: *09 December 2020*
Published: *08 January 2021*

Citation:
Fallah MS, Szarics D, Robson CM and Eubanks JH (2021) Impaired Regulation of Histone Methylation and Acetylation Underlies Specific Neurodevelopmental Disorders. Front. Genet. 11:613098. doi: 10.3389/fgene.2020.613098

INTRODUCTION

The "neurodevelopmental disorders" represent a group of conditions in which altered brain development leads to cognitive, neurological, and/or psychiatric impairments in children (Thapar et al., 2017). While multiple causes exist, this broad group of disorders houses a number of genetic conditions, in which spontaneous or inherited genetic variations are the specific cause for the neurodevelopmental phenotypes (Niemi et al., 2018). Of these, there are mutations, including histone modifiers, that affect epigenetic-related cell machinery (Millan, 2013).

An early definition of epigenetics by Russo et al. states that it is "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Russo et al., 1996). This broad definition was later refined as "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). These modifications can be inherited or may develop during the lifetime (Morgan et al., 1999; Daxinger and Whitelaw, 2012). Epigenetic modifications are thought to involve multiple processes,

including an epigenator signal, an initiator, and a maintainer (Berger et al., 2009). Briefly, epigenetic modifications often begin when an extracellular signal, or epigenator, interacts with a host cell. This transient event (for example, a post-translational protein modification) activates the epigenetic initiator. This initiator (for example, a DNA binding protein or non-coding RNA) in turn facilitates a location specific modification of chromatin structure. The epigenetic maintainer (for example, histone modifiers or DNA methylators), which may have been recruited by the initiator, then stabilizes the epigenetic signal to maintain the chromatin modification (Berger et al., 2009).

Histone modifications are one example of epigenetic alteration, and represent a key mechanism through which local gene expression is regulated (Cedar and Bergman, 2009). Histones play a major role in determining chromatin structure. The vast majority of genomic DNA is found associated with nucleosomes; octameric groups of histone proteins comprised largely of H2A, H2B, H3, and H4 proteins. Covalent modifications of these core histone proteins have been well-demonstrated to impact several aspects of nuclear function, such as transcription and DNA repair (Karlic et al., 2010). There are several known types of modifications that can occur, such as methylation, acetylation, phosphorylation, and ubiquitination; although not all have been found to be causal for neurological disease at this time. The modification of histones through methylation and acetylation, and the neurodevelopmental consequences of their aberrant activity in specific disorders, will be the focus of this review. While a number of genetic alterations affecting distinct encoded products have been implicated in different neurological conditions (**Table 1**), this review will focus on a select group of neurodevelopmental conditions linked to altered histone methylation or acetylation processes.

HISTONE METHYLATION

Histone proteins can be methylated on arginine or lysine residues (Bannister et al., 2002; Bannister and Kouzarides, 2005). These modifications likely allow for the binding of specific regulatory proteins, which affect chromatin structure. This code is complex: Arginine residues may be mono- or di- methylated, while lysine residues may be mono-, di-, or tri-methylated (Bannister et al., 2002; Bannister and Kouzarides, 2005). This patterning allows for a vast array of different methylation state signatures (Bannister et al., 2002; Bannister and Kouzarides, 2005).

Arginine methylation typically occurs at amino acid residues 2, 8, 17, and 26 of H3, and residue 3 of H4 proteins (Klose and Zhang, 2007). Members of the family of arginine methyltransferase (PRMT1) enzymes carry out these modifications; their effects on chromatin structure can lead to either transcriptional activating or repressing consequences depending on context and code (Chen, 1999; Wang, 2001; Klose and Zhang, 2007). While histone methylation at arginine residues is recognized, neurodevelopmental conditions have not been identified to date that arise from systems involved in normal histone arginine modification. Rather, different neurodevelopmental conditions have been identified that stem from systems affecting histone lysine modifications.

Lysine methylation is carried out by the enzymes belonging to the disruptor of telomeric silencing 1-like (DOT1L) family, or to the family of SET domain containing proteins (Martin and Zhang, 2005; Klose and Zhang, 2007). Lysine methylation occurring at residues 4, 36, and 79 of histone H3 are generally associated with active chromatin, while methylation at residues 9 and 27 of histone H3, and/or lysine residue 20 of histone H4, are typically associated with inactive or repressed chromatin regions (Klose and Zhang, 2007). Methylation of histone proteins is not sufficient to alter their charge, and therefore, likely acts to recruit other effector proteins to specific chromatin sites rather than directly disrupting the contact between the histone complex and its associated DNA (**Figure 1**) (Klose and Zhang, 2007).

Histone lysine methylation is a dynamic process, and families of demethylase enzymes allow for the regulation of different histone methylation states (Black et al., 2012). Originally reported by Shi and colleagues, lysine specific demethylase 1 (LSD1) was the first histone demethylase described (Shi et al., 2003, 2004). LSD1 functions as a transcriptional corepressor by catalyzing oxidative demethylation of mono- and di-methylated lysine residues at positions 4 and 9 of H3 proteins (Shi et al., 2004; Metzger et al., 2005). LSD1 demethylase activity can also be influenced by other protein complexes, including the restin corepressor (CoREST) complex (Lee et al., 2005; Shi et al., 2005). Subsequently, the Jumonji C (JmjC)-domain-containing histone demethylase (JHDM) family of demethylase enzymes was identified, which contains the largest number of histone demethylases (Klose et al., 2006). These enzymes require a different set of cofactors [i.e., alpha-ketoglutarate, iron (Fe II)] (Tsukada et al., 2006). Unlike LSD1, these are able to modify all three histone methylation states (Klose et al., 2006). JHDMs have shown activity at a number of histone lysine residues, including H3K9 (Yamane et al., 2006) and H3K36 (Tsukada et al., 2006).

This dynamic process of methylating/demethylating histones can be disrupted by mutations in genes that encode the necessary catalytic proteins to produce these alterations. Mutations in histone methyltransferase and demethylase genes have been shown to be causal for several neurodevelopmental disorders, including, but not limited to, Wiedemann-Steiner, Kabuki, and Sotos syndromes (Kim et al., 2017).

Conditions Associated With Impaired Histone Methyltransferase Function

Wiedemann-Steiner Syndrome

Wiedemann-Steiner syndrome (WDSTS; OMIM #605130) is a rare congenital malformation and neurodevelopmental disorder first described by Wiedemann et al. in 1989 and later by Steiner and Marques in 2000 (Wiedemann et al., 1989; Steiner and Marques, 2000). WDSTS is characterized by intellectual disability, language and motor delays, hypertrichosis cubiti, delayed bone age, and distinct craniofacial features (Koenig et al., 2010; Jones et al., 2012; Miyake et al., 2016; Li et al., 2018).

Through whole-exome-sequencing, Jones et al. determined that heterozygous mutations in the *KMT2A* gene (OMIM

TABLE 1 | Developmental Disorders Caused by Mutations in Histone Modifying Enzymes.

Gene	OMIM #	Condition	Histone modification
Histone methyltransferases			
KMT2A	605130	Wiedemann-Steiner syndrome	Mutant mice display normal global histone methylation; but decreased H4K5, 8, 12, 16 acetylation
KMT2D	147920	Kabuki syndrome 1	Mutant mice display decreased trimethylation of H3K4 in dentate granule cell layer
NSD1	117550	Sotos syndrome	Global reduction of H3K36 di-methylation in male germ cells
NSD2	194190	Wolf-Hirschhorn syndrome*	Reduced H3K36 methylation in heterozygous and homozygous KO mouse ESCs and embryonic zebrafish tissues
EHMT1	610253	Kleefstra syndrome*	In mouse hippocampus, cortex, cerebellum, and olfactory bulb increased H3K9 tri-methylation
EZH2	277590	Weaver syndrome	Mutant embryonic mouse tissues display decreased di- and tri-methylation of H3K27
Histone Demethylases			
KDM6A	300867	Kabuki syndrome 2	Mutant mouse neural crest cells display increased methylation of H3K27
KDM5C	300534	X-linked intellectual disability, Claes-Jensen type	No global changes in H3K4 methylation from *Kdm5c*-KO mouse neurons
PHF8	300263	X-linked intellectual disability, Siderius type	Not Determined
Histone Acetyltransferases			
CBP	180849	Rubinstein-Taybi syndrome (Type 1)	Hypoacetylated at H2A K5 (patient lymphoblastoid cells), H2B K5, K12, K15, K20 (mice and patient lymphoblastoid cells), H3 K14, K27 (CaMKIIα-Cre mice), and H4 K8 (CaMKIIα-Cre mice)
p300	613684	Rubinstein-Taybi syndrome (Type 2)	No observed alterations
KAT6A	616268	KAT6A syndrome	Hypoacetylated at H3K9
KAT6B	606170	Genitopatellar syndrome	Not determined
	603736	Say-Barber-Biesecker-Young-Simpson syndrome	Not determined
Histone Deacetylases			
HDAC4	600430	Brachydactyly mental retardation syndrome	No observed alterations

**indicates disorders in which multiple genes have been implicated. References: (Nimura et al., 2009; Gervasini et al., 2013; Lopez-Atalaya et al., 2014; Iwase et al., 2016; Shpargel et al., 2017; Yu et al., 2017; Iacono et al., 2018; Lui et al., 2018; Shirane et al., 2020).*

#159555) were causal for WDSTS in five patients (Jones et al., 2012). The *KMT2A* gene resides at chromosome 11q23 and encodes the KMT2A histone lysine methyltransferase (Jones et al., 2012). KMT2A is a SET domain-containing enzyme that catalyzes mono-, di-, and tri-methylation of H3K4. Known targets regulated by KMT2A include genes encoding several Hox and Wnt factors (Yu et al., 1995; Milne et al., 2002; Cosgrove and Patel, 2010; Jones et al., 2012). Methylated H3K4 is typically found at enhancers and promoters of genes being actively transcribed, with both methylation status (mono-, di-, or tri-) and density correlating with the level of transcriptional activity (Heintzman et al., 2007, 2009; Kim et al., 2017).

Several mutations of *KMT2A* have been reported, which include single nucleotide missense mutations, non-sense mutations, insertion/deletion alterations that cause frame-shift mutations, and mutations affecting splice site sequences (Li et al., 2018). Most of the known mutations alter the reading frame and introduce termination codons in a non-terminal exon. Such inappropriate termination codons are likely to activate non-sense-mediated mRNA decay pathways (Jones et al., 2012), and if indeed non-sense mediated decay does target these non-sense mutation transcripts, the resulting condition would likely arise due to haploinsufficiency (Yu et al., 1995; Jones et al., 2012; Li et al., 2018). The mechanism associated with missense mutations of *KMT2A* remain unclear, however, and both loss of function and dominant negative mechanisms have been hypothesized to cause WDSTS (Stellacci et al., 2016; Lebrun et al., 2018; Li et al., 2018).

A mouse model harboring a lacZ insertion into exon 3 of *Kmt2a* has been generated for preclinical assessments of WDSTS (Yu et al., 1995; Kim et al., 2007; Gupta et al., 2010). *Kmt2a* null mutations are embryonic lethal in mice (Yu et al., 1995), although heterozygous mice are viable and display phenotypes that phenocopy clinically-relevant issues often seen in patients. *Kmt2a* heterozygous mice were found to have skeletal malformations, reductions in growth, and haematopoietic abnormalities, including anemia and thrombocytopenia (Yu et al., 1995). These mice also showed substantial deficits in long-term contextual fear learning (Gupta et al., 2010). Targeted ablation of *Kmt2a* in the postnatal forebrain and adult prefrontal cortex (floxed exon 3 and 4) has been reported to produce profound cognitive deficits, increase anxiety-like behaviors,

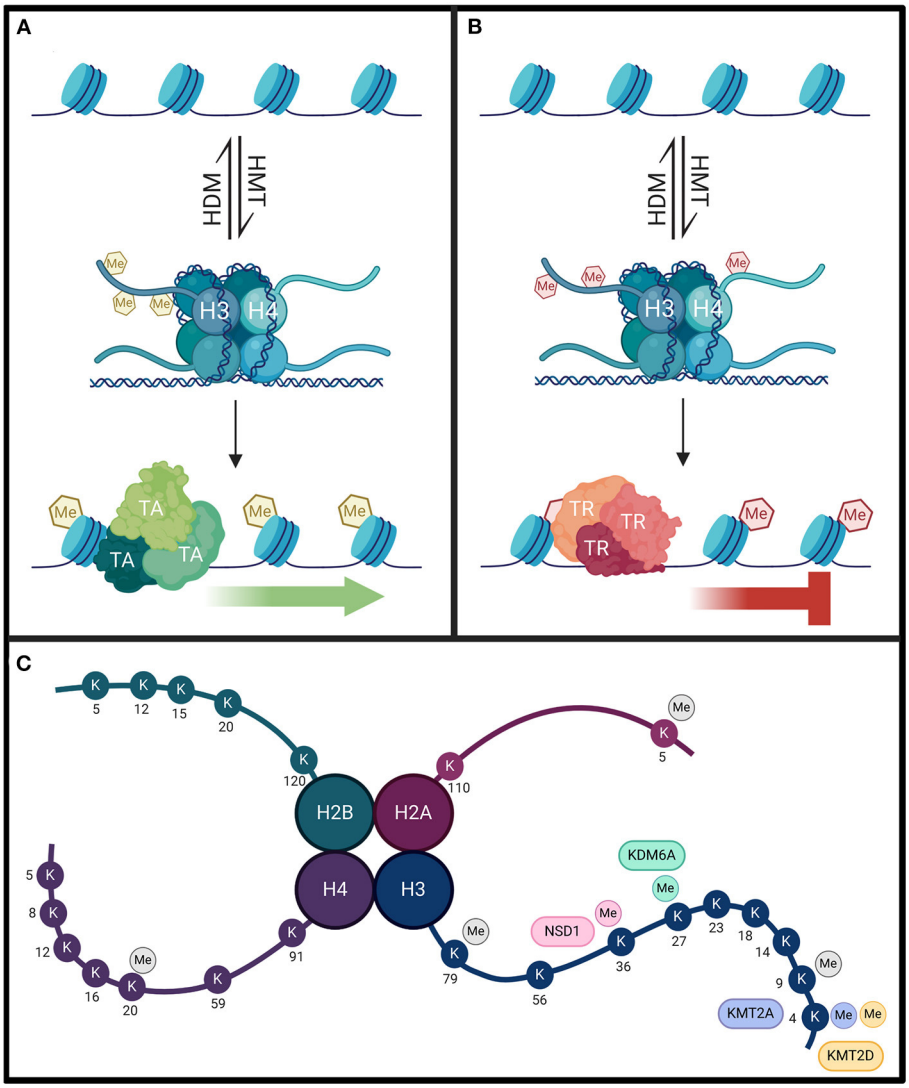

FIGURE 1 | Histone methylation regulates gene expression through recruitment of different transcription factors rather than by directly altering chromatin structure. **(A)** When histone methyltransferases (HMT) methylate lysine or arginine residues on H3 and H4 associated with active chromatin (yellow hexagons), transcriptional activators (TA) can be recruited to those sites to promote gene expression. **(B)** However, when HMTs target different residues on H3 and H4 that are associated with chromatin repression, these methylated regions (red hexagons) can recruit transcriptional repressors to silence the gene. The transcriptional effects of histone methylation can be reversed by histone demethylases (HDM). Proper gene expression is dependent on the homeostasis of HMT and HDM enzyme activity. **(C)** Lysine methylation occurs primarily on H3 and H4. Each KMT/KDM will only target select residues, and the enzymes discussed have high specificity for these targets. Mutations that impair the function of these enzymes would be expected to affect the methylation status of these residues. Each enzyme and their lysine methylation sites are color coded. Other lysine methylation sites not specifically targeted by these enzymes are indicated in gray. Arginine methylation sites are not indicated in this figure.

deficits in working memory, and impaired synaptic plasticity (Jakovcevski et al., 2015). Perhaps surprisingly, neither global histone methylation nor H3K4 methylation were found to be altered in brain of *Kmt2a* heterozygous mice at either postnatal day 0 or in adulthood (Kim et al., 2007). However, these mice do display decreases in histone acetylation at H4K5, 8, 12, and 16 (Kim et al., 2007). As Kmt2a normally interacts with HDAC1/2 complexes (van der Vlag and Otte, 1999; Xia et al., 2003), the altered H4K5 acetylation could stem from the lack of proper Kmt2a-directed HDAC1/2 activity. In addition, the expression of the homeodomain transcription factor Meis2 is reduced in *Kmt2a*-ablated neurons. This alteration may be linked to pathogenesis, as the selective knockdown of *Meis2* in the prefrontal cortex of mice using RNA interference produced similar deficits in working memory as those seen in *Kmt2a* mutants (Jakovcevski et al., 2015). While the precise role of Meis2 in WDSTS pathogenesis remains unclear, restoring Meis2 function has been proposed as a potential strategy for translational investigation of WDSTS (Kim et al., 2017).

Kabuki Syndrome

Kabuki syndrome (KS; OMIM #147920) is an intellectual disability and multiple congenital malformation disorder first described by Niikawa et al. (1981) and Kuroki et al. (1981). An international consensus panel has outlined the diagnostic criteria for KS to include an individual with infantile hypotonia, developmental delay or intellectual disability, and one or both of the following: (1) a pathogenic, or likely pathogenic mutation of *KMT2D* or *KDM6*A; (2) dysmorphic features associated with KS (for full details see Adam et al., 2019) (Adam et al., 2019). KS has an estimated incidence rate of ∼1 in 32,000 in Japan (Niikawa et al., 1988).

Heterozygous mutations in *KMT2D* (OMIM #602113) have been identified as causal for KS through whole-exome and sanger sequencing (Ng et al., 2010). KMT2D is a SET domain containing histone lysine methyltransferase first cloned in 1997, whose encoding gene resides on chromosome 12q13.12 (Prasad et al., 1997). KMT2D is a H3K4 tri-methyltransferase, whose activities are typically associated with enhancing local gene expression and are critical for proper cell differentiation (Lee et al., 2013; Van Laarhoven et al., 2015). A vast number of KS-associated mutations in *KMT2D* have been identified; the majority being non-sense and frameshift mutations (Ng et al., 2010; Hannibal et al., 2011; Li et al., 2011; Miyake et al., 2013a; Micale et al., 2014). These mutations were predicted to produce truncated proteins that do not include the SET domain necessary for the protein's methylation abilities (Ng et al., 2010; Hannibal et al., 2011; Miyake et al., 2013a), or may activate non-sense-mediated mRNA decay pathways (as discussed previously). Either mechanism would support the hypothesis of KMT2D haploinsufficiency as a causal mechanism for KS (Hannibal et al., 2011).

Knockdown of *Kmt2d* in zebrafish using morpholino antisense oligonucleotides resulted in substantial craniofacial defects and viscerocranial hypoplasia (Van Laarhoven et al., 2015). The effects of this knockdown on neurodevelopment were assessed in the embryonic stage, and brain cross-sections showed global volume reductions compared to wildtype. Cell layer thickness was found to be reduced in the optic tectum, midbrain tegmentum, hypothalamus, and, to a lesser extent, the medulla oblongata and hindbrain (Van Laarhoven et al., 2015). The presence of morphological abnormalities, such as elongated nuclei, in cells found in the central forebrain and midbrain regions has also been reported. These areas were also shown to contain much larger populations of cells expressing sox2 (a marker of neural precursor cells) and very diminished number of cells expressing the neuronal post-mitotic marker huc. Together, these observations suggest the differentiation of neural precursor cells is inhibited by the absence of Kmt2d (Van Laarhoven et al., 2015). In mice, the complete knockout of *Kmt2d* was found to be embryonically lethal (Lee et al., 2013). However, a mouse model has been generated in which the SET domain of *Kmt2d* has been substituted in frame with a beta-Geo cassette including its own termination codon, 3' untranslated region, and poly-adenylation signal. This mouse expresses a truncated protein that lacks methyltransferase activity but should retain its more proximal amino terminal domains (Bjornsson et al., 2014). Mice homozygous for this mutation become non-viable during embryonic development by ED12. Mice heterozygous for this mutation show a reduction in neurogenesis of the granule cell layer of the dentate gyrus and a decrease of hippocampal dentate gyrus volume. Consistent with this observation, these mice display deficits in the Morris water maze, contextual fear learning, and novel object recognition tests, suggesting impairments in hippocampal memory (Bjornsson et al., 2014). At the cellular level, these phenotypic alterations correlated with significant decreases in H3K4 trimethylation levels in the hippocampal dentate granule cells of these mice (Bjornsson et al., 2014). To date, H3K4 methylation status in patient-derived material or cells has yet to be investigated.

A less frequent cause of KS, seen in <5% of patients (Banka et al., 2015), stems from mutations in the *KDM6A* (also known as *UTX*) gene (OMIM #300128) (Lederer et al., 2012; Van Laarhoven et al., 2015). The *KDM6A* gene resides at chromosome Xp11.3 and encodes a histone-demethylase that targets mono-, di-, and tri- methylated H3K27 (Hong et al., 2007; Lan et al., 2007; Lederer et al., 2012). Interestingly, *Kdm6a* has been shown to escape X-inactivation (Greenfield et al., 1998), and display some sex dependent differences in magnitude of expression in mice (Xu et al., 2008). The activity of Kdm6a is primarily associated with gene silencing (Hübner and Spector, 2010; Margueron and Reinberg, 2011). Its regulation of methylation state has been shown to influence gene expression and developmental processes, such as transitions in cell lineage (Miller et al., 2010; Wang et al., 2010), and is critical for neural tube development in mice (Shpargel et al., 2012). Several *KDM6A* mutations have been identified in KS patients, which include deletions, frameshifts, mutations affecting splice site junctions, and non-sense mutations (Lederer et al., 2012, 2014; Miyake et al., 2013b; Banka et al., 2015; Van Laarhoven et al., 2015). Each of these mutations would likely generate a non-functional product or promote non-sense-mediated decay, suggesting pathogenesis likely stems from KDM6A haploinsufficiency (Lindgren et al., 2013).

Similar to what was observed in *Kmt2d* models, morpholino knockdown of *Kdm6a* resulted in significant craniofacial abnormalities in zebrafish (Lindgren et al., 2013; Van Laarhoven et al., 2015), as well as reductions in global brain volume and cell layer thickness of the optic tectum, hypothalamus, midbrain tegmentum, and medulla oblongata (Van Laarhoven et al., 2015). The co-administration of wild-type *hKDM6A* mRNA was found to partially reverse these cell layer reductions in the morpholino knockdown mutants (Van Laarhoven et al., 2015). In normal embryonic development in mice, the expression of *Kdm6a* is high in the ventricular zone of the caudal neural tube, the anterior region of the neural tube, in neural crest cells, and in somites at E8.5, and in the cortex at E11.5 (Lee et al., 2012). Knockout of *Kdm6a* in embryonic female mice is lethal at mid-gestational periods, with severe defects in early developmental patterning being observed (including failure of neural tube closure). However, male embryos expressing the same mutation develop to term, but are underweight at birth with only 25% surviving into adulthood. Females heterozygous for this mutation were found to be viable and fertile (Lee et al., 2012; Shpargel et al., 2012). Consistent with expectations that would

stem from impaired Kdm6a function, increased methylation at H3K27 has been reported in neural crest cells isolated from these mutant mice (Shpargel et al., 2017).

Sotos Syndrome

Sotos syndrome (OMIM #117550) is an autosomal dominant, neurologic disorder first described by Sotos et al. (1964). Sotos syndrome is characterized by excessive pre- and post-natal growth, advanced bone age, distinct facial features, macrocephaly, neurodevelopmental and intellectual delay, and in some instances seizures (Sotos et al., 1964; Kurotaki et al., 2002). In light of the discovery of causal mutations associated with Sotos syndrome, the Childhood Overgrowth Collaboration Consortium found an overgrowth in occipitofrontal head circumference, facial dysmorphism, and learning disability were defining characteristics, and an array of associated features, such as advanced bone age, macrocephaly, hypotonia, seizures, scoliosis, cardiac defects, and neonatal jaundice (Tatton-Brown et al., 2005). Neuroimaging studies have identified common ventricular abnormalities, including prominence of the trigone and occipital horns, and ventriculomegaly. Other imaging findings include enlarged supratentorial and posterior fossa extracerebral fluid space, anomalies of the corpus callosum and midline structures, and in a small percentage of cases gray matter heterotopias (Schaefer et al., 1997). The prevalence of Sotos syndrome is currently unclear and under investigation. Treatment of Sotos syndrome currently involves symptomatic management (Baujat and Cormier-Daire, 2007).

Mutations in the nuclear receptor-binding SET domain protein 1 (*NSD1*; OMIM #606681) were found to be causal for Sotos syndrome (Kurotaki et al., 2002) and mutations of *NSD1* may account for more than 75% of the reported cases (Baujat and Cormier-Daire, 2007). Different types of *NSD1* mutations have been reported, which include non-allelic homologous recombination, large and small scale deletions, frameshifts, as well as missense and non-sense mutations (Kurotaki et al., 2002; Douglas et al., 2003; Türkmen et al., 2003; van Haelst et al., 2005; Visser et al., 2005; Kaminsky et al., 2011). The *NSD1* gene is located at chromosome 5q35 (Jaju et al., 2001) and encodes a SET domain containing histone lysine methyltransferase (Huang et al., 1998). NSD1 is a highly specific mono- or di-methyltransferase of H3K36 (Li et al., 2009; Lucio-Eterovic et al., 2010), and has been shown to associate with the promoter regions of a number of genes across the genome to regulate their regional methylation signatures (Lucio-Eterovic et al., 2010). The majority of mutations identified are predicted to disrupt the reading frame in a way that causes early translational termination and/or activates non-sense-mediated decay (Tatton-Brown et al., 2005), suggesting haploinsufficiency is the likely cause of pathogenesis (Kurotaki et al., 2002). Mutations known to be causal for Sotos syndrome that localized to the plant homeodomain (PHD) regions of the encoded protein diminished NSD1 binding to specific methylated sites (H3K4 and H3K9), and abrogated cofactor recruitment (e.g., Nizp1), which collectively lead to impairments in normal transcriptional regulation (Pasillas et al., 2011). NSD1 has also been found to interact with RNA polymerase II, possibly through NSD1-methylation dependent recruitment to promoter regions. The loss of NSD1 also leads to reductions in methylated H3K36 and gene expression *in vitro* (Lucio-Eterovic et al., 2010).

The *NSD1* gene is highly conserved between human and mouse (83% amino acid identity) (Kurotaki et al., 2001) and mouse models have been developed for targeted knockdown using Cre-loxP recombination (Rayasam et al., 2003). Complete *Nsd1* knockout in mice is embryonic lethal (by E10.5), with the *Nsd1* knockout embryos displaying significant alterations in endoderm, mesoderm, and neurectoderm pattern formation at E8.0, and a significant increase in terminal deoxynuleotidyltransferase-mediated dUTP-biotin nick-end labeling (TUNEL)-positive cell labeling to suggest increased levels of apoptosis (Rayasam et al., 2003). Unlike the consequences of heterozygous *NSD1* mutations in patients, however, mice heterozygous for a *Nsd1* null allele are viable and fertile, and display normal growth rates (Rayasam et al., 2003). A second mouse model has also been generated that houses a microdeletion that ablated 36 genes including *Nsd1* on mouse chromosome 13 (which is syntenic with the *NSD1* region on human chromosome 5q35.2 – q35.3). Mice heterozygous for this mutation displayed growth reductions and impairments in long-term memory retention (Migdalska et al., 2012), although the specific role played by Nsd1 in these phenotypic consequences remains unclear. Germ cells isolated from male Nsd1-deficient mice display a global reduction in di-methylation of H3K36 (Shirane et al., 2020).

Although less common, mutations in the adenomatous polyposis coli 2 (*APC2*) gene (OMIM #612034) – a gene expressed in post-mitotic neurons that regulates cytoskeletal structure, axon guidance and neuronal migration – can also cause a Sotos syndrome variant condition (Almuriekhi et al., 2015). The *APC2* gene is normally targeted by NSD1 regulation, and knockdown of Nsd1 using shRNA significantly reduced the expression of *Apc2* mRNA and protein in primary mouse cortical cells (Almuriekhi et al., 2015), suggesting there is a direct link between Nsd1 and Apc2 cooperative function. Consistently, transfection of Apc2-miRNA into cortical progenitor cells (E15.5) resulted in abnormal migration, with a large proportion of neurons remaining in the lower cortical layers (Almuriekhi et al., 2015). This was highly similar to the pattern seen in the same system after treatment with *Nsd1*-miRNA (Almuriekhi et al., 2015), and in the *Apc2* null mouse brain (Shintani et al., 2012). Importantly, these deficits arising from *Nsd1*-miRNA expression were rescued by the co-administration of an *Apc2*-expression plasmid (Almuriekhi et al., 2015). This further indicates the critical role played by Nsd1 in neuronal migration, and proper neurodevelopment requires its regulation of *Apc2* expression. Mutations affecting either factor disrupt this normal signaling cascade and result in Sotos syndrome.

HISTONE ACETYLATION

Histone acetylation is another critical epigenetic process that defines chromatin structure. Histone acetylation regulates

numerous cell processes including cell-cycle progression, differentiation, and metabolism, largely by its role in orchestrating transcriptional responsiveness (Podobinska et al., 2017). Histone acetylation is facilitated by histone acetyltransferases (HATs), and conversely these modifications are removed by histone deacetylases (HDACs) (Gräff and Tsai, 2013). Alternative nomenclature for these enzymes can include lysine acetyltransferases and deacetylases (KAT and KDAC, respectively), as their enzymatic activities are not limited to histone proteins (Seto and Yoshida, 2014).

HATs can be classified into three main groups based largely on homology. The first group is the Gcn5-related N-acetyltransferase (GNAT) family, which is arguably the best understood and includes Gcn5 – the first HAT discovered – its close relatives (ex. PCAF), and the related HATs HAT1, ELP3, and HPA2 (Sterner and Berger, 2000; Bonnaud et al., 2016). The second group of HATs is referred to as CBP/p300, and consists, as the name suggests, of the ubiquitously expressed CREB-binding protein (CBP) and its close relative p300 (Sterner and Berger, 2000; Bonnaud et al., 2016). The third group of HATs is the MYST family, which is an acronym of its founding members MOZ, Ybf2/Sas3, Sas2, and Tip60 enzymes (Sterner and Berger, 2000; Bonnaud et al., 2016). In addition to targeting histones, HATs can act as scaffolds for promoter-binding transcription factors, and acetylate these factors to alter their activity (Sterner and Berger, 2000). Additional proteins with HAT activity also exist, and include nuclear receptor co-activators, such as SLC-1, NCoA-3, and CLOCK (Sterner and Berger, 2000; Bonnaud et al., 2016). Lastly, certain transcription factors can also possess HAT activities, with TFIID and TFIIIC being known examples (Sterner and Berger, 2000; Bonnaud et al., 2016). HATs typically act either as nuclear modifiers of nucleosomes (A-type) or cytoplasmic modifiers of newly synthesized histones (B-type) (Sterner and Berger, 2000; Bannister and Kouzarides, 2011).

HDACs can be subdivided into four classes. Class I HDACs consists of HDACs 1–3, and 8, whose functions are restricted exclusively to the nucleus (Kouzarides, 2007; Bonnaud et al., 2016). Class IIa HDACs (4, 5, 7, and 9) are shuttled between the nucleus and cytoplasm, but possess no intrinsic deacetylase activity and are thought to act as scaffolds for other co-repressor systems (Mielcarek et al., 2015; Bonnaud et al., 2016). Class IIb enzymes (6 and 10) largely function outside the nucleus, where they mediate deacetylation of cytosolic proteins (Seto and Yoshida, 2014; Bonnaud et al., 2016). Class III HDACs consists of the sirtuin (SIRT) subfamily of enzymes that reside within the nucleus, cytosol, and mitochondria (Kouzarides, 2007; Bonnaud et al., 2016). Class IV HDACs presently consists of only the nuclear HDAC11 (Seto and Yoshida, 2014; Bonnaud et al., 2016). Deacetylation reactions are zinc dependent for Class I, II, and IV enzymes, while Class III enzymes are dependent on NAD^+ as a cofactor (Seto and Yoshida, 2014; Bonnaud et al., 2016).

Histone acetylation patterns influence transcription by altering chromatin structure. Mechanistically, lysine residues residing toward the amino terminus of histone proteins are positively charged at physiological pH. These charges form an electrostatic interaction with the negative charge of DNA, which tightens the association of the DNA with the histone protein and encourages chromatin compaction. When these lysine sites are targeted for acetylation, the result is the neutralization of the positive charge, a weakening of the electrostatic interactions between histones and DNA, and a relaxation of chromatin (**Figure 2**) (Gräff and Tsai, 2013; Podobinska et al., 2017). Deacetylation, conversely, exposes the positive charges, promotes stronger electrostatic interactions, and favors condensed chromatin (Gräff and Tsai, 2013; Podobinska et al., 2017). Histone acetylation promotes transcription factor recruitment to the exposed DNA, whereas deacetylation restricts access (Bannister and Kouzarides, 2011). While HATs primarily modify histones H3 and H4 on the exposed N-terminal tail, acetylation can occur at core and tail regions of all histones (Podobinska et al., 2017). The hallmark lysine residues acetylated on H3 are K9, K14, K18, K23, and K56, while on H4 the hallmark residues are K5, K8, K12, and K16 (Podobinska et al., 2017; Sheikh and Akhtar, 2019). In addition, the pattern of acetylation established can provide a recognition site for certain transcription factors that also facilitate transcription (Kouzarides, 2007). Ultimately, the balance of HAT and HDAC activity throughout genomic segments determines transcriptional accessibility and activity of many dynamically expressed genes (Kouzarides, 2007).

Much like histone methylation, when an imbalance in these systems occurs it can have detrimental effects on neurodevelopment. This can arise from mutations in key acetylating or deacetylating enzymes. Several gene mutations have been identified that affect these systems and are causal for neurodevelopmental diseases. Examples include Rubinstein-Taybi Syndrome, KAT6A syndrome, Genitopatellar syndrome, Say-Barber-Biesecker-Young-Simpson syndrome and Brachydactyly mental retardation syndrome.

Conditions Associated With Impaired Histone Acetyltransferase Function

Rubinstein-Taybi Syndrome

Rubinstein-Taybi syndrome (RTS; OMIM #180849) is a neurodevelopmental disorder affecting 1 in ∼125,000 live births (Hutchinson and Sullivan, 2015). RTS is caused by mutations in the KAT3 subfamily member *CREBBP,* which encodes CREB binding protein (CBP) (Type 1 RTS) or in subfamily member *EP300,* which encodes p300 (Type 2 RTS) (Korzus, 2017; López et al., 2018). Approximately 55% of RTS cases stem from *CREBBP* mutations, while ∼10% of RTS patients result from *EP300* mutations (Korzus, 2017). Both of these mutations cause deficiencies in histone acetylation activity/efficiency. The autosomal dominant mutations of either gene largely occur *de novo* in germline cells, and the cellular consequences arise either through enzymatic haploinsufficiency or via a dominant negative mechanism (Barco, 2007; Park et al., 2014). The genetic cause of the remaining third of RTS cases is currently unknown (López et al., 2018). The diagnosis of RTS is based predominantly on clinical presentation, often within the 1st year of life (Korzus, 2017). Features of RTS include broad thumbs and halluces, severe cognitive impairment, facial abnormalities, and psychomotor delay (Park et al., 2014; Hutchinson and Sullivan,

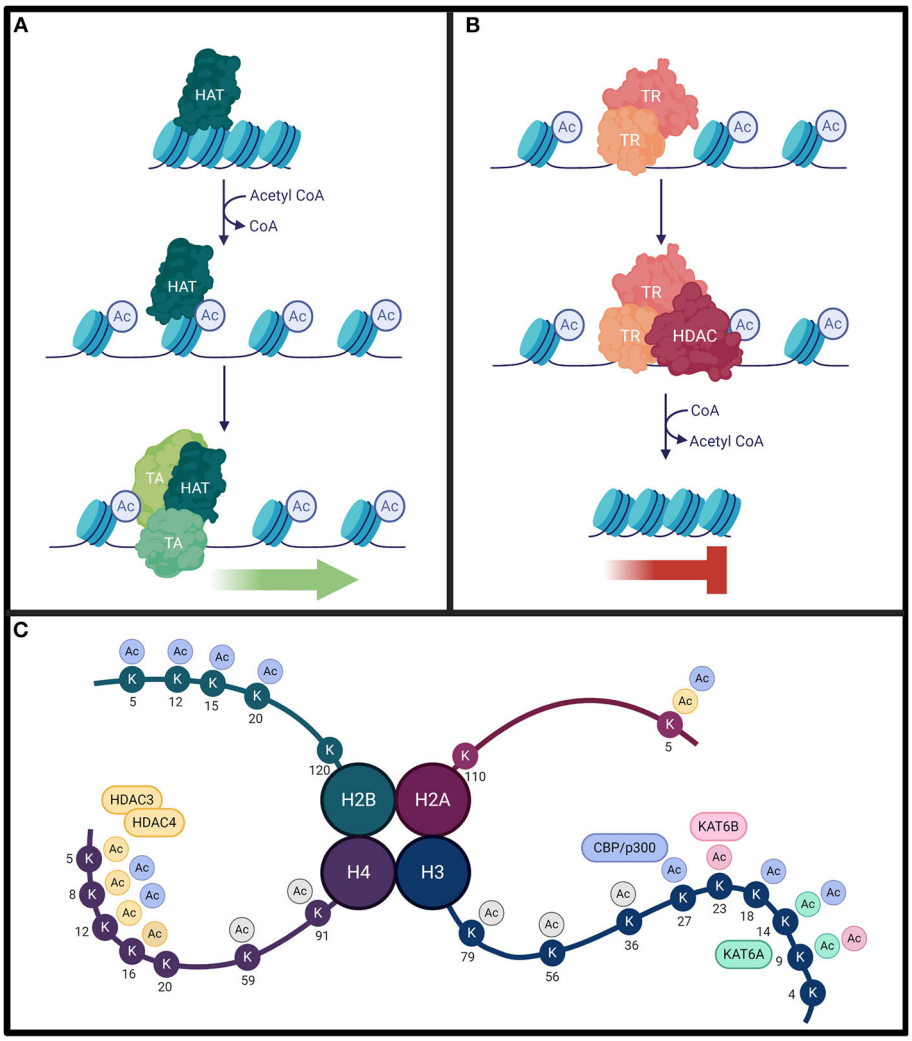

FIGURE 2 | (A) Histone acetyltransferases (HAT) will acetylate histone lysine residues using acetyl CoA cofactor (blue circles). This weakens the electrostatic interactions between positively-charged histones and negatively-charged DNA to loosen chromatin structure. This results in DNA exposure, allowing for the recruitment of transcriptional activators (TA). **(B)** Conversely, transcriptional repressor (TR) complexes can interact with histone deacetylases (HDAC) to remove these modifications. This strengthens the electrostatic interactions between the DNA and the histones, resulting in compact chromatin, inhibiting transcription. The homeostasis between histone acetylation and deacetylation is critical for proper gene expression. **(C)** Lysine acetylation can occur on all four histone subunits. Each HAT/HDAC has preferred target sites. While KAT6A and KAT6B acetylate residues with high specificity, CBP and HDAC4/HDAC3 complexes have a broader range of targets. The lysine residues targeted by each enzyme are color coded. There is some target redundancy between HAT/HDACs, and other histone (de)acetylating enzymes can also target common sites (not indicated on figure). Additional lysine acetylation sites not specifically targeted by these enzymes are indicated in gray.

2015). Some patients may also present with seizures, spinal deformities, syndactyly, and congenital heart abnormalities (Lopez-Atalaya et al., 2014; Hutchinson and Sullivan, 2015). RTS patients have also been found to be more susceptible to the development of CNS cancers later in life (Park et al., 2014). Cognitive impairments occur in 99% of patients harboring *CREBBP* mutations and 94% of those with *EP300* mutations, with IQ scores ranging from <25 to 80 (Korzus, 2017; López et al., 2018). The remaining phenotypes discussed above tend to be observed more frequently and with greater severity in patients with *CREBBP* mutations as compared to those with *EP300* mutations (Korzus, 2017).

CREBBP is located on chromosome 16p13.3 and mutations consist of either chromosomal rearrangements within this region or mutations within the *CREBBP* gene itself (Petrij et al., 1995; Barco, 2007). Translocations, breakpoints, insertions, and microdeletions of chromosome 16 in this region can all contribute to RTS, and larger rearrangements generally associate with an increased severity of the condition (Petrij et al., 1995; Lopez-Atalaya et al., 2014). Various mutations within the *CREBBP* gene have been reported including missense, non-sense, frameshift, insertion, deletions, and mutations located near the splice-site junctions (Lopez-Atalaya et al., 2014; Korzus, 2017). Deletion mutations can occur along the entire length of the

gene or affect specific parts, although mutations in reading frame exons, that encode the histone acetyltransferase domain, are most common (Barco, 2007; Korzus, 2017). The same is true for *EP300*, which is located on chromosome 22q13.2; both with regards to chromosome aberrations and coding sequence mutations (Korzus, 2017). CBP and p300 target numerous lysine residues on all four histones: H2A (K5), H2B (K5, K12, K15, K20), H3 (K14, K18, K27), and H4 (K8, K12) (Bedford and Brindle, 2012; Valor et al., 2013; Lipinski et al., 2020). CBP and p300 are critical in many nuclear processes, due both to their respective HAT activities, and their ability to interact with over 400 transcription factors that effectively compete for the proportionally limited amount of CBP/p300 present in the cell (Dyson and Wright, 2016). While there is some overlap between the target genes of CBP and p300, these two HATS also display some distinct targets and functions. Thus, mutations in either protein cannot be completely accommodated by the other (Lopez-Atalaya et al., 2014; Korzus, 2017).

Various mouse models of RTS have been developed which target both *Crebbp* and *Ep300*. The first models created were knockout mice in which specific domains of either gene were targeted for ablation (e.g., their respective KAT, KIX, and CH1 domains) (Barco, 2007; Lopez-Atalaya et al., 2014). For both *Crebbp* and *Ep300*, homozygous mutants expressing these disruptions were embryonic lethal (Barco, 2007; Lopez-Atalaya et al., 2014). Interestingly, mice jointly heterozygous for both *Crebbp* and *Ep300* were also not viable (Lopez-Atalaya et al., 2014). The *Crebbp* or *Ep300* heterozygous mice mirrored many phenotypes observed in patients, including inhibited growth, select skeletal deformities, and memory impairments (Barco, 2007), although the characteristic broad halluces seen in RTS patients are not evident in these mice (Lopez-Atalaya et al., 2014). Examination of histone acetylation status in *Crebbp* heterozygous mice revealed global hypoacetylation of H2B, whereas the other histone subunits were unaffected (Alarcón et al., 2004). In contrast, *EP300* heterozygous mice showed no change in acetylation status (Oliveira et al., 2007; Viosca et al., 2010). Knock-in mice have also been generated that express specific point mutations for both *Crebbp* and *Ep300*, permitting examination of the role specific domains of each protein have on overall function (Lopez-Atalaya et al., 2014). Conditional knockouts have also been generated using Cre-loxP systems that target specific regions of the hippocampal formations, such as the dentate gyrus and CA1 using the CamKIIα promoter (Barco, 2007; Barrett et al., 2011; Lopez-Atalaya et al., 2014). CaMKIIα-Cre/*Crebbp* mice display impaired short- and long-term memory, and acetylation deficiencies at H2BK12, H3K14, H3K27, and H4K8 (Barrett et al., 2011; Lipinski et al., 2020). A tetracycline-induced transgenic mouse was developed to allow for temporal control of *Crebbp* expression (Barco, 2007; Lopez-Atalaya et al., 2014). In these mice, the HAT domain was inactive, but Cbp could still interact with its partner transcription factors (Barco, 2007; Lopez-Atalaya et al., 2014). These mice also presented with long-term memory impairments (Barco, 2007). These region-specific ablations resulted in learning deficits and diminished LTP, which mirror what is seen in RTS patients (Gräff and Tsai, 2013; Lopez-Atalaya et al., 2014). Overall, these mouse models indicate that H2B acetylation is the site most predominantly affected by CBP deficiency, but that intact p300 cannot fully compensate for the CBP deficiency at H2B. This diminished H2B acetylation pattern was also found in patient-derived lymphoblastoid cultures expressing *CREBBP* mutations (Lopez-Atalaya et al., 2012), although these cells also displayed hypoacetylated H2A (but not H3 and H4). Intriguingly, this hypoacetylation in H2A and H2B was rescued by treating the cells with an HDAC inhibitor (Lopez-Atalaya et al., 2012), suggesting a possible avenue for translational development.

KAT6A Syndrome

KAT6A syndrome (OMIM # 616268) is a rare autosomal dominant disorder first described in 2015 by Arboleda et al. and Tham et al. (Arboleda et al., 2015; Tham et al., 2015). Intellectual disability and global developmental delay are observed in patients, concomitant with notable speech delays (Arboleda et al., 2015; Tham et al., 2015). This oromotor dyspraxia is universal in KAT6A syndrome, with more marked impairments in expressive language than receptive language (Kennedy et al., 2019). Clinical features are variable between patients, however, hypotonia, microcephaly, craniofacial dysmorphisms, congenital cardiac defects, gastrointestinal problems, feeding difficulties, strabismus, and sleep disturbances have been reported (Arboleda et al., 2015; Millan et al., 2016; Kennedy et al., 2019). Delayed myelination has also been observed in some patients, although was found to resolve over time (Millan et al., 2016; Alkhateeb and Alazaizeh, 2019).

KAT6A syndrome arises from mutations in the gene encoding the lysine acetyltransferase *KAT6A* (also known as *MYST3*, *MOZ*), which is located on chromosome 8p11.21 (Tham et al., 2015). The *KAT6A* gene locus consists of 18 exons, which encodes for a protein of 2004 amino acids that houses five domains: H15 nuclear localization, PHD, KAT, acidic Glu/Asp-rich, and Ser/Met-rich transactivation domain (Klein et al., 2014; Tham et al., 2015). Key residues in the KAT domain have been identified as E680 and C646, wherein the former actively interacts with the acetyl CoA cofactor (Klein et al., 2014). KAT6A primarily targets H3K9, and to a lesser extent H3K14, to loosen chromatin structure as described above (Yang, 2015). In addition, KAT6A can complex with factors, such as BRPF1/2/3, hEAF6, and ING5, or interact with RUNX1/2 to activate the transcription of a number of genes, such as Hox (Voss et al., 2009; Yang, 2015). Truncating, frameshift, and non-sense mutations are the most common mutations causal for KAT6A syndrome, although missense mutations, and mutations occurring at splice junction sites have been identified in some patients (Tham et al., 2015; Millan et al., 2016; Kennedy et al., 2019). Mutations in the KAT domain itself are infrequent; mutations are rather predominantly found in the more distal exons 16 to 18, that encode for the acidic Glu/Asp-rich domain and the Ser/Met-rich domain (Tham et al., 2015; Kennedy et al., 2019). Specific non-sense mutation hotspots at positions 1,019, 1,024, and 1,129 have been identified that account for about 20% of the total KAT6A syndrome cases (Tham et al., 2015; Kennedy et al., 2019). Interestingly, patients with distal exon mutations exhibit more severe symptoms than those with mutations in more proximal exons 1–15 (Kennedy et al.,

2019). This is postulated to be a result of KAT6A transcript fate. If the mutation is proximal, the transcript undergoes nonsense-mediated decay, and symptoms arise through haploinsufficiency (Kennedy et al., 2019). It has been hypothesized that transcripts with mutations in exons 16, 17, and terminal exon 18, however, may not be degraded, and consequently encode a truncated protein that acts in a dominant-negative or gain-of-function manner (Kennedy et al., 2019). At present, however, experiments to test this hypothesis remain to be conducted.

Various animal models have been generated for investigating the role of Kat6a in development. Initially, Kat6a deficiency was examined in mutant zebrafish obtained through ENU mutagenesis, which resulted in aberrant craniofacial development (Miller et al., 2004; Crump, 2006). As mentioned above, Kat6a is a key player in the proper expression of the *Hox* genes that define segmental identity (Crump, 2006; Vanyai et al., 2019). In these Kat6a-deficient zebrafish, the second pharyngeal arch of the cranial neural crest was replaced by a duplicated jaw structure (Miller et al., 2004; Crump, 2006). Loss of Kat6a function in zebrafish appears to exclusively affect the head, whereas rodents are more globally affected (Tham et al., 2015). Two *Kat6a* targeted mouse models have also been generated (Voss et al., 2012). To model early exon deletions in *Kat6a*, exons 3-7 were flanked with loxP sites and removed globally via Cre recombinase (Voss et al., 2009). In contrast to findings in humans, early exon mutations were more severe than C-terminal mutations in mice, with complete null mice displaying embryonic lethality between E14.5 and birth (Voss et al., 2012). These homozygous early exon *Kat6a* knockout mice aged E10.5 days had elongated necks with an additional eighth cervical vertebra, and one fewer thoracic vertebra than heterozygotes and wild-types (Voss et al., 2009). This was attributed to *Hox* gene repression, as the histones of those loci were hypoacetylated instead of being tri-methylated at H3K9 (Voss et al., 2009). Consistently, this same histone modification pattern was observed in fibroblasts from KAT6A syndrome patients (Arboleda et al., 2015). The other model was generated by an in-frame insertion of a neomycin phosphotransferase sequence in exon 16, near the carboxyl terminus (Thomas et al., 2006). These mice survived throughout gestation, but remained severely affected and died within hours after birth (Voss et al., 2012). Both mouse models presented with cardiac defects, caused by *Tbx1* repression in the absence of Kat6a (Voss et al., 2012). In addition, both early and late exon *Kat6a* knockouts exhibited a loss of hematopoietic stem cells, and craniofacial dysmorphisms, including cleft palates (Voss et al., 2012; Vanyai et al., 2019). Except for mild craniofacial dysmorphisms and heart defects, these mouse models do not appear to phenocopy the human condition.

Genitopatellar Syndrome and Say-Barber-Biesecker-Young-Simpson Syndrome

Genitopatellar syndrome (GPS; OMIM #606170), first reported in 1988 (Goldblatt et al., 1988), is a condition characterized by neurological, skeletal, and genital abnormalities. Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS; OMIM#603736) was first described in 1986 (Ohdo et al., 1986; Campeau et al., 2012) and shares many overlapping clinical features with GPS, although somewhat less severe (Campeau et al., 2012). Both of these autosomal dominant conditions result from mutations in the *KAT6B* gene (previously named *MORF* and *MYST4*) that resides on chromosome 10q22.2 (Lonardo et al., 2019). However, the mutations in GPS and SBBYSS occur at different loci on the *KAT6B* gene (Lonardo et al., 2019). Intellectual disability, developmental delay, congenital heart defects, thyroid and dental anomalies, hypotonia, feeding difficulties, and hearing loss are present in both GPS and SBBYSS (Campeau et al., 2012; Vlckova et al., 2015; Lonardo et al., 2019). Genital and anal anomalies are common for both sexes in GPS, whereas they are rare, mild, and restricted to males in SBBYSS (Campeau et al., 2012; Lonardo et al., 2019). Patients with GPS also exhibit patellar hypoplasia or agenesis, flexion contractures of the hips and knees, club-foot, hydronephrosis, renal cysts, agenesis of the corpus callosum, and microcephaly (Campeau et al., 2012; Vlckova et al., 2015; Lonardo et al., 2019) While both conditions present with cleft palates, prominent cheeks and bulbous noses, the craniofacial dysmorphisms of SBBYSS also include mask-like, immobile facies, blepharophimosis, and ptosis (Vlckova et al., 2015; Lonardo et al., 2019). In addition, SBBYSS patients typically have long fingers and toes (Campeau et al., 2012; Vlckova et al., 2015). Clinical studies, however, have reported ambiguities between diagnosis and presentations. The term "KAT6B spectrum disorder" has been proposed to encompass both GPS and SBBYSS (Gannon et al., 2015; Lundsgaard et al., 2017; Radvanszky et al., 2017).

The mutations associated with GPS and SBBYSS largely occur at distinct positions on the *KAT6B* gene, and for classification purposes the *KAT6B* mutations can be divided into four groups (Campeau et al., 2012; Sheikh and Akhtar, 2019). Group 1 mutations promote an SBBYSS phenotype and cluster within exons 15, 16, and in the proximal region of 17 of the *KAT6B* gene (although rare cases with early exon mutations also exist in this group) (Vlckova et al., 2015; Lundsgaard et al., 2017; Radvanszky et al., 2017). Group 2 mutations, which promote GPS phenotypes, are located in distal exon 17 and the proximal portion of exon 18 (Vlckova et al., 2015; Radvanszky et al., 2017). Group 3 mutations occur in the medial portion of exon 18, and give rise to GPS, SBBYSS, and mixed GPS/SBBYSS phenotypes (Vlckova et al., 2015; Radvanszky et al., 2017). Group 4 mutations are found in the distal section of exon 18 that encodes for the transactivation domain of KAT6B, and are associated with SBBYSS phenotypes (Vlckova et al., 2015; Radvanszky et al., 2017). The majority of mutations that cause both the GPS and SBBYSS clinical presentations are non-sense or frameshift, although some missense mutations and in-frame deletions have also been identified (Gannon et al., 2015; Lundsgaard et al., 2017). The postulated mechanism of pathogenesis is haploinsufficiency or loss-of-function for SBBYSS, and gain-of-function or a dominant-negative mechanisms for GPS (Campeau et al., 2012; Gannon et al., 2015). This would account for the milder phenotypes generally associated with SBBYSS (Campeau et al., 2012; Gannon et al., 2015). However, direct experimental evidence to support this assertion is lacking to date. Normally, KAT6B can integrate into complexes that regulate transcription,

and is known to also partner with RUNX2 and form a tetrameric complex with BRPF1/2/3, hEAF6, and ING5 (Yang, 2015). In mice, *KAT6B* expression in the brain is dynamic during early development, after which it remains highly expressed in neural stem cells in the subventricular zone throughout adulthood (Thomas et al., 2000; Sheikh et al., 2012). KAT6B predominantly acetylates H3K9 and H3K23 and is critical for maintaining the multipotent and self-renewing properties of neural stem cells (Campeau et al., 2012; Sheikh et al., 2012; Sheikh and Akhtar, 2019).

There have been few animal models of KAT6B deficiency generated to date. Thomas et al. using random gene trap methodology generated a mouse model in which only 10% of the normal mRNA expression of Querkopf (*Qkf*), the mouse ortholog of *KAT6B*, was observed (Thomas et al., 2000; Thomas and Voss, 2004). Mice homozygous for the *Qkf* gene trap displayed craniofacial dysmorphisms, decreased body mass, and failure to thrive, with two-thirds not surviving to weaning (Thomas et al., 2000; Thomas and Voss, 2004). The cortex and olfactory bulb of these mice were smaller than normal, possessing fewer cells overall, GAD67-positive interneurons and large pyramidal neurons in the cortex (Thomas et al., 2000). Impairments in adult neurogenesis have also been identified in these mice, with reduced numbers of neural stem cells in the subventricular zone and granule and periglomerular interneurons in the olfactory bulb (Merson et al., 2006).

Conditions Associated With Impaired Histone Deacetylase Activity

Brachydactyly Mental Retardation Syndrome

Brachydactyly mental retardation syndrome (BDMR; OMIM #600430) is a developmental condition that shares a number of similar clinical features with Albright hereditary osteodystrophy (Leroy et al., 2013; Wheeler et al., 2014). Cognitive impairment and developmental delay were originally thought to be universal, although closer evaluations suggest developmental delay is clearly evident in ∼79% of patients (Wheeler et al., 2014; Le et al., 2019). Autistic-like behaviors are also a common feature in BDMR (Williams et al., 2010). Additional neurological impairments, such as hypotonia and seizures, are present in ∼46% of patients (Williams et al., 2010; Le et al., 2019). In addition, brachydactyly type E was observed in approximately half of patients, with the fourth and fifth metacarpal being the most severely affected (Le et al., 2019). Other features observed include distinctive craniofacial dysmorphisms, obesity, sleep disturbances, and cardiac defects (Morris et al., 2012; Le et al., 2019).

BDMR is associated with variably sized deletions of the telomeric region of chromosome 2q37. The boundaries of the commonly deleted region encompass over 197 genes and the identity of the primary causal gene(s) for the condition was initially unclear (Leroy et al., 2013). However, deletion mapping identified *HDAC4* as a common deleted gene in most patients, including those with the smaller microdeletions of 2q37.3 (Villavicencio-Lorini et al., 2013; Le et al., 2019). Moreover, a recent report identified a BDMR patient whose causal mutation was a one base insertion into the HDAC4 gene open reading frame (Williams et al., 2010). Collectively, these observations implicate *HDAC4* as critical gene in BDMR, and suggest that HDAC4 haploinsufficiency represents a pathogenic mechanism for BDMR (Williams et al., 2010; Wheeler et al., 2014). Further support for this possibility comes from genotype-phenotype studies, where common features seen in BDMR patients that include skeletal development, obesity, and behavioral tendencies have been reported (Leroy et al., 2013).

HDAC4 is a Class IIa HDAC, and like all members of this subfamily, contains a specific alteration - a His976Tyr in its catalytic domain (Mielcarek et al., 2015). This alteration abolishes the deacetylase activity of HDAC4, and thus its effects on histone acetylation are exerted through its transcription binding domain (TBD), allowing it to form specific complexes to repress gene expression (Williams et al., 2010; Mielcarek et al., 2015). Mutations in *HDAC4* alter its ability to recruit required enzymes. Notably, it has been postulated that HDAC4 can act as a scaffold for the N-CoR/HDAC3 complex and enhance HDAC3-mediated H3 and H4 lysine deacetylation, at H4K5, K8, K12, and K16 (McQuown and Wood, 2011; Seto and Yoshida, 2014; Mielcarek et al., 2015). HDAC3/4 complexes will also deacetylate H2A at K5 (Seto and Yoshida, 2014). An *Hdac4*-null mouse, however, showed no change in the acetylation status of histones at age P3 (Mielcarek et al., 2013). Examination of histone acetylation in older *Hdac4*-null mice have yet to be reported and could prove informative.

Hdac4 is highly expressed in the brain, with expression levels peaking in early postnatal life during periods of heightened synaptogenesis (Sando et al., 2012; Mielcarek et al., 2013). The importance of Hdac4 during this stage of development was investigated using mice, in which an inserted lacZ cassette caused the deletion of the TBD, nuclear localization signal, and catalytic domains of the encoded protein (Vega et al., 2004). These mice were severely undersized, presented skeletal defects, and did not survive to weaning (Vega et al., 2004). A separate study found the brains of *Hdac4* knockout mice were 40% smaller than wild-types, and had enlarged ventricles (Majdzadeh et al., 2008). Progressive degeneration of Purkinje neurons was also noted P3 to P10, with surviving neurons displaying stunted dendrite growth (Majdzadeh et al., 2008). It is worth noting, though, that *Thy1-Cre/Hdac4* and *Nes-Cre/Hdac4* conditional knockout mice have normal locomotor activity and display normal brain morphology (Price et al., 2013). Both of these transgenic lines ablated *Hdac4* from large populations of neurons, thereby raising the possibility that Hdac4 impairments could promote effects in neurons through non-cell autonomous mechanisms. However, when Hdac4 was absent in forebrain excitatory neurons in another conditional knockout mouse model, *CamKII-Cre/Hdac4*, the mice were hyperactive, with deficits in memory and motor coordination, and reduced anxiety-like behaviors (Kim et al., 2012). Furthermore, a transgenic mouse model harboring truncated Hdac4 had attenuated excitatory synaptic strength, and impaired spatial memory (Sando et al., 2012). Taken together, these studies illustrate that Hdac4 deficiency alone is sufficient to induce neurological and peripheral consequences that have commonality with BDMR patients. Despite these data,

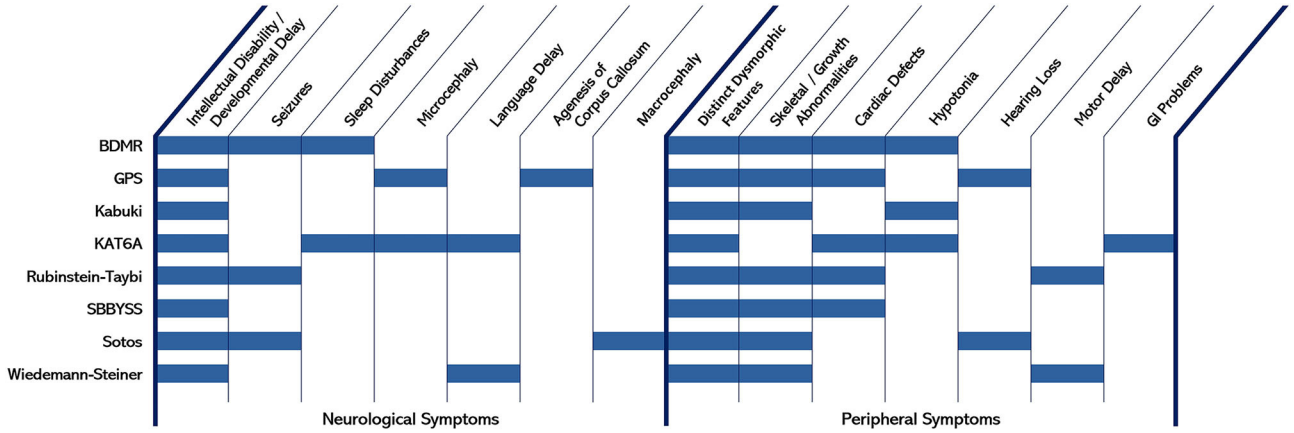

FIGURE 3 | Neurological and peripheral symptoms common to different epigenetic-linked developmental conditions.

the potential involvement of other genes deleted on chromosome 2q37 in some patients cannot be ruled out for a potential contributing role in pathogenesis.

DISCUSSION

Epigenetic modifications, and the functional fidelity of systems directly involved in making these modifications, clearly play important roles in neurodevelopment. Epigenetic modifications of histones specifically, whether through methylation or acetylation, can have affects on individual gene regulation or effects on multiple genes, since large regions of chromatin can be influenced by these histone modifications. Inappropriate regulation can lead to impairments in constitutive or cue-dependent activation or deactivation of genes. Indeed, model systems investigating specific enzymes involved in coordinated histone methylation or acetylation have revealed complete ablation of the enzyme frequently results in embryonic lethality, with defects in germ layer formation and/or gross impairments in early developmental ontogenesis often observed. In haploinsufficient or functionally-hypomorphic systems, alterations in the normal homeostatic activity range for the pathways encoded by these factors can be sufficient to cause severe conditions that manifest in early childhood. For a number of these cases, similar or overlapping clinical features can be observed even though the specific gene mutated may encode a factor that seemingly regulates opposite aspects of epigenetic control. Indeed, mutations affecting several enzymes and involving multiple systems can result in widespread consequences that collectively impair neurological, skeletal, and cardiac development (amongst others) from properly occurring. Although the expression profiles of these enzymes vary, the convergence of phenotypes that affect the central nervous system and peripheral systems is clear (**Figure 3**). Development requires complex but orchestrated spatiotemporal changes in gene expression, which is strongly mediated by cues provided through histone modifications. If the fidelity of this mechanism is impaired, the normal orchestrated response fails, cue-dependent developmental patterning does not properly execute, and the ontogeny of normal synaptic and network formation is negatively impacted in the developing brain.

The effect of histone methylation and acetylation on general chromatin structure and gene expression is well-documented (Bannister and Kouzarides, 2011). It is also important to note, however, that the specific enzymes catalyzing these histone modifications may also possess dual or multi-modal functionality and regulate other non-histone systems. Certain tumor suppressor factors have been found to be acetylated by various HATs; for instance, both CBP and p300 target p53, p73 E2F, and Rb to affect their respective binding to promoter regions (Sterner and Berger, 2000; Iyer et al., 2004). PCAF shares many of these same targets, and KAT6A also is known to acetylate p53 (Sterner and Berger, 2000). These modifications can influence transcriptional responses in conjunction with, or independently of, MOZ-dependent histone acetylation (Huang et al., 2016). Thus, while mutations of histone acetylation or methylation factors alter normal epigenetic histone coding, other targets may also be affected whose functions may extend to systems not classically viewed as "epigenetic." Likewise, gene regulation within different regions of the genome that are subject to the same modification may not respond equally. While histone modifications affect chromatin structure and enable/restrict transcriptional responsiveness, the actions of specific transcription factors and transcriptional complexes dictate the dynamics of the response. Moreover, the activity of one histone modification system can alter specific histones in a manner that better enables the recruitment of additional systems. For example, the methylation of H3K4 by KMT2 family members allows for a higher-affinity recruitment of MYST acetyl-transferases to those sites. The ensuing acetylation of the histone H3 is associated with increased local transcriptional activation (Sheikh and Akhtar, 2019). If the initial histone modification is impaired, the subsequent recruitment that directly effects transcription will not properly occur. Thus, altering the activity of one histone modifier can have broad impact on a multitude of transcriptional regulatory complexes

that normally function in a cooperative manner to achieve a proper homeostatic response. In addition, several histone methyltransferase and acetyltransferase proteins include motifs that bind other transcriptional regulators to allow complexes to assemble. In many cases, these additional motifs reside downstream of the catalytic motif in the expressed transcript. Since they reside distally, these domains would also be affected by mutations that induce nonsense-mediated decay. Thus, not only is the catalytic activity of the transferase lost, but critical interactions dictated by these more distal motifs are also compromised. This is especially important for Class IIa HDACs, for instance HDAC4, which do not have intrinsic deacetylase activity, but regulate gene expression through the complexes they form with other transcription factors.

In summary, epigenetic processes conveyed by post-translational modifications of histone proteins represent a key component in the overall regulation of static and dynamic gene expression. The neurodevelopmental disorders discussed in this review highlight the neurological (along with peripheral and/or embryonic development) consequences that can arise from impaired histone acetylation or methylation coding. A failure in normal cue-dependent execution of epigenetic signaling at any of several key developmental windows can lead to the same endpoint, which may explain why common features can emerge in patients affected by different causal mutations. Further research into these post-translational histone modification diseases, and developing additional model systems to complement those currently available, will be important to delineate a clearer understanding of the molecular, cellular and pathological mechanisms underlying these conditions, and ultimately the development of translationally relevant therapeutics.

AUTHOR CONTRIBUTIONS

MF: wrote sections of manuscript, generated table, and compiled bibliography. DS: wrote sections of manuscript, generated figures, and proof-read manuscript. CR: generated figures, edited manuscript, and contributed to discussion section. JE: selected topic of review, oversaw completion of text, edited and proof-read manuscript, compiled, and submitted manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Work toward this review was supported by the Ontario Brain Institute and an operating grant from the Canadian Institutes of Health Research (CIHR grant MOP125909). MF and DS were supported by graduate awards from the Department of Pharmacology and Toxicology at the University of Toronto.

ACKNOWLEDGMENTS

The authors thank members of the Eubanks lab for helpful discussion with this review, and Dr. Mojgan Rastegar (University of Manitoba) for comments and suggestions. Figures were constructed using licensing from BioRender.com.

REFERENCES

Adam, M. P., Banka, S., Bjornsson, H. T., Bodamer, O., Chudley, A. E., Harris, J., et al. (2019). Kabuki syndrome: international consensus diagnostic criteria. *J. Med. Genet.* 56, 89–95. doi: 10.1136/jmedgenet-2018-105625

Alarcón, J. M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E. R., et al. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP+/– mice: a model for the cognitive deficit in rubinstein-taybi syndrome and its amelioration. *Neuron* 42, 947–959. doi: 10.1016/j.neuron.2004.05.021

Alkhateeb, A., and Alazaizeh, W. (2019). A novel de novo frameshift mutation in KAT6A identified by whole exome sequencing. *J. Pediatr. Genet.* 8, 10–14. doi: 10.1055/s-0038-1676649

Almuriekhi, M., Shintani, T., Fahiminiya, S., Fujikawa, A., Kuboyama, K., Takeuchi, Y., et al. (2015). Loss-of-function mutation in APC2 causes sotos syndrome features. *Cell Rep.* 10, 1585–1598. doi: 10.1016/j.celrep.2015.02.011

Arboleda, V. A., Lee, H., Dorrani, N., Zadeh, N., Willis, M., Macmurdo, C. F., et al. (2015). De novo nonsense mutations in KAT6A, a lysine acetyl-transferase gene, cause a syndrome including microcephaly and global developmental delay. *Am. J. Hum. Genet.* 96, 498–506. doi: 10.1016/j.ajhg.2015.01.017

Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S., et al. (2015). Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin. Genet.* 87, 252–258. doi: 10.1111/cge.12363

Bannister, A. J., and Kouzarides, T. (2005). Reversing histone methylation. *Nature* 436, 1103–1106. doi: 10.1038/nature04048

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

Bannister, A. J., Schneider, R., and Kouzarides, T. (2002). Histone methylation. *Cell* 109, 801–806. doi: 10.1016/S0092-8674(02)00798-5

Barco, A. (2007). The rubinstein? Taybi syndrome: modeling mental impairment in the mouse. *Genes Brain Behav.* 6, 32–39. doi: 10.1111/j.1601-183X.2007.00320.x

Barrett, R. M., Malvaez, M., Kramar, E., Matheos, D. P., Arrizon, A., Cabrera, S. M., et al. (2011). Hippocampal focal knockout of CBP affects specific histone modifications, long-term potentiation, and long-term memory. *Neuropsychopharmacology* 36, 1545–1556. doi: 10.1038/npp.2011.61

Baujat, G., and Cormier-Daire, V. (2007). Sotos syndrome. *Orphanet J. Rare Dis.* 2:36. doi: 10.1186/1750-1172-2-36

Bedford, D. C., and Brindle, P. K. (2012). Is histone acetylation the most important physiological function for CBP and p300? *Aging* 4, 247–255. doi: 10.18632/aging.100453

Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes Dev.* 23, 781–783. doi: 10.1101/gad.1787609

Bird, A. (2007). Perceptions of epigenetics. *Nature* 447, 396–398. doi: 10.1038/nature05913

Bjornsson, H. T., Benjamin, J. S., Zhang, L., Weissman, J., Gerber, E. E., Chen, Y.-C., et al. (2014). Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of kabuki syndrome. *Sci. Transl. Med.* 6:256ra135. doi: 10.1126/scitranslmed.3009278

Black, J. C., Van Rechem, C., and Whetstine, J. R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol. Cell* 48, 491–507. doi: 10.1016/j.molcel.2012.11.006

Bonnaud, E. M., Suberbielle, E., and Malnou, C. E. (2016). Histone acetylation in neuronal (dys) function. *Biomol. Concepts* 7, 103–116. doi: 10.1515/bmc-2016-0002

Campeau, P. M., Lu, J. T., Dawson, B. C., Fokkema, I. F. A. C., Robertson, S. P., Gibbs, R. A., et al. (2012). The KAT6B -related disorders genitopatellar syndrome and Ohdo/SBBYS syndrome have distinct clinical features reflecting distinct molecular mechanisms. *Hum. Mutat.* 33, 1520–1525. doi: 10.1002/humu.22141

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10, 295–304. doi: 10.1038/nrg2540

Chen, D. (1999). Regulation of transcription by a protein methyltransferase. *Science* 284, 2174–2177. doi: 10.1126/science.284.5423.2174

Cosgrove, M. S., and Patel, A. (2010). Mixed lineage leukemia: a structure-function perspective of the MLL1 protein. *FEBS J.* 277, 1832–1842. doi: 10.1111/j.1742-4658.2010.07609.x

Crump, J. G. (2006). Moz-dependent Hox expression controls segment-specific fate maps of skeletal precursors in the face. *Development* 133, 2661–2669. doi: 10.1242/dev.02435

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat. Rev. Genet.* 13, 153–162. doi: 10.1038/nrg3188

Douglas, J., Hanks, S., Temple, I. K., Davies, S., Murray, A., Upadhyaya, M., et al. (2003). NSD1 mutations are the major cause of sotos syndrome and occur in some cases of weaver syndrome but are rare in other overgrowth phenotypes. *Am. J. Hum. Genet.* 72, 132–143. doi: 10.1086/345647

Dyson, H. J., and Wright, P. E. (2016). Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J. Biol. Chem.* 291, 6714–6722. doi: 10.1074/jbc.R115.692020

Gannon, T., Perveen, R., Schlecht, H., Ramsden, S., Anderson, B., Kerr, B., et al. (2015). Further delineation of the KAT6B molecular and phenotypic spectrum. *Eur. J. Hum. Genet.* 23, 1165–1170. doi: 10.1038/ejhg.2014.248

Gervasini, C., Parenti, I., Picinelli, C., Azzollini, J., Masciadri, M., Cereda, A., et al. (2013). Molecular characterization of a mosaic NIPBL deletion in a cornelia de lange patient with severe phenotype. *Eur. J. Med. Genet.* 56, 138–143. doi: 10.1016/j.ejmg.2012.12.009

Goldblatt, J., Wallis, C., and Zieff, S. (1988). A syndrome of hypoplastic patellae, mental retardation, skeletal and genitourinary anomalies with normal chromosomes. *Dysmorph. Clin. Genet.* 2, 91–93.

Gräff, J., and Tsai, L.-H. (2013). Histone acetylation: molecular mnemonics on the chromatin. *Nat. Rev. Neurosci.* 14, 97–111. doi: 10.1038/nrn3427

Greenfield, A., Carrel, L., Pennisi, D., Philippe, C., Quaderi, N., Siggers, P., et al. (1998). The UTX gene escapes X inactivation in mice and humans. *Hum. Mol. Genet.* 7, 737–742. doi: 10.1093/hmg/7.4.737

Gupta, S., Kim, S. Y., Artis, S., Molfese, D. L., Schumacher, A., Sweatt, J. D., et al. (2010). Histone methylation regulates memory formation. *J. Neurosci.* 30, 3589–3599. doi: 10.1523/JNEUROSCI.3732-09.2010

Hannibal, M. C., Buckingham, K. J., Ng, S. B., Ming, J. E., Beck, A. E., McMillin, M. J., et al. (2011). Spectrum of MLL2 (ALR) mutations in 110 cases of kabuki syndrome. *Am. J. Med. Genet. A* 155, 1511–1516. doi: 10.1002/ajmg.a.34074

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112. doi: 10.1038/nature07829

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39, 311–318. doi: 10.1038/ng1966

Hong, S., Cho, Y.-W., Yu, L.-R., Yu, H., Veenstra, T. D., and Ge, K. (2007). Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18439–18444. doi: 10.1073/pnas.0707292104

Huang, F., Abmayr, S. M., and Workman, J. L. (2016). Regulation of KAT6 acetyltransferases and their roles in cell cycle progression, stem cell maintenance, and human disease. *Mol. Cell. Biol.* 36, 1900–1907. doi: 10.1128/MCB.00055-16

Huang, N., vom Baur, E., Garnier, J. M., Lerouge, T., Vonesch, J. L., Lutz, Y., et al. (1998). Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J.* 17, 3398–3412. doi: 10.1093/emboj/17.12.3398

Hübner, M. R., and Spector, D. L. (2010). Role of H3K27 demethylases Jmjd3 and UTX in transcriptional regulation. *Cold Spring Harb. Symp. Quant. Biol.* 75, 43–49. doi: 10.1101/sqb.2010.75.020

Hutchinson, D. T., and Sullivan, R. (2015). Rubinstein-taybi syndrome. *J. Hand Surg.* 40, 1711–1712. doi: 10.1016/j.jhsa.2014.08.043

Iacono, G., Dubos, A., Méziane, H., Benevento, M., Habibi, E., Mandoli, A., et al. (2018). Increased H3K9 methylation and impaired expression of Protocadherins are associated with the cognitive dysfunctions of the Kleefstra syndrome. *Nucleic Acids Res.* 46, 4950–4965. doi: 10.1093/nar/gky196

Iwase, S., Brookes, E., Agarwal, S., Badeaux, A. I., Ito, H., Vallianatos, C. N., et al. (2016). A mouse model of X-linked intellectual disability associated with impaired removal of histone methylation. *Cell Rep.* 14, 1000–1009. doi: 10.1016/j.celrep.2015.12.091

Iyer, N. G., Özdag, H., and Caldas, C. (2004). p300/CBP and cancer. *Oncogene* 23, 4225–4231. doi: 10.1038/sj.onc.1207118

Jaju, R. J., Fidler, C., Haas, O. A., Strickson, A. J., Watkins, F., Clark, K., et al. (2001). A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in *de novo* childhood acute myeloid leukemia. *Blood* 98, 1264–1267. doi: 10.1182/blood.V98.4.1264

Jakovcevski, M., Ruan, H., Shen, E. Y., Dincer, A., Javidfar, B., Ma, Q., et al. (2015). Neuronal Kmt2a/Mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory. *J. Neurosci.* 35, 5097–5108. doi: 10.1523/JNEUROSCI.3004-14.2015

Jones, W. D., Dafou, D., McEntagart, M., Woollard, W. J., Elmslie, F. V., Holder-Espinasse, M., et al. (2012). *De novo* mutations in MLL cause wiedemann-steiner syndrome. *Am. J. Hum. Genet.* 91, 358–364. doi: 10.1016/j.ajhg.2012.06.008

Kaminsky, E. B., Kaul, V., Paschall, J., Church, D. M., Bunke, B., Kunig, D., et al. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 13, 777–784. doi: 10.1097/GIM.0b013e31822c79f9

Karlic, R., Chung, H.-R., Lasserre, J., Vlahovicek, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2926–2931. doi: 10.1073/pnas.0909344107

Kennedy, J., Goudie, D., Blair, E., Chandler, K., Joss, S., McKay, V., et al. (2019). KAT6A Syndrome: genotype-phenotype correlation in 76 patients with pathogenic KAT6A variants. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 21, 850–860. doi: 10.1038/s41436-018-0259-2

Kim, J.-H., Lee, J. H., Lee, I.-S., Lee, S. B., and Cho, K. S. (2017). Histone lysine methylation and neurodevelopmental disorders. *Int. J. Mol. Sci.* 18:1404. doi: 10.3390/ijms18071404

Kim, M.-S., Akhtar, M. W., Adachi, M., Mahgoub, M., Bassel-Duby, R., Kavalali, E. T., et al. (2012). An essential role for histone deacetylase 4 in synaptic plasticity and memory formation. *J. Neurosci. Off. J. Soc. Neurosci.* 32, 10879–10886. doi: 10.1523/JNEUROSCI.2089-12.2012

Kim, S. Y., Levenson, J. M., Korsmeyer, S., Sweatt, J. D., and Schumacher, A. (2007). Developmental regulation of eed complex composition governs a switch in global histone modification in brain. *J. Biol. Chem.* 282, 9962–9972. doi: 10.1074/jbc.M608722200

Klein, B. J., Lalonde, M.-E., Côté, J., Yang, X.-J., and Kutateladze, T. G. (2014). Crosstalk between epigenetic readers regulates the MOZ/MORF HAT complexes. *Epigenetics* 9, 186–193. doi: 10.4161/epi.26792

Klose, R. J., Kallin, E. M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* 7, 715–727. doi: 10.1038/nrg1945

Klose, R. J., and Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell Biol.* 8, 307–318. doi: 10.1038/nrm2143

Koenig, R., Meinecke, P., Kuechler, A., Schäfer, D., and Müller, D. (2010). Wiedemann–steiner syndrome: three further cases. *Am. J. Med. Genet. A.* 152A, 2372–2375. doi: 10.1002/ajmg.a.33587

Korzus, E. (2017). Rubinstein-taybi syndrome and epigenetic alterations. *Adv. Exp. Med. Biol.* 978, 39–62. doi: 10.1007/978-3-319-53889-1_3

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. doi: 10.1016/j.cell.2007.02.005

Kuroki, Y., Suzuki, Y., Chyo, H., Hata, A., and Matsui, I. (1981). A new malformation syndrome of long palpebralfissures, large ears, depressed nasal

tip, and skeletal anomalies associated with postnatal dwarfism and mental retardation. *J. Pediatr.* 99, 570–573. doi: 10.1016/S0022-3476(81)80256-9

Kurotaki, N., Harada, N., Yoshiura, K., Sugano, S., Niikawa, N., and Matsumoto, N. (2001). Molecular characterization of NSD1, a human homologue of the mouse Nsd1 gene. *Gene* 279, 197–204. doi: 10.1016/S0378-1119(01)00750-8

Kurotaki, N., Imaizumi, K., Harada, N., Masuno, M., Kondoh, T., Nagai, T., et al. (2002). Haploinsufficiency of NSD1 causes sotos syndrome. *Nat. Genet.* 30, 365–366. doi: 10.1038/ng863

Lan, F., Bayliss, P. E., Rinn, J. L., Whetstine, J. R., Wang, J. K., Chen, S., et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 449, 689–694. doi: 10.1038/nature06192

Le, T. N., Williams, S. R., Alaimo, J. T., and Elsea, S. H. (2019). Genotype and phenotype correlation in 103 individuals with 2q37 deletion syndrome reveals incomplete penetrance and supports HDAC4 as the primary genetic contributor. *Am. J. Med. Genet. A* 179, 782–791. doi: 10.1002/ajmg.a.61089

Lebrun, N., Giurgea, I., Goldenberg, A., Dieux, A., Afenjar, A., Ghoumid, J., et al. (2018). Molecular and cellular issues of KMT2A variants involved in wiedemann-steiner syndrome. *Eur. J. Hum. Genet.* 26, 107–116. doi: 10.1038/s41431-017-0033-y

Lederer, D., Grisart, B., Digilio, M. C., Benoit, V., Crespin, M., Ghariani, S. C., et al. (2012). Deletion of KDM6A, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. *Am. J. Hum. Genet.* 90, 119–124. doi: 10.1016/j.ajhg.2011.11.021

Lederer, D., Shears, D., Benoit, V., Verellen-Dumoulin, C., and Maystadt, I. (2014). A three generation X-linked family with Kabuki syndrome phenotype and a frameshift mutation in KDM6A. *Am. J. Med. Genet. A.* 164, 1289–1292. doi: 10.1002/ajmg.a.36442

Lee, J.-E., Wang, C., Xu, S., Cho, Y.-W., Wang, L., Feng, X., et al. (2013). H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* 2:e01503. doi: 10.7554/eLife.01503.027

Lee, M. G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* 437, 432–435. doi: 10.1038/nature04021

Lee, S., Lee, J. W., and Lee, S.-K. (2012). UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. *Dev. Cell* 22, 25–37. doi: 10.1016/j.devcel.2011.11.009

Leroy, C., Landais, E., Briault, S., David, A., Tassy, O., Gruchy, N., et al. (2013). The 2q37-deletion syndrome: an update of the clinical spectrum including overweight, brachydactyly and behavioural features in 14 new patients. *Eur. J. Hum. Genet.* 21, 602–612. doi: 10.1038/ejhg.2012.230

Li, N., Wang, Y., Yang, Y., Wang, P., Huang, H., Xiong, S., et al. (2018). Description of the molecular and phenotypic spectrum of Wiedemann-Steiner syndrome in Chinese patients. *Orphanet J. Rare Dis.* 13:178. doi: 10.1186/s13023-018-0909-0

Li, Y., Bögershausen, N., Alanay, Y., Simsek Kiper, P. Ö., Plume, N., Keupp, K., et al. (2011). A mutation screen in patients with Kabuki syndrome. *Hum. Genet.* 130, 715–724. doi: 10.1007/s00439-011-1004-y

Li, Y., Trojer, P., Xu, C.-F., Cheung, P., Kuo, A., Drury, W. J., et al. (2009). The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem.* 284, 34283–34295. doi: 10.1074/jbc.M109.034462

Lindgren, A. M., Hoyos, T., Talkowski, M. E., Hanscom, C., Blumenthal, I., Chiang, C., et al. (2013). Haploinsufficiency of KDM6A is associated with severe psychomotor retardation, global growth restriction, seizures and cleft palate. *Hum. Genet.* 132, 537–552. doi: 10.1007/s00439-013-1263-x

Lipinski, M., Muñoz-Viana, R., del Blanco, B., Marquez-Galera, A., Medrano-Relinque, J., Caramés, J. M., et al. (2020). KAT3-dependent acetylation of cell type-specific genes maintains neuronal identity in the adult mouse brain. *Nat. Commun.* 11:2588. doi: 10.1038/s41467-020-16246-0

Lonardo, F., Lonardo, M. S., Acquaviva, F., Della Monica, M., Scarano, F., and Scarano, G. (2019). Say-barber-biesecker-young-simpson syndrome and genitopatellar syndrome: lumping or splitting?: LONARDO et al. *Clin. Genet.* 95, 253–261. doi: 10.1111/cge.13127

López, M., García-Oguiza, A., Armstrong, J., García-Cobaleda, I., García-Miñaur, S., Santos-Simarro, F., et al. (2018). Rubinstein-Taybi 2 associated to novel EP300 mutations: deepening the clinical and genetic spectrum. *BMC Med. Genet.* 19:36. doi: 10.1186/s12881-018-0548-2

Lopez-Atalaya, J. P., Gervasini, C., Mottadelli, F., Spena, S., Piccione, M., Scarano, G., et al. (2012). Histone acetylation deficits in lymphoblastoid cell lines from patients with Rubinstein–Taybi syndrome. *J. Med. Genet.* 49, 66–74. doi: 10.1136/jmedgenet-2011-100354

Lopez-Atalaya, J. P., Valor, L. M., and Barco, A. (2014). "Epigenetic factors in intellectual disability," in *Progress in Molecular Biology and Translational Science* eds Akbarian and Lubin (Academic Press), 139–176.

Lucio-Eterovic, A. K., Singh, M. M., Gardner, J. E., Veerappan, C. S., Rice, J. C., and Carpenter, P. B. (2010). Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16952–16957. doi: 10.1073/pnas.1002653107

Lui, J. C., Barnes, K. M., Dong, L., Yue, S., Graber, E., Rapaport, R., et al. (2018). Ezh2 Mutations found in the weaver overgrowth syndrome cause a partial loss of H3K27 histone methyltransferase activity. *J. Clin. Endocrinol. Metab.* 103, 1470–1478. doi: 10.1210/jc.2017-01948

Lundsgaard, M., Le, V. Q., Ernst, A., Laugaard-Jacobsen, H. C., Rasmussen, K., Pedersen, I. S., et al. (2017). *De novo* KAT6B mutation identified with whole-exome sequencing in a girl with say-barber/biesecker/young-simpson syndrome. *Mol. Syndromol.* 8, 24–29. doi: 10.1159/000452258

Majdzadeh, N., Wang, L., Morrison, B. E., Bassel-Duby, R., Olson, E. N., and D'Mello, S. R. (2008). HDAC4 inhibits cell-cycle progression and protects neurons from cell death. *Dev. Neurobiol.* 68, 1076–1092. doi: 10.1002/dneu.20637

Margueron, R., and Reinberg, D. (2011). The polycomb complex PRC2 and its mark in life. *Nature* 469, 343–349. doi: 10.1038/nature09784

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838–849. doi: 10.1038/nrm1761

McQuown, S. C., and Wood, M. A. (2011). HDAC3 and the molecular brake pad hypothesis. *Neurobiol. Learn. Mem.* 96, 27–34. doi: 10.1016/j.nlm.2011.04.005

Merson, T. D., Dixon, M. P., Collin, C., Rietze, R. L., Bartlett, P. F., Thomas, T., et al. (2006). The transcriptional coactivator querkopf controls adult neurogenesis. *J. Neurosci. Off. J. Soc. Neurosci.* 26, 11359–11370. doi: 10.1523/JNEUROSCI.2247-06.2006

Metzger, E., Wissmann, M., Yin, N., Müller, J. M., Schneider, R., Peters, A. H. F. M., et al. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436–439. doi: 10.1038/nature04020

Micale, L., Augello, B., Maffeo, C., Selicorni, A., Zucchetti, F., Fusco, C., et al. (2014). Molecular analysis, pathogenic mechanisms, and readthrough therapy on a large cohort of kabuki syndrome patients. *Hum. Mutat.* 35, 841–850. doi: 10.1002/humu.22547

Mielcarek, M., Seredenina, T., Stokes, M. P., Osborne, G. F., Landles, C., Inuabasi, L., et al. (2013). HDAC4 does not act as a protein deacetylase in the postnatal murine brain *in vivo*. *PLoS ONE* 8:e80849. doi: 10.1371/journal.pone.0080849

Mielcarek, M., Zielonka, D., Carnemolla, A., Marcinkowski, J. T., and Guidez, F. (2015). HDAC4 as a potential therapeutic target in neurodegenerative diseases: a summary of recent achievements. *Front. Cell. Neurosci.* 9:42. doi: 10.3389/fncel.2015.00042

Migdalska, A. M., van der Weyden, L., Ismail, O., Rust, A. G., Rashid, M., White, J. K., et al. (2012). Generation of the sotos syndrome deletion in mice. *Mamm. Genome* 23, 749–757. doi: 10.1007/s00335-012-9416-0

Millan, F., Cho, M. T., Retterer, K., Monaghan, K. G., Bai, R., Vitazka, P., et al. (2016). Whole exome sequencing reveals *de novo* pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. *Am. J. Med. Genet. A* 170, 1791–1798. doi: 10.1002/ajmg.a.37670

Millan, M. J. (2013). An epigenetic framework for neurodevelopmental disorders: from pathogenesis to potential therapy. *Neuropharmacology* 68, 2–82. doi: 10.1016/j.neuropharm.2012.11.015

Miller, C. T., Maves, L., and Kimmel, C. B. (2004). moz regulates hox expression and pharyngeal segmental identity in zebrafish. *Development* 131, 2443–2461. doi: 10.1242/dev.01134

Miller, S. A., Mohn, S. E., and Weinmann, A. S. (2010). Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol. Cell* 40, 594–605. doi: 10.1016/j.molcel.2010.10.028

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., et al. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* 10, 1107–1117. doi: 10.1016/S1097-2765(02)00741-4

Miyake, N., Koshimizu, E., Okamoto, N., Mizuno, S., Ogata, T., Nagai, T., et al. (2013a). MLL2 and KDM6A mutations in patients with Kabuki syndrome. *Am. J. Med. Genet. A.* 161, 2234–2243. doi: 10.1002/ajmg.a.36072

Miyake, N., Mizuno, S., Okamoto, N., Ohashi, H., Shiina, M., Ogata, K., et al. (2013b). KDM6A point mutations cause kabuki syndrome. *Hum. Mutat.* 34, 108–110. doi: 10.1002/humu.22229

Miyake, N., Tsurusaki, Y., Koshimizu, E., Okamoto, N., Kosho, T., Brown, N. J., et al. (2016). Delineation of clinical features in wiedemann–steiner syndrome caused by KMT2A mutations. *Clin. Genet.* 89, 115–119. doi: 10.1111/cge.12586

Morgan, H. D., Sutherland, H. G. E., Martin, D. I. K., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23, 314–318. doi: 10.1038/15490

Morris, B., Etoubleau, C., Bourthoumieu, S., Reynaud-Perrine, S., Laroche, C., Lebbar, A., et al. (2012). Dose dependent expression of HDAC4 causes variable expressivity in a novel inherited case of brachydactyly mental retardation syndrome. *Am. J. Med. Genet. A* 158A, 2015–2020. doi: 10.1002/ajmg.a.35463

Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., et al. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* 42, 790–793. doi: 10.1038/ng.646

Niemi, M. E. K., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M., et al. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* 562, 268–271. doi: 10.1038/s41586-018-0566-4

Niikawa, N., Kuroki, Y., Kajii, T., Matsuura, N., Ishikiriyama, S., Tonoki, H., et al. (1988). Kabuki make-up (Niikawa-Kuroki) syndrome: a study of 62 patients. *Am. J. Med. Genet.* 31, 565–589. doi: 10.1002/ajmg.1320310312

Niikawa, N., Matsuura, N., Fukushima, Y., Ohsawa, T., and Kajii, T. (1981). Kabuki make-up syndrome: a syndrome of mentalretardation, unusual facies, large and protruding ears, and postnatal growth deficiency. *J. Pediatr.* 99, 565–569. doi: 10.1016/S0022-3476(81)80255-7

Nimura, K., Ura, K., Shiratori, H., Ikawa, M., Okabe, M., Schwartz, R. J., et al. (2009). A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to wolf–hirschhorn syndrome. *Nature* 460, 287–291. doi: 10.1038/nature08086

Ohdo, S., Madokoro, H., Sonoda, T., and Hayakawa, K. (1986). Mental retardation associated with congenital heart disease, blepharophimosis, blepharoptosis, and hypoplastic teeth. *J. Med. Genet.* 23, 242–244. doi: 10.1136/jmg.23.3.242

Oliveira, A. M. M., Wood, M. A., McDonough, C. B., and Abel, T. (2007). Transgenic mice expressing an inhibitory truncated form of p300 exhibit long-term memory deficits. *Learn. Mem.* 14, 564–572. doi: 10.1101/lm.656907

Park, E., Kim, Y., Ryu, H., Kowall, N. W., Lee, J., and Ryu, H. (2014). Epigenetic mechanisms of rubinstein–taybi syndrome. *NeuroMol. Med.* 16, 16–24. doi: 10.1007/s12017-013-8285-3

Pasillas, M. P., Shah, M., and Kamps, M. P. (2011). NSD1 PHD domains bind methylated H3K4 and H3K9 using interactions disrupted by point mutations in human sotos syndrome. *Hum. Mutat.* 32, 292–298. doi: 10.1002/humu.21424

Petrij, F., Giles, R. H., Dauwerse, H. G., Saris, J. J., Hennekam, R. C. M., Masuno, M., et al. (1995). Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* 376, 348–351. doi: 10.1038/376348a0

Podobinska, M., Szablowska-Gadomska, I., Augustyniak, J., Sandvig, I., Sandvig, A., and Buzanska, L. (2017). Epigenetic modulation of stem cells in neurodevelopment: the role of methylation and acetylation. *Front. Cell. Neurosci.* 11:23. doi: 10.3389/fncel.2017.00023

Prasad, R., Zhadanov, A. B., Sedkov, Y., Bullrich, F., Druck, T., Rallapalli, R., et al. (1997). Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to drosophila trithorax. *Oncogene* 15, 549–560. doi: 10.1038/sj.onc.1201211

Price, V., Wang, L., and D'Mello, S. R. (2013). Conditional deletion of histone deacetylase-4 in the central nervous system has no major effect on brain architecture or neuronal viability. *J. Neurosci. Res.* 91, 407–415. doi: 10.1002/jnr.23170

Radvanszky, J., Hyblova, M., Durovcikova, D., Hikkelova, M., Fiedler, E., Kadasi, L., et al. (2017). Complex phenotypes blur conventional borders between say-barber-biesecker-young-simpson syndrome and genitopatellar syndrome: complex phenotypes blur conventional borders. *Clin. Genet.* 91, 339–343. doi: 10.1111/cge.12840

Rayasam, G. V., Wendling, O., Angrand, P.-O., Mark, M., Niederreither, K., Song, L., et al. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* 22, 3153–3163. doi: 10.1093/emboj/cdg288

Russo, V. E. A., Martienssen, R. A., and Riggs, A. D. (eds.). (1996). *Epigenetic Mechanisms of Gene Regulation.* Plainview, NY: Cold Spring Harbor Laboratory Press.

Sando, R., Gounko, N., Pieraut, S., Liao, L., Yates, J., and Maximov, A. (2012). HDAC4 Governs a transcriptional program essential for synaptic plasticity and memory. *Cell* 151, 821–834. doi: 10.1016/j.cell.2012.09.037

Schaefer, G. B., Bodensteiner, J. B., Buehler, B. A., Lin, A., and Cole, T. R. P. (1997). The neuroimaging findings in Sotos syndrome. *Am. J. Med. Genet.* 68, 462–465. doi: 10.1002/(SICI)1096-8628(19970211)68:4<462::AID-AJMG18>3.0.CO;2-Q

Seto, E., and Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb. Perspect. Biol.* 6:a018713. doi: 10.1101/cshperspect.a018713

Sheikh, B. N., and Akhtar, A. (2019). The many lives of KATs — detectors, integrators and modulators of the cellular environment. *Nat. Rev. Genet.* 20, 7–23. doi: 10.1038/s41576-018-0072-4

Sheikh, B. N., Dixon, M. P., Thomas, T., and Voss, A. K. (2012). Querkopf is a key marker of self-renewal and multipotency of adult neural stem cells. *J. Cell Sci.* 125, 295–309. doi: 10.1242/jcs.077271

Shi, Y., Lan, F., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., et al. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941–953. doi: 10.1016/j.cell.2004.12.012

Shi, Y., Sawada, J., Sui, G., Affar, E. B., Whetstine, J. R., Lan, F., et al. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* 422, 735–738. doi: 10.1038/nature01550

Shi, Y.-J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell* 19, 857–864. doi: 10.1016/j.molcel.2005.08.027

Shintani, T., Takeuchi, Y., Fujikawa, A., and Noda, M. (2012). Directional neuronal migration is impaired in mice lacking adenomatous polyposis coli 2. *J. Neurosci.* 32, 6468–6484. doi: 10.1523/JNEUROSCI.0590-12.2012

Shirane, K., Miura, F., Ito, T., and Lorincz, M. C. (2020). NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts polycomb-associated silencing. *Nat. Genet.* 52, 1088–1098. doi: 10.1038/s41588-020-0689-z

Shpargel, K. B., Sengoku, T., Yokoyama, S., and Magnuson, T. (2012). UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. *PLoS Genet.* 8:e1002964. doi: 10.1371/journal.pgen.1002964

Shpargel, K. B., Starmer, J., Wang, C., Ge, K., and Magnuson, T. (2017). UTX-guided neural crest function underlies craniofacial features of Kabuki syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9046–E9055. doi: 10.1073/pnas.1705011114

Sotos, J. F., Dodge, P. R., Muirhead, D., Crawford, J. D., and Talbot, N. B. (1964). Cerebral gigantism in childhood. A syndrome of excessively rapid growth and acromegalic features and a nonprogressive neurologic disorder. *N. Engl. J. Med.* 271, 109–116. doi: 10.1056/NEJM196407162710301

Steiner, C. E., and Marques, A. P. (2000). Growth deficiency, mental retardation and unusual facies. *Clin. Dysmorphol.* 9, 155–156. doi: 10.1097/00019605-200009020-00021

Stellacci, E., Onesimo, R., Bruselles, A., Pizzi, S., Battaglia, D., Leoni, C., et al. (2016). Congenital immunodeficiency in an individual with wiedemann-steiner syndrome due to a novel missense mutation in KMT2A. *Am. J. Med. Genet. A* 170, 2389–2393. doi: 10.1002/ajmg.a.37681

Sterner, D. E., and Berger, S. L. (2000). Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* 64, 435–459. doi: 10.1128/MMBR.64.2.435-459.2000

Tatton-Brown, K., Douglas, J., Coleman, K., Baujat, G., Cole, T. R. P., Das, S., et al. (2005). Genotype-phenotype associations in sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am. J. Hum. Genet.* 77, 193–204. doi: 10.1086/432082

Tham, E., Lindstrand, A., Santani, A., Malmgren, H., Nesbitt, A., Dubbs, H. A., et al. (2015). Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. *Am. J. Hum. Genet.* 96, 507–513. doi: 10.1016/j.ajhg.2015.01.016

Thapar, A., Cooper, M., and Rutter, M. (2017). Neurodevelopmental disorders. *Lancet Psychiatry* 4, 339–346. doi: 10.1016/S2215-0366(16)30376-5

Thomas, T., Corcoran, L. M., Gugasyan, R., Dixon, M. P., Brodnicki, T., Nutt, S. L., et al. (2006). Monocytic leukemia zinc finger protein is essential for the development of long-term reconstituting hematopoietic stem cells. *Genes Dev.* 20, 1175–1186. doi: 10.1101/gad.1382606

Thomas, T., and Voss, A. K. (2004). Querkopf, a histone acetyltransferase, is essential for embryonic neurogenesis. *Front. Biosci. J. Virtual Libr.* 9, 24–31. doi: 10.2741/1208

Thomas, T., Voss, A. K., Chowdhury, K., and Gruss, P. (2000). Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development. *Dev. Camb. Engl.* 127, 2537–2548

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., et al. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811–816. doi: 10.1038/nature04433

Türkmen, S., Gillessen-Kaesbach, G., Meinecke, P., Albrecht, B., Neumann, L. M., Hesse, V., et al. (2003). Mutations in NSD1 are responsible for Sotos syndrome, but are not a frequent finding in other overgrowth phenotypes. *Eur. J. Hum. Genet.* 11, 858–865. doi: 10.1038/sj.ejhg.5201050

Valor, M. L., Viosca, J., P., Lopez-Atalaya, J., and Barco, A. (2013). Lysine acetyltransferases CBP and p300 as therapeutic targets in cognitive and neurodegenerative disorders. *Curr. Pharm. Des.* 19, 5051–5064. doi: 10.2174/13816128113199990382

van der Vlag, J., and Otte, A. P. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat. Genet.* 23, 474–478. doi: 10.1038/70602

van Haelst, M. M., Hoogeboom, J. J. M., Baujat, G., Brüggenwirth, H. T., Van de Laar, I., Coleman, K., et al. (2005). Familial gigantism caused by an NSD1 mutation. *Am. J. Med. Genet. A.* 139, 40–44. doi: 10.1002/ajmg.a.30973

Van Laarhoven, P. M., Neitzel, L. R., Quintana, A. M., Geiger, E. A., Zackai, E. H., Clouthier, D. E., et al. (2015). Kabuki syndrome genes KMT2D and KDM6A: functional analyses demonstrate critical roles in craniofacial, heart and brain development. *Hum. Mol. Genet.* 24, 4443–4453. doi: 10.1093/hmg/ddv180

Vanyai, H. K., Garnham, A., May, R. E., McRae, H. M., Collin, C., Wilcox, S., et al. (2019). MOZ directs the distal-less homeobox gene expression program during craniofacial development. *Development* 146:dev175042. doi: 10.1242/dev.175042

Vega, R. B., Matsuda, K., Oh, J., Barbosa, A. C., Yang, X., Meadows, E., et al. (2004). Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* 119, 555–566. doi: 10.1016/j.cell.2004.10.024

Villavicencio-Lorini, P., Klopocki, E., Trimborn, M., Koll, R., Mundlos, S., and Horn, D. (2013). Phenotypic variant of brachydactyly-mental retardation syndrome in a family with an inherited interstitial 2q37.3 microdeletion including HDAC4. *Eur. J. Hum. Genet.* 21, 743–748. doi: 10.1038/ejhg.2012.240

Viosca, J., Lopez-Atalaya, J. P., Olivares, R., Eckner, R., and Barco, A. (2010). Syndromic features and mild cognitive impairment in mice with genetic reduction on p300 activity: differential contribution of p300 and CBP to Rubinstein–Taybi syndrome etiology. *Neurobiol. Dis.* 37, 186–194. doi: 10.1016/j.nbd.2009.10.001

Visser, R., Shimokawa, O., Harada, N., Kinoshita, A., Ohta, T., Niikawa, N., et al. (2005). Identification of a 3.0-kb major recombination hotspot in patients with Sotos syndrome who carry a common 1.9-Mb microdeletion. *Am. J. Hum. Genet.* 76, 52–67. doi: 10.1086/426950

Vlckova, M., Simandlova, M., Zimmermann, P., Stranecky, V., Hartmannova, H., Hodanova, K., et al. (2015). A patient showing features of both SBBYSS and GPS supports the concept of a KAT6B-related disease spectrum, with mutations in mid-exon 18 possibly leading to combined phenotypes. *Eur. J. Med. Genet.* 58, 550–555. doi: 10.1016/j.ejmg.2015.09.004

Voss, A. K., Collin, C., Dixon, M. P., and Thomas, T. (2009). Moz and retinoic acid coordinately regulate H3K9 acetylation, hox gene expression, and segment identity. *Dev. Cell* 17, 674–686. doi: 10.1016/j.devcel.2009.10.006

Voss, A. K., Vanyai, H. K., Collin, C., Dixon, M. P., McLennan, T. J., Sheikh, B. N., et al. (2012). MOZ regulates the Tbx1 locus, and moz mutation partially phenocopies digeorge syndrome. *Dev. Cell* 23, 652–663. doi: 10.1016/j.devcel.2012.07.010

Wang, H. (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* 293, 853–857. doi: 10.1126/science.1060781

Wang, J. K., Tsai, M.-C., Poulin, G., Adler, A. S., Chen, S., Liu, H., et al. (2010). The histone demethylase UTX enables RB-dependent cell fate control. *Genes Dev.* 24, 327–332. doi: 10.1101/gad.1882610

Wheeler, P. G., Huang, D., and Dai, Z. (2014). Haploinsufficiency of HDAC4 does not cause intellectual disability in all affected individuals. *Am. J. Med. Genet. A* 164, 1826–1829. doi: 10.1002/ajmg.a.36542

Wiedemann, H., Kunze, J., and Dibbern, H. (1989). *Atlas of Clinical Syndromes. 2nd Edn*. London: Wolfe Publishing Ltd.

Williams, S. R., Aldred, M. A., Der Kaloustian, V. M., Halal, F., Gowans, G., McLeod, D. R., et al. (2010). Haploinsufficiency of HDAC4 causes brachydactyly mental retardation syndrome, with brachydactyly type E, developmental delays, and behavioral problems. *Am. J. Hum. Genet.* 87, 219–228. doi: 10.1016/j.ajhg.2010.07.011

Xia, Z.-B., Anderson, M., Diaz, M. O., and Zeleznik-Le, N. J. (2003). MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8342–8347. doi: 10.1073/pnas.1436338100

Xu, J., Deng, X., Watkins, R., and Disteche, C. M. (2008). Sex-specific differences in expression of histone demethylases utx and uty in mouse brain and neurons. *J. Neurosci.* 28, 4521–4527. doi: 10.1523/JNEUROSCI.5382-07.2008

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J., et al. (2006). JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* 125, 483–495. doi: 10.1016/j.cell.2006.03.027

Yang, X.-J. (2015). MOZ and MORF acetyltransferases: molecular interaction, animal development and human disease. *Biochim. Biophys. Acta BBA Mol. Cell Res.* 1853, 1818–1826. doi: 10.1016/j.bbamcr.2015.04.014

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J., and Korsmeyer, S. J. (1995). Altered hox expression and segmental identity in Mll -mutant mice. *Nature* 378, 505–508. doi: 10.1038/378505a0

Yu, C., Yao, X., Zhao, L., Wang, P., Zhang, Q., Zhao, C., et al. (2017). Wolf–hirschhorn syndrome candidate 1 (whsc1) functions as a tumor suppressor by governing cell differentiation. *Neoplasia* 19, 606–616. doi: 10.1016/j.neo.2017.05.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 21 January 2021
doi: 10.3389/fgene.2021.620859

MeCP2: The Genetic Driver of Rett Syndrome Epigenetics

*Katrina V. Good[1], John B. Vincent[2,3,4] and Juan Ausió[1]**

[1] Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada, [2] Molecular Neuropsychiatry & Development (MiND) Lab, Centre for Addiction and Mental Health, Campbell Family Mental Health Research Institute, Toronto, ON, Canada, [3] Institute of Medical Science, University of Toronto, Toronto, ON, Canada, [4] Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Mutations in methyl CpG binding protein 2 (MeCP2) are the major cause of Rett syndrome (RTT), a rare neurodevelopmental disorder with a notable period of developmental regression following apparently normal initial development. Such MeCP2 alterations often result in changes to DNA binding and chromatin clustering ability, and in the stability of this protein. Among other functions, MeCP2 binds to methylated genomic DNA, which represents an important epigenetic mark with broad physiological implications, including neuronal development. In this review, we will summarize the genetic foundations behind RTT, and the variable degrees of protein stability exhibited by MeCP2 and its mutated versions. Also, past and emerging relationships that MeCP2 has with mRNA splicing, miRNA processing, and other non-coding RNAs (ncRNA) will be explored, and we suggest that these molecules could be missing links in understanding the epigenetic consequences incurred from genetic ablation of this important chromatin modifier. Importantly, although MeCP2 is highly expressed in the brain, where it has been most extensively studied, the role of this protein and its alterations in other tissues cannot be ignored and will also be discussed. Finally, the additional complexity to RTT pathology introduced by structural and functional implications of the two MeCP2 isoforms (MeCP2-E1 and MeCP2-E2) will be described. Epigenetic therapeutics are gaining clinical popularity, yet treatment for Rett syndrome is more complicated than would be anticipated for a purely epigenetic disorder, which should be taken into account in future clinical contexts.

Keywords: methyl CpG binding protein 2, Rett syndrome, mutations, protein stability, RNA binding

OPEN ACCESS

Edited by:
Mojgan Rastegar,
University of Manitoba, Canada

Reviewed by:
Walter Erwin Kaufmann,
Emory University, United States
Nathalie Berube,
Western University, Canada

****Correspondence:***
Juan Ausió
jausio@uvic.ca

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *23 October 2020*
Accepted: *05 January 2021*
Published: *21 January 2021*

Citation:
Good KV, Vincent JB and Ausió J (2021) MeCP2: The Genetic Driver of Rett Syndrome Epigenetics. Front. Genet. 12:620859. doi: 10.3389/fgene.2021.620859

INTRODUCTION

The term epigenetics has gained much popularity and has gathered the attention of many researchers in recent years. Yet, the term has, at times, been loosely used and quite often in an ambiguous way (Greally, 2018).

In the right context, epi (beyond)-genetics is defined as gene expression alterations resulting from a change in the DNA/chromatin structure which does not involve a change in the underlying DNA nucleotide sequence (i.e., mutations). At the molecular level, this can be elicited by chemical post-replication/post-translational "tags" that mark DNA (Bird, 1993; Greenberg and Bourc'his, 2019), histones (Bannister and Kouzarides, 2011) (the main protein component of chromatin) or other chromosomal and non-chromosomal proteins. These "tags" (for instance DNA methylation

(Greenberg and Bourc'his, 2019) or protein modifications) are written and erased through the action of different enzymes (writers and erasers) and read by transcriptional regulatory cofactors (readers) (Marmorstein and Zhou, 2014; Seto and Yoshida, 2014). Such is the case for the methyl CpG binding protein 2 (MeCP2), a DNA methylation reader protein. This protein, initially thought to have repressor activity (Nan et al., 1997), is now recognized to have both transcriptionally repressive and activating functions through its interaction with different cofactors (Yasui et al., 2007; Chahrour et al., 2008). The protein is able to recognize (or read) DNA and histone methylation marks (Lewis et al., 1992; Thambirajah et al., 2012; Lee et al., 2020) and, hence, it acts as a methylation-dependent transcriptional modulator within the context of chromatin (Li et al., 2020). DNA methylation dysregulation is one of the hallmarks of diseases such as cancer (Baylin and Jones, 2016), and MeCP2 mutations can alter the reading of this mark, as in RTT (Kriaucionis and Bird, 2003), where it can impact the normal activity of cells. Interestingly, MeCP2 has been recognized as a *bona fide* oncogene and has been involved in many cancers (Neupane et al., 2015). In any such instance, these diseases often have a genetic origin with downstream epigenetic effects (Yoon et al., 2020).

From structural and functional perspectives, MeCP2 is a good example of an intrinsically disordered protein (IDP) (Dunker et al., 2001; Hite et al., 2012), given its relatively low contents of secondary and tertiary structure organization in solution. Despite this, the protein can be divided into several well-defined structural/functional domains (Ghosh et al., 2010): NTD; N-terminal; MBD, methyl binding; ID, intervening; TRD, transcription repression; NID, NCoR interaction; and CTD, C-terminal; domains (see **Figure 1**). The TRD includes the nuclear localization sequence (NLS) [amino acids 253–271 in MeCP2-E2 nomenclature (**Figure 1B**)]. Genetic mutations in the coding region of the X-chromosome-linked *MECP2* gene alter the ability with which its encoded protein MeCP2 binds to DNA within the context of chromatin. In particular, those mutations affecting the MBD of the protein (**Figure 1**) which affect the stability (Kucukkal et al., 2015) and affinity (Yang et al., 2016) of its DNA binding and which represents an important aspect of this review.

Besides its ability to bind methylated DNA and histones, MeCP2 was earlier recognized (Jeffery and Nakielny, 2004) and more recently confirmed to be (Castello et al., 2016) an RNA binding protein. This less studied facet of MeCP2 will be described next. Following this, we will focus on an equally less understood role of this protein, namely its function in tissues other than those within the brain, and will finally conclude this review with the controversial potential physiological relevance of the two isoforms, MeCP2-E1 and MeCP2-E2 (**Figure 1**), in RTT (Liyanage and Rastegar, 2014; Martínez de Paz et al., 2019). Because the MeCP2-E2 isoform was the first to be discovered (Lewis et al., 1992), the mutations observed in RTT originally referred to this isoform. Therefore, the amino acid numbers (mutations) referring to the protein sequence of MeCP2 used in the following sections will be those of this isoform (unless otherwise indicated).

Abbreviations: ATRX, α-thalassemia, mental retardation, X-linked protein; AS, alternative splicing; BDNF, brain-derived neurotrophic factor; BRG1, brahma-related gene 1; CD44, cluster of differentiation 44; Cdk10, cyclin-dependent kinase 10; ChIP, chromatin immune-precipitation; CTD, C-terminal domain; DGCR8, DiGeorge syndrome Critical Region 8; DNMT1, DNA methyltransferase 1; DLX1, distal-less homeobox 1; Dlx5, distal-less homeobox 5; eAT-hook, extended AT-hook; FBP11, formin-binding protein 11; FOXG1, fork head box G1; FOXP3, fork head box P3; FRG1, FSHD region gene 1; GABA, gamma aminobutyric acid; EHMT, euchromatic histone-lysine N-methyltransferase; ERK, extracellular signal-regulated kinase; EVF2, *Rattus norvegicus* non-coding RNA; HIF1a, hypoxia inducible factor 1 alpha; hmC, hydroxymethyl cytosine; HD, Huntington's disease; HDAC, histone deacetylase; HMG, high mobility group; HSATII, human satellite II; HTT, huntingtin; HYPC, huntingtin yeast partner C; ID, intervening domain; IDP, intrinsically disordered protein; INDEL, insertions and deletion; lncRNA, long non-coding RNA; Malat1, metastasis-associated lung adenocarcinoma transcript 1; MBD, methyl biding domain; MeCP2, methyl CpG binding protein 2; mTOR, mechanistic target of rapamycin; NCoR, nuclear receptor co-repressor; ncRNA, non-coding RNA; NEAT1, nuclear enriched abundant transcript 1; NFκB1, nuclear factor kappa B subunit 1; NID, NCoR interaction; NLS, nuclear localization sequence; NME, N-methionine excision; NTD, N-terminal domain; NSC, neural stem cell; PDE4D, cAMP-specific 3′-,5′-cyclic phosphodiesterase 4D; PEST, enriched in proline, glutamate, serine, threonine; PPARG, peroxisome proliferator activated receptor-γ; PRC1, polycomb group complex 1; PRMT6, protein arginine methyltransferase 6; PRPF3, pre-mRNA processing factor 3; PTM, post-translational modification; RANKL, receptor activator of nuclear factor-κB ligand; RBD, RNA binding domain; RBP, RNA binding protein; RIP, RNA immunoprecipitation; RNCR3, retinal non-coding RNA 3; RTT, Rett syndrome; Sin3A, switch independent 3 gene encoded protein a; SDCCAG1, serologically defined colon cancer antigen gene 1; SIRT1, Sirtuin 1; SMRT, silencing mediator of retinoic acid and thyroid hormone receptor; SWI/SNF, switch of the mating type/sucrose non-fermenting; TBLR1, transducin beta-Like 1X-Related protein 1; TRD, transcription repression; UPS, ubiquitin–proteasome system; UTR, untranslated region; WWDR, WW domain binding region; YB-1, Y box binding protein 1.

SUBSECTIONS

Genetic Origin of Rett Syndrome

Most of the nucleotide transition mutations in Rett syndrome are the C > T type that take place at CpG hotspots (Wan et al., 1999), and likely reflect variable site methylation in the male germline (Cheadle et al., 2000). Hence, all the sporadic mutations, which represent more than 99% of the individuals affected by this syndrome (Chahrour and Zoghbi, 2007), and which involve *de novo* mutations (Comings, 1986) of the *MECP2* gene, are of paternal origin (Trappe et al., 2001). This paternal origin may be explained by a combination of the elevated levels of methylation and mitotic divisions in the male germline (Driscoll and Migeon, 1990; Shahbazian and Zoghbi, 2002). As in the case of other C > T transition mutations, this is likely to involve methyl cytosine oxidative deamination of abnormally methylated cytosines (Tomatsu et al., 2004).

Rett syndrome is almost exclusively a disease that affects girls (XX), yet is not a disease with epigenetic inheritance, such as Prader-Willi syndrome and Angelman syndrome, where the clinical outcome depends on whether a mutation is transmitted from a paternal or maternal chromosome, and RTT mutations are not epigenetic mutations (epimutations) *per se*. Rather, RTT mutations have epigenetic consequences, as MeCP2 is considered to be a reader of epigenetic signals. Although considered a disease affecting girls, this is not completely exclusive. The vast majority of mutations that lead to RTT occur *de novo* in paternal germline cells (Cheadle et al., 2000), and these can only be transmitted to

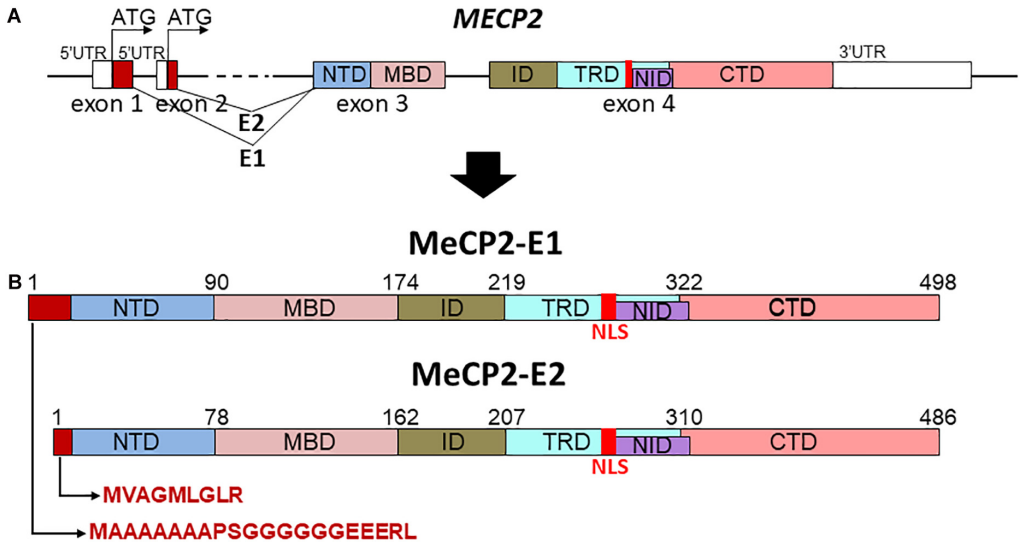

FIGURE 1 | Schematic representation of *MECP2* gene organization **(A)** and different protein domains of the MeCP2-E1 and MeCP2-E2 isoforms **(B)**. NTD, N-terminal; MBD, methyl binding; ID, intervening; TRD, transcription repression; NID, NCoR interaction; CTD, C-terminal domains; NLS, nuclear localization signal.

female offspring and never to males. *De novo MECP2* mutations can occasionally be transmitted from mothers, or inherited from mothers who either have mild cognitive impairment or are asymptomatic, due to skewed X-inactivation favoring expression from their wild-type allele.

Rett syndrome clinical features include regression of motor and communicative skills after 6–18 months of apparently normal development [the reader is referred to Einspieler and Marchik (2019) and Banerjee et al. (2019)] for comprehensive descriptions and review of typical and atypical RTT clinical features. While females born with a *de novo* (from mother's or father's germ cells) or inherited (from a carrier mother) *MECP2* RTT mutation may have a wide spectrum of severity, males with the same *MECP2* mutations typically have much severer consequences, with a more rapid progression of symptoms and lower average age of death (Neul et al., 2019). However, these may be ameliorated in the presence of an additional X chromosome (Klinefelter's syndrome), or where the mutation is a somatic mosaic rather than germline. Also, there are reports of males with *MECP2* mutations that are not known pathogenic RTT mutations who are affected, but not with classical RTT (Neul et al., 2019), and some where the clinical presentation includes psychiatric disorders such as schizophrenia (Cohen et al., 2002; Villard, 2007; McCarthy et al., 2014; Curie et al., 2017; Sheikh et al., 2018), bipolar disorder (Sheikh et al., 2016), and Asperger's (Curie et al., 2017).

MeCP2 Mutations and the High Complexity of MeCP2 Stability

Rett syndrome can arise from a number of missense, nonsense, frame shift, splice site, and start codon mutations as well as larger deletions that can lead to a range of phenotypes with varying degrees of severity (Chahrour and Zoghbi, 2007). From the protein structural point of view (**Figure 1B**), MeCP2 mutations can be grouped into three main broad categories. The first corresponds to mutations that affect the NTD, a second, and corresponding to a very significant group of RTT phenotypes, are those that affect the MBD, and a third, those affecting the rest of the molecule. This classification is not arbitrary as the NTD has been shown to modulate the ability of MeCP2 (through the MBD) to interact with DNA (Martínez de Paz et al., 2019) as well as to influence the turn-over rate of the protein (Sheikh et al., 2017; Martínez de Paz et al., 2019), and hence mutations within this region can affect these parameters. With the MBD being the only structurally ordered portion of MeCP2, mutations within the MBD can affect the tertiary structure (folding) (**Figure 2B**; Kucukkal et al., 2015) of this region and hence its binding affinity (Yang et al., 2016). Many of the remaining mutations are located in the C-terminal domain (Moncla et al., 2002; Bebbington et al., 2010) and can affect the interactions of MeCP2 with many of its diverse interaction partners (Lyst et al., 2013), including the chromatin (Nikitina et al., 2007a) itself and RNA (see following section).

The first attempts to study the functional correlation between RTT MeCP2 mutations and the impairment to DNA methyl binding and transcriptional regulation activities were carried out in the late Alan Wolffe's lab (Yusufzai and Wolffe, 2000) soon after the discovery of their involvement in this disease (Amir et al., 1999). These initial results were subsequently followed by a detailed characterization of the binding affinity alterations caused by several missense mutations within MBD (Ballestar et al., 2000) and set the framework for the type of work which will be described in the following sections.

MeCP2 N-Terminal Mutations

Although not as frequent as the mutations affecting the MBD or TRD at their C-termini (Shah and Bird, 2017; Spiga et al., 2019),

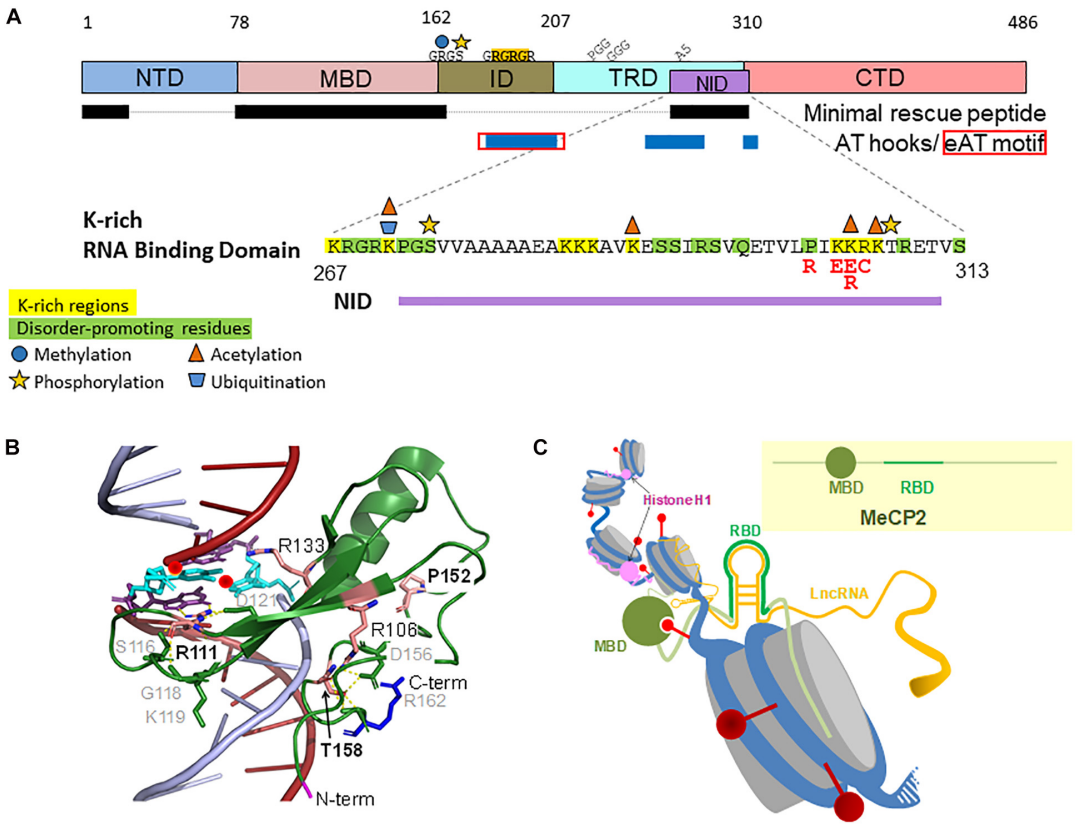

FIGURE 2 | Placement and amino acid sequence and PTMs of the RNA Binding Domain (RBD) within the MeCP2-E2 domains. The main missense RTT mutations are in red **(A)**. Detail of the DNA-bound MBD three dimensional structure (Ho et al., 2008) (Protein Data Bank PDB ID 5BT2) (Berman et al., 2000) **(B)**. Cartoon representation of RNA-bound MeCP2 within a chromatin context **(C)**. Red spheres, DNA methylation; Gray cylinders, nucleosomes.

several NTD mutations have been described to date and are more often than not associated with a typical RTT phenotype (Saunders et al., 2009). However, as with mutations elsewhere in MeCP2, other factors, skewing of X-inactivation, for example, play a role. Without the ability to compare clinical severity across a larger number of RTT girls with the same N-terminal mutation, and without identification of males with these mutations, it is difficult to draw any firm conclusions. However, a study of clinical severity, albeit with only five cases with N-terminal mutations, of which four displayed typical RTT, suggests on average lower clinical severity in comparison to the common nonsense mutations and missense mutations such as T158M (Cuddapah et al., 2014). It is important to note that the two isoforms differ at the N-terminal sequences, with the MeCP2-E2 N-terminus encoded by exon 2, and the slightly longer MeCP2-E1 N-terminus coming from exon 1 (**Figure 1**). It is also worth emphasizing that, until recently, no unequivocal RTT mutations have been reported within exon 2 (or indeed for any clinical entity). Very recently, however, there is a report of a NM_004992:c.7G > C; p.Ala3Pro Rett mutation in exon 2 (Wen et al., 2020). Given its rarity, it will be important to assess the molecular effects of this variant to confirm its true pathogenicity.

The MeCP2-E1 N-terminus contains polyGGC and polyGGA stretches that encode stretches of alanine and glycine residues, respectively. In-frame insertions and deletions within the polyalanine and polyglycine regions of exon 1 (**Figure 1B**) have been identified. Although initially they were suggested to be a relatively frequent cause of intellectual disability or developmental delay (Harvey et al., 2007), with the current availability of exome sequence data from large control populations[1] it is likely that these in-frame indels are unrelated to disease.

Although genuine disease-causing mutations within exon 1 would appear to affect the MeCP2-E1 isoform exclusively, this is not necessarily the case. The first mutation reported in exon 1, an 11 bp frameshifting deletion, was described when the MeCP2-E1 isoform itself was first reported (Mnatzakanian et al., 2004). Since then, the same mutations have been reported in multiple studies (Saunders et al., 2009), and in one study it was shown that, while there was no disruption of transcription of the MeCP2-E2 mRNA, there was interference with, and reduction of translation of MeCP2-E2 protein (Saxena et al., 2006), leading to the possibility that mutations within the MeCP2-E1 N-terminus affect both major isoforms.

However, subsequently, several classic Rett patients were identified with mutations affecting the start codon in exon

[1]gnomad.broadinstitute.org

1 (Gauthier et al., 2005; Saunders et al., 2009). Levels of mRNA for MeCP2-E1 and E2 were unaffected, and peripheral blood lymphocytes were still positive for MeCP2 antibodies, and thus presumably only able to generate the MeCP2-E2 protein (Gianakopoulos et al., 2012). Additionally, a genuine RTT missense mutation (A2V) within the same N-terminal region of MeCP2-E1 was reported in two patients (Fichou et al., 2009; Saunders et al., 2009). The mutation resulted in an RTT phenotype characterized by severe epilepsy, cognitive impairment and developmental delay, in addition to microcephaly and no language in one patient (Fichou et al., 2009), and is described as classic Rett syndrome in the other (Saunders et al., 2009). Importantly, neither the transcriptional nor the translational properties of MeCP2-E2 were affected as observed in fibroblasts or lymphocytes obtained from the patients (Fichou et al., 2009; Gianakopoulos et al., 2012). An apparent synonymous or silent mutation in exon 1, p.Gly16Gly, that was shown to trigger usage of a cryptic splice donor resulting in a frameshift and premature truncation for the MeCP2-E1 isoform but with no predicted effect on the MeCP2-E2 isoform has also been documented (Sheikh et al., 2013). A study on the cellular and molecular effects caused by the A2V mutation showed that, while it neither impacted the localization of the MeCP2-E1 isoform nor its co-localization with chromatin, it affected the N-terminal co- and post-translational modifications that regulate the physiological turnover of the protein. Complete N-methionine excision (NME) and evidence of excision of multiple alanine residues from the N-terminal poly-alanine stretch of wild type (WT) MeCP2-E1 was observed, whereas the A2V mutant exhibited only partial NME of either methionine or valine and reduced N-acetylation (NA). This resulted in different *in vitro* protein degradation rates between the WT and the mutant. Indeed, a higher proteasomal degradation activity was observed for MeCP2-E1-A2V compared with that of WT MeCP2-E1 (Sheikh et al., 2017). Hence, the etiopathology of this mutation is likely due to a reduced bio-availability of MeCP2 resulting from the defective co-post-translational N-terminal modifications that lead to a faster degradation of the A2V mutant (Sheikh et al., 2017).

Apart from A2V, there are no other published reports of exon 1 missense mutations. In fact, there are remarkably few examples of missense mutations within the N-terminal region leading up to the MBD. One of these few rarities is A59P. The A59P mutation was described in three Tunisian RTT patients with variable scores of clinical severity (Kharrat et al., 2015). Despite the intrinsically disordered organization of the MeCP2 NTD, such an amino acid change was predicted to have an important structural effect on the overall conformation of the protein backbone. However, the structural role and protein stability properties affected remain to be determined. This also applies to all the other NTD mutations described above, with the exception of A2V, and hints to the complexity of the molecular mechanisms probably involved.

MeCP2 Missense Mutations

Missense mutations represent the most abundant mutations in RTT [over 70% (Spiga et al., 2019)] and mainly affect the MBD (residues 78 to 162, **Figure 1B**) where they make up to approximately 45% of the cases (Ghosh et al., 2008), underscoring the primary role of the MBD in the function of this protein. This domain corresponds to the main structured part within this intrinsically disordered protein (Dunker et al., 2001; Ausió et al., 2014), and is the only region that has been amenable to crystallization (Ho et al., 2008). The MBD crystal structure has provided an excellent resource for the analysis of the structural alterations caused by mutations within this region.

Because of their high occurrence, these mutations have been studied extensively using *in vitro* structural, *in situ* cell culture and *in vivo* mouse model approaches (Tillotson and Bird, 2019). Within the first category, we have already referred to the early pioneering work on R106W, R133C, F155S, and T158M by Ballestar et al. (2000). This structural work was followed by a study of R106W, R111G, R133C, F155S, and T158M, which, in the absence of crystallographic MBD information (Ho et al., 2008), used NMR and provided a more detailed molecular characterization of the structural changes resulting from these mutants (Free et al., 2001). It was noticed that the R133C mutation affected DNA binding without changing the MBD structure, thus highlighting, for the first time, the relevance for proper DNA binding of the basic amino acids at the MBD-DNA interface (Spiga et al., 2019). These studies were ensued by a later characterization of the same mutants using a combination of biophysical techniques that included fluorescence spectroscopy and circular dichroism. They allowed the authors to correlate the magnitude of the structural changes elicited by each mutant to the severity of the associated RTT phenotypes (Ghosh et al., 2008). Due to the complexity of the structural work, initial studies focused on some of the most prevalent RTT mutations and only more recently have been extended to other mutants such as the Y120D and to their binding to mCH (where H = A, T, or C) (Sperlazza et al., 2017), and in particular to mCA (Gabel et al., 2015). Although no structural differences were observed between the binding of the MeCP2 mutants to mCG versus mCA (Sperlazza et al., 2017), this study was very relevant as, immediately after birth, during neuron differentiation, mCH (Lister et al., 2013) and MeCP2 (Kaufmann et al., 2005b; Olson et al., 2014) concomitantly increase during a changing methylation landscape that may account for the onset of RTT (Lavery and Zoghbi, 2019). From the structural point of view, the different missense MBD-RTT associated mutations can change either the stability of the MBD, its DNA binding affinity, or both to a different extent. In this regard, a couple of recent exhaustive structural studies summarize this in a way that clusters these mutations into three different groups (Kucukkal et al., 2015; Yang et al., 2016) (see **Table 1** and **Figure 3**).

The information from the structural studies has been complemented by *in situ* experiments in different cell culture settings. A few representative papers covering a wide spectrum of mutations have been published by Kudo et al. (2003), and by Agarwal et al. (2007). In addition to stability and affinity of binding, these studies have also focused on the clustering ability of MeCP2 around the pericentromeric heterochromatin. Importantly, the work has also provided insight to the altered distribution within the nuclear/cytoplasmic compartments caused by the MeCP2 mutations (**Figure 3C**). As with the *in vitro*

TABLE 1 | Classification of MeCP2 MBD missense mutations according to their structural characteristics (Yang et al., 2016).

Group (Cluster)	Mutation (No. of cases) (Krishnaraj et al., 2017; Spiga et al., 2019)	Structural characteristics
Group 1	L100V (7) S134C (21) P152R (71) D156E (15)	MBD propensity to unfold and reduced binding affinity for C methylated DNA
Group 2	R106W (132) R106Q (21) R133H (8) R133C (217) F155S (2) T158M (419) T158A (2)	No major changes in MBD structure but various binding affinities for C methylated and unmethylated DNA
Group 3	R111G (1) A140V (29)	Loss of MBD flexibility leads to reduced binding affinity for either C methylated or unmethylated DNA

structural work, an attempt has been made to correlate the observations made with clinical severity (Sheikh et al., 2016; **Figure 3C**). Interestingly, while some of the MeCP2 protein mutants resulting from MBD mutations (i.e., R133C and A140V) are still able to bind to chromatin, their interaction with ATRX is fully compromised (Nan et al., 2007). ATRX is a protein member of the ATP-dependent SWI/SNF family of chromatin remodeling complexes (Pazin and Kadonaga, 1997). This additional disruption of a functionally relevant protein-protein interaction underscores the molecular mechanistic complexity of some of these mutations.

In more recent years, a few knock-in mice models have been produced for the Y120D, R133C and T158M/A mutations as well as two transgenic models for the R111G and R306C mutations (Heckman et al., 2014). These models have provided useful information from an *in vivo* perspective. This work has revealed that, in instances such as T158M, where the mutation significantly decreases the amount of MeCP2 in the nucleus, the RTT phenotype can be rescued by increasing the expression of the T158M mutant (Lamonica et al., 2017). Moreover, the decrease in the mutated MeCP2 was shown to be due to proteasomal degradation (Lamonica et al., 2017). This decrease is reminiscent of that observed for the truncated form of MeCP2 expressed in the Jaenisch (*Mecp2* $^{TM1.1Jae}$/Mmcd) mouse model (Stuss et al., 2013), in which exon 3 of MeCP2 is deleted (**Figure 1A**) such that most the MBD is lacking (Chen et al., 2001).

MeCP2 Nonsense and C-Terminal Mutations

In this section, we include all the mutants affecting MeCP2 beyond its MBD. These include mutations affecting the ID, TRD, and CTD domains that encompass the NID and RNA binding domain (RBD) (**Figure 3A**). The C-terminal region defined in this way, is also where MeCP2 mutations pertaining to other brain disorders such as schizophrenia, which involves the ID, (Sheikh et al., 2018; Chen et al., 2020) take place. Many of the RTT nonsense mutations occur within this region and its most significant missense mutations take place in the TRD (**Figure 3A**). From a structural perspective, mutations within this region have been less extensively studied, in contrast to those of the MBD. This is particularly true as it pertains to the nonsense and frameshift mutations leading to early

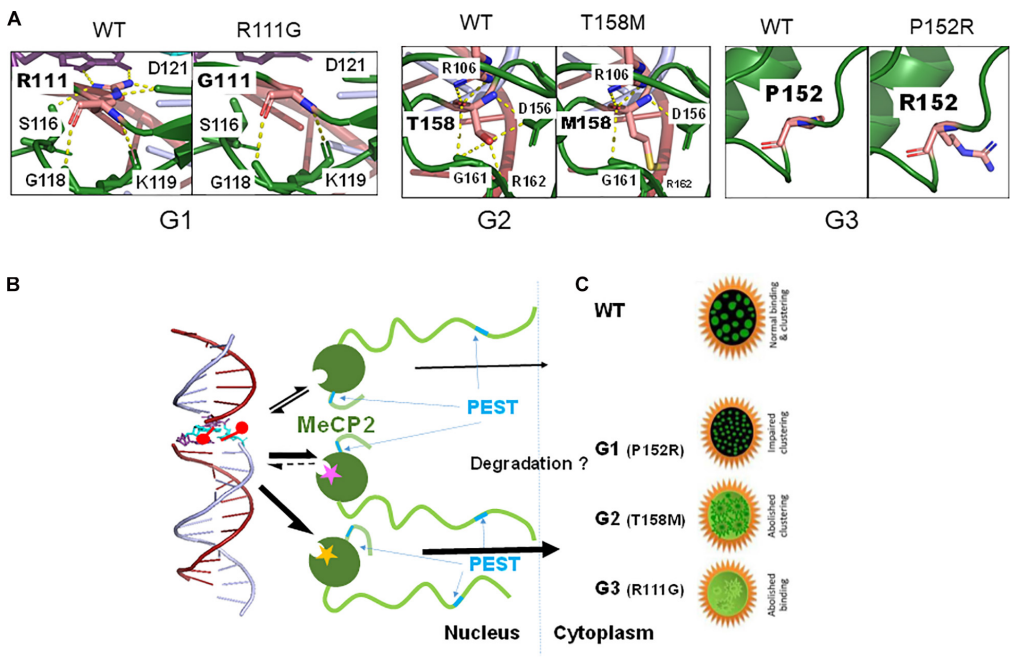

FIGURE 3 | Effect of a few MBD missense mutations [groups G1 to G3 (Yang et al., 2016)] on MeCP2 structure (stability) **(A)**, binding affinity (magenta star: G1; yellow star: G3) **(B)** and nuclear distribution **(C)**.

termination. In this regard, a hint into some of their potential molecular effects can be envisaged from the early structural work carried out in the late Alan Wolffe's lab (Chandler et al., 1999; Yusufzai and Wolffe, 2000) which used C-terminally truncated versions of MeCP2 to show that the CTD facilitates its binding to nucleosomes (Chandler et al., 1999) and hence to chromatin. A more detailed follow up, carried out by Nikitina et al., highlighted the role of residues 295–486 for chromatin interaction (Nikitina et al., 2007b), and showed that R294X failed to produce the nucleosome–nucleosome interactions (Nikitina et al., 2007a) that are observed with the native form of the protein. Disruption of the inter-nucleosome interactions may play an important role in the intrinsic ability of MeCP2 to organize chromatin into chromocenters (Brero et al., 2005; Agarwal et al., 2011; Ausio, 2016; Wang et al., 2020). The RTT nonsense mutations, R168X, R255X, R270X, and R294X, which together account for about 90% of all nonsense mutations (Krishnaraj et al., 2017) were recently shown to disrupt the ability of MeCP2 to cluster heterochromatin (Li et al., 2020) in a way that progressively decreased with the proximity of the MeCP2 truncation to the C-terminus (Wang et al., 2020). These results agree with the clinical observations which indicate that severity of the C-terminal truncations decreases with its proximity to the carboxy-terminus of MeCP2, with individuals with the R168X mutation being more severely affected than those with R294X (a mild RTT mutation) or other more C-terminal mutations (Neul et al., 2008; Bebbington et al., 2010; Cuddapah et al., 2014). They also agree well with the recently described critical role of the unstructured MeCP2 CTD in conjunction with the MBD to form heterochromatin condensates (Li et al., 2020). Of note, the R255X and R270X mutations fall within the NLS region, however, the molecular relevance of this is not clear, especially since NLS inactivation does not affect the progression of the disease in a RTT mouse model (Lyst et al., 2018).

Given the confounding effects resulting from the mosaic expression of MeCP2 in females (XX), boys (XY) with RTT allow for a better correlation between mutation and phenotype severity. In this regard, boys with truncation or frameshift mutations before or including residue R270 exhibit neonatal encephalopathy and death, whereas males with the same type of mutations beyond G273 survive. Using R270X and G273X mouse models, the breaking point was shown to be due to the disruption of an AT-hook 2 HMGA1 (high mobility group)-like domain in MeCP2 that was found to be critical for chromatin maintenance and α-thalassemia mental retardation X-linked protein (ATRX) localization in the nervous system (Baker et al., 2013), underscoring again the multifaceted role of MeCP2 in chromatin organization. As with the chromatin architectural HMGA1 non-histone protein (Reeves, 2001), MeCP2 contains three homologous AT hook domains (Ausió et al., 2014) that provide the molecule with DNA binding properties involved in chromatin clustering and heterochromatin organization. Interestingly, HMGAs play an important role in the regulation of the neurogenic potential of neuron precursor cells, and their expression is lost during neuron differentiation (Tyssowski et al., 2014).

In addition to all the above, the possibility exists that the deleterious consequences of the nonsense mutations giving rise to the truncated forms of MeCP2 may also be partially indirect in nature. Indeed, analyses of the histone PTMs in lymphocytes from RTT patients showed a decrease in the levels of acetylation of lysines 9 and 14 of histone H3 (Kaufmann et al., 2005a). These analyses are interesting, and add to the promise of biomarker discovery in RTT patient lymphocytes, but unfortunately they do not provide insight into the potential molecular mechanisms involved. In experiments carried out on clonal cell cultures from an RTT female with the R168X mutant and cells from a male hemizygote for the frameshift mutation 803delG (V288X), both sets of mutant cells exhibited histone H4 hyperacetylation specifically associated with increased acetylation of lysine 16 (H4K16ac) (Wan et al., 2001). In a different study using a mouse model of Rett syndrome expressing MeCP2 truncated by introducing a stop codon after codon 308 ($Mecp2^{308/y}$), a 2–3 fold increase in histone H3 acetylation was observed in cortex (Shahbazian M. et al., 2002). The changes in global acetylation observed in these studies might have important alterations in gene expression and in both instances had been attributed to the inability of these truncated versions of the expressed protein to recruit the MeCP2-associated histone deacetylase (HDAC) complexes (Nan et al., 1998; Jones et al., 2001). However, the relation of this acetylation to the MeCP2-dependent HDAC recruitment is surprising, as the null mouse model lacking the expression of MeCP2 does not exhibit any differences in histone H3 or H4 acetylation (Urdinguio et al., 2007). A more plausible explanation would be that MeCP2 might have a developmental-dependent downstream effect on gene expression (Thatcher and Lasalle, 2006) which is altered in different ways in the presence of different truncated MeCP2 forms and in different tissues.

The most important C-terminal missense mutations (in terms of frequency): P302R, K304E, K305R, and R306C occur in the TRD within the NID (**Figure 3A**). In this regard, R306C represents one of the most frequent mutations observed in RTT with 245 (5.1%) RTT cases reported (Krishnaraj et al., 2017). R306C suppresses MeCP2 binding to the nuclear receptor co-repressor (NCoR)-mediated recruitment of HDACs, which is also severely compromised by any of these four RTT mutations (Kruusvee et al., 2017). Yet, despite the functional relevance to RTT attributed to the MeCP2 NID (Tillotson et al., 2017), and despite the main functional role of MeCP2 in the brain being to recruit the NCoR1/2 co-repressor complex to methylated DNA sites in this tissue (Tillotson and Bird, 2019), mutations within this region correspond to some of the clinically milder RTT forms reported (Schanen et al., 2004; Cuddapah et al., 2014; Neul et al., 2014). Also, the genomic sites to which the NCoR1/2 complex is recruited (Connolly and Zhou, 2019) and their relative abundance are still unknown.

Other molecular implications for the mutations within the MeCP2 C-terminus could arise from the fact that the CTD had been shown early on to contain a WW domain binding region (WDR) encompassing amino acids 325–498. This region is responsible for the interaction of MeCP2 with group II WW-containing domains in splicing factors FBP11 and HYPC (Buschdorf and Stratling, 2004). Several missense mutations

and small C-terminal INDELs causing truncations occur within this region (Kyle et al., 2018; Lavery and Zoghbi, 2019; Spiga et al., 2019), the most prominent being E395K (Krishnaraj et al., 2017). However, structural information for any of these is still lacking.

Proteasomal Degradation and the PEST Sequences

As has already been mentioned above for various MeCP2 mutations, several of the anomalous cellular levels and pathological aspects involved arise from alterations in the proteasomal degradation processing of these mutants (Lamonica et al., 2017; Sheikh et al., 2017). Whether the proteasome activity takes place in the nucleus or in the cytoplasm is not yet clear and it may be dependent on the type of mutation. While both the N terminus-dependent pathway and the lysine-dependent (PEST) degradation pathways can occur in the nucleus, the latter occurs more actively in the cytoplasm (Lingbeck et al., 2003). Hence, the cellular localization of the MeCP2 protein molecules to be degraded might be mutation dependent. Indeed, the MBD by itself is important for nuclear localization (Lyst et al., 2018) and hence MeCP2 partitioning within the cell might be affected by impairment of its ability to bind to its methylated DNA target.

Regardless of the cellular compartment where the degradation of MeCP2 mutants takes place, the ubiquitin proteasome system (UPS) plays a very important role in neurological diseases, including cognitive disorders like RTT (Lehman, 2009). Mesenchymal stromal cells from a heterozygous RTT female mouse model null for MeCP2 (Mecp2tm1.1Bird), which mimics partial MeCP2 loss of function, have been shown to exhibit increased proteasome activity (Squillaro et al., 2019). Similar observations associated with changes in cellular ubiquitination have been described for peripheral blood lymphomonocytes from RTT patients (Pecorelli et al., 2013).

In view of all of this, we propose here a mechanism for MeCP2 degradation of missense mutations that relies on the presence of two PEST domains in MeCP2 (Thambirajah et al., 2009; **Figure 3B**). These domains consist of consensus sequences enriched in proline, glutamate, serine, and threonine (PEST) residues, which act as a recognition signal for rapid degradation by the 26S UPS (Rogers et al., 1986; Rechsteiner and Rogers, 1996). The two PEST domains of MeCP2 are N-terminally and C-terminally located at amino acid residues 73–94 and 389–426, respectively (Thambirajah et al., 2009). However, whether this mechanism would apply to all the RTT mutations it is not clear and, as in the case of the N-terminal mutations, additional mechanisms may also apply (Sheikh et al., 2017). While limited detailed information is available for some of the N-terminal (Sheikh et al., 2017) and MBD missense (Lamonica et al., 2017) mutations, information in this regard on mutations occurring at other MeCP2 domain locations is significantly lacking, and information on stability and binding affinity it is only known in a few instances. Such is the case of the reduced DNA binding affinity of the R306C mutation described in the previous section, which affects the TRD/NID (**Figure 3A**). Using a mouse model for this mutation, Heckman et al. have conclusively shown that, beyond the impairment of binding the repressive NCoR complex (Kruusvee et al., 2017), alteration of the basic cluster of basic amino acids (304–309) within the RNA binding domain (**Figure 3A**) by R306C lowers the binding affinity of the mutant protein by MeCP2 binding sequences *in vivo* (Heckman et al., 2014). Importantly, it might also disrupt the interaction of the protein with RNA as will be described next.

MeCP2 as an RNA Interacting Protein

MeCP2 RNA Binding Domain(s)

Direct interaction between MeCP2 and RNA was first shown by *in vitro* electrophoretic mobility assays (Jeffery and Nakielny, 2004). MeCP2 shifted mouse immunoglobulin mRNA and *Xenopus* U1 spliceosomal small nuclear (sn)RNA, but not human tRNA or *Xenopus* 5 S rRNA, suggesting that RNA binding is not promiscuous. Removal of an RG repeat motif C-terminal to the MBD abolished RNA binding (**Figure 2A**). Of note, double stranded (ds), but not single stranded (ss) RNA was shown to compete with methylated DNA, implying mutually exclusive binding to methylated DNA or dsRNA. Despite these promising but preliminary results, very little further research has been performed to characterize MeCP2-RNA interactions. It is becoming increasingly clear that protein interactions with both DNA and RNA are key to almost all nuclear processes, particularly because of the emerging regulatory roles played by lncRNAs (Hudson and Ortlund, 2014). RG repeat motifs are well established RNA binding modules (Thandapani et al., 2013). The binding preference for these motifs is debated, but evidence tends toward affinity for G-quadruplexes or GC-rich dsRNA, where arginine residues form hydrogen bonds with guanines (Jarvelin et al., 2016). This allows for almost transcriptome-wide binding possibilities, where specificity might be modulated by structural context and/or residues commonly occurring near RG repeats, such as the PGG, GGG, and polyalanine residues found in MeCP2 (**Figure 2A**; Chong et al., 2018). MeCP2 has a relatively short RG repeat, which is generally associated with low RNA binding affinity. An AT-hook domain overlaps the RG repeat, which was later characterized as a non-canonical extended AT-hook (eAT-hook) (**Figure 2A**; Filarsky et al., 2015). eAT-hook proteins have an order of magnitude greater affinity for dsRNA than DNA, and long stem-loop RNA structures are preferred over short hairpins (**Figure 2C**).

Advanced proteome-wide screens of RNA binding proteins (RBPs) have identified MeCP2 as an important RBP (He et al., 2016; Trendel et al., 2019). Recent comprehensive *in vivo* capture of RBPs identified a lysine-rich non-canonical RNA binding motif within the TRD of MeCP2 (Castello et al., 2016). RNA binding domains (RBDs) that lack sequence homology to known RBDs are being found with increasing frequency, and are presently characterized by basic (R and K) and disorder-promoting (R, G, P, S, and Q in MeCP2) residues. K-rich regions in DNA binding domains are thought to allow "hopping" or "sliding" to specific sequences; however, K-rich RNA binding domains characterized to date imply RNA structure over sequence-specific binding (Wilson et al., 1998; Takeuchi et al., 2009; Castello et al., 2016; **Figure 2C**). Most non-canonical RBDs are also enriched in having DNA and protein interaction surfaces, suggesting competition between RNA binding and other molecular interactions at these regions. This agrees with

the K-rich RBD of MeCP2 overlapping its NID, a region within which RTT-causing missense mutations are also enriched (**Figure 2A**). RTT-like phenotype rescue in knock-in *Mecp2*-null mice expressing a minimal MeCP2 protein lacking the N- and C-terminal regions as well as the intervening region suggest that just the MBD and NID are necessary and sufficient for MeCP2 function (**Figure 2A**; Tillotson et al., 2017). The CTD probably plays a minor functional role, as its absence caused mild intellectual phenotypes. The discrepancy between a minimal peptide and diverse functionality of the protein has been reconciled by concluding that the dominant role of MeCP2 is to mediate transcriptional repression via NCoR complex recruitment to methylated DNA. An alternative explanation could be that the regions containing the MBD and/or NID mediate some degree of protein multifunctionality (Ghosh et al., 2010). Surface plasmon resonance assays show that MeCP2-TBLR1 (the direct binding subunit of the NCoR complex) interaction is relatively weak (K_D 9.5 $\pm$ 0.5 μM), which could be due to a lack of physiological context *in vitro*, or it could indicate transient binding (Kruusvee et al., 2017). Evidence for NCoR mutations specifically causing Rett syndrome is lacking (Sakaguchi et al., 2018; Zaghlula et al., 2018). Instead, they can cause intellectual disabilities, rather than RTT, *per se*, similar to misregulation of other MeCP2-associated processes such as mRNA splicing or miRNA biogenesis, suggesting comparable importance of MeCP2's differing roles (Young et al., 2005; Ha and Kim, 2014). Given a growing catalog of diverse regulatory RNAs, the presence of a flexible RBD within the indispensable NID region of MeCP2 is a good candidate to explain its functional multiplicity. Moreover, combinatorial action between the K-rich RBD and the eAT-hook/RG repeat region of MeCP2 could add to the complexity.

Post-translational modifications (PTMs) are another regulatory mechanism of RBPs, some of which MeCP2 may share (Xu et al., 2019). Conservation analysis of PTMs found within RBDs showed that phosphoserine is often immediately preceded by a conserved glycine, and phosphothreonine can often be found 5 amino acids upstream from a conserved serine such as that seen on MeCP2 residues S274 and T308 (**Figure 2A**; Castello et al., 2016). Both PTMs have been experimentally determined (Bellini et al., 2014). Differential activity-dependent mechanisms determine phosphorylation of S274 (protein kinase A) and T308 (membrane depolarization), where phosphorylated T308 abrogates NCoR binding, potentially also representing differential activity-dependent mechanisms of RNA binding regulation (Ebert et al., 2013). MeCP2 is also ubiquitinated (K271) and acetylated (K271, K289, K305 or 307) at sites within the RBD (Gonzales et al., 2012; Pandey et al., 2015). The roles of these PTMs are unclear, but ubiquitination can influence protein conformation, and acetylated K305 could be important for protein function, as indicated by the acetyl-defective RTT mutations, K305E/R (Ausió et al., 2014). Methylation at R162 within the RG repeat motif could affect affinity for RNA, as is common for other RGG/RG proteins (Blackwell and Ceman, 2012; Guo et al., 2014). PTMs outside of RBDs can also modulate RBP function, allowing for further positive or negative regulation of RNA processing and fate (Lovci et al., 2016). This may also be true for MeCP2, similar to how activity-dependent phosphorylation outside the MBD bi-directionally regulates DNA binding (Tillotson and Bird, 2019). Altogether, the available data support MeCP2-RNA binding, but the role(s) RNA may play, and how the K-rich and/or RG repeat motifs are involved, can only be speculated as of now.

mRNA Splicing Regulation

Methyl CpG binding protein 2 has been shown to increase exon inclusion of a CD44 minigene reporter through RNA-dependent interaction with the YB-1 splice factor in HeLa and Neuro2A cells (Young et al., 2005). MeCP2 pulls-down YB-1 through TRD residues, but a C-terminal RTT-causing truncation, MeCP2-308X, binds less efficiently to YB-1, reducing exon inclusion. MeCP2 also immunoprecipitates CD44 precursor (pre-) mRNA, and MeCP2-308X abrogates this binding. Given that MeCP2's putative RBDs do not overlap with the CTD (see above), the pre-mRNA interaction may be indirect, or direct pre-mRNA binding could require CTD-protein binding, PTMs, or structural context. Wild type and *Mecp2*$^{308/Y}$ mice have significantly altered genome-wide alternative splicing (AS), including that of *Dlx5* and *Cdk10* – direct targets of MeCP2-mediated repression and activation, respectively. Activity-dependent dephosphorylation at Serine pS80 enhances MeCP2-YB-1 interaction, suggesting MeCP2-dependent splicing regulation occurs in the brain (Gonzales et al., 2012).

In addition to YB-1, MeCP2 binds many splicing factors in different contexts, primarily through the CTD, and to a lesser extent the transcriptional repression domain (TRD) (**Table 2**; Buschdorf and Stratling, 2004; Long et al., 2011; Maxwell et al., 2013). The RNA-binding or RTT relevance of these interactions also vary, or are unknown. MeCP2 assembles with pre-mRNA processing factor 3 (Prpf3) and serologically defined colon cancer antigen gene 1 (Sdccag1) to pre- and mature mRNA of the MeCP2 gene targets *Cdk10* and *Frg1* (Long et al., 2011). The number of documented MeCP2 splice factor interactions continues to grow, with one article even reporting that the majority of MeCP2-bound proteins are involved in RNA splicing and processing (Cheng et al., 2017). The spliceosome is a massive macromolecular complex with many auxiliary proteins providing context-specific AS, so it is unsurprising that such variation exists and that concrete ties to RTT pathology have been difficult to make.

C-terminal truncations account for ∼10% of RTT cases, yet significant functional relevance has yet to be attributed to this domain. Several truncations tested in the articles above coincide with known Rett syndrome genotypes, suggesting correlation with splice factor binding.

Most reported MeCP2-mediated AS events are cassette exon inclusion and intron exclusion (Young et al., 2005; Wong et al., 2017; Osenberg et al., 2018). Intron exclusion events were aberrant for MeCP2 target gene transcripts, *Dlx5* and *Cdk10*, in *Mecp2*$^{308/Y}$ mice (Young et al., 2005). RNA-seq analysis in *Mecp2* KO mouse cortex, however, suggests bidirectional roles in several types of AS (Li et al., 2016). In addition to protein-RNA-mediated regulation, AS is intimately tied with DNA methylation, and MeCP2 binds methylated exonic DNA, stalling RNA Polymerase

TABLE 2 | MeCP2 interacting splicing factors.

Splicing Factor	Tissue/Cell type	Methods	MeCP2 interaction site	RTT mutations abrogate binding?	RNA dependent?	Ref
FBP11, HYPC	HEK293	GST pull-down and Co-IP	From 325 (CTD)	Yes, C-terminal truncations	Untested Untested	Buschdorf and Stratling, 2004
YB-1	HeLa, Neuro2a	GST pull-down and Co-IP	195–329 (TRD)	Yes, C-terminal truncations	Yes	Young et al., 2005
Prpf3 Sdccag1	Whole rat brain nuclei	GST pull-down and Co-IP	104–141 (MBD), 207–294 (TRD) From 311 (CTD)	Yes, C-terminal truncations	No No	Long et al., 2011
Prp8 Top2b DXH9	Whole mouse brain nuclei	Co-IP, MS	Unmapped	Untested	No Yes Yes	Maxwell et al., 2013
LEDGF, DXH9	*MeCP2-Flag* KI whole mouse brain nuclei	Co-IP	163–270 (TRD) Unmapped	Yes, TRD truncations Untested	No No	Li et al., 2016
TDP-43, FUS, hnRNP F					No No No	

II (RNAPII), thus reducing the chance of skipping alternative exons (Maunakea et al., 2013). These initial findings did not distinguish methylation (5mC) from hydroxymethylation (5hmC). However, more discriminatory experiments reveal enrichment of 5hmC at exon-intron boundaries in neurons, whereas 5mC exon-intron enrichment is prevalent in non-neuronal cells, supporting a role for 5hmC in MeCP2-mediated AS in the brain (Khare et al., 2012; Wen et al., 2014; Li et al., 2016). This agrees with MeCP2 enrichment on exon-intron gene boundaries and on 5hmC at active neuronal genes during postnatal development (Kinde et al., 2015; Li et al., 2016).

Alternative splicing is a highly conserved process allowing for greater protein and ncRNA diversity than provided by individual genes (Weyn-Vanhentenryck et al., 2018). A plethora of splice factors with different expression profiles is essential for correct AS during development. Cassette exon inclusion increases in mouse brain during development, which correlates with the increase in MeCP2 expression (Olson et al., 2014; Weyn-Vanhentenryck et al., 2018). MeCP2 interaction with the spliceosome and with pre-mRNA occur primarily through the CTD, suggesting some RTT pathologies from C-terminal truncations could, to some extent, derive from aberrant AS. Of note, during the writing of this manuscript, two research articles related to MeCP2's role in AS were published. In the first, quantitative assessment of high-quality sequencing datasets found little variation in global AS as a result of differential MeCP2 and/or DNA methylation levels (Chhatbar et al., 2020). However, in the second article, MeCP2 was found to be required for maintaining mature hippocampal AS profiles, and to regulate splicing of specific neuronal genes in the hippocampus during memory consolidation (Brito et al., 2020). These two recent papers underscore the still long road ahead in understanding MeCP2's role in AS, as well as the importance of careful context-specific interpretation of MeCP2 studies moving forward.

miRNA Biogenesis and Binding

MicroRNAs (miRNAs) represent an important class of ∼22 nucleotide molecules with key roles in regulating the translation of the proteome (Ha and Kim, 2014). Genome-wide miRNA expression levels are aberrant in the brains of Rett syndrome patients, and offer a potential tool to measure RTT disease progression and treatment response (Wu et al., 2010; Sheinerman et al., 2019). Processing primary (pri-) miRNA into precursor (pre-) miRNA by the nuclear microprocessor complex before export to the cytoplasm is the key regulatory step in determining mature miRNA levels in the cell (Conrad et al., 2014). The core microprocessor proteins are Drosha, which cleaves the pri-miRNA, and DiGeorge syndrome critical region 8 (DGCR8), which provides RNA binding affinity to Drosha (Ha and Kim, 2014). Additional development and cell-type specific co-factors regulate microprocessor activity. In addition to MeCP2, a major atypical RTT-pathogenic protein, FOXG1, is recruited to Drosha to influence miRNA biogenesis, implicating the importance of this process to RTT pathology (Weise et al., 2019).

At a resting state, phosphorylated MeCP2 pS80 inhibits miRNA biogenesis in cultured rat cortical neurons by binding and sequestering DGCR8; activity-dependent dephosphorylation reduces this interaction, allowing miRNA processing to proceed (Cheng et al., 2014). In *Mecp2*-KO mice, miR-134 increases, resulting in decreased levels of its targets involved in neuronal development and plasticity, in addition to reduced dendritic growth. Another group corroborated MeCP2 regulation of miRNA processing through microprocessor interaction, but with some key differences that require reconciliation (Tsujimura et al., 2015). Here, MeCP2 was found to positively regulate miRNA levels in neurons and neural stem cells (NSCs). The authors posited that the different state of MeCP2 phosphorylation, which varies between neuron types and brain regions, could explain the seemingly opposite observations. A global screen of significantly reduced miRNAs in *Mecp2*-KO neurons and NSCs identified miR-199a as important to RTT pathophysiology due to its positive regulation of mechanistic target of rapamycin (mTOR) signaling. MeCP2-mediated miR-199a biogenesis results in targeted inhibition of mTOR inhibitors SIRT1, HIF1a, and PDE4D. SIRT1 deacetylates MeCP2, adding the possibility of feedback regulation by acetylation level, in addition to

phosphorylation (Zocchi and Sassone-Corsi, 2013). Furthermore, phosphorylation of DGCR8 by the mTOR-kinase ERK increases its stability (Herbert et al., 2013). RNA immunoprecipitation (RIP) shows either direct or indirect *in vivo* interaction of MeCP2 with pri-miR199a-1 and pri-miR199a-2. Similar to a proposed mechanism of alternative splicing (see above), MeCP2 binds to methylated miRNA gene boundaries, stalling RNAPII, and enhancing miRNA biogenesis by permitting access to processing machinery (Glaich et al., 2019). It could be that methylation level and MeCP2-DNA binding promote miRNA biogenesis, such as with miR199a, whereas unmethylated miRNA genes are subject to alternative MeCP2-mediated miRNA biogenesis suppression, like in Cheng et al. (2014). The above data offer tenuous support of direct MeCP2-RNA binding in miRNA processing regulation. It is intriguing to speculate a correlation between MeCP2 interaction with paraspeckle lncRNA, NEAT1, which can scaffold Drosha and DGCR8 to peripheral paraspeckle proteins, resulting in the regulation of miRNA biogenesis (Jiang et al., 2017). MeCP2 is known to bind the long isoform of NEAT1 in the brain (see below) (Cheng et al., 2018).

In addition to pri-miRNAs, RIP identifies 87 mature nuclear miRNAs associated with MeCP2 in primary mouse cortical cells (Khan et al., 2017). All MeCP2-interacting miRNA target gene sets are inhibited in *Mecp2*-null mouse cerebellum, implying an inhibitory role of MeCP2 on mature miRNAs, thus positively targeting gene expression. In addition to the canonical role of miRNAs in mRNA decay, nuclear miRNAs are associated with transcriptional repression and activation, as well as alternative splicing (Roberts, 2014).

miRNAs are essential to mammalian cell function, and their aberrant regulation as a result of MeCP2 mutations likely contributes to RTT phenotypes, but the exact interplay of molecular interactions, and whether RNA is directly involved, is complex and remains unclear.

lncRNA Interactions

Long non-coding RNAs (lncRNAs) are >200 base molecules with low coding potential, the varying species of which are involved at every processing step in the nucleus (Zhang et al., 2019). Tissue-specific expression patterns of lncRNAs during development are highly dynamic, allowing diverse outcomes, and are thus unsurprisingly aberrant in RTT (Petazzi et al., 2013; Hosseini et al., 2019). Protein-lncRNA interactions occur with all major classes of epigenetic modifying complexes (Betancur, 2016). Notably, all known RNA-binding subunits of lncRNA-interacting epigenetic complexes lack a canonical RNA binding region, and have at least some level of disorder, similar to MeCP2. Currently, there are four lncRNAs whose interaction with MeCP2 has been reported: *Evf2*, *RNCR3*, *Neat1L*, and *HSATII*, as discussed below.

During embryonic GABAergic neuron development, *Evf2* recruits MeCP2 and the transcriptional activator distal-less homeobox 1 (DLX1) to the *Dlx5/6* homeotic gene cluster, and inhibits DNA methylation there to facilitate antagonism between the two proteins (Berghoff et al., 2013). *Evf2* deletion leads to impaired synaptic connectivity, and *Mecp2*-KO as well as common RTT mutations present GABAergic defects (Horike et al., 2005; Schule et al., 2007). *Evf2* also recruits the SWI/SNF-like chromatin remodeling complex protein brahma-related gene 1 (BRG1) to *Dlx5/6*, which has overlapping protein and RNA binding motifs, similar to MeCP2 (Cajigas et al., 2015). These data led to an important speculation that global DNA binding proteins that paradoxically cause specific intellectual phenotypes when dysregulated, such as MeCP2 in RTT or BRG1 in Coffin-Siris syndrome, may be regulated by specific lncRNAs like *Evf2*. RNA immunoprecipitation of mouse cerebellum robustly pulled down the retinal non-coding RNA (*Rncr3*), followed by metastasis associated lung adenocarcinoma transcript 1 (*Malat1*) (Maxwell et al., 2013). *Rncr3* is upregulated during retinal development, whereas its expression is reduced in *Mecp2*-null mice (Blackshaw et al., 2004). *Rncr3*$^{-/-}$ mice display hindlimb clasping, small brain size, and aberrant axonal sprouting, similar to RTT phenotypes, but these phenotypes are attributed to miR-124a, which is expressed from the *Rncr3* gene (Sanuki et al., 2011).

Aberrant expression of pericentric *HSATII* RNA occurs in several cancers, and recruits polycomb group complex 1 (PRC1) as well as MeCP2 and its protein partner Sin3a into large nuclear condensates (Landers et al., 2020). These epigenetic factors are sequestered from regular function, facilitating genomic instability and cancer development. *HSATII* RNA is also aberrantly enriched in Parkinson's disease patient blood samples, and MeCP2 regulates pericentric heterochromatin regions in neurons in an RNA-dependent manner, suggesting a role for *HSATII* RNA with MeCP2 in intellectual disorders (Billingsley et al., 2019; Marano et al., 2019). A paper published during the preparation of this manuscript showed that MeCP2 and major satellite RNA cooperate to organize pericentric heterochromatin, and that RNA interaction depends on the TRD (Fioriniello et al., 2020). This is consistent with the presence of an RBD overlapping MeCP2's NID, as mentioned earlier.

Huntington's disease (HD) studies found that MeCP2 binds and inhibits the long isoform of nuclear enriched abundant transcript 1 (*Neat1L*) lncRNA in various neuronal and brain tissue types (Cheng et al., 2018). MeCP2 inhibits *NEAT1L* through RNA rather than DNA interaction in wild type cells, whereas MeCP2 is reduced in HD, and increased *NEAT1L* levels protect against the mutant HTT gene. *NEAT1L* increases expression of anti-inflammatory and growth factors including peroxisome proliferator activated receptor-γ (PPARG), Nuclear Factor Kappa B Subunit 1 (NFκB1), and brain-derived neurotrophic factor (BDNF), which are also targets of MeCP2-mediated repression (Martinowich et al., 2003; Mann et al., 2010; Kishi et al., 2016). The short (3,735 nt, *Neat1S*) and long (22,741 nt, *Neat1L*) NEAT1 transcripts are ubiquitously expressed in mammalian cells, and the long isoform is required for the formation of massive ribonucleoprotein paraspeckles, which are heavily implicated in transcriptional regulation (Yamazaki et al., 2020). *Neat1*, *Malat1*, and *Evf2* are upregulated upon neuronal differentiation, similar to MeCP2 (Olson et al., 2014; Roberts et al., 2014). Overall, the data point toward important corollary roles for lncRNAs and MeCP2 during brain development. Further research of their interactions with MeCP2 are integral to understanding how they may relate to RTT. Of note, RNA has been shown

to promote the formation of spatial compartments in the nucleus (Quinodoz et al., 2020), and might assist MeCP2 in its formation of heterochromatin condensates (Sheikh et al., 2016; Li et al., 2020).

Histone PTMs and lncRNAs

Given that MeCP2 repression occurs primarily through HDAC recruitment, it would be expected that MeCP2 deletion invariably increases histone acetylation levels. However, as we discussed earlier (section "MeCP2 Non-sense and C-Terminal Mutations"), studies to date have been conflicting (Wan et al., 2001; Balmer et al., 2002; Kaufmann et al., 2005a; Thatcher and Lasalle, 2006; Lilja et al., 2013). As explored in the previous section, lncRNAs are a promising means to explain disparate functions of the same protein complexes. Proteome-wide analysis of RNA-dependent protein complex formation shows that the Sin3a complex, including HDACs 1 and 2, requires RNA, whereas the NCoR1 complex, including HDAC3, forms independent of RNA (Caudron-Herger et al., 2019). This is consistent with mutually exclusive MeCP2 binding to RNA or NCoR, as suggested by their overlapping binding domains (**Figure 1**). Moreover, Sin3a has been shown to bind lncRNAs in the brain (Dharap et al., 2013). Whether there is any specific binding and regulation of MeCP2 together with Sin3a and its associated HDACs by lncRNA, however, has yet to be determined.

In addition to methylated DNA, MeCP2 has been shown to interact with histone methylation marks associated with constitutive and facultative heterochromatin: di-methylated histone H3 lysine 9 (H3K9me2) and tri-methylated histone H3 lysine 27 (H3K27me3), respectively, in mouse brain nuclear extracts (Thambirajah et al., 2012). A recent report shows that MeCP2 preferentially binds nucleosomes with H3K27me3 via the MBD (Lee et al., 2020; **Figure 2A**). Also, the genomic distribution of the DNA and histone methylation marks overlap, and MeCP2 differentially regulates transcription depending on H3K27me3 and H3K9ac profiles. MeCP2 also associates with histone methyltransferases G9a, protein arginine methyltransferase 6 (PRMT6), and euchromatic histone-lysine N-methyltransferase 1 (EHMT-1) (Dhawan et al., 2011; Xue et al., 2013; Subbanna et al., 2014). MeCP2 increases H3K9 methylation in mouse fibroblasts (Fuks et al., 2003). The *Neat1* lncRNA interacts with EHMT-1 at select genes in neuronal cells, and *Neat1* knockdown decreases H3K9me2 and is associated with increased memory formation (Butler et al., 2019). Little is known about the relationship between MeCP2 and H3K27me3 in the context of RTT, except the previously mentioned findings in the brain, and contrasting findings by Zachariah et al. (2012) where MeCP2 overlaps more with constitutive than facultative heterochromatin marks in primary mouse cortical neurons. These contradictory findings may be explained by the different contexts, or the fact that H3K27me3 levels were compared to different normalizers. Both constitutive (H3K9me3 marked) and facultative (H3K27me3 marked) heterochromatin domains have been shown to be regulated or stabilized to some extent by lncRNA (Yang et al., 2015; Thakur et al., 2020). Despite convoluted results, MeCP2 has continued to be found associated with various histone PTMs and chromatin-modifying enzymes over the years. Whether they are relevant to altered gene expression profiles in RTT patients is still unclear, and will require scrupulous context-specific examination in the future to form conclusions. Potential regulation by previously unconsidered factors like lncRNAs adds to the complexity of the issue.

Functional Roles of MeCP2 Beyond the Brain

Because of MeCP2 multi-functionality as well as its high abundance in the brain, alterations of MeCP2 have been involved in almost every single neurodevelopmental and neurodegenerative disorder of this organ (Ausio, 2016). Nevertheless, besides the brain, the protein is quite abundant in several other tissues, for instance in the lungs (Shahbazian M.D. et al., 2002), and the implications of MeCP2 mutations for RTT within this context represent one of the less studied areas in RTT research. It is thus highly possible that several symptoms observed in RTT do not simply arise from neurological disorders, but are also caused in part by disfunctional cellular regulation (Kyle et al., 2018) in organs other than the brain. Indeed, RTT patients often develop breathing issues and one of the most abundant causes of death in RTT is related to respiratory failure (Ehrhart et al., 2016).

In what follows, we will provide a few examples where MeCP2 has been shown to have an involvement that transcends the neural system and which might be of relevance to RTT.

MeCP2 plays an important role in the modulation of the immune system by influencing the expression of the transcription factor FOXP3 (fork head box P3) – a master regulator of T-helper and T-reg cells (Li et al., 2014). Thus, MeCP2 mutations can contribute to the pathogenesis of inflammatory disease in RTT. Moreover, intestinal isolates from RTT subjects show the presence of an altered microbiota and altered production of short chain fatty acids (Strati et al., 2016), and the presence of proinflammatory strains of *Candida parapsilosis* (Strati et al., 2018). The alteration in the microbiota may also contribute to the gastrointestinal pathophysiology such as constipation status (Motil et al., 2012; Strati et al., 2016). MeCP2 can also contribute to the development of rheumatoid arthritis (Miao et al., 2013).

Cardiac arrhythmia is one of the factors contributing to the greater than expected occurrence of sudden death in RTT individuals (Acampa and Guideri, 2006). Although the molecular involvement of MeCP2 is not completely understood, it appears that dysregulation of Rho GTPase cytoskeletal and inflammation mediated chemokine and cytokine signaling pathway genes are involved (Wang et al., 2018).

Osteopenia is another early symptom that RTT patients are at a risk of developing, and which is dependent on the MeCP2 mutation type (Caffarelli et al., 2020). Although the use of RTT murine models suggest an epigenetic regulation of bone (O'Connor et al., 2009) which might involve RANKL (receptor activator of nuclear factor-κB ligand) (Kitazawa and Kitazawa, 2007), which negatively regulates osteoblast differentiation and bone formation in bone marrow mesenchymal stem cells (Cao, 2018). Yet, the detailed molecular mechanism(s) by which MeCP2 is involved are not understood. This example underscores how the studies on the effects of MeCP2 mutations in non-neural

cell types are still in their relative infancy, as this area of MeCP2 research remains understudied.

Metabolic dysfunction also represents an important component of RTT (Kyle et al., 2018). In this regard, it was recently shown that MeCP2 plays an important role in the regulation of liver homeostasis through a molecular mechanism that involves the targeting the NCoR1/HDAC3 complex to lipogenic gene targets in hepatocytes (Kyle et al., 2016). This underscores the relevance of the MeCP2/HDAC complex outside the neuronal realm. Moreover, lipid metabolism is a more approachable therapeutic target, offering the potential to alleviate the symptoms associated with altered metabolism in RTT patients (Kyle et al., 2016).

Although some of the RTT peripheral organ-related symptoms described above might also have a neuronal component (Cronk et al., 2016), the transcriptional regulatory role of MeCP2 of specific genes within the context of the cell types of the particular organs affected indicates an important role of the MeCP2 mutations within each specific tissue. As in the case of the impaired response to stimuli and stressors observed in RTT (Pillion et al., 2003; Rose et al., 2019), such specificity might also be MeCP2 isoform-dependent, as will be discussed in the next section.

Do MeCP2-E1 and MeCP2-E2 Isoforms Play a Role in RTT?

Not only is the function of MeCP2 important in tissues other than the brain, but also within the brain its role transcends (Ausio, 2016, 2018) that of mere involvement in neurodevelopmental and neurodegenerative disorders (Tan and Zoghbi, 2019). For example, under healthy conditions, the levels of MeCP2 in mouse have been shown to change in a circadian cycle-dependent way (Martinez de Paz et al., 2015; **Figure 4**) – a mechanism which is likely regulated by miR-132/212 (Mendoza-Viveros et al., 2017) in response to the metabolism-dependent circadian cycle changes of the epigenome (Haws et al., 2020). This might have consequences for RTT (Tsuchiya et al., 2015). Indeed, a circadian rhythm disruption has been described in a mouse model of RTT, and disruption of the cycle was observed in fibroblasts from RTT patients (Li et al., 2015). RTT patients are known to frequently experience sleep disorders (McArthur and Budden, 1998). However, the implications of MeCP2 in the circadian cycle are undoubtedly much broader, and the system has allowed us to gain some insight into the different functionality of the E1 and E2 isoforms (Martínez de Paz et al., 2019).

Ironically, for several years most of the research on MeCP2 was carried out with the E2 isoform, which was the first to be identified (Lewis et al., 1992). However, it was not until almost 12 years later that a previously unknown MeCP2 isoform, which is much more highly expressed in human brain than its MeCP2-E2 counterpart, was discovered first in humans (Mnatzakanian et al., 2004), and then several months later in mice (Kriaucionis and Bird, 2004). The two isoforms are the product of alternative splicing. Despite an initially conflicting nomenclature, it was agreed that the longer E1 isoform corresponds to the encoded form starting at exon 1 (skipping exon 2) whereas the E2 isoform was the one encoded starting at exon 2 (**Figure 1A**). The physiological relevance of the two different isoforms of MeCP2 [MeCP2-E1 and MeCP2-E2 (**Figure 1B**)] and in particular, the relevance of the E2 form to RTT have been very controversial (Itoh et al., 2012). However, it is worth emphasizing that RTT mutations have never been identified within exon 2. In all these considerations, however, it is important to recognize that the ratio between the two isoforms and their overall abundance varies significantly from tissue to tissue (Mnatzakanian et al., 2004), and particularly in mature brain, MeCP2-E2 is present at a much lower ratio (approximately 15 fold less) than MeCP2-E1 (Martínez de Paz et al., 2019) as a result of their differential gene expression (Mnatzakanian et al., 2004).

In what follows, we will discuss the information available in support of a different specialized functional involvement of the two isoforms. As we mentioned in the previous sections, the occurrence so far of mutations in the amino acid distinctive MeCP2-E1 NTD suggests that only this isoform is relevant to

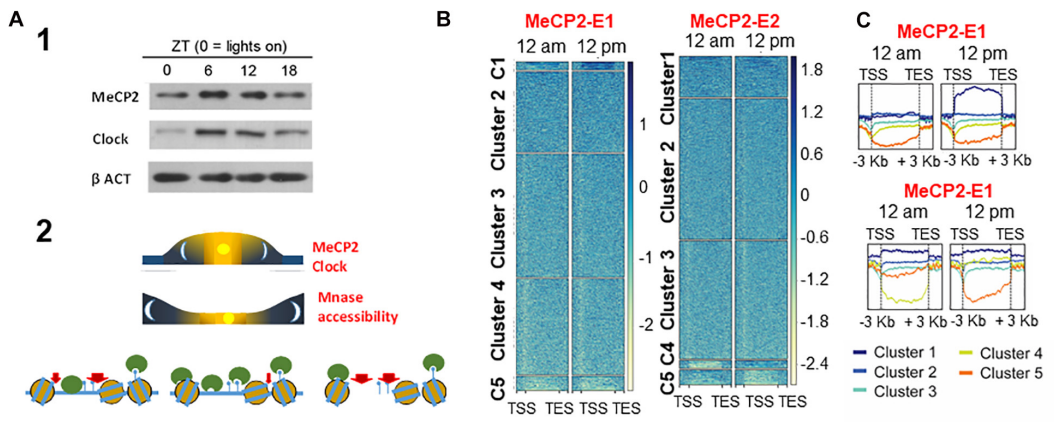

FIGURE 4 | Circadian-dependent MeCP2 **(A-1)** and resulting chromatin changes **(A-2)** and diurnal dynamic change heatmaps for MeCP2-E1 and MecP2-E2 isoform gene occupancy **(B)**, divided into 5 clusters, with isoform-specific enrichment on the different clustered genes over time shown in **(C)**. MeCP2-E1 and MeCP2 enrichment differences in the different clusters in **(A)** (Martinez de Paz et al., 2015; Martínez de Paz et al., 2019).

RTT. However, as it was also mentioned earlier, mutations in exon 1 interfere with the translation of the MeCP2-E2 isoform (Saxena et al., 2006) and the possibility exists for other mutations along the MeCP2 protein to have a similar effect. At the gene level, the 5′ and 3′ UTRs of each isoform have been shown to be differently regulated in a cell type and development-dependent way (Liyanage et al., 2019; Rodrigues et al., 2020). At the 3′UTR, the transcripts undergo alternative polyadenylation affecting the length of these regions, ranging from 0.1 to 8.6 kb, with a preferential association of MeCP2-E1 with the longest 3′UTR (Rodrigues et al., 2020). Although how these differences regulate the expression of the isoforms is not clearly understood, they represent important targets for the binding of regulatory miRNAs and RBPs (Rodrigues et al., 2016). At the 5′ UTRs, it was reported that DNA methylation is significantly correlated with the differential expression of the two isoforms in neurons and astrocytes in a sex-dependent way, with higher levels of DNA methylation corresponding to lower levels of their expression (Liyanage et al., 2019).

As it was mentioned at the beginning of this section, recently, we took advantage of the circadian oscillation of MeCP2 to gain a functional insight on the role of the MeCP2-E1 and MeCP2-E2 isoforms. ChIP-seq analysis, taking advantage of the availability of isoform-specific antibodies, showed a differential binding site preference. MeCP2 isoform-specific enrichments were found to be mainly involved in ligand-receptor interaction in E1 and ribosomal proteins in E2 (Martínez de Paz et al., 2019). Of note, analysis of brains from RTT patients carrying MeCP2 mutations showed abnormal ribosome biogenesis (Olson et al., 2018).

At the protein level, a biophysical analysis using isothermic titration calorimetry and fluorescence spectroscopy analyses of the interaction of the E1 and E2 NTD-MBD protein region (**Figure 1B**) with methylated and unmethylated DNA showed a 10-fold higher affinity and higher structural stability of this region for the E2 isoform compared to the E1 counterpart. Half-life, MS, and FRAP (fluorescence recovery after photobleaching) analysis consistently reported a higher dynamic turnover of MeCP2-E1 compared to MeCP2-E2 (Sheikh et al., 2017; Martínez de Paz et al., 2019). Moreover, using isoform-specific antibodies, a proteomic analysis of the proteins interacting with the two isoforms revealed that, while both isoforms appear to be involved in similar processes, they act through different sets of protein partners. Of interest was the enriched association found between E1 and β-tubulin and microtubule-associated proteins (Martínez de Paz et al., 2019). The extent of overlap observed is, to a certain degree, unsurprising. It has been recently shown that MeCP2-E2 is able to partially compensate for the lack of the E1 isoform in a male case of RTT phenotype (Takeguchi et al., 2020). However, the association of E1 with tubulin remains intriguing. MeCP2 deficiency and mutations have been shown to affect microtubule stability (Delepine et al., 2013) and ciliogenesis (Frasca et al., 2020), respectively, through an indirect association between MeCP2 alteration and HDAC6 deacetylation of tubulin (Gold et al., 2015), though the molecular mechanisms are not yet clearly understood.

In conclusion, while the two MeCP2 isoforms may have a significant extent of generic overlapping functionality as a result of their extensive overlapping primary structure (**Figure 1**), they nevertheless exhibit important distinctive functional traits. Their effect(s) may depend on their different stoichiometry and overall abundance in different tissues as well as on the alteration of the mechanisms regulating their gene expression (Rodrigues et al., 2020), and their potential implications for RTT should not to be overlooked. Indeed, it has been recently shown that, in human brain, the MeCP2E1/E2-BDNF-*miR132* homeostasis regulatory network is region-dependent and is altered in RTT patients (Pehjan et al., 2020).

Epigenetic Therapeutics

The area of therapeutics for RTT has become quite crowded over recent years. There are two main approaches, firstly that of addressing downstream effects of the MECP2 mutation, for instance attempting to upregulate genes under MeCP2's control, such as BDNF or IGF1 [reviewed in Vashi and Justice (2019)], or neurotransmitter pathways such as NMDA receptors, or downstream target K^+/Cl^- co-transporter 2 (KCC2) (Tang et al., 2019). The second approach is to target *MECP2* directly, either through gene therapy, delivering a functional version of the *MECP2* gene exogenously, for example using adenoviral delivery systems [e.g., (Tillotson et al., 2017)], through correcting the mutation at the level of genomic DNA, for example using CRISPR/cas9 editing, or at the RNA level by programmable RNA editing (e.g., Sinnamon et al., 2020), or using compounds such as aminoglycosides to enable "read-through" of MECP2 nonsense mutations [e.g., (Merritt et al., 2020)].

Attempts at epigenetic therapies for RTT, also within this second category of direct targeting of *MECP2*, would aim at the upregulation of *MECP2* expression. This in itself is a somewhat hazardous approach, as *MECP2* over-expression may also have severe developmental repercussions, as witnessed in *MECP2*-duplication syndrome (MIM # 300260). Since RTT is almost exclusively in females, who carry one normal copy of the *MECP2* gene along with a mutated copy, one approach that is being considered is reactivating the silent X [reviewed in Vashi and Justice (2019)]. Naturally occurring X-chromosomal inactivation (XCI) randomly silences one or other of the two X-chromosomes possessed by females epigenetically. This process allows dosage compensation of X-linked genes, which helps maintain the expression of most X-linked genes at a similar level to males.

Skewed XCI may favor the wild-type (WT) allele and hence expression of WT *MECP2*, in which cases RTT symptoms are milder, or, with extreme skewing, asymptomatic. If XCI skewing favors the mutant allele, expression of mutant *MECP2* is favored, and RTT symptoms would be more severe. XCI is an epigenetic process that occurs through the expression of an X-linked non-coding RNA, *Xist*. One potential therapeutic strategy for RTT (and other X-linked disorders) involves reactivating the inactive X chromosome in order to increase expression of WT *MECP2*, which should compensate for the loss of function (and/or expression) of the mutant *MECP2*. The drawback here is that X reactivation could potentially increase dosage of other X-linked genes to pathogenic levels, and so the challenge is to

reactivate only *MECP2* or *MECP2* and its immediate genomic neighbors. Studies are still at an early stage, and there have been a number of high-throughput screens to identify molecules that can reactivate *MECP2* expression from the inactive X chromosome [e.g., (Minkovsky et al., 2015; Lessing et al., 2016; Sripathy et al., 2017)]. Subsequently, in one study, researchers used a small-molecule inhibitor of DNA methylation, 5-aza-2′-deoxycytidine, together with an antisense oligonucleotide knock-down of *Xist* RNA *in vitro*, that significantly upregulated *MECP2* expression, and *in vivo* using *Xist* knockout mice together with the 5-aza-2′-deoxycytidine-induced inhibition of DNA methylation successfully reactivated the inactive (Carrette et al., 2018a). In heterozygous *Mecp2* knockout mice with a mutation in *Tsix*, the antisense regulator of *Xist*, the phenotype observed resembled that of severely affected knockout null male mice, and demonstrated that small increases (5–10%) in WT MeCP2 protein expression can have dramatic improvements on the phenotype (Carrette et al., 2018b). The Tsix/Mecp2 mouse model generated in this study may prove to be an excellent preclinical model for evaluating the effects of XCI-based epigenetic therapeutic compounds.

AUTHOR CONTRIBUTIONS

KG, JV, and JA contributed equally to the preparation, writing, and revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant of the Ontario Rett Syndrome Association (ORSA) to JV and JA, and by a grant from the Canadian Institutes of Health Research; (CIHR grant MOP-130417) to JA, CVG is the recipient of a Natural Sciences and Engineering Research Council of Canada (NSERC) CGS-M fellowship.

REFERENCES

Acampa, M., and Guideri, F. (2006). Cardiac disease and Rett syndrome. *Arch. Dis. Child.* 91, 440–443. doi: 10.1136/adc.2005.090290

Agarwal, N., Becker, A., Jost, K. L., Haase, S., Thakur, B. K., Brero, A., et al. (2011). MeCP2 Rett mutations affect large scale chromatin organization. *Hum. Mol. Genet.* 20, 4187–4195. doi: 10.1093/hmg/ddr346

Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., et al. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res.* 35, 5402–5408. doi: 10.1093/nar/gkm599

Amir, R. E., Van Den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* 23, 185–188. doi: 10.1038/13810

Ausio, J. (2016). MeCP2 and the enigmatic organization of brain chromatin. Implications for depression and cocaine addiction. *Clin. Epigenetics* 8:58.

Ausio, J. (2018). Role of MeCP2 in neurological disorders: current status and future perspectives. *Epigenomics* 10, 5–8. doi: 10.2217/epi-2017-0128

Ausió, J., De Paz, A., and Esteller, M. (2014). MeCP2: the long trip from a chromatin protein to neurological disorders. *Trends Mol. Med.* 20, 487–498. doi: 10.1016/j.molmed.2014.03.004

Baker, S. A., Chen, L., Wilkins, A. D., Yu, P., Lichtarge, O., and Zoghbi, H. Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell* 152, 984–996. doi: 10.1016/j.cell.2013.01.038

Ballestar, E., Yusufzai, T. M., and Wolffe, A. P. (2000). Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry* 39, 7100–7106. doi: 10.1021/bi0001271

Balmer, D., Arredondo, J., Samaco, R. C., and Lasalle, J. M. (2002). MECP2 mutations in Rett syndrome adversely affect lymphocyte growth, but do not affect imprinted gene expression in blood or brain. *Hum. Genet.* 110, 545–552. doi: 10.1007/s00439-002-0724-4

Banerjee, A., Miller, M. T., Li, K., Sur, M., and Kaufmann, W. (2019). Towards a better diagnosis and treatment of Rett syndrome: a model synaptic disorder. *Brain* 142, 239–248. doi: 10.1093/brain/awy323

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Nat. Publish. Group* 21, 381–395. doi: 10.1038/cr.2011.22

Baylin, S. B., and Jones, P. A. (2016). Epigenetic determinants of cancer. *Cold Spring Harb. Perspect. Biol.* 8:a019505. doi: 10.1101/cshperspect.a019505

Bebbington, A., Percy, A., Christodoulou, J., Ravine, D., Ho, G., Jacoby, P., et al. (2010). Updating the profile of C-terminal MECP2 deletions in Rett syndrome. *J. Med. Genet.* 47, 242–248. doi: 10.1136/jmg.2009.072553

Bellini, E., Pavesi, G., Barbiero, I., Bergo, A., Chandola, C., Nawaz, M. S., et al. (2014). MeCP2 post-translational modifications: a mechanism to control its involvement in synaptic plasticity and homeostasis? *Front. Cell. Neurosci.* 8:236. doi: 10.3389/fncel.2014.00236

Berghoff, E. G., Clark, M. F., Chen, S., Cajigas, I., Leib, D. E., and Kohtz, J. D. (2013). Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development* 140, 4407–4416. doi: 10.1242/dev.099390

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242.

Betancur, J. G. (2016). Pervasive lncRNA binding by epigenetic modifying complexes–The challenges ahead. *Biochim. Biophys. Acta* 1859, 93–101. doi: 10.1016/j.bbagrm.2015.10.009

Billingsley, K. J., Lattekivi, F., Planken, A., Reimann, E., Kurvits, L., Kadastik-Eerme, L., et al. (2019). Analysis of repetitive element expression in the blood and skin of patients with Parkinson's disease identifies differential expression of satellite elements. *Sci. Rep.* 9:4369.

Bird, A. P. (1993). Functions for DNA methylation in vertebrates. *Cold. Spring Harb. Symp. Quant. Biol.* 58, 281–285. doi: 10.1101/sqb.1993.058.01.033

Blackshaw, S., Harpavat, S., Trimarchi, J., Cai, L., Huang, H., Kuo, W. P., et al. (2004). Genomic analysis of mouse retinal development. *PLoS Biol.* 2:E247. doi: 10.1371/journal.pbio.0020247

Blackwell, E., and Ceman, S. (2012). Arginine methylation of RNA-binding proteins regulates cell function and differentiation. *Mol. Reprod. Dev.* 79, 163–175.

Brero, A., Easwaran, H. P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., et al. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J. Cell Biol.* 169, 733–743. doi: 10.1083/jcb.200502062

Brito, D. V. C., Karaca, K. G., Kupke, J., Frank, L., and Oliveira, A. M. M. (2020). MeCP2 gates spatial learning-induced alternative splicing events in the mouse hippocampus. *Mol. Brain* 13:156.

Buschdorf, J. P., and Stratling, W. H. (2004). A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J. Mol. Med.* 82, 135–143. doi: 10.1007/s00109-003-0497-9

Butler, A. A., Johnston, D. R., Kaur, S., and Lubin, F. D. (2019). Long noncoding RNA NEAT1 mediates neuronal histone methylation and age-related memory impairment. *Sci. Signal.* 12:eaaw9277. doi: 10.1126/scisignal.aaw9277

Caffarelli, C., Gonnelli, S., Pitinca, M. D. T., Camarri, S., Al Refaie, A., Hayek, J., et al. (2020). Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with bone disease severity in Rett syndrome. *BMC Med. Genet.* 21:21. doi: 10.1186/s12881-020-0960-2

Cajigas, I., Leib, D. E., Cochrane, J., Luo, H., Swyter, K. R., Chen, S., et al. (2015). Evf2 lncRNA/BRG1/DLX1 interactions reveal RNA-dependent inhibition of chromatin remodeling. *Development* 142, 2641–2652. doi: 10.1242/dev.126318

Cao, X. (2018). RANKL-RANK signaling regulates osteoblast differentiation and bone formation. *Bone Res.* 6:35.

Carrette, L. L. G., Blum, R., Ma, W., Kelleher, R. J. III, and Lee, J. T. (2018a). Tsix-Mecp2 female mouse model for Rett syndrome reveals that low-level MECP2 expression extends life and improves neuromotor function. *Proc. Natl. Acad. Sci. U.S.A.* 115, 8185–8190. doi: 10.1073/pnas.1800931115

Carrette, L. L. G., Wang, C. Y., Wei, C., Press, W., Ma, W., Kelleher, R. J. III, et al. (2018b). A mixed modality approach towards Xi reactivation for Rett syndrome and other X-linked disorders. *Proc. Natl. Acad. Sci. U.S.A.* 115, E668–E675.

Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A. M., Foehr, S., et al. (2016). Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell.* 63, 696–710. doi: 10.1016/j.molcel.2016.06.029

Caudron-Herger, M., Rusin, S. F., Adamo, M. E., Seiler, J., Schmid, V. K., Barreau, E., et al. (2019). R-DeeP: proteome-wide and quantitative identification of RNA-dependent proteins by density gradient ultracentrifugation. *Mol. Cell.* 75, 184.e10–199.e10.

Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T., Qin, J., et al. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* 320, 1224–1229. doi: 10.1126/science.1153252

Chahrour, M., and Zoghbi, H. Y. (2007). The story of Rett syndrome: from clinic to neurobiology. *Neuron* 56, 422–437. doi: 10.1016/j.neuron.2007.10.001

Chandler, S. P., Guschin, D., Landsberger, N., and Wolffe, A. P. (1999). The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry* 38, 7008–7018. doi: 10.1021/bi990224y

Cheadle, J. P., Gill, H., Fleming, N., Maynard, J., Kerr, A., Leonard, H., et al. (2000). Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum. Mol. Genet.* 9, 1119–1129. doi: 10.1093/hmg/9.7.1119

Chen, C. H., Cheng, M. C., Huang, A., Hu, T. M., Ping, L. Y., and Chang, Y. S. (2020). Detection of rare Methyl-CpG binding protein 2 gene missense mutations in patients with schizophrenia. *Front. Genet.* 11:476. doi: 10.3389/fgene.2020.00476

Chen, R. Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* 27, 327–331. doi: 10.1038/85906

Cheng, C., Spengler, R. M., Keiser, M. S., Monteys, A. M., Rieders, J. M., Ramachandran, S., et al. (2018). The long non-coding RNA NEAT1 is elevated in polyglutamine repeat expansion diseases and protects from disease gene-dependent toxicities. *Hum. Mol. Genet.* 27, 4303–4314.

Cheng, T. L., Chen, J., Wan, H., Tang, B., Tian, W., Liao, L., et al. (2017). Regulation of mRNA splicing by MeCP2 via epigenetic modifications in the brain. *Sci. Rep.* 7:42790.

Cheng, T. L., Wang, Z., Liao, Q., Zhu, Y., Zhou, W. H., Xu, W., et al. (2014). MeCP2 suppresses nuclear microRNA processing and dendritic growth by regulating the DGCR8/Drosha complex. *Dev. Cell* 28, 547–560. doi: 10.1016/j.devcel.2014.01.032

Chhatbar, K., Cholewa-Waclaw, J., Shah, R., Bird, A., and Sanguinetti, G. (2020). Quantitative analysis questions the role of MeCP2 as a global regulator of alternative splicing. *PLoS Genet.* 16:e1009087. doi: 10.1371/journal.pgen.1009087

Chong, P. A., Vernon, R. M., and Forman-Kay, J. D. (2018). RGG/RG motif regions in RNA binding and phase separation. *J. Mol. Biol.* 430, 4650–4665. doi: 10.1016/j.jmb.2018.06.014

Cohen, D., Lazar, G., Couvert, P., Desportes, V., Lippe, D., Mazet, P., et al. (2002). MECP2 mutation in a boy with language disorder and schizophrenia. *Am. J. Psychiatry* 159, 148–149. doi: 10.1176/appi.ajp.159.1.148-a

Comings, D. E. (1986). The genetics of Rett syndrome: the consequences of a disorder where every case is a new mutation. *Am. J. Med. Genet. Suppl.* 1, 383–388. doi: 10.1002/ajmg.1320250540

Connolly, D. R., and Zhou, Z. (2019). Genomic insights into MeCP2 function: a role for the maintenance of chromatin architecture. *Curr. Opin. Neurobiol.* 59, 174–179. doi: 10.1016/j.conb.2019.07.002

Conrad, T., Marsico, A., Gehre, M., and Orom, U. A. (2014). Microprocessor activity controls differential miRNA biogenesis in vivo. *Cell. Rep.* 9, 542–554. doi: 10.1016/j.celrep.2014.09.007

Cronk, J. C., Derecki, N. C., Litvak, V., and Kipnis, J. (2016). Unexpected cellular players in Rett syndrome pathology. *Neurobiol. Dis.* 92, 64–71. doi: 10.1016/j.nbd.2015.05.005

Cuddapah, V. A., Pillai, R. B., Shekar, K. V., Lane, J. B., Motil, K. J., Skinner, S. A., et al. (2014). Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *J. Med. Genet.* 51, 152–158. doi: 10.1136/jmedgenet-2013-102113

Curie, A., Lesca, G., Bussy, G., Manificat, S., Arnaud, V., Gonzalez, S., et al. (2017). Asperger syndrome and early-onset schizophrenia associated with a novel MECP2 deleterious missense variant. *Psychiatr. Genet.* 27, 105–109. doi: 10.1097/ypg.0000000000000165

Delepine, C., Nectoux, J., Bahi-Buisson, N., Chelly, J., and Bienvenu, T. (2013). MeCP2 deficiency is associated with impaired microtubule stability. *FEBS Lett.* 587, 245–253. doi: 10.1016/j.febslet.2012.11.033

Dharap, A., Pokrzywa, C., and Vemuganti, R. (2013). Increased binding of stroke-induced long non-coding RNAs to the transcriptional corepressors Sin3A and coREST. *ASN Neuro* 5, 283–289.

Dhawan, S., Georgia, S., Tschen, S. I., Fan, G., and Bhushan, A. (2011). Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. *Dev. Cell* 20, 419–429. doi: 10.1016/j.devcel.2011.03.012

Driscoll, D. J., and Migeon, B. R. (1990). Sex difference in methylation of single-copy genes in human meiotic germ cells: implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations. *Somat. Cell. Mol. Genet.* 16, 267–282. doi: 10.1007/bf01233363

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., et al. (2001). Intrinsically disordered protein. *J. Mol. Graph. Modell.* 19, 26–59.

Ebert, D. H., Gabel, H. W., Robinson, N. D., Kastan, N. R., Hu, L. S., Cohen, S., et al. (2013). Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. *Nature* 499, 341–345. doi: 10.1038/nature12348

Ehrhart, F., Coort, S. L., Cirillo, E., Smeets, E., Evelo, C. T., and Curfs, L. M. (2016). Rett syndrome - biological pathways leading from MECP2 to disorder phenotypes. *Orphanet. J. Rare Dis.* 11:158.

Einspieler, C., and Marschik, P. B. (2019). Regression in Rett syndrome: developmental pathways to its onset. *Neurosci. Biobehav. Rev.* 98, 320–332. doi: 10.1016/j.neubiorev.2019.01.028

Fichou, Y., Nectoux, J., Bahi-Buisson, N., Rosas-Vargas, H., Girard, B., Chelly, J., et al. (2009). The first missense mutation causing Rett syndrome specifically affecting the MeCP2_e1 isoform. *Neurogenetics* 10, 127–133. doi: 10.1007/s10048-008-0161-1

Filarsky, M., Zillner, K., Araya, I., Villar-Garea, A., Merkl, R., Langst, G., et al. (2015). The extended AT-hook is a novel RNA binding motif. *RNA Biol.* 12, 864–876. doi: 10.1080/15476286.2015.1060394

Fioriniello, S., Csukonyi, E., Marano, D., Brancaccio, A., Madonna, M., Zarrillo, C., et al. (2020). MeCP2 and major satellite forward RNA cooperate for pericentric heterochromatin organization. *Stem Cell Rep.* 15, 1317–1332. doi: 10.1016/j.stemcr.2020.11.006

Frasca, A., Spiombi, E., Palmieri, M., Albizzati, E., Valente, M. M., Bergo, A., et al. (2020). MECP2 mutations affect ciliogenesis: a novel perspective for Rett syndrome and related disorders. *EMBO Mol. Med.* 12:e10270.

Free, A., Wakefield, R. I., Smith, B. O., Dryden, D. T., Barlow, P. N., and Bird, A. P. (2001). DNA recognition by the methyl-CpG binding domain of MeCP2. *J. Biol. Chem.* 276, 3353–3360.

Fuks, F., Hurd, P. J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* 31, 2305–2312. doi: 10.1093/nar/gkg332

Gabel, H. W., Kinde, B., Stroud, H., Gilbert, C. S., Harmin, D. A., Kastan, N. R., et al. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* 522, 89–93. doi: 10.1038/nature14319

Gauthier, J., De Amorim, G., Mnatzakanian, G. N., Saunders, C., Vincent, J. B., Toupin, S., et al. (2005). Clinical stringency greatly improves mutation detection in Rett syndrome. *Can. J. Neurol. Sci.* 32, 321–326. doi: 10.1017/s0317167100004200

Ghosh, R. P., Horowitz-Scherer, R. A., Nikitina, T., Gierasch, L. M., and Woodcock, C. L. (2008). Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. *J. Biol. Chem.* 283, 20523–20534. doi: 10.1074/jbc.m803021200

Ghosh, R. P., Nikitina, T., Horowitz-Scherer, R. A., Gierasch, L. M., Uversky, V. N., Hite, K., et al. (2010). Unique physical properties and interactions of the

domains of methylated DNA binding protein 2. *Biochemistry* 49, 4395–4410. doi: 10.1021/bi9019753

Gianakopoulos, P. J., Zhang, Y., Pencea, N., Orlic-Milacic, M., Mittal, K., Windpassinger, C., et al. (2012). Mutations in MECP2 exon 1 in classical Rett patients disrupt MECP2_e1 transcription, but not transcription of MECP2_e2. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 159B, 210–216. doi: 10.1002/ajmg.b.32015

Glaich, O., Parikh, S., Bell, R. E., Mekahel, K., Donyo, M., Leader, Y., et al. (2019). DNA methylation directs microRNA biogenesis in mammalian cells. *Nat Commun* 10:5657.

Gold, W. A., Lacina, T. A., Cantrill, L. C., and Christodoulou, J. (2015). MeCP2 deficiency is associated with reduced levels of tubulin acetylation and can be restored using HDAC6 inhibitors. *J. Mol. Med.* 93, 63–72. doi: 10.1007/s00109-014-1202-x

Gonzales, M. L., Adams, S., Dunaway, K. W., and Lasalle, J. M. (2012). Phosphorylation of distinct sites in MeCP2 modifies cofactor associations and the dynamics of transcriptional regulation. *Mol. Cell. Biol.* 32, 2894–2903. doi: 10.1128/mcb.06728-11

Greally, J. M. (2018). A user's guide to the ambiguous word 'epigenetics'. *Nat. Rev. Mol. Cell Biol.* 19, 207–208. doi: 10.1038/nrm.2017.135

Greenberg, M. V. C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* 20, 590–607. doi: 10.1038/s41580-019-0159-6

Guo, A., Gu, H., Zhou, J., Mulhern, D., Wang, Y., Lee, K. A., et al. (2014). Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell. Proteomics* 13, 372–387. doi: 10.1074/mcp.o113.027870

Ha, M., and Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524.

Harvey, C. G., Menon, S. D., Stachowiak, B., Noor, A., Proctor, A., Mensah, A. K., et al. (2007). Sequence variants within exon 1 of MECP2 occur in females with mental retardation. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 355–360. doi: 10.1002/ajmg.b.30425

Haws, S. A., Leech, C. M., and Denu, J. M. (2020). Metabolism and the epigenome: a dynamic relationship. *Trends Biochem. Sci.* 45, 731–747. doi: 10.1016/j.tibs.2020.04.002

He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D. C., Wilusz, J. E., Garcia, B. A., et al. (2016). High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. *Mol. Cell.* 64, 416–430. doi: 10.1016/j.molcel.2016.09.034

Heckman, L. D., Chahrour, M. H., and Zoghbi, H. Y. (2014). Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. *eLife* 3:e02676.

Herbert, K. M., Pimienta, G., Degregorio, S. J., Alexandrov, A., and Steitz, J. A. (2013). Phosphorylation of DGCR8 increases its intracellular stability and induces a progrowth miRNA profile. *Cell Rep.* 5, 1070–1081. doi: 10.1016/j.celrep.2013.10.017

Hite, K. C., Kalashnikova, A. A., and Hansen, J. C. (2012). Coil-to-helix transitions in intrinsically disordered methyl CpG binding protein 2 and its isolated domains. *Protein Sci.* 21, 531–538. doi: 10.1002/pro.2037

Ho, K. L., Mcnae, I. W., Schmiedeberg, L., Klose, R. J., Bird, A. P., and Walkinshaw, M. D. (2008). MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell.* 29, 525–531. doi: 10.1016/j.molcel.2007.12.028

Horike, S., Cai, S., Miyano, M., Cheng, J. F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* 37, 31–40. doi: 10.1038/ng1491

Hosseini, E., Bagheri-Hosseinabadi, Z., De Toma, I., Jafarisani, M., and Sadeghi, I. (2019). The importance of long non-coding RNAs in neuropsychiatric disorders. *Mol. Aspects Med.* 70, 127–140. doi: 10.1016/j.mam.2019.07.004

Hudson, W. H., and Ortlund, E. A. (2014). The structure, function and evolution of proteins that bind DNA and RNA. *Nat. Rev. Mol. Cell Biol.* 15, 749–760. doi: 10.1038/nrm3884

Itoh, M., Tahimic, C. G., Ide, S., Otsuki, A., Sasaoka, T., Noguchi, S., et al. (2012). Methyl CpG-binding protein isoform MeCP2_e2 is dispensable for Rett syndrome phenotypes but essential for embryo viability and placenta development. *J. Biol. Chem.* 287, 13859–13867. doi: 10.1074/jbc.m111.309864

Jarvelin, A. I., Noerenberg, M., Davis, I., and Castello, A. (2016). The new (dis)order in RNA regulation. *Cell Commun. Signal.* 14:9.

Jeffery, L., and Nakielny, S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. *J. Biol. Chem.* 279, 49479–49487. doi: 10.1074/jbc.m409070200

Jiang, L., Shao, C., Wu, Q. J., Chen, G., Zhou, J., Yang, B., et al. (2017). NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. *Nat. Struct. Mol. Biol.* 24, 816–824. doi: 10.1038/nsmb.3455

Jones, P. L., Wade, P. A., and Wolffe, A. P. (2001). Purification of the MeCP2/histone deacetylase complex from *Xenopus laevis*. *Methods Mol. Biol.* 181, 297–307. doi: 10.1385/1-59259-211-2:297

Kaufmann, W. E., Jarrar, M. H., Wang, J. S., Lee, Y.-J. M., Reddy, S., Bibat, G., et al. (2005a). Histone modifications in Rett syndrome lymphocytes: a preliminary evaluation. *Brain Dev.* 27, 331–339. doi: 10.1016/j.braindev.2004.09.005

Kaufmann, W. E., Johnston, M. V., and Blue, M. E. (2005b). MeCP2 expression and function during brain development: implications for Rett syndrome's pathogenesis and clinical evolution. *Brain Dev.* 27(Suppl. 1), S77–S87.

Khan, A. W., Ziemann, M., Rafehi, H., Maxwell, S., Ciccotosto, G. D., and El-Osta, A. (2017). MeCP2 interacts with chromosomal microRNAs in brain. *Epigenetics* 12, 1028–1037. doi: 10.1080/15592294.2017.1391429

Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., et al. (2012). 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat. Struct. Mol. Biol.* 19, 1037–1043. doi: 10.1038/nsmb.2372

Kharrat, M., Hsairi, I., Fendri-Kriaa, N., Kenoun, H., Othmen, H. B., Ben Mahmoud, A., et al. (2015). A novel mutation p.A59P in N-terminal domain of methyl-CpG-binding protein 2 confers phenotypic variability in 3 cases of tunisian rett patients: clinical evaluations and in silico investigations. *J. Child Neurol.* 30, 1715–1721. doi: 10.1177/0883073815578529

Kinde, B., Gabel, H. W., Gilbert, C. S., Griffith, E. C., and Greenberg, M. E. (2015). Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6800–6806. doi: 10.1073/pnas.1411269112

Kishi, N., Macdonald, J. L., Ye, J., Molyneaux, B. J., Azim, E., and Macklis, J. D. (2016). Reduction of aberrant NF-kappaB signalling ameliorates Rett syndrome phenotypes in Mecp2-null mice. *Nat. Commun.* 7:10520.

Kitazawa, R., and Kitazawa, S. (2007). Methylation status of a single CpG locus 3 bases upstream of TATA-box of receptor activator of nuclear factor-kappaB ligand (RANKL) gene promoter modulates cell- and tissue-specific RANKL expression and osteoclastogenesis. *Mol. Endocrinol.* 21, 148–158. doi: 10.1210/me.2006-0205

Kriaucionis, S., and Bird, A. (2003). DNA methylation and Rett syndrome. *Hum. Mol. Genet.* 2, R221–R227.

Kriaucionis, S., and Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res.* 32, 1818–1823. doi: 10.1093/nar/gkh349

Krishnaraj, R., Ho, G., and Christodoulou, J. (2017). RettBASE: Rett syndrome database update. *Hum. Mutat.* 38, 922–931. doi: 10.1002/humu.23263

Kruusvee, V., Lyst, M. J., Taylor, C., Tarnauskaite, Z., Bird, A. P., and Cook, A. G. (2017). Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders. *Proc. Natl. Acad. Sci. U.S.A.* 114, E3243–E3250.

Kucukkal, T. G., Yang, Y., Uvarov, O., Cao, W., and Alexov, E. (2015). Impact of Rett Syndrome Mutations on MeCP2 MBD Stability. *Biochemistry* 54, 6357–6368. doi: 10.1021/acs.biochem.5b00790

Kudo, S., Nomura, Y., Segawa, M., Fujita, N., Nakao, M., Schanen, C., et al. (2003). Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain. *J. Med. Genet.* 40, 487–493. doi: 10.1136/jmg.40.7.487

Kyle, S. M., Saha, P. K., Brown, H. M., Chan, L. C., and Justice, M. J. (2016). MeCP2 co-ordinates liver lipid metabolism with the NCoR1/HDAC3 corepressor complex. *Hum. Mol. Genet.* 25, 3029–3041.

Kyle, S. M., Vashi, N., and Justice, M. J. (2018). Rett syndrome: a neurological disorder with metabolic components. *Open Biol.* 8:170216. doi: 10.1098/rsob.170216

Lamonica, J. M., Kwon, D. Y., Goffin, D., Fenik, P., Johnson, B. S., Cui, Y., et al. (2017). Elevating expression of MeCP2 T158M rescues DNA binding and Rett syndrome-like phenotypes. *J. Clin. Invest.* 127, 1889–1904. doi: 10.1172/jci90967

Landers, C. C., Rabeler, C. A., Ferrari, E. K., Alessandro, R. L. D., Kang, D. D., Malisa, J., et al. (2020). Ectopic expression of pericentric HSATII RNA results in nuclear RNA accumulation, MeCP2 recruitment and cell division deffects. *bioRxiv* [Preprint]. doi: 10.1101/2020.04.30.064329v1.full

Lavery, L. A., and Zoghbi, H. Y. (2019). The distinct methylation landscape of maturing neurons and its role in Rett syndrome pathogenesis. *Curr. Opin. Neurobiol.* 59, 180–188. doi: 10.1016/j.conb.2019.08.001

Lee, W., Kim, J., Yun, J. M., Ohn, T., and Gong, Q. (2020). MeCP2 regulates gene expression through recognition of H3K27me3. *Nat. Commun.* 11:3140.

Lehman, N. L. (2009). The ubiquitin proteasome system in neuropathology. *Acta Neuropathol.* 118, 329–347. doi: 10.1007/s00401-009-0560-x

Lessing, D., Dial, T. O., Wei, C., Payer, B., Carrette, L. L., Kesner, B., et al. (2016). A high-throughput small molecule screen identifies synergism between DNA methylation and Aurora kinase pathways for X reactivation. *Proc. Natl. Acad. Sci. U.S.A.* 113, 14366–14371. doi: 10.1073/pnas.1617597113

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., et al. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905–914. doi: 10.1016/0092-8674(92)90610-o

Li, C., Jiang, S., Liu, S. Q., Lykken, E., Zhao, L. T., Sevilla, J., et al. (2014). MeCP2 enforces Foxp3 expression to promote regulatory T cells' resilience to inflammation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2807–E2816.

Li, C. H., Coffey, E. L., Dall'agnese, A., Hannett, N. M., Tang, X., Henninger, J. E., et al. (2020). MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature* 5867, 440–444. doi: 10.1038/s41586-020-2574-4

Li, Q., Loh, D. H., Kudo, T., Truong, D., Derakhshesh, M., Kaswan, Z. M., et al. (2015). Circadian rhythm disruption in a mouse model of Rett syndrome circadian disruption in RTT. *Neurobiol. Dis.* 77, 155–164. doi: 10.1016/j.nbd.2015.03.009

Li, R., Dong, Q., Yuan, X., Zeng, X., Gao, Y., Chiao, C., et al. (2016). Misregulation of alternative splicing in a mouse model of rett syndrome. *PLoS Genet.* 12:e1006129. doi: 10.1371/journal.pgen.1006129

Lilja, T., Wallenborg, K., Bjorkman, K., Albage, M., Eriksson, M., Lagercrantz, H., et al. (2013). Novel alterations in the epigenetic signature of MeCP2-targeted promoters in lymphocytes of Rett syndrome patients. *Epigenetics* 8, 246–251. doi: 10.4161/epi.23752

Lingbeck, J. M., Trausch-Azar, J. S., Ciechanover, A., and Schwartz, A. L. (2003). Determinants of nuclear and cytoplasmic ubiquitin-mediated degradation of MyoD. *J. Biol. Chem.* 278, 1817–1823. doi: 10.1074/jbc.m208815200

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science* 341:1237905. doi: 10.1126/science.1237905

Liyanage, V., and Rastegar, M. (2014). Rett syndrome and MeCP2. *Neuromol. Med.* 16, 231–264.

Liyanage, V. R. B., Olson, C. O., Zachariah, R. M., Davie, J. R., and Rastegar, M. (2019). DNA methylation contributes to the differential expression levels of Mecp2 in male mice neurons and astrocytes. *Int. J. Mol. Sci.* 20:1845. doi: 10.3390/ijms20081845

Long, S. W., Ooi, J. Y., Yau, P. M., and Jones, P. L. (2011). A brain-derived MeCP2 complex supports a role for MeCP2 in RNA processing. *Biosci. Rep.* 31, 333–343. doi: 10.1042/bsr20100124

Lovci, M. T., Bengtson, M. H., and Massirer, K. B. (2016). "Post-translational modifications and RNA-binding proteins," in *RNA Processing. Advances in Experimental Medicine and Biology*, ed. G. E. Yeo (Cham: Springer), 297–317. doi: 10.1007/978-3-319-29073-7_12

Lyst, M. J., Ekiert, R., Ebert, D. H., Merusi, C., Nowak, J., Selfridge, J., et al. (2013). Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat. Neurosci.* 16, 898–902. doi: 10.1038/nn.3434

Lyst, M. J., Ekiert, R., Guy, J., Selfridge, J., Koerner, M. V., Merusi, C., et al. (2018). Affinity for DNA contributes to NLS independent nuclear localization of MeCP2. *Cell Rep.* 24, 2213–2220. doi: 10.1016/j.celrep.2018.07.099

Mann, J., Chu, D. C., Maxwell, A., Oakley, F., Zhu, N. L., Tsukamoto, H., et al. (2010). MeCP2 controls an epigenetic pathway that promotes myofibroblast transdifferentiation and fibrosis. *Gastroenterology* 138, 705–714. doi: 10.1053/j.gastro.2009.10.002

Marano, D., Fioriniello, S., Fiorillo, F., Gibbons, R. J., D'esposito, M., and Della Ragione, F. (2019). ATRX contributes to MeCP2-mediated pericentric heterochromatin organization during neural differentiation. *Int. J. Mol. Sci.* 20:5371. doi: 10.3390/ijms20215371

Marmorstein, R., and Zhou, M. M. (2014). Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harb. Perspect. Biol.* 6:a018762. doi: 10.1101/cshperspect.a018762

Martínez de Paz, A., Khajavi, L., Martin, H., Claveria-Gimeno, R., Tom Dieck, S., Cheema, M. S., et al. (2019). MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. *Epigenet. Chromatin* 12:63.

Martinez de Paz, A., Vicente Sanchez-Mut, J., Samitier-Marti, M., Petazzi, P., Saez, M., Szczesna, K., et al. (2015). Circadian cycle-dependent MeCP2 and brain chromatin changes. *PLoS One* 10:e0123693. doi: 10.1371/journal.pone.0123693

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., et al. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* 302, 890–893. doi: 10.1126/science.1090842

Maunakea, A. K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* 23, 1256–1269. doi: 10.1038/cr.2013.110

Maxwell, S. S., Pelka, G. J., Tam, P. P., and El-Osta, A. (2013). Chromatin context and ncRNA highlight targets of MeCP2 in brain. *RNA Biol.* 10, 1741–1757. doi: 10.4161/rna.26921

McArthur, A. J., and Budden, S. S. (1998). Sleep dysfunction in Rett syndrome: a trial of exogenous melatonin treatment. *Dev. Med. Child Neurol.* 40, 186–192. doi: 10.1111/j.1469-8749.1998.tb15445.x

McCarthy, S. E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., et al. (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol. Psychiatry* 19, 652–658. doi: 10.1038/mp.2014.29

Mendoza-Viveros, L., Chiang, C. K., Ong, J. L. K., Hegazi, S., Cheng, A. H., Bouchard-Cannon, P., et al. (2017). miR-132/212 modulates seasonal adaptation and dendritic morphology of the central circadian clock. *Cell Rep.* 19, 505–520. doi: 10.1016/j.celrep.2017.03.057

Merritt, J. K., Collins, B. E., Erickson, K. R., Dong, H., and Neul, J. L. (2020). Pharmacological read-through of R294X Mecp2 in a novel mouse model of Rett syndrome. *Hum. Mol. Genet.* 29, 2461–2470. doi: 10.1093/hmg/ddaa102

Miao, C. G., Yang, Y. Y., He, X., and Li, J. (2013). New advances of DNA methylation and histone modifications in rheumatoid arthritis, with special emphasis on MeCP2. *Cell. Signal.* 25, 875–882. doi: 10.1016/j.cellsig.2012.12.017

Minkovsky, A., Sahakyan, A., Bonora, G., Damoiseaux, R., Dimitrova, E., Rubbi, L., et al. (2015). A high-throughput screen of inactive X chromosome reactivation identifies the enhancement of DNA demethylation by 5-aza-2'-dC upon inhibition of ribonucleotide reductase. *Epigenet. Chromatin* 8:42.

Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., Macleod, P. J., et al. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.* 36, 339–341. doi: 10.1038/ng1327

Moncla, A., Kpebe, A., Missirian, C., Mancini, J., and Villard, L. (2002). Polymorphisms in the C-terminal domain of MECP2 in mentally handicapped boys: implications for genetic counselling. *Eur. J. Hum. Genet.* 10, 86–89. doi: 10.1038/sj.ejhg.5200761

Motil, K. J., Caeg, E., Barrish, J. O., Geerts, S., Lane, J. B., Percy, A. K., et al. (2012). Gastrointestinal and nutritional problems occur frequently throughout life in girls and women with Rett syndrome. *J. Pediatr. Gastroenterol. Nutr.* 55, 292–298. doi: 10.1097/mpg.0b013e31824b6159

Nan, X., Campoy, F. J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88, 471–481. doi: 10.1016/s0092-8674(00)81887-5

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M. J., Shu, X., et al. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2709–2714. doi: 10.1073/pnas.0608056104

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., et al. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386–389. doi: 10.1038/30764

Neul, J. L., Benke, T. A., Marsh, E. D., Skinner, S. A., Merritt, J., Lieberman, D. N., et al. (2019). The array of clinical phenotypes of males with mutations

in Methyl-CpG binding protein 2. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 180, 55–67.

Neul, J. L., Fang, P., Barrish, J., Lane, J., Caeg, E. B., Smith, E. O., et al. (2008). Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* 70, 1313–1321. doi: 10.1212/01.wnl.0000291011.54508.aa

Neul, J. L., Lane, J. B., Lee, H. S., Geerts, S., Barrish, J. O., Annese, F., et al. (2014). Developmental delay in Rett syndrome: data from the natural history study. *J. Neurodev. Disord.* 6, 20. doi: 10.1186/1866-1955-6-20

Neupane, M., Clark, A. P., Landini, S., Birkbak, N. J., Eklund, A. C., Lim, E., et al. (2015). MECP2 is a frequently amplified oncogene with a novel epigenetic mechanism that mimics the role of activated RAS in malignancy. *Cancer Discov.* 6, 45–58. doi: 10.1158/2159-8290.cd-15-0341

Nikitina, T., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., Grigoryev, S. A., and Woodcock, C. L. (2007a). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. *J. Biol. Chem.* 282, 28237–28245. doi: 10.1074/jbc.m704304200

Nikitina, T., Shi, X., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., and Woodcock, C. L. (2007b). Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin. *Mol. Cell. Biol.* 27, 864–877. doi: 10.1128/mcb.01593-06

O'Connor, R. D., Zayzafoon, M., Farach-Carson, M. C., and Schanen, N. C. (2009). Mecp2 deficiency decreases bone formation and reduces bone volume in a rodent model of Rett syndrome. *Bone* 45, 346–356. doi: 10.1016/j.bone.2009.04.251

Olson, C. O., Pejhan, S., Kroft, D., Sheikholeslami, K., Fuss, D., Buist, M., et al. (2018). MECP2 mutation interrupts nucleolin-mTOR-P70S6K signaling in rett syndrome patients. *Front. Genet.* 9:635. doi: 10.3389/fgene.2018.00635

Olson, C. O., Zachariah, R. M., Ezeonwuka, C. D., Liyanage, V. R., and Rastegar, M. (2014). Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements. *PLoS One* 9:e90645. doi: 10.1371/journal.pone.0090645

Osenberg, S., Karten, A., Sun, J., Li, J., Charkowick, S., Felice, C. A., et al. (2018). Activity-dependent aberrations in gene expression and alternative splicing in a mouse model of Rett syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 115, E5363–E5372.

Pandey, S., Simmons, G. E. Jr., Malyarchuk, S., Calhoun, T. N., and Pruitt, K. (2015). A novel MeCP2 acetylation site regulates interaction with ATRX and HDAC1. *Genes Cancer* 6, 408–421. doi: 10.18632/genesandcancer.84

Pazin, M. J., and Kadonaga, J. T. (1997). SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein-DNA interactions? *Cell* 88, 737–740. doi: 10.1016/s0092-8674(00)81918-2

Pecorelli, A., Leoni, G., Cervellati, F., Canali, R., Signorini, C., Leoncini, S., et al. (2013). Genes related to mitochondrial functions, protein degradation, and chromatin folding are differentially expressed in lymphomonocytes of Rett syndrome patients. *Mediators Inflamm.* 2013:137629.

Pehjan, S., Del Bigio, M. R., and Rastegar, M. (2020). The MeCP2E1/E2-BDNF-miR132 homeostasis regulatory network is region-dependent in the human brain and is impaired in rett syndrome patients. *Front. Cell Dev. Biol.* 8:763. doi: 10.3389/fcell.2020.00763

Petazzi, P., Sandoval, J., Szczesna, K., Jorge, O. C., Roa, L., Sayols, S., et al. (2013). Dysregulation of the long non-coding RNA transcriptome in a Rett syndrome mouse model. *RNA Biol.* 10, 1197–1203. doi: 10.4161/rna.24286

Pillion, J. P., Rawool, V. W., Bibat, G., and Naidu, S. (2003). Prevalence of hearing loss in Rett syndrome. *Dev. Med. Child Neurol.* 45, 338–343. doi: 10.1111/j.1469-8749.2003.tb00405.x

Quinodoz, S. A., Bhat, P., Ollikainen, N., Jachowicz, J. W., Banerjee, A. K., Peter, C., et al. (2020). RNA promotes the formation of spatial compartments in the nucleus. *BioRxiv* [Preprint]. doi: 10.1101/2020.1108.1125.267435

Rechsteiner, M., and Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* 21, 267–271. doi: 10.1016/s0968-0004(96)10031-1

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* 277, 63–81. doi: 10.1016/s0378-1119(01)00689-8

Roberts, T. C. (2014). The MicroRNA biology of the mammalian nucleus. *Mol. Ther. Nucleic Acids* 3:e188. doi: 10.1038/mtna.2014.40

Roberts, T. C., Morris, K. V., and Wood, M. J. (2014). The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130507. doi: 10.1098/rstb.2013.0507

Rodrigues, D. C., Kim, D. S., Yang, G., Zaslavsky, K., Ha, K. C., Mok, R. S., et al. (2016). MECP2 Is Post-transcriptionally Regulated during Human Neurodevelopment by Combinatorial Action of RNA-Binding Proteins and miRNAs. *Cell. Rep.* 17, 720–734. doi: 10.1016/j.celrep.2016.09.049

Rodrigues, D. C., Mufteev, M., and Ellis, J. (2020). Regulation, diversity and function of MECP2 Exon and 3'UTR isoforms. *Hum. Mol. Genet.* 29, R89–R99.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234, 364–368. doi: 10.1126/science.2876518

Rose, S. A., Wass, S., Jankowski, J. J., Feldman, J. F., and Djukic, A. (2019). Impaired visual search in children with rett syndrome. *Pediatr. Neurol.* 92, 26–31. doi: 10.1016/j.pediatrneurol.2018.10.002

Sakaguchi, Y., Uehara, T., Suzuki, H., Sakamoto, Y., Fujiwara, M., Kosaki, K., et al. (2018). Haploinsufficiency of NCOR1 associated with autism spectrum disorder, scoliosis, and abnormal palatogenesis. *Am. J. Med. Genet. A* 176, 2466–2469. doi: 10.1002/ajmg.a.40354

Sanuki, R., Onishi, A., Koike, C., Muramatsu, R., Watanabe, S., Muranishi, Y., et al. (2011). miR-124a is required for hippocampal axogenesis and retinal cone survival through Lhx2 suppression. *Nat. Neurosci.* 14, 1125–1134. doi: 10.1038/nn.2897

Saunders, C. J., Minassian, B. E., Chow, E. W., Zhao, W., and Vincent, J. B. (2009). Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. *Am. J. Med. Genet. A* 149A, 1019–1023. doi: 10.1002/ajmg.a.32776

Saxena, A., De Lagarde, D., Leonard, H., Williamson, S. L., Vasudevan, V., Christodoulou, J., et al. (2006). Lost in translation: translational interference from a recurrent mutation in exon 1 of MECP2. *J. Med. Genet.* 43, 470–477. doi: 10.1136/jmg.2005.036244

Schanen, C., Houwink, E. J., Dorrani, N., Lane, J., Everett, R., Feng, A., et al. (2004). Phenotypic manifestations of MECP2 mutations in classical and atypical Rett syndrome. *Am. J. Med. Genet. A* 126A, 129–140. doi: 10.1002/ajmg.a.20571

Schule, B., Li, H. H., Fisch-Kohl, C., Purmann, C., and Francke, U. (2007). DLX5 and DLX6 expression is biallelic and not modulated by MeCP2 deficiency. *Am. J. Hum. Genet.* 81, 492–506. doi: 10.1086/520063

Seto, E., and Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb. Perspect. Biol.* 6:a018713. doi: 10.1101/cshperspect.a018713

Shah, R. R., and Bird, A. P. (2017). MeCP2 mutations: progress towards understanding and treating Rett syndrome. *Genome Med.* 9:17.

Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., et al. (2002). Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* 35, 243–254. doi: 10.1016/s0896-6273(02)00768-7

Shahbazian, M. D., Antalffy, B., Armstrong, D. L., and Zoghbi, H. Y. (2002). Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum. Mol. Genet.* 11, 115–124. doi: 10.1093/hmg/11.2.115

Shahbazian, M. D., and Zoghbi, H. Y. (2002). Rett syndrome and MeCP2: linking epigenetics and neuronal function. *Am. J. Hum. Genet.* 71, 1259–1272. doi: 10.1086/345360

Sheikh, T. I., Ausio, J., Faghfoury, H., Silver, J., Lane, J. B., Eubanks, J. H., et al. (2016). From Function to Phenotype: Impaired DNA Binding and Clustering Correlates with Clinical Severity in Males with Missense Mutations in MECP2. *Sci. Rep.* 6:38590.

Sheikh, T. I., De Paz, A. M., Akhtar, S., Ausio, J., and Vincent, J. B. (2017). MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation. *Hum. Mol. Genet.* 26, 4132–4141. doi: 10.1093/hmg/ddx300

Sheikh, T. I., Harripaul, R., Ayub, M., and Vincent, J. B. (2018). MeCP2 AT-Hook1 mutations in patients with intellectual disability and/or schizophrenia disrupt DNA binding and chromatin compaction in vitro. *Hum. Mutat.* 39, 717–728. doi: 10.1002/humu.23409

Sheikh, T. I., Mittal, K., Willis, M. J., and Vincent, J. B. (2013). A synonymous change, p.*Gly*16Gly in MECP2 Exon 1, causes a cryptic splice event in a Rett syndrome patient. *Orphanet. J. Rare Dis.* 8:108. doi: 10.1186/1750-1172-8-108

Sheinerman, K., Djukic, A., Tsivinsky, V. G., and Umansky, S. R. (2019). Brain-enriched microRNAs circulating in plasma as novel biomarkers for Rett syndrome. *PLoS One* 14:e0218623. doi: 10.1371/journal.pone.0218623

Sinnamon, J. R., Kim, S. Y., Fisk, J. R., Song, Z., Nakai, H., Jeng, S., et al. (2020). *In vivo* repair of a protein underlying a neurological disorder by programmable RNA editing. *Cell Rep.* 32:107878. doi: 10.1016/j.celrep.2020.107878

Sperlazza, M. J., Bilinovich, S. M., Sinanan, L. M., and Javier, F. R. (2017). Structural Basis of MeCP2 Distribution on Non-CpG Methylated and Hydroxymethylated DNA. *J. Mol. Biol.* 429, 1581–1594. doi: 10.1016/j.jmb.2017.04.009

Spiga, O., Gardini, S., Rossi, N., Cicaloni, V., Pettini, F., Niccolai, N., et al. (2019). Structural investigation of Rett-inducing MeCP2 mutations. *Genes Dis.* 6, 31–34. doi: 10.1016/j.gendis.2018.09.005

Squillaro, T., Alessio, N., Capasso, S., Di Bernardo, G., Melone, M. A. B., Peluso, G., et al. (2019). Senescence phenomena and metabolic alteration in mesenchymal stromal cells from a mouse model of rett syndrome. *Int. J. Mol. Sci.* 20:2508. doi: 10.3390/ijms20102508

Sripathy, S., Leko, V., Adrianse, R. L., Loe, T., Foss, E. J., Dalrymple, E., et al. (2017). Screen for reactivation of MeCP2 on the inactive X chromosome identifies the BMP/TGF-beta superfamily as a regulator of XIST expression. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1619–1624. doi: 10.1073/pnas.1621356114

Strati, F., Calabro, A., Donati, C., De Felice, C., Hayek, J., Jousson, O., et al. (2018). Intestinal Candida parapsilosis isolates from Rett syndrome subjects bear potential virulent traits and capacity to persist within the host. *BMC Gastroenterol.* 18:57. doi: 10.1186/s12876-018-0785-z

Strati, F., Cavalieri, D., Albanese, D., De Felice, C., Donati, C., Hayek, J., et al. (2016). Altered gut microbiota in Rett syndrome. *Microbiome* 4:41.

Stuss, D. P., Cheema, M., Ng, M. K., Martinez de Paz, A., Williamson, B., Missiaen, K., et al. (2013). Impaired in vivo binding of MeCP2 to chromatin in the absence of its DNA methyl-binding domain. *Nucleic Acids Res.* 41, 4888–4900. doi: 10.1093/nar/gkt213

Subbanna, S., Nagre, N. N., Shivakumar, M., Umapathy, N. S., Psychoyos, D., and Basavarajappa, B. S. (2014). Ethanol induced acetylation of histone at G9a exon1 and G9a-mediated histone H3 dimethylation leads to neurodegeneration in neonatal mice. *Neuroscience* 258, 422–432. doi: 10.1016/j.neuroscience.2013.11.043

Takeguchi, R., Takahashi, S., Kuroda, M., Tanaka, R., Suzuki, N., Tomonoh, Y., et al. (2020). MeCP2_e2 partially compensates for lack of MeCP2_e1: a male case of Rett syndrome. *Mol. Genet. Genomic Med.* 8:e1088.

Takeuchi, A., Schmitt, D., Chapple, C., Babaylova, E., Karpova, G., Guigo, R., et al. (2009). A short motif in *Drosophila* SECIS Binding Protein 2 provides differential binding affinity to SECIS RNA hairpins. *Nucleic Acids Res.* 37, 2126–2141. doi: 10.1093/nar/gkp078

Tan, Q., and Zoghbi, H. Y. (2019). Mouse models as a tool for discovering new neurological diseases. *Neurobiol. Learn. Mem.* 165:106902. doi: 10.1016/j.nlm.2018.07.006

Tang, X., Drotar, J., Li, K., Clairmont, C. D., Brumm, A. S., Sullins, A. J., et al. (2019). Pharmacological enhancement of KCC2 gene expression exerts therapeutic effects on human Rett syndrome neurons and Mecp2 mutant mice. *Sci. Transl. Med.* 11:eaau0164. doi: 10.1126/scitranslmed.aau0164

Thakur, J., Fang, H., Lagas, T., Disteche, C. M., and Henikoff, S. (2020). Architectural RNA isequired for heterochromatin organization. *bioRxiv* 3:784835.

Thambirajah, A. A., Eubanks, J. H., and Ausio, J. (2009). MeCP2 post-translational regulation through PEST domains: two novel hypotheses: potential relevance and implications for Rett syndrome. *Bioessays* 31, 561–569. doi: 10.1002/bies.200800220

Thambirajah, A. A., Ng, M. K., Frehlick, L. J., Li, A., Serpa, J. J., Petrotchenko, E. V., et al. (2012). MeCP2 binds to nucleosome free (linker DNA) regions and to H3K9/H3K27 methylated nucleosomes in the brain. *Nucleic Acids Res.* 40, 2884–2897. doi: 10.1093/nar/gkr1066

Thandapani, P., O'Connor, T. R., Bailey, T. L., and Richard, S. (2013). Defining the RGG/RG motif. *Mol. Cell.* 50, 613–623. doi: 10.1016/j.molcel.2013.05.021

Thatcher, K. N., and Lasalle, J. M. (2006). Dynamic changes in Histone H3 lysine 9 acetylation localization patterns during neuronal maturation require MeCP2. *Epigenetics* 1, 24–31.

Tillotson, R., and Bird, A. (2019). The molecular basis of MeCP2 function in the brain. *J. Mol. Biol.* [Epub ahead of print].

Tillotson, R., Selfridge, J., Koerner, M. V., Gadalla, K. K. E., Guy, J., De Sousa, D., et al. (2017). Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. *Nature* 550, 398–401. doi: 10.1038/nature24058

Tomatsu, S., Orii, K. O., Bi, Y., Gutierrez, M. A., Nishioka, T., Yamaguchi, S., et al. (2004). General implications for CpG hot spot mutations: methylation patterns of the human iduronate-2-sulfatase gene locus. *Hum. Mutat.* 23, 590–598. doi: 10.1002/humu.20046

Trappe, R., Laccone, F., Cobilanschi, J., Meins, M., Huppke, P., Hanefeld, F., et al. (2001). MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am. J. Hum. Genet.* 68, 1093–1101. doi: 10.1086/320109

Trendel, J., Schwarzl, T., Horos, R., Prakash, A., Bateman, A., Hentze, M. W., et al. (2019). The human RNA-binding proteome and its dynamics during translational arrest. *Cell* 176:e319.

Tsuchiya, Y., Minami, Y., Umemura, Y., Watanabe, H., Ono, D., Nakamura, W., et al. (2015). Disruption of MeCP2 attenuates circadian rhythm in CRISPR/Cas9-based Rett syndrome model mouse. *Genes Cells* 20, 992–1005. doi: 10.1111/gtc.12305

Tsujimura, K., Irie, K., Nakashima, H., Egashira, Y., Fukao, Y., Fujiwara, M., et al. (2015). miR-199a Links MeCP2 with mTOR signaling and its dysregulation leads to rett syndrome phenotypes. *Cell Rep.* 12, 1887–1901. doi: 10.1016/j.celrep.2015.08.028

Tyssowski, K., Kishi, Y., and Gotoh, Y. (2014). Chromatin regulators of neural development. *Neuroscience* 264, 4–16. doi: 10.1016/j.neuroscience.2013.10.008

Urdinguio, R. G., Pino, I., Ropero, S., Fraga, M. F., and Esteller, M. (2007). Histone H3 and H4 modification profiles in a Rett syndrome mouse model. *Epigenetics* 2, 11–14. doi: 10.4161/epi.2.1.3698

Vashi, N., and Justice, M. J. (2019). Treating Rett syndrome: from mouse models to human therapies. *Mamm. Genome* 30, 90–110. doi: 10.1007/s00335-019-09793-5

Villard, L. (2007). MECP2 mutations in males. *J. Med. Genet.* 44, 417–423. doi: 10.1136/jmg.2007.049452

Wan, M., Lee, S. S., Zhang, X., Houwink-Manville, I., Song, H. R., Amir, R. E., et al. (1999). Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. *Am. J. Hum. Genet.* 65, 1520–1529. doi: 10.1086/302690

Wan, M., Zhao, K., Lee, S. S., and Francke, U. (2001). MECP2 truncating mutations cause histone H4 hyperacetylation in Rett syndrome. *Hum. Mol. Genet.* 10, 1085–1092. doi: 10.1093/hmg/10.10.1085

Wang, C., Wang, F., Cao, Q., Li, Z., Huang, L., and Chen, S. (2018). The effect of Mecp2 on heart failure. *Cell Physiol. Biochem.* 47, 2380–2387. doi: 10.1159/000491610

Wang, L., Hu, M., Zuo, M. Q., Zhao, J., Wu, D., Huang, L., et al. (2020). Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Res.* 30, 393–407. doi: 10.1038/s41422-020-0288-7

Weise, S. C., Arumugam, G., Villarreal, A., Videm, P., Heidrich, S., Nebel, N., et al. (2019). FOXG1 regulates PRKAR2B transcriptionally and posttranscriptionally via miR200 in the adult hippocampus. *Mol. Neurobiol.* 56, 5188–5201. doi: 10.1007/s12035-018-1444-7

Wen, L., Li, X., Yan, L., Tan, Y., Li, R., Zhao, Y., et al. (2014). Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol.* 15:R49.

Wen, Y., Wang, J., Zhang, Q., Chen, Y., Wu, X., and Bao, X. (2020). MECP2 mutation spectrum and its clinical characteristics in a Chinese cohort. *Clin. Genet.* 98, 240–250. doi: 10.1111/cge.13790

Weyn-Vanhentenryck, S. M., Feng, H., Ustianenko, D., Duffie, R., Yan, Q., Jacko, M., et al. (2018). Precise temporal regulation of alternative splicing during neural development. *Nat. Commun.* 9:2189.

Wilson, S. A., Brown, E. C., Kingsman, A. J., and Kingsman, S. M. (1998). TRIP: a novel double stranded RNA binding protein which interacts with the leucine rich repeat of flightless I. *Nucleic Acids Res.* 26, 3460–3467. doi: 10.1093/nar/26.15.3460

Wong, J. J., Gao, D., Nguyen, T. V., Kwok, C. T., Van Geldermalsen, M., Middleton, R., et al. (2017). Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. *Nat. Commun.* 8:15134.

Wu, H., Tao, J., Chen, P. J., Shahab, A., Ge, W., Hart, R. P., et al. (2010). Genome-wide analysis reveals methyl-CpG-binding protein 2-dependent regulation of

microRNAs in a mouse model of Rett syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18161–18166. doi: 10.1073/pnas.1005595107

Xu, Y., Wu, W., Han, Q., Wang, Y., Li, C., Zhang, P., et al. (2019). Post-translational modification control of RNA-binding protein hnRNPK function. *Open Biol.* 9:180239. doi: 10.1098/rsob.180239

Xue, J., Wijeratne, S. S., and Zempleni, J. (2013). Holocarboxylase synthetase synergizes with methyl CpG binding protein 2 and DNA methyltransferase 1 in the transcriptional repression of long-terminal repeats. *Epigenetics* 8, 504–511. doi: 10.4161/epi.24449

Yamazaki, T., Nakagawa, S., and Hirose, T. (2020). Architectural RNAs for membraneless nuclear body formation. *Cold Spring Harb. Symp. Quant. Biol.* 84, 227–237. doi: 10.1101/sqb.2019.84.039404

Yang, F., Deng, X., Ma, W., Berletch, J. B., Rabaia, N., Wei, G., et al. (2015). The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol.* 16:52.

Yang, Y., Kucukkal, T. G., Li, J., Alexov, E., and Cao, W. (2016). Binding analysis of Methyl-CpG binding domain of MeCP2 and Rett syndrome mutations. *ACS Chem. Biol.* 11, 2706–2715. doi: 10.1021/acschembio.6b00450

Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., et al. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19416–19421. doi: 10.1073/pnas.0707442104

Yoon, S. H., Choi, J., Lee, W. J., and Do, J. T. (2020). Genetic and epigenetic etiology underlying autism spectrum disorder. *J. Clin. Med.* 9:966. doi: 10.3390/jcm9040966

Young, J. I., Hong, E. P., Castle, J. C., Crespo-Barreto, J., Bowman, A. B., Rose, M. F., et al. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17551–17558. doi: 10.1073/pnas.0507856102

Yusufzai, T. M., and Wolffe, A. P. (2000). Functional consequences of Rett syndrome mutations on human MeCP2. *Nucleic Acids Res.* 28, 4172–4179. doi: 10.1093/nar/28.21.4172

Zachariah, R. M., Olson, C. O., Ezeonwuka, C., and Rastegar, M. (2012). Novel MeCP2 isoform-specific antibody reveals the endogenous MeCP2E1 expression in murine brain, primary neurons and astrocytes. *PLoS One* 7:e49763. doi: 10.1371/journal.pone.0049763

Zaghlula, M., Glaze, D. G., Enns, G. M., Potocki, L., Schwabe, A. L., and Suter, B. (2018). Current clinical evidence does not support a link between TBL1XR1 and Rett syndrome: description of one patient with Rett features and a novel mutation in TBL1XR1, and a review of TBL1XR1 phenotypes. *Am. J. Med. Genet. A* 176, 1683–1687. doi: 10.1002/ajmg.a.38689

Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., et al. (2019). Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* 20:5573. doi: 10.3390/ijms20225573

Zocchi, L., and Sassone-Corsi, P. (2013). SIRT1-mediated deacetylation of MeCP2 contributes to BDNF expression. *Epigenetics* 7, 695–700. doi: 10.4161/epi.20733

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 15 February 2021
doi: 10.3389/fgene.2021.624581

Epigenetics in Prader-Willi Syndrome

*Aron Judd P. Mendiola and Janine M. LaSalle**

Department of Medical Microbiology and Immunology, Genome Center, MIND Institute, University of California, Davis, Davis, CA, United States

Prader-Willi Syndrome (PWS) is a rare neurodevelopmental disorder that affects approximately 1 in 20,000 individuals worldwide. Symptom progression in PWS is classically characterized by two nutritional stages. Stage 1 is hypotonia characterized by poor muscle tone that leads to poor feeding behavior causing failure to thrive in early neonatal life. Stage 2 is followed by the development of extreme hyperphagia, also known as insatiable eating and fixation on food that often leads to obesity in early childhood. Other major features of PWS include obsessive-compulsive and hoarding behaviors, intellectual disability, and sleep abnormalities. PWS is genetic disorder mapping to imprinted 15q11.2-q13.3 locus, specifically at the paternally expressed *SNORD116* locus of small nucleolar RNAs and noncoding host gene transcripts. *SNORD116* is processed into several noncoding components and is hypothesized to orchestrate diurnal changes in metabolism through epigenetics, according to functional studies. Here, we review the current status of epigenetic mechanisms in PWS, with an emphasis on an emerging role for *SNORD116* in circadian and sleep phenotypes. We also summarize current ongoing therapeutic strategies, as well as potential implications for more common human metabolic and psychiatric disorders.

Keywords: epigenetic, imprinting, neurodevelopment, metabolic, circadian, diurnal, genetic, obesity

OPEN ACCESS

Edited by:
Mojgan Rastegar,
University of Manitoba, Canada

Reviewed by:
Merlin G. Butler,
University of Kansas Medical Center, United States
Christian P. Schaaf,
Heidelberg University, Germany
Marc Lalande,
University of Connecticut Health Center, United States

***Correspondence:**
Janine M. LaSalle
jmlasalle@ucdavis.edu

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *31 October 2020*
Accepted: *18 January 2021*
Published: *15 February 2021*

Citation:
Mendiola AJP and LaSalle JM (2021) Epigenetics in Prader-Willi Syndrome. Front. Genet. 12:624581. doi: 10.3389/fgene.2021.624581

INTRODUCTION

Clinical Features and Metabolic Phases of PWS

Prader-Willi Syndrome (PWS) is initially characterized by infantile hypotonia, failure to thrive due to poor suck, small hands and feet, and hypogonadism due to growth hormone deficiencies (Holm et al., 1993; Cassidy et al., 2012; Butler, 2020). During childhood, the development of extreme hyperphagia leads to obesity if not controlled is a major clinical feature of PWS. Other PWS features include obsessive-compulsive disorders, behavioral difficulties, intellectual disability, and sleep abnormalities.

PWS clinical characteristics are classically divided into two nutritional stages; however, it was recently identified that the stages are more complex and can be subdivided into five stages as described in **Table 1** (Miller et al., 2011; Butler et al., 2019b). The first stage (phase 0) occurs *in utero*, characterized by decreased movement in the womb and a low birth weight and size. Generally undiagnosed until birth, infants are assessed for PWS through a series of physical tests that determine the state of reflex and musculature (Holm et al., 1993; Miller et al., 2011; Cassidy et al., 2012). The next stage (phase 1a) of PWS is characterized by hypotonia, which leads to poor feeding and a resultant failure to thrive. Eventually, feeding normalizes entering phase 1b, but difficulty in feeding remains, and PWS infants often lag in

TABLE 1 | Clinical characteristics of nutritional phases.

Phase 0	Decreased fetal movement and growth restriction	*In utero*
Phase 1a	Infant becomes hypotonic and can develop failure to thrive	~0–9 months
Phase 1b	Infant begins to feed and grows steadily along a growth curve	~9–25 months
Phase 2a	Weight increase, without significant change in appetite or caloric intake	~2–4 years of age
Phase 2b	Continuous weight gain with increased food interest	~4–8 years of age
Phase 3	Development of hyperphagia, increased food seeking, and lack of satiety	~8 years of age
Phase 4	Loss of insatiable appetite and can feel full	Adulthood

Miller et al. (2011) and Butler et al. (2019b).

meeting standard developmental milestones. In the more severe cases of PWS, cranial and skeletal features are also apparent (Kindler et al., 2015). Although development is altered and delayed at infancy, patients feeding normalizes resulting in a steady increase in weight. However, stage 2 of nutritional development persists through early childhood, characterized by extreme fixation on food and development of hyperphagia (Holm et al., 1993; Cassidy and Driscoll, 2009; Miller et al., 2011). Stage 2 is divided into two phases in which phase 2a is an increase in weight that occurs without changes in appetite or feeding followed by phase 2b, characterized by fixation on food leading to phase 3, hyperphagia. In PWS, hyperphagia is developed at 2 years of age on average, and the severity of hyperphagia varies between children (Miller et al., 2011; Kim et al., 2012; Relkovic and Isles, 2013). Food intake and presence can be controlled by caretakers through proper rationing, reinforcement, and care which is most effective in the early PWS nutritional stages. However, hyperphagia continues to be a life-long struggle that is difficult to control with mitigation efforts. As PWS enters later stages of childhood and into adolescence, some patients enter the final stage (phase 4) and are able to feel full due to increased satiety and decreased behavioral difficulties related to food. It is unclear whether all PWS patients enter phase 4. Severity of clinical features is attributed to the size of deletions and may impact the recovery from hyperphagia (Kim et al., 2012).

Although abnormal sleep patterns are not featured in the nutritional PWS stages, disrupted REM sleep is a severe clinical feature in PWS. Patients with PWS exhibit a disrupted sleep pattern, which is similar to narcolepsy, including increased daytime sleepiness coupled to alterations to REM sleep at night. It is possible that the disrupted REM sleep is directly linked to the other clinical features in PWS. The importance of sleep is critical to the establishment of epigenetic patterns that solidify a diurnal pattern of feeding and metabolism. Once established, this diurnal rhythm is responsible for timing mechanisms regulating development from infancy through adulthood. Disruption of these rhythmic patterns may be causing the delay in development, resulting in the PWS clinical features including hyperphagia, inability to communicate, intellectual disabilities, behavioral difficulties, and obsessive-compulsive tendencies. Abnormal sleep patterns have been well-established in PWS, however, the molecular outcomes and downstream effects are not well understood. In this article, we will review what is known, delve into promising research findings, as well as discuss some therapeutic strategies for PWS that are either encouraging or controversial.

Molecular Genetics of PWS

PWS is both a genetic and epigenetic disorder, mapping the imprinted chromosomal domain of 15q11.2-13.3. Common to all cases of PWS is the absence of an expressed paternal copy of the *SNORD116* locus. Due to parental imprinting of the locus, outlined in more detail in the next section, loss of *SNORD116* can occur through deletion, uniparental disomy, or imprinting error. Most cases of PWS are caused by a large 6 Mb deletion of the entire 15q11.2-q13.3 locus (Holm et al., 1993; Cassidy et al., 2012). Two major large deletion classes include those with breakpoints at BP1 vs. BP2 combined with the downstream BP3 common deletion (Butler, 2020). However, microdeletions of the imprinting control region upstream of *SNRPN* (**Figure 1**) also result in loss of expression of *SNORD116* due to loss of the promoter. Rare microdeletions that only encompass *SNORD116*, but not *SNRPN* or *SNORD115*, have also been found in patients with PWS (Sahoo et al., 2008; de Smith et al., 2009; Duker et al., 2010). Approximately 60% of patients have paternal deletions, 36% are a result of maternal uniparental disomy, 4% are due to imprinting mutations that lead to a maternal imprinting status, and <1% are microdeletions of *SNORD116* (Butler et al., 2019a). What is common to all causes of PWS is the absence of *SNORD116* expression (Sahoo et al., 2008; Duker et al., 2010; Bieth et al., 2015; Rozhdestvensky et al., 2016).

While these findings establish that the lack of paternally expressed *SNORD116* is the likely predominant cause of PWS, there are a greater number of genes in the locus that may contribute to phenotypes of PWS. Both PWS large deletions include *MRKN3*, *MAGEL2*, *NDN*, *NPAP1*, *SNRPN*, *SNORD repeats*, *UBE3A*, *ATP10A*, *GABRB3*, *GABRA5*, *GABRG3*, *OCA2*, and *HERC2*. Additional genes between the proximal 15q11.2 breakpoints BP1 and BP2 include *TUBGCP5*, *CYFIP1*, *NIPA1*, and *NIPA2*. Genotype-phenotype investigations between the major molecular subtypes have been somewhat revealing at improving understanding of the genes involved in specific PWS phenotypes. In deletion compared to non-deletion etiologies of PWS, sleep abnormalities were more common (Torrado et al., 2007). Adaptive behavior scores were worse in PWS individuals with BP1-BP3 compared to BP2-BP3 or UPD and obsessive-compulsive behaviors more common in BP1-BP3 compared to UPD (Butler et al., 2004). In the Reiss Screen for maladaptive behaviors, deletion PWS patients showed higher self-injury and stealing scores compared to UPD (Hartley et al., 2005). Together, these studies indicate that gene expression patterns of one or more of these genes may contribute to variable phenotypes within PWS between the molecular subclasses. Below, the imprinted genes in the locus that have been implicated in PWS phenotypes will be discussed in more

detail, as well as the cluster of biallelically expressed $GABA_A$ receptor genes (*GABRB3*, *GABRA5*, and *GABRG3*), which are implicated in some of the neuropsychiatric phenotypes that are more severe in the deletion PWS molecular subclass.

SNORD116 is processed through a long noncoding transcript that initiates at the imprinting control region upstream of *SNRPN*, followed by two repeat clusters of small nucleolar RNAs (snoRNAs *SNORD116* and *SNORD115*) and terminating at the *UBE3A* antisense transcript (**Figure 1**; Sutcliffe et al., 1994; Buiting et al., 1995; Runte et al., 2001; Landers et al., 2004; Vitali et al., 2010; Chamberlain, 2013). In humans, *SNORD115*, but not *SNORD116* or *UBE3A-ATS*, is exclusively expressed in neurons, while *Snord116*, *Snord115*, and *Ube3a-ats* are all neuron-specific transcripts in mouse. *SNORD115* and *SNORD116* encompass clusters of repeated subunits of sequences encoding a C/D box snoRNAs embedded within intronic regions of the noncoding exons encoding the snoRNA host transcript *SNHG14* (Cavaillé et al., 2000; de los Santos et al., 2000; Bortolin-Cavaillé and Cavaillé, 2012; Stanurova et al., 2018). C/D box snoRNAs have known functions in regulating 2-O methylation rRNA modifications by recruiting ribonucleoprotein complexes including fibrillarin, which catalyzes methylation (Dupuis-Sandoval et al., 2015; Bratkovič et al., 2020).

SnoRNAs are processed from introns of the *SNORD116* and *SNORD115* within the *SNHG14* host gene subunits, called as *116HG* and *115HG* (**Figure 2**; Cavaillé et al., 2000; Leung et al., 2009; Vitali et al., 2010). Unlike other C/D box snoRNAs, *SNORD116* and *SNORD115* are classified as "orphan snoRNAs" because their targets and functions are unknown (Bratkovič et al., 2020). Previous studies have shown that *SNORD116* localizes in the nucleolus and may participate in splicing and RNA modifications (Bazeley et al., 2008; Leung et al., 2009). In contrast, *116HG* and *115HG* localize in the form of RNA "clouds" at the site of their own transcription in the nucleus (**Figure 2**), and dynamically regulate many additional genes across the genome (Powell et al., 2013; Coulson et al., 2018b). *SNORD115* is also shown to be involved in the alternative splicing specifically of the serotonin receptor *5-HT2C* mRNA (Bazeley et al., 2008; Raabe et al., 2019). Although, both loci are potentially implicated in PWS, microdeletion of only the *SNORD115* cluster does not lead to the PWS phenotype in humans (Runte et al., 2005). To date, the precise mechanisms of how *Snord116* functions are critical for neurodevelopment remain elusive, however, advancements in sequencing technology have provided new insights and will be covered in more detail in the section below. In addition to *SNORD116*, other genes in the 15q11.2-13.3 locus, including *NECDIN*, *MAGEL2*, and a cluster of GABA receptor genes are implicated in the phenotypes observed in most cases of PWS.

NECDIN (*NDN*) is an imprinted gene that is paternally expressed and encodes for the protein NECDIN, which belongs to the melanoma antigen-encoding gene (*MAGE*) family of proteins that are enriched in differentiated cells. *NDN* is one of several protein coding genes deleted from the large 6 Mb chromosomal deletion observed in PWS patients and is implicated in neuronal maturation (Ren et al., 2003). Other than its role in cellular differentiation and neuronal maturation, *NDN* is also involved in neurite and axonal growth, arborization, migration, and fasciculation, which are important for normal neurological signaling and development (MacDonald and Wevrick, 1997; Kuwajima et al., 2006; Davies et al., 2008; Miller et al., 2009; Bervini and Herzog, 2013). Mouse models of *Ndn* deficiency have been instrumental for studying abnormal brain development and cognitive impairments in PWS. However, studies using *Ndn* deficient mice did not exhibit any morphological differences in brain development, but led to respiratory failure causing apneas and irregular breathing patterns that are caused by increased activity in serotonin transporter (SERT/slc6a4; Matarazzo et al., 2017). Furthermore, *Ndn* knockout mice exhibit a higher pain threshold due to a decrease in nerve growth factor sensory neurons (Kuwako et al., 2005). Respiratory failure and higher pain thresholds are also observed in patients with PWS (Rittinger, 2001; Butler et al., 2002; Angulo et al., 2015). Specifically, irregularities in breathing may be a large proponent to sleep abnormalities in PWS.

MAGEL2 is another imprinted gene, that is, paternally expressed and encodes for the protein MAGEL2 that belongs to the MAGE family of proteins. Truncated *MAGEL2* mutations cause PWS-like phenotypes observed in patients (Schaaf et al., 2013; Fountain and Schaaf, 2016), but these

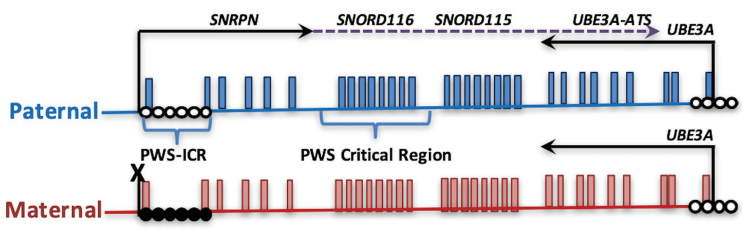

FIGURE 1 | Parental imprinting in the heart of the Prader-Willi syndrome (PWS) locus. The PWS on human chromosome 15q11.2-q13.3 is shown, depicting transcripts specifically expressed from the paternal (blue) or maternal (red) alleles. PWS patients with rare paternal microdeletions have defined the critical region over *SNORD116*. DNA methylation (closed circles) on the maternal allele of the PWS imprinting control region (PWS-ICR) silences the expression of *SNRPN* (solid arrow) and the long noncoding transcript expressed in neurons (dotted arrow) that encompasses repeated snoRNA clusters (including *SNORD116* and *SNORD115*) and the antisense transcript to *UBE3A* (*UBE3A-ATS*). *UBE3A* encodes an E3 ubiquitin ligase protein that regulates protein turnover of multiple cytoplasmic and nuclear factors. Since the paternal *UBE3A* allele is silenced by the expression of *UBE3A-ATS* in neurons, deletion or mutation of the maternal copy of *UBE3A* causes Angelman syndrome.

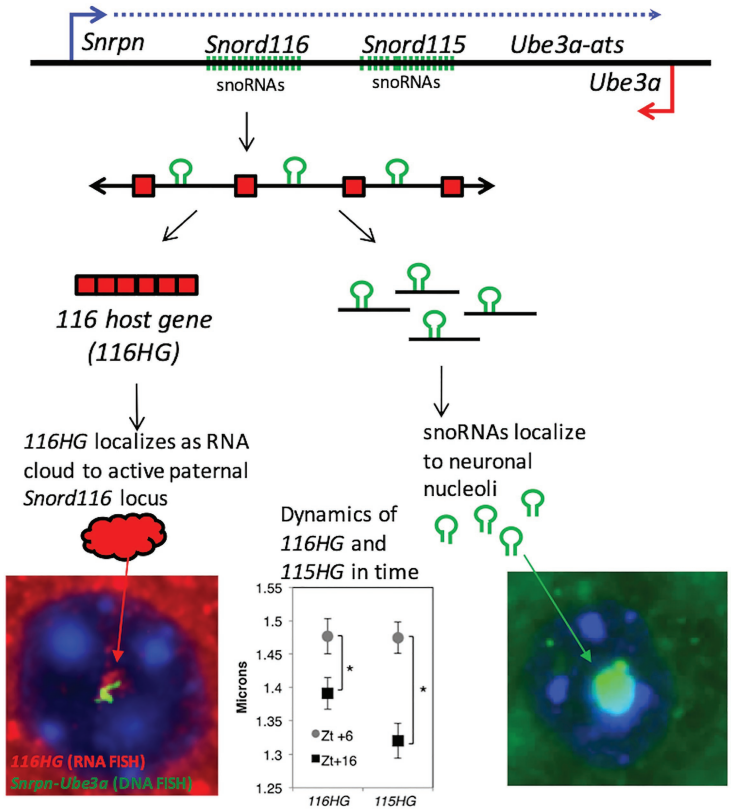

FIGURE 2 | PWS noncoding RNA summary. **(Top panel)** Individual components of the processed PWS snoRNA-lncRNA region between *Snrpn* and *Ube3a*. Within the *Snord116* and *Snord115* loci are repeated units of snoRNAs (green), lncRNA exons (red boxes), and introns with G-C skew. Processing results in spliced *116HG* and *115HG* lncRNAs that localize to their sites of transcription, the snoRNAs that localize to nucleoli, and R-loops that displace histones and promote locus chromatin decondensation. **(Bottom left panel)** Seen by RNA-FISH, *116HG* forms a large RNA cloud (red) localized to the decondensed paternal allele (green) in nuclei (blue), associated with 2,403 genes enriched for metabolic function. *116HG* and *115HG* RNA clouds are significantly larger at diurnal time ZT6 (sleep) than ZT16 (wake), corresponding to gene dysregulation in *Snord116+/–* specifically at ZT6. **(Bottom right panel)** Processed *Snord116* snoRNAs (green) localize to a single nucleolus in mature cortical neurons.

cases have been recently distinguished from PWS in a new classification of Schaaf-Yang syndrome (SYS). SYS shares phenotypic overlap with PWS, but also exhibit distinct behavioral and metabolic phenotypes including autism spectrum disorder (Fountain and Schaaf, 2016). In mouse embryogenesis, *Magel2* is highly expressed in non-neuronal (placenta, midgut turbucle, and midgut region) and neuronal tissue types (dorsal root ganglia and peripheral neurons surrounding limb and trunk muscles (Bervini and Herzog, 2013). In adult mouse brain, *Magel2* is highly enriched in hypothalamic regions and extends to the superchiasmic nucleus, specific regions that regulate feeding and circadian rhythms, respectively (Kozlov et al., 2007; Mercer et al., 2009). The prevalence of *MAGEL2* in the hypothalamus initially identified it as strong candidate for the hyperphagia phenotype of PWS. However, SYS patients and mouse models with *MAGEL2* mutations show a lower prevalence of overeating and obesity. Instead, it was determined that *MAGEL2* functions as a ubiquitin transporter that localizes in SCN neurons and acts as a direct regulator of circadian clock proteins through ubiquitination (Mercer et al., 2009; Tacer and Potts, 2017; Vanessa Carias et al., 2020).

The 15q11-q13 PWS region also contains a cluster of three genes encoding subunits of receptors for the neurotransmitter, $GABA_A$. GABA is the major inhibitory neurotransmitter in the postnatal brain, so loss of these GABA receptors in the large deletion cases of PWS is expected to be involved in some of the phenotypes of PWS. 15q11.2-13.3 genes *GABRB3*, *GABRA5*, and *GABRG3* encode for β3, α5, and γ3 subunits, respectively. $GABA_A$ receptors are assembled into hexameric protein complexes made up of combinations of a1–6, b1–3, g1–3, and other subunits, with α5 containing receptors making up ~5% of $GABA_A$ receptors in human brain (Mohamad and Has, 2019). Unlike the imprinted genes in the PWS locus, these 15q11.2-13.3 $GABA_A$ receptor genes are biallelically expressed in the brain. However, monoallelic expression and decreased protein expression of each $GABA_A$ receptor subunits have been observed in autism postmortem brain (Samaco et al., 2005; Hogart et al., 2007). Furthermore, both transcript and protein levels of GABRB3 were not correlated with copy

number in an analysis of PWS, AS, and 15q11.2-13.3 duplication syndrome postmortem brain (Scoles et al., 2011). A recent study on phenotypes and gene expression patterns in a *Gabrb3* deletion mouse model is also consistent with complex gene regulation, as neighboring *Oca2* expression was reduced and ocular hypopigmentation observed (Delahanty et al., 2016). Dysregulated gene expression of the 15q11.2-13.3 $GABA_A$ receptors is expected to have consequences for the balance of inhibitory and excitatory signals that regulate sleep, metabolism, and mood in PWS. Recently, it has been shown that levels of GABA metabolites vary between different molecular subclasses of PWS (Lucignani et al., 2004; Rice et al., 2016; Brancaccio et al., 2017). Since there are major targets for therapeutic intervention in multiple neurodevelopmental disorders, understanding their altered expression in PWS is expected to be important for the treatment of other neurodevelopmental disorders (Braat and Kooy, 2015).

EPIGENETIC MECHANISMS IN PWS

Epigenetic Regulation of the Imprinting Control Region in PWS

As mentioned in the previous section on molecular genetics, small deletions of the imprinting control region (PWS-ICR) are sufficient to cause PWS when inherited on the paternal allele. Interestingly, the ICR at 15q11.2-13.1 is actually bipartite, because maternal microdeletions of a region called as the AS-ICR are found in rare cases of Angelman syndrome (Buiting et al., 1995; Smith et al., 2011). Subsequent studies in a variety of mammals have demonstrated that the AS-ICR contains alternate 5' noncoding exon for *SNRPN* that are uniquely expressed in oocytes, but not sperm or other tissues (Smith et al., 2011; Lewis et al., 2015, 2019). It is the oocyte-specific transcription that leads to methylation and transcriptional silencing of the maternal allele specifically on the maternal but not the paternal allele of the PWS-ICR. A more recent study of individuals with AS imprinting mutations have identified a more common haplotype that deletes a binding site for the transcription factor SOX2 (Beygo et al., 2020). Together, these studies have demonstrated that this upstream region, defined as the AS-ICR, is critical for establishing silencing of the maternal allele of the imprinted genes within the PWS locus.

In addition to being characterized by allele-specific DNA methylation, several additional epigenetic marks are differential by parental origin at the PWS-ICR. Specifically, the histone H3 lysine 9 (H3K9) methyltransferase SETDB1 associates with the transcription factor ZNF274 bound to sites within the 5' cluster of SNORD116 repeats, resulting in the deposition of maternal-specific H3K9me3 marks (Cruvinel et al., 2014). Knockdown or inhibition of either SETDB1 or ZNF274 was sufficient to induce a low level of *SNORD116* transcript expression from the normally silent maternal allele (Cruvinel et al., 2014; Wu et al., 2019; Langouët et al., 2020). Together, these results suggest some promise for possible epigenetic therapies that will be discussed at the end of this review.

Epigenetics and Imprinting in PWS and Related Human Neurodevelopmental Disorders

In addition to PWS, loss of imprinting is involved in related neurodevelopmental disorders: Angelman (AS), 15q duplication (Dup15q), Kagami-Ogata (KOS14), and Temple (TS14) syndromes (Schanen, 2006; Kagami et al., 2015; Briggs et al., 2016). Unlike the default state of biallelic expression, imprinted genes are selectively silenced on either the maternal or paternal allele by epigenetic differences including DNA methylation and repressive chromatin modifications. Imprinted genes are clustered in discrete chromosomal loci and are regulated by a central imprinting control region (ICR), such as the PWS-ICR, in which methylation is diagnostic for AS, PWS, and Dup15q disorders (**Figure 1**). Some imprinted genes exhibit tissue-specific or developmental-specific imprinting patterns regulated by long noncoding RNAs. Furthermore, the largest conserved cluster of microRNA (miRNA) in the mammalian genome is found within the KOS14 imprinted locus and is responsible for regulating neuronal maturation and mTOR growth pathways (Winter, 2015). Experimental evidence is emerging for regulatory cross-talk between different imprinted gene loci (Stelzer et al., 2014; Jung and Nolta, 2016; Martinet et al., 2016; Vincent et al., 2016; Lopez et al., 2017), but this emerging "imprinted gene network" hypothesis (Fauque et al., 2010; Haga and Phinney, 2012; Monnier et al., 2013; Ribarska et al., 2014) has been understudied in the context of the developing nervous system.

RNA FISH has shown that *116HG* localizes in the nucleus, where it forms an RNA cloud that is absent in *Snord116* deletion brain. *116HG* was also found to colocalize with metabolic, circadian, and epigenetic gene loci including *Mtor*, *Clock*, *Cry1/2*, *Per1/2/3*, *Dnmt1/3b*, *Tet1/2/3*, *Mecp2*, and others at ZT6, the time point with the largest effect of *Snord116* deletion on transcription globally (Powell et al., 2013). *Snord116*'s involvement in transcriptional regulation, therefore, prompted an investigation of epigenetic differences that may explain the interaction of *Snord116* with diurnal light cycles. Whole genome bisulfite sequencing (WGBS) was performed on cortex samples from wild-type (WT) and PWS mice sacrificed every 3 h starting from Zt0–Zt16 and showed that *Snord116* is involved in regulating a dynamic rhythm of diurnal methylation (Coulson et al., 2018b). Rhythmically methylated CpG dinucleotides were identified (<1% of all CpGs) within enhancers and promoters of genes that were undergoing a pattern of reduced methylation during sleep (light hours) in wild-type mouse cortex, a pattern that was lost upon *Snord116* deletion. The differentially methylated regions mapped to genes involved in circadian rhythms, metabolism, and epigenetic regulation, similar to the prior genes identified associated with *116HG*. **Table 2** gives examples of specific genes in each of these categories that were identified by multiple unbiased genomic approaches in both studies. A large portion of genes identified are involved in adding, removing, and

TABLE 2 | Examples of *Snord116* associated and impacted genes and predicted functions.

Category	Function	Gene name	Gene binding to *116HG*[a]	*Snord116*-dependent transcriptional change[a]	*Snord116*-dependent DNA methylation change[b]	
					Mouse	Human
Epigenetic	Methyl binding protein critical to neurodevelopment	*Mecp2*	Yes	Increased at Zt6	No	No
	Binds DNA:RNA hybrids	*Setx*	No	Increased at Zt6	Yes	Yes
	DNA demethylases	*Tet1*	No	Increased at Zt6	Yes	Yes
		Tet2	No	Increased at Zt6	Yes	No
		Tet3	No	Increased at Zt6	Yes	Yes
	Histone deacetylases	*Hdac3*	No	Increased at Zt6	No	No
		Hdac4	No	Increased at Zt6	Yes	Yes
		Hdac5	No	Increased at Zt6	Yes	Yes
	DNA methyltransferases	*Dnmt1*	No	Increased at Zt6	No	Yes
		Dnmt3a	Yes	Increased at Zt6	No	Yes
Circadian	Establishes phases and periods	*Per2*	No	Increased at Zt6	No	Yes
		Per3	No	Increased at Zt6	No	No
		Arntl	Yes	Increased at Zt6	Yes	Yes
Metabolic	Kinase involved in regulating cellular energy homeostasis	*Mtor*	Yes	Increased at Zt6	No	Yes
Transcription	Transcriptional regulator of E-box motif containing genes	*Neurod1*	No	Increased at Zt6 & Zt16	Yes	No

[a]*Full gene lists are included in Powell et al. (2013).*
[b]*Full gene lists are included in Coulson et al. (2018b).*

recognizing DNA methylation while other genes are important transcriptional regulators for development. Further integration of promoter methylation and RNA-seq data revealed that genes being diurnally dysregulated were central to the body weight, behavior, and metabolic phenotypes of PWS (Coulson et al., 2018b).

The Coulson et al. study also demonstrated a molecular connection between the *116HG* and the KOS14 locus, building upon a prior study showing a connection between *IPW* (part of the *116HG* transcript) and *DLK1* regulation at the KOS14/TS14 locus in human neuronal culture (Stelzer et al., 2014). In this case, DNA FISH was used to examine chromosome decondensation, a measurement of neuronal activation of the paternal allele resulting from histone displacement, at both PWS and TS14 loci in adult mouse brain at six different diurnal time points. Interestingly, the TS14 locus only showed evidence of active chromatin decondensation in *Snord116* deletion mouse cortex. Furthermore, chromatin decondensation at the PWS locus did occur in *Snord116* deletion, but the timing was shifted from light to dark cycle, similar to the effects observed on DNA methylation. Together, these results suggest that the ancestrally older imprinted TS14/KOS14 locus may become more active as a compensatory mechanism to fill in for loss of *Snord116*, but this comes at a cost of proper timing of these epigenetic events.

Epigenetics and Imprinting of Mammalian Imprinted Loci and the Emerging Importance in Circadian Rhythmicity and Sleep

Daily and seasonal cycles of light, temperature, and feeding govern energy and activity of organisms from all branches of life. These environmental and metabolic inputs play an important role in the synchronization of the core circadian clock with the rhythmic patterns of many physiological and behavioral processes in peripheral tissues (Wright et al., 2013; Legates et al., 2014; Mukherji et al., 2015; Blasiak et al., 2017). The genetically encoded circadian cycle and the environmentally regulated diurnal/nocturnal cycle are integrated by a complex regulatory feedback network, which acts at the chromatin, transcriptional, and translational levels to coordinate biological and environmental rhythms (Koike et al., 2012; Papazyan et al., 2016; Takahashi, 2017). In mammals, the core circadian clock resides in the suprachiasmatic nucleus of the hypothalamus; however, almost half of all transcripts, both protein-coding and non-coding, exhibit diurnal rhythms in one or more peripheral tissues (Yan et al., 2008; Zhang R. et al., 2014). While most studies on circadian biology focus on the suprachiasmatic nucleus, investigations into diurnal rhythms of cerebral cortex are relevant to the cognitive deficits in PWS and to energy expenditure. For instance, circadian and metabolic genes showed light-cycle-specific dysregulation in the *Snord116del* mouse model, corresponding to cyclical dynamics of *Snord116* expression (Powell et al., 2013). Rhythmic epigenetic dynamics within the cerebral cortex are less well characterized; however, increasing evidence indicates a role for DNA methylation in these rhythms. Approximately 6% (25,476) of CpG sites assayed by 450k array are dynamically regulated throughout diurnal and seasonal cycles in human cortex (Lim et al., 2017). This epigenetic plasticity plays an important role in circadian entrainment and the resiliency of the circadian clock to changes in the diurnal environment (Stevenson and Prendergast, 2013; Azzi et al., 2014; Lim et al., 2014).

The 14q32.2 imprinted locus bears striking similarity to the PWS locus, as it encodes the only other repetitive cluster of snoRNAs in the mammalian genome (*SNORD113* and *SNORD114*), which are maternally expressed and exhibit allele-specific chromatin decondensation in neurons, similar to *SNORD116* and *SNORD115* (Cavaillé et al., 2002; Tierling et al., 2006; Leung et al., 2009). TS and KOS are reciprocally imprinted disorders, with TS caused by maternal uniparental disomy 14 [UPD(14)mat], and KOS caused by paternal uniparental disomy 14 [UPD(14)pat]. Loss of paternal gene expression at this locus in TS, results in aberrantly high expression of maternal non-coding RNAs, including *SNORD113* and *SNORD114*, whereas KOS results from the loss of maternally expressed, non-coding RNAs and the upregulation of paternally expressed *DLK1*. Interestingly, TS phenocopies PWS suggesting that these two imprinted loci may perform similar functions and share common pathways (Temple et al., 1991; Hosoki et al., 2009; Kagami et al., 2015). The loss of *Snord116* in PWS increases gene expression in the TS locus, indicating that the two loci may interact through a cross-regulatory network. In support of this hypothesis, *IPW* from the PWS locus has been shown to regulate the TS locus in an induced pluripotent stem cell line of PWS (Stelzer et al., 2014). Though both PWS and TS loci show circadian oscillations, the mechanism of this regulation and the impact of circadian rhythms on their cross-regulation suggests that a balance between the two loci is critical for sleep and metabolism (Labialle et al., 2008a,b; Powell et al., 2013).

Most imprinted loci, such as *IGF2*, *PEG1/MEST*, and *IGR2R*, are imprinted in marsupials as well as eutherian (placental) mammals (**Figure 3**). In contrast, *Snrpn* and *Ube3a* are not imprinted in marsupials and are on distinct chromosomes (Rapkins et al., 2006). Interestingly, the ancestral eutherian mammal tenrec (*Echinops telfairi*) lacks the *Snord116* and *Snord115* genes and *Snrpn* and *Ube3a* are on separate chromosomes (Rapkins et al., 2006; Yasui et al., 2011; Zhang Y. J. et al., 2014). Humans (and chimps) have 22 *SNORD116* and 44 *SNORD115* copies, while mouse has 27 detectable *Snord116* and 130 *Snord115* copies. Potentially relevant for the PWS phenotype, tenrecs have a unique metabolic and sleep structure among mammals adapted to long periods of reduced activity and body temperature called torpor (Lovegrove and Génin, 2008; Lovegrove et al., 2014a,b). Non-REM sleep and periods of torpor are thought to be ancestrally adaptive to conserve energy and escape predation. Eutherian mammals have distinct adaptations for daily sleep and activity patterns based on diet, body size, and brain size (Siegel, 2005; Gerhart-Hines and Lazar, 2015).

Unlike the PWS/AS locus, the chromosomal arrangement of the *SNORD113/SNORD114* cluster at the KOS/TS locus is similar in monotremes, marsupials, and placental mammals, and the miRNAs at this cluster are evolutionarily stable (Zhang Y. J. et al., 2014). Both imprinted snoRNA loci exhibit neuron-specific chromatin decondensation (Leung et al., 2009) and also show evidence for diurnally expressed transcripts, many of which are also dysregulated in *Snord116del* mice (Powell et al., 2013; Coulson et al., 2018b). Interestingly, circadian rhythmicity of the *Dlk1/Dio3* (Labialle et al., 2008a,b) and *Magel2* (Kozlov et al., 2007; Devos et al., 2011; Tennese and Wevrick, 2011) loci and cross-regulation between PWS/AS and *DLK1* loci have been described previously (Stelzer et al., 2014) but are poorly understood at a mechanistic level.

Despite its function being fully known, loss of *Snord116* in PWS mouse models has been demonstrated to dysregulate sleep, feeding, and temperature cycles (Lassi et al., 2016a,b). These studies have demonstrated the importance of hypothalamic *Snord116* expression on temporally regulated behavior. Interestingly, *Snord116* deficient mice exhibited disrupted feeding cues induced by erratic behavior due to

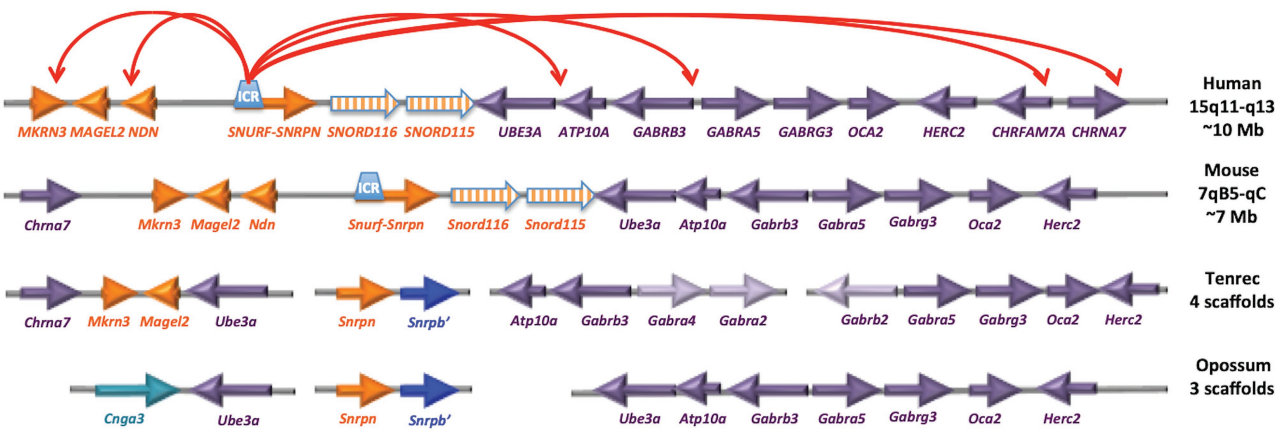

FIGURE 3 | The PWS/AS imprinted locus has emerged recently within placental mammals. The gene orientation and linear organization is shown for human and mouse, as well as the earliest placental mammal (tenrec) and marsupial (opossum). The red arrows on top represent results from neuronal 4C analysis of chromatin looping (Yasui et al., 2011). Interestingly, the tenrec arrangement of *Chrna7-Mkrn3-Magel2-Ube3a* (spanning ~500 kb) is similar to the human 4C long-range interactions spanning ~10 Mb, despite the lack of evidence for *Snrpn* or *Snord* clusters at the locus. Humans (and chimps) have 22 *SNORD116* and 44 *SNORD115* copies, mouse has 27 detectable *Snord116* and 130 *Snord115* copies.

increased activity prompted by foraging. Reminiscent of humans with PWS, *Snord116* deficient mice exhibited a strong fixation on food and high food intake irrespective of weight gain (Lassi et al., 2016a). Furthermore, *Snord116* deficient mice also exhibited a prolonged REM phase that was uncoupled with normal circadian patterning (Lassi et al., 2016b). Together, these studies have demonstrated the importance of *Snord116* on temporally regulated behaviors including sleep, feeding, foraging, and temperature regulation that are consistent with the recent evolutionary selection of the imprinted PWS locus in mammalian-specific diurnal cycles.

Multiple studies have also explored the role of *Snord116* in hypothalamic regulation of hormones linked to diurnal behaviors. Orexin neurons in the hypothalamus facilitate sleep-wake cycles by regulating hormones that promote wakefulness (noradrenaline, histamine, and acetylcholine) and rest [melanine-concentrating hormone (MCH); Pace et al., 2020]. Loss of orexin neurons are widely implicated in dysregulated sleep in patients with narcolepsy and has been observed in patients with PWS as well (Vgontzas et al., 1996; Chemelli et al., 1999; Mignot et al., 2002; Omokawa et al., 2016). However, it was not until recently that loss of *Snord116* was demonstrated to decrease orexin neuron levels in the lateral hypothalamus without altering levels of MCH and MCH neurons in mice (Pace et al., 2020). A decrease in orexin neurons may facilitate the prolonged REM sleep characteristic of PWS due to the imbalance in orexin/MCH ratio with a higher MCH concentration during wake cycles promoting more rest (Pace et al., 2020). This phenomenon is not unique to *Snord116* deletion mice, however, as this orexin/MCH imbalance was also observed in *Magel2* deficient models (Kozlov et al., 2007).

Patients with PWS are characterized as having reduced levels of growth hormone, but elevated levels of ghrelin (Tauber et al., 2019). Ghrelin is the endogenous ligand of growth hormone secretagogue receptor 1a. Ghrelin is peptide produced by the gut with a diversity of physiological effects, including appetite stimulation and lipid accumulation. Subsequent studies have demonstrated that it is actually the acylated form of ghrelin that is elevated in PWS children and young adults, while nonacylated ghrelin levels are indistinguishable from controls (Kuppens et al., 2015). However, while both growth hormone and ghrelin are known to have clear diurnal patterns of secretion, with nocturnal levels being higher than daytime levels in humans, there have been a surprising lack of investigation into the possibility growth hormone abnormalities in PWS may be due to altered diurnal rhythms (Kyung et al., 2004; Stawerska et al., 2020).

Despite orexin neurons being a critical cell type for *Snord116* regulation on hormonal regulation from the hypothalamus, loss of *Snord116* in other brain regions, such as cerebral cortex, also appear to contribute to the proper expression of core circadian clock regulators such as *Per* and *Bmal* genes (Powell et al., 2013; Coulson et al., 2018a). These findings reinforce the role of *Snord116* in establishing multiple aspects of circadian rhythms that are lost upon deletion. Studying *Snord116* and identifying its targets can contribute to the development of therapeutic interventions that target sleep and metabolism which are critical to development.

PWS Mouse Models for Preclinical Testing of Therapeutic Interventions

Mouse models of *Snord116* deficiency that recapitulate some features of PWS have been created as useful models for testing possible therapeutic interventions. Like in humans, *Snord116* is a maternally imprinted gene in mouse and localizes to a syntenic loci chromosome 7qC. The first generation of mouse models generated were designed with large deletions mimicking those observed in humans with PWS (Yang et al., 1998). These mouse models exhibited extreme hypotonia and failure to thrive, leading to death 1 week after birth. The high lethality rate was caused by the loss of protein coding genes *Snrpn* and *Ube3a-ats*, which are hypothesized to be important to alternative splicing (Tsai et al., 1999; Bressler et al., 2001; Bervini and Herzog, 2013). As in humans with PWS, these mouse models exhibited a dysregulation of major endocrine hormones including growth hormone, glucose, and insulin, which are necessary for cellular homeostasis and proliferation. Disruption of each hormone lead to metabolic dysregulation which results in extreme hypotonia that leads to the failure to thrive.

Today, the most commonly used PWS mouse models were originally generated by two separate labs using *cre*-mediated deletion of *Snord116* (Skryabin et al., 2007; Ding et al., 2008). These mouse models were designed by a targeted insertion of *loxP* cassettes flanking the *Snord116* (Ding et al., 2008) cluster or *Snord116* and *IPW* (Skryabin et al., 2007) through homologous recombination in embryonic stem (ES) cells derived from male blastocytes. The *2-loxP* ES cells were then injected into C57Bl/6J mice that gave birth to male mice with a *2-loxP* (+/−) genotype. These mice were mated with a transgenic strain expressing *Cre* recombinase under an ovary specific promoter producing 1-loxP mice with a *Snord116* (+/−) genotype (Ding et al., 2008). For ES cells targeted with *loxP* cassettes flanking *Snord116* and *IPW*, CRE recombinase were expressed then injected into blastocytes to produce *PWScre*(+/−) (Skryabin et al., 2007). These mouse models have a 150 kb deletion of the *Snord116* cluster or a deletion that encompasses *Snord116* and *IPW*. Like previous models, both mice develop hypotonia and failure to thrive with low to no post-natal lethality. Although these mouse models do not consistently exhibit the hyperphagia phenotype, they do exhibit a significant deficiency in cognition and energy expenditure (Powell et al., 2013; Adhikari et al., 2019) making these phenotypes useful in preclinical therapeutic strategies. Furthermore, development of *2-loxp*(−/+) and *PWScre*(+/−) mice enabled the generation of several new mouse models that are able to recapitulate the hyperphagia phenotype in adult mice through *Cre*-mediated and tamoxifen induced *Snord116* deletion in the hypothalamus (Qi et al., 2016; Purtell et al., 2017; Polex-Wolf et al., 2018) and identified the disrupted REM sleep phenotypes (Lassi et al., 2016b), respectively. Previous studies have shown that *Snord116* expression in the

hypothalamus is developmentally regulated and is enriched postnatally at weaning and early adulthood (Zhang et al., 2012), implicating its involvement in regulating metabolism and circadian rhythms.

Genetic Therapies

While most genetic diseases are amenable to genetic complementation and standard gene therapy design and delivery, there are unique challenges to gene therapy in PWS because of the epigenetic and molecular complexities of the *SNORD116* locus. In the original characterization of a *Snord116* deletion mouse model of PWS, it was mentioned that a transgene containing a single snoRNA from *Snord116* was insufficient to rescue the metabolic phenotypes (Ding et al., 2008). Since it remained possible that the limitations of using either a single copy and/or an already processed snoRNA were the reason for the lack of complementation, a new transgenic mouse was created and reported by our group using the *Snord*(+/−) model (Coulson et al., 2018a). This *Snord116* transgene contained the complete subunits of *116HG* exons, introns, and snoRNAs repeated in a total of 27 copies was expressed broadly at the transcript level in all tissues, but was only spliced and processed into snoRNAs in brain. The neuron-specific splicing was attributed to the splicing factor RBFOX3, which is also known as the neuron-specific marker NeuN. In wild-type neurons, the extra copies of *Snord116* contributed to the nucleolar accumulation of processed snoRNAs as well as the size of the *116HG* RNA cloud. However, in the *Snord116* deletion PWS model, the *Snord116* transgene did not become processed or localized to these locations, indicating that an active allele was needed for correct processing and localization. In addition, the body weight phenotype of the *Snord116* mice was similar to that of the *Snord116* deletion mouse, and there was no complementation of this phenotype in the cross.

In another study, a mouse model was generated with a *5'HPRT-LoxP-Neo*^{R} insertion upstream of the maternally imprinted *Snord116* using the *PWScre*(+/−) model (Rozhdestvensky et al., 2016). The cassette insertion did not affect the PWS imprinting center methylation status, but disrupted the imprinting effect enabling expression of *Snord116* from the maternal allele, a result that was not observed in WT and KO mice without the cassette. Like the Coulson et al. study in 2018, *Snord116* was expressed across all tissue types, but in this case, the body weight phenotype was rescued in KO mice with the cassette insertion. The differences in results may depend on the imprinting mechanism of the PWS region as well as the genomic location of *Snord116*. For instance, when the *Snord116* transgene is introduced outside of the imprinted region, as would be the case for most gene therapy strategies, the ability to complement the missing paternal allele is expected to be challenging. These results demonstrate the complexities of this locus and suggest that gene therapy for PWS using conventional complementation strategies will be problematic. Despite the issues, these results also highlight the importance of targeting imprinting regulation for therapeutic interventions.

Epigenetic Therapies

In contrast to gene therapy, epigenetic therapy for PWS has a stronger potential for clinical relevance, since PWS is an inherently epigenetic disorder. The general strategy for epigenetic strategies for PWS involves de-repressing the maternal silent PWS-ICR to activate SNRPN and *Snord116* transcription (Crunkhorn, 2017; Chung et al., 2020). Recent successes using high throughput screening of small molecule libraries identified several inhibitors of EHMT2/G9a, a histone 3 lysine 9 methyltransferase, that were capable of reactivating the expression of paternally expressed *SNRPN* and *SNORD116* from the maternal chromosome, both in cultured PWS cell lines and in a PWS mouse model (Kim et al., 2017, 2019). Similarly, inhibitor of SETDB1 using shRNA knockdown resulted in partial reactivation of *SNORD116* and *116HG* in PWS-derived iPSC cell lines and neurons (Cruvinel et al., 2014). The main differences in the epigenetic changes resulting between these two epigenetic therapies was that EHMT2/G9a did not alter DNA methylation at the PWS-ICR, while SETDB1 did not show a change in H3K9me3 at the PWS-ICR. Potentially more completely, the inactivation of ZNF274 using CRISPR/Cas9 in PWS-derived iPSC lines resulted in reactivation of both *SNRPN* and *SNORD116* as well as a reduction of H3K9me3 at the PWS-ICR (Langouët et al., 2020). Together, these studies suggest that combinations of targeted epigenetic strategies for unsilencing maternal *SNORD116* hold promise for future treatments of PWS.

POTENTIAL DEVELOPMENTS: RELEVANCE OF *SNORD116*-MEDIATED EPIGENETIC MECHANISMS TOWARD COMMON HUMAN DISEASES

While this review has focused on the relevance of epigenetic regulation of and by *SNORD116* and other genes within the locus to the pathogenesis of PWS, we expect that understanding the interactions between imprinted genes and metabolism at this locus will have relevance to other more common metabolic and neuropsychiatric human disorders. Because of the hypothalamic network alterations in PWS associated with satiety and food reward systems, this locus is considered to be a model for understanding food addictions as well as other addictive behaviors (Salles et al., 2020). The molecular mechanisms leading hyperphagia and overeating in PWS could be informative for understanding the intersections of epigenetics, diurnal rhythms, and metabolism in more common causes of overweight and obesity. Food addictions in PWS may be similar in mechanisms to those establishing other addictions. Interestingly, "morphine addiction" and "circadian entrainment" were among the gene pathway terms identified by the unbiased search for gene promoters that showed both rhythmic demethylation and increased expression during sleep in *Snord116* deletion mice (Coulson et al., 2018a), suggesting that further characterization of these pathways could be relevant to improved treatments for opioid use disorders. In addition, there are emerging links between circadian disruptions

and the exacerbation of psychiatric disorders such as bipolar disorder and depression. Chronotherapy involving sleep deprivation followed by the re-entrainment of diurnal cycles has shown effectiveness in treating these common mood disorders (Gottlieb et al., 2019; D'Agostino et al., 2020).

In conclusion, the PWS locus epigenetically regulated *SNORD116* transcripts that have evolved to become parentally imprinted within mammals, in turn serve to regulate a large number of additional genes through the genome that are related to circadian rhythms, metabolic and nutritional cycles, and brain functions. Future studies designed to better understand the genomic impacts of *SNORD116* regulation is expected to have far-reaching impacts beyond the scope of PWS.

AUTHOR CONTRIBUTIONS

Both authors contributed to the literature review, writing, and editing of the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We are grateful for the support of research on epigenetics in Prader-Willi syndrome from the NIH/NICHD (R01HD098038) and the Foundation for Prader-Willi Research. We would like to thank the home department of Medical Microbiology and Immunology for covering open access fees for this publication.

REFERENCES

Adhikari, A., Copping, N. A., Onaga, B., Pride, M. C., Coulson, R. L., Yang, M., et al. (2019). Cognitive deficits in the Snord116 deletion mouse model for Prader-Willi syndrome. *Neurobiol. Learn. Mem.* 165:106874. doi: 10.1016/j.nlm.2018.05.011

Angulo, M. A., Butler, M. G., and Cataletto, M. E. (2015). Prader-Willi syndrome: a review of clinical, genetic, and endocrine findings. *J. Endocrinol. Investig.* 38, 1249–1263. doi: 10.1007/s40618-015-0312-9

Azzi, A., Dallmann, R., Casserly, A., Rehrauer, H., Patrignani, A., Maier, B., et al. (2014). Circadian behavior is light-reprogrammed by plastic DNA methylation. *Nat. Neurosci.* 17, 377–382. doi: 10.1038/nn.3651

Bazeley, P. S., Shepelev, V., Talebizadeh, Z., Butler, M. G., Fedorova, L., Filatov, V., et al. (2008). SnoTARGET shows that human orphan SnoRNA targets locate close to alternative splice junctions. *Gene* 408, 172–179. doi: 10.1016/j.gene.2007.10.037

Bervini, S., and Herzog, H. (2013). Mouse models of Prader-Willi syndrome: a systematic review. *Front. Neuroendocrinol.* 34, 107–119. doi: 10.1016/j.yfrne.2013.01.002

Beygo, J., Grosser, C., Kaya, S., Mertel, C., Buiting, K., and Horsthemke, B. (2020). Common genetic variation in the Angelman syndrome imprinting centre affects the imprinting of chromosome 15. *Eur. J. Hum. Genet.* 28, 835–839. doi: 10.1038/s41431-020-0595-y

Bieth, E., Eddiry, S., Gaston, V., Lorenzini, F., Buffet, A., Auriol, F. C., et al. (2015). Highly restricted deletion of the SNORD116 region is implicated in Prader-Willi syndrome. *Eur. J. Hum. Genet.* 23, 252–255. doi: 10.1038/ejhg.2014.103

Blasiak, A., Gundlach, A. L., Hess, G., and Lewandowski, M. H. (2017). Interactions of circadian rhythmicity, stress and orexigenic neuropeptide systems: implications for food intake control. *Front. Neurosci.* 11:127. doi: 10.3389/fnins.2017.00127

Bortolin-Cavaillé, M. L., and Cavaillé, J. (2012). The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader-Willi locus generate canonical box C/D SnoRNAs. *Nucleic Acids Res.* 40, 6800–6807. doi: 10.1093/nar/gks321

Braat, S., and Kooy, R. F. (2015). The GABAA receptor as a therapeutic target for neurodevelopmental disorders. *Neuron* 86, 1119–1130. doi: 10.1016/j.neuron.2015.03.042

Brancaccio, M., Patton, A. P., Chesham, J. E., Maywood, E. S., and Hastings, M. H. (2017). Astrocytes control circadian timekeeping in the suprachiasmatic nucleus via glutamatergic signaling. *Neuron* 93, 1420–1435. doi: 10.1016/j.neuron.2017.02.030

Bratkovič, T., Božič, J., and Rogelj, B. (2020). Functional diversity of small nucleolar RNAs. *Nucleic Acids Res.* 48, 1627–1651. doi: 10.1093/nar/gkz1140

Bressler, J., Tsai, T. F., Wu, M. Y., Tsai, S. F., Ramirez, M. A., Armstrong, D., et al. (2001). The SNRPN promoter is not required for genomic imprinting of the Prader-Willi/Angelman domain in mice. *Nat. Genet.* 28, 232–240. doi: 10.1038/90067

Briggs, T. A., Lokulo-Sodipe, K., Chandler, K. E., Mackay, D. J. G., and Karen Temple, I. (2016). Temple syndrome as a result of isolated hypomethylation of the 14q32 imprinted DLK1/MEG3 region. *Am. J. Med. Genet. A* 170A, 170–175. doi: 10.1002/ajmg.a.37400

Buiting, K., Saitoh, S., Gross, S., Dittrich, B., Schwartz, S., Nicholls, R. D., et al. (1995). Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. *Nat. Genet.* 9, 395–400. doi: 10.1038/ng0495-395

Butler, M. G. (2020). Imprinting disorders in humans: a review. *Curr. Opin. Pediatr.* 32, 719–729. doi: 10.1097/MOP.0000000000000965

Butler, M. G., Bittel, D. C., Kibiryeva, N., Talebizadeh, Z., and Thompson, T. (2004). Behavioral differences among subjects with Prader-Willi syndrome and type I or type II deletion and maternal disomy. *Pediatrics* 113, 565–573. doi: 10.1542/peds.113.3.565

Butler, M. G., Hartin, S. N., Hossain, W. A., Manzardo, A. M., Kimonis, V., Dykens, E., et al. (2019a). Molecular genetic classification in Prader-Willi syndrome: a multisite cohort study. *J. Med. Genet.* 56, 149–153. doi: 10.1136/jmedgenet-2018-105301

Butler, M. G., Miller, J. L., and Forster, J. L. (2019b). Prader-Willi syndrome—clinical genetics, diagnosis and treatment approaches: an update. *Curr. Pediatr. Rev.* 15, 207–244. doi: 10.2174/1573396315666190716120925

Butler, J. V., Whittington, J. E., Holland, A. J., Boer, H., Clarke, D., and Webb, T. (2002). Prevalence of, and risk factors for, physical ill-health in people with Prader-Willi syndrome: a population-based study. *Dev. Med. Child Neurol.* 44, 248–255. doi: 10.1017/S001216220100202X

Cassidy, S. B., and Driscoll, D. J. (2009). Prader-Willi syndrome. *Eur. J. Hum. Genet.* 17, 3–13. doi: 10.1038/ejhg.2008.165

Cassidy, S. B., Schwartz, S., Miller, J. L., and Driscoll, D. J. (2012). Prader-Willi syndrome. *Genet. Med.* 14, 10–26. doi: 10.1038/gim.0b013e31822bead0

Cavaillé, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C. I., Horsthemke, B., et al. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14311–14316. doi: 10.1073/pnas.250426397

Cavaillé, J., Seitz, H., Paulsen, M., Ferguson-Smith, A. C., and Bachellerie, J. P. (2002). Identification of tandemly-repeated C/D SnoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum. Mol. Genet.* 11, 1527–1538. doi: 10.1093/hmg/11.13.1527

Chamberlain, S. J. (2013). RNAs of the human chromosome 15q11-Q13 imprinted region. *Wiley Interdiscip. Rev. RNA* 4, 155–166. doi: 10.1002/wrna.1150

Chemelli, R. M., Willie, J. T., Sinton, C. M., Elmquist, J. K., Scammell, T., Lee, C., et al. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 98, 437–451. doi: 10.1016/S0092-8674(00)81973-X

Chung, M. S., Langouët, M., Chamberlain, S. J., and Carmichael, G. G. (2020). Prader-Willi syndrome: reflections on seminal studies and future therapies. *Open Biol.* 10:200195. doi: 10.1098/rsob.200195

Coulson, R. L., Powell, W. T., Yasui, D. H., Dileep, G., Resnick, J., and LaSalle, J. M. (2018a). Prader-Willi locus Snord116 RNA processing requires an active endogenous allele and neuron-specific splicing by Rbfox3/NeuN. *Hum. Mol. Genet.* 27, 4051–4060. doi: 10.1093/hmg/ddy296

Coulson, R. L., Yasui, D. H., Dunaway, K. W., Laufer, B. I., Ciernia, A. V., Zhu, Y., et al. (2018b). Snord116-dependent diurnal rhythm of DNA

methylation in mouse cortex. *Nat. Commun.* 9:1616. doi: 10.1038/s41467-018-03676-0

Crunkhorn, S. (2017). Steps towards epigenetic therapy for PWS. *Nat. Rev. Drug Discov.* 16:85. doi: 10.1038/nrd.2017.3

Cruvinel, E., Budinetz, T., Germain, N., Chamberlain, S., Lalande, M., and Martins-Taylor, K. (2014). Reactivation of maternal SNORD116 cluster via SETDB1 knockdown in Prader-Willi syndrome IPSCs. *Hum. Mol. Genet.* 23, 4674–4685. doi: 10.1093/hmg/ddu187

D'Agostino, A., Ferrara, P., Terzoni, S., Ostinelli, E. G., Carrara, C., Prunas, C., et al. (2020). Efficacy of triple chronotherapy in unipolar and bipolar depression: a systematic review of the available evidence. *J. Affect. Disord.* 276, 297–304. doi: 10.1016/j.jad.2020.07.026

Davies, W., Lynn, P. M. Y., Relkovic, D., and Wilkinson, L. S. (2008). Imprinted genes and neuroendocrine function. *Front. Neuroendocrinol.* 29, 413–427. doi: 10.1016/j.yfrne.2007.12.001

de los Santos, T., Schweizer, J., Rees, C. A., and Francke, U. (2000). Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader-Willi deletion region, which is highly expressed in brain. *Am. J. Hum. Genet.* 67, 1067–1082. doi: 10.1086/303106

de Smith, A. J., Purmann, C., Walters, R. G., Ellis, R. J., Holder, S. E., Van Haelst, M. M., et al. (2009). A deletion of the HBII-85 class of small nucleolar RNAs (SnoRNAs) is associated with hyperphagia, obesity and hypogonadism. *Hum. Mol. Genet.* 18, 3257–3265. doi: 10.1093/hmg/ddp263

Delahanty, R. J., Zhang, Y., Bichell, T. J., Shen, W., Verdier, K., Macdonald, R. L., et al. (2016). Beyond epilepsy and autism: disruption of GABRB3 causes ocular hypopigmentation. *Cell Rep.* 17, 3115–3124. doi: 10.1016/j.celrep.2016.11.067

Devos, J., Weselake, S. V., and Wevrick, R. (2011). Magel2, a Prader-Willi syndrome candidate gene, modulates the activities of circadian rhythm proteins in cultured cells. *J. Circadian Rhythms* 9:12. doi: 10.1186/1740-3391-9-12

Ding, F., Li, H. H., Zhang, S., Solomon, N. M., Camper, S. A., Cohen, P., et al. (2008). SnoRNA Snord116 (Pwcr1/MBll-85) deletion causes growth deficiency and hyperphagia in mice. *PLoS One* 3:1709. doi: 10.1371/journal.pone.0001709

Duker, A. L., Ballif, B. C., Bawle, E. V., Person, R. E., Mahadevan, S., Alliman, S., et al. (2010). Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box SnoRNA cluster in Prader-Willi syndrome. *Eur. J. Hum. Genet.* 18, 1196–1201. doi: 10.1038/ejhg.2010.102

Dupuis-Sandoval, F., Poirier, M., and Scott, M. S. (2015). The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip. Rev. RNA* 6, 381–397. doi: 10.1002/wrna.1284

Fauque, P., Ripoche, M. A., Tost, J., Journot, L., Gabory, A., Busato, F., et al. (2010). Modulation of imprinted gene network in placenta results in normal development of in vitro manipulated mouse embryos. *Hum. Mol. Genet.* 19, 1779–1790. doi: 10.1093/hmg/ddq059

Fountain, M., and Schaaf, C. (2016). Prader-Willi syndrome and Schaaf-Yang syndrome: neurodevelopmental diseases intersecting at the MAGEL2 gene. *Diseases* 4:2. doi: 10.3390/diseases4010002

Gerhart-Hines, Z., and Lazar, M. A. (2015). Circadian metabolism in the light of evolution. *Endocr. Rev.* 36, 289–304. doi: 10.1210/er.2015-1007

Gottlieb, J. F., Benedetti, F., Geoffroy, P. A., Henriksen, T. E. G., Lam, R. W., Murray, G., et al. (2019). The chronotherapeutic treatment of bipolar disorders: a systematic review and practice recommendations from the ISBD task force on chronotherapy and chronobiology. *Bipolar Disord.* 8, 741–773. doi: 10.1111/bdi.12847

Haga, C. L., and Phinney, D. G. (2012). MicroRNAs in the imprinted DLK1-DIO3 region repress the epithelial-to-mesenchymal transition by targeting the TWIST1 protein signaling network. *J. Biol. Chem.* 287, 42695–42707. doi: 10.1074/jbc.M112.387761

Hartley, S. L., MacLean, W. E., Butler, M. G., Zarcone, J., and Thompson, T. (2005). Maladaptive behaviors and risk factors among the genetic subtypes of Prader-Willi syndrome. *Am. J. Med. Genet.* 136 A, 140–145. doi: 10.1002/ajmg.a.30771

Hogart, A., Nagarajan, R. P., Patzel, K. A., Yasui, D. H., and LaSalle, J. M. (2007). 15q11-13 GABAA receptor genes are normally biallelically expressed in brain yet are subject to epigenetic dysregulation in autism-spectrum disorders. *Hum. Mol. Genet.* 16, 691–703. doi: 10.1093/hmg/ddm014

Holm, V. A., Cassidy, S. B., Butler, M. G., Hanchett, J. M., Greenswag, L. R., Whitman, B. Y., et al. (1993). Prader-Willi syndrome: consensus diagnostic criteria. *Pediatrics* 91, 398–402.

Hosoki, K., Kagami, M., Tanaka, T., Kubota, M., Kurosawa, K., Kato, M., et al. (2009). Maternal uniparental disomy 14 syndrome demonstrates Prader-Willi syndrome-like phenotype. *J. Pediatr.* 155, 900–903. doi: 10.1016/j.jpeds.2009.06.045

Jung, Y., and Nolta, J. A. (2016). BMI1 regulation of self-renewal and multipotency in human mesenchymal stem cells. *Curr. Stem Cell Res. Ther.* 11, 131–140. doi: 10.2174/1574888X1102160107171432

Kagami, M., Kurosawa, K., Miyazaki, O., Ishino, F., Matsuoka, K., and Ogata, T. (2015). Comprehensive clinical studies in 34 patients with molecularly defined UPD(14)pat and related conditions (Kagami-Ogata syndrome). *Eur. J. Hum. Genet.* 23, 1488–1498. doi: 10.1038/ejhg.2015.13

Kim, Y., Lee, H. M., Xiong, Y., Sciaky, N., Hulbert, S. W., Cao, X., et al. (2017). Targeting the histone methyltransferase G9a activates imprinted genes and improves survival of a mouse model of Prader-Willi syndrome. *Nat. Med.* 23, 213–222. doi: 10.1038/nm.4257

Kim, S. J., Miller, J. L., Kuipers, P. J., German, J. R., Beaudet, A. L., Sahoo, T., et al. (2012). Unique and atypical deletions in Prader-Willi syndrome reveal distinct phenotypes. *Eur. J. Hum. Genet.* 20, 283–290. doi: 10.1038/ejhg.2011.187

Kim, Y., Wang, S. E., and Jiang, Y. -H. (2019). Epigenetic therapy of Prader-Willi syndrome. *Transl. Res.* 205, 105–118. doi: 10.1016/j.trsl.2019.02.012

Kindler, J. M., Lewis, R. D., and Hamrick, M. W. (2015). Skeletal muscle and pediatric bone development. *Curr. Opin. Endocrinol. Diabetes Obes.* 22, 467–474. doi: 10.1097/MED.0000000000000201

Koike, N., Yoo, S. H., Huang, H. C., Kumar, V., Lee, C., Kim, T. K., et al. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* 338, 349–354. doi: 10.1126/science.1226339

Kozlov, S. V., Bogenpohl, J. W., Howell, M. P., Wevrick, R., Panda, S., Hogenesch, J. B., et al. (2007). The imprinted gene Magel2 regulates normal circadian output. *Nat. Genet.* 39, 1266–1272. doi: 10.1038/ng2114

Kuppens, R. J., Diène, G., Bakker, N. E., Molinas, C., Faye, S., Nicolino, M., et al. (2015). Elevated ratio of acylated to unacylated ghrelin in children and young adults with Prader-Willi syndrome. *Endocrine* 50, 633–642. doi: 10.1007/s12020-015-0614-x

Kuwajima, T., Nishimura, I., and Yoshikawa, K. (2006). Necdin promotes GABAergic neuron differentiation in cooperation with Dlx homeodomain proteins. *J. Neurosci.* 26, 5383–5392. doi: 10.1523/JNEUROSCI.1262-06.2006

Kuwako, K. I., Hosokawa, A., Nishimura, I., Uetsuki, T., Yamada, M., Nada, S., et al. (2005). Disruption of the paternal necdin gene diminishes TrkA signaling for sensory neuron survival. *J. Neurosci.* 25, 7090–7099. doi: 10.1523/JNEUROSCI.2083-05.2005

Kyung, H. P., Jin, D. K., Sang, Y. S., Ji, E. L., Si, H. K., Seng, M. S., et al. (2004). Correlation between fasting plasma ghrelin levels and age, body mass index (BMI), BMI percentiles, and 24-hour plasma ghrelin profiles in Prader-Willi syndrome. *J. Clin. Endocrinol. Metab.* 89, 3885–3889. doi: 10.1210/jc.2003-032137

Labialle, S., Croteau, S., Bélanger, V., McMurray, E. N., Ruan, X., Moussette, S., et al. (2008a). Novel imprinted transcripts from the Dlk1-Gtl2 intergenic region, mico1 and mico1os, show circadian oscillations. *Epigenetics* 3, 322–329. doi: 10.4161/epi.3.6.7109

Labialle, S., Yang, L., Ruan, X., Villemain, A., Schmidt, J. V., Hernandez, A., et al. (2008b). Coordinated diurnal regulation of genes from the Dlk1-Dio3 imprinted domain: implications for regulation of clusters of non-paralogous genes. *Hum. Mol. Genet.* 17, 15–26. doi: 10.1093/hmg/ddm281

Landers, M., Bancescu, D. L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., et al. (2004). Regulation of the large (∼1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res.* 32, 3480–3492. doi: 10.1093/nar/gkh670

Langouët, M., Gorka, D., Orniacki, C., Dupont-Thibert, C. M., Chung, M. S., Glatt-Deeley, H. R., et al. (2020). Specific ZNF274 binding interference at SNORD116 activates the maternal transcripts in Prader-Willi syndrome neurons. *Hum. Mol. Genet.* 29, 3285–3295. doi: 10.1093/hmg/ddaa210

Lassi, G., Maggi, S., Balzani, E., Cosentini, I., Garcia-Garcia, C., and Tucci, V. (2016a). Working-for-food behaviors: a preclinical study in Prader-Willi mutant mice. *Genetics* 204, 1129–1138. doi: 10.1534/genetics.116.192286

Lassi, G., Priano, L., Maggi, S., Garcia-Garcia, C., Balzani, E., El-Assawy, N., et al. (2016b). Deletion of the Snord116/SNORD116 alters sleep in mice and patients with Prader-Willi syndrome. *Sleep* 39, 637–644. doi: 10.5665/sleep.5542

Legates, T. A., Fernandez, D. C., and Hattar, S. (2014). Light as a central modulator of circadian rhythms, sleep and affect. *Nat. Rev. Neurosci.* 15, 443–454. doi: 10.1038/nrn3743

Leung, K. N., Vallero, R. O., Dubose, A. J., Resnick, J. L., and Lasalle, J. M. (2009). Imprinting regulates mammalian SnoRNA-encoding chromatin decondensation and neuronal nucleolar size. *Hum. Mol. Genet.* 18, 4227–4238. doi: 10.1093/hmg/ddp373

Lewis, M. W., Brant, J. O., Kramer, J. M., Moss, J. I., Yang, T. P., Hansen, P. J., et al. (2015). Angelman syndrome imprinting center encodes a transcriptional promoter. *Proc. Natl. Acad. Sci. U. S. A.* 112, 6871–6875. doi: 10.1073/pnas.1411261111

Lewis, M. W., Vargas-Franco, D., Morse, D. A., and Resnick, J. L. (2019). A mouse model of Angelman syndrome imprinting defects. *Hum. Mol. Genet.* 28, 220–229. doi: 10.1093/hmg/ddy345

Lim, A. S. P., Klein, H. U., Yu, L., Chibnik, L. B., Ali, S., Xu, J., et al. (2017). Diurnal and seasonal molecular rhythms in human neocortex and their relation to Alzheimer's disease. *Nat. Commun.* 8:14931. doi: 10.1038/ncomms14931

Lim, A. S. P., Srivastava, G. P., Yu, L., Chibnik, L. B., Xu, J., Buchman, A. S., et al. (2014). 24-hour rhythms of DNA methylation and their relation with rhythms of RNA expression in the human dorsolateral prefrontal cortex. *PLoS Genet.* 10:e1004792. doi: 10.1371/journal.pgen.1004792

Lopez, S. J., Dunaway, K., Saharul Islam, M., Mordaunt, C., Ciernia, A. V., Meguro-Horike, M., et al. (2017). UBE3A-mediated regulation of imprinted genes and epigenome-wide marks in human neurons. *Epigenetics* 12, 982–990. doi: 10.1080/15592294.2017.1376151

Lovegrove, B. G., Canale, C., Levesque, D., Fluch, G., Řeháková-Petrů, M., and Ruf, T. (2014a). Are tropical small mammals physiologically vulnerable to arrhenius effects and climate change? *Physiol. Biochem. Zool.* 87, 30–45. doi: 10.1086/673313

Lovegrove, B. G., and Génin, F. (2008). Torpor and hibernation in a basal placental mammal, the Lesser Hedgehog Tenrec *Echinops telfairi*. *J. Comp. Physiol. B, Biochem. Syst. Environ. Physiol.* 178, 691–698. doi: 10.1007/s00360-008-0257-9

Lovegrove, B. G., Lobban, K. D., and Levesque, D. L. (2014b). Mammal survival at the cretaceous–palaeogene boundary: metabolic homeostasis in prolonged tropical hibernation in tenrecs. *Proc. R. Soc. B Biol. Sci.* 281:20141304. doi: 10.1098/rspb.2014.1304

Lucignani, G., Panzacchi, A., Bosio, L., Moresco, R. M., Ravasi, L., Coppa, I., et al. (2004). GABAA receptor abnormalities in Prader-Willi syndrome assessed with positron emission tomography and [11C] flumazenil. *NeuroImage* 22, 22–28. doi: 10.1016/j.neuroimage.2003.10.050

MacDonald, H. R., and Wevrick, R. (1997). The necdin gene is deleted in Prader-Willi syndrome and is imprinted in human and mouse. *Hum. Mol. Genet.* 6, 1873–1878. doi: 10.1093/hmg/6.11.1873

Martinet, C., Monnier, P., Louault, Y., Benard, M., Gabory, A., and Dandolo, L. (2016). H19 controls reactivation of the imprinted gene network during muscle regeneration. *Development* 143, 962–971. doi: 10.1242/dev.131771

Matarazzo, V., Caccialupi, L., Schaller, F., Shvarev, Y., Kourdougli, N., Bertoni, A., et al. (2017). Necdin shapes serotonergic development and sert activity modulating breathing in a mouse model for Prader-Willi syndrome. *elife* 6:e32640. doi: 10.7554/eLife.32640

Mercer, R. E., Kwolek, E. M., Bischof, J. M., Van Eede, M., Henkelman, R. M., and Wevrick, R. (2009). Regionally reduced brain volume, altered serotonin neurochemistry, and abnormal behavior in mice null for the circadian rhythm output gene Magel2. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 150B, 1085–1099. doi: 10.1002/ajmg.b.30934

Mignot, E., Lammers, G. J., Ripley, B., Okun, M., Nevsimalova, S., Overeem, S., et al. (2002). The role of cerebrospinal fluid hypocretin measurement in the diagnosis of narcolepsy and other hypersomnias. *Arch. Neurol.* 59, 1553–1562. doi: 10.1001/archneur.59.10.1553

Miller, J. L., Lynn, C. H., Driscoll, D. J. C., Goldstone, A. P., Gold, J. A., Kimonis, V., et al. (2011). Nutritional phases in Prader-Willi syndrome. *Am. J. Med. Genet. A* 155A, 1040–1049. doi: 10.1002/ajmg.a.33951

Miller, N. L. G., Wevrick, R., and Mellon, P. L. (2009). Necdin, a Prader-Willi syndrome candidate gene, regulates gonadotropin-releasing hormone neurons during development. *Hum. Mol. Genet.* 18, 248–260. doi: 10.1093/hmg/ddn344

Mohamad, F. H., and Has, A. T. C. (2019). The α5-containing GABA A receptors—a brief summary. *J. Mol. Neurosci.* 67, 343–351. doi: 10.1007/s12031-018-1246-4

Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., and Dandolo, L. (2013). H19 LncRNA controls gene expression of the imprinted gene network by recruiting MBD1. *Proc. Natl. Acad. Sci. U. S. A.* 110, 20693–20698. doi: 10.1073/pnas.1310201110

Mukherji, A., Kobiita, A., Damara, M., Misra, N., Meziane, H., Champy, M. F., et al. (2015). Shifting eating to the circadian rest phase misaligns the peripheral clocks with the master SCN clock and leads to a metabolic syndrome. *Proc. Natl. Acad. Sci. U. S. A.* 112, E6691–E6698. doi: 10.1073/pnas.1519807112

Omokawa, M., Ayabe, T., Nagai, T., Imanishi, A., Omokawa, A., Nishino, S., et al. (2016). Decline of CSF orexin (hypocretin) levels in Prader-Willi syndrome. *Am. J. Med. Genet. A* 170, 1181–1186. doi: 10.1002/ajmg.a.37542

Pace, M., Falappa, M., Freschi, A., Balzani, E., Berteotti, C., Martire, V. L., et al. (2020). Loss of Snord116 impacts lateral hypothalamus, sleep, and food-related behaviors. *JCI Insight* 5:e137495. doi: 10.1172/jci.insight.137495

Papazyan, R., Zhang, Y., and Lazar, M. A. (2016). Genetic and epigenomic mechanisms of mammalian circadian transcription. *Nat. Struct. Mol. Biol.* 23, 1045–1052. doi: 10.1038/nsmb.3324

Polex-Wolf, J., Lam, B. Y. H., Larder, R., Tadross, J., Rimmington, D., Bosch, F., et al. (2018). Hypothalamic loss of Snord116 recapitulates the hyperphagia of Prader-Willi syndrome. *J. Clin. Investig.* 128, 960–969. doi: 10.1172/JCI97007

Powell, W. T., Coulson, R. L., Crary, F. K., Wong, S. S., Ach, R. A., Peter, T., et al. (2013). A Prader-Willi locus LncRNA cloud modulates diurnal genes and energy expenditure. *Hum. Mol. Genet.* 22, 4318–4328. doi: 10.1093/hmg/ddt281

Purtell, L., Qi, Y., Campbell, L., Sainsbury, A., and Herzog, H. (2017). Adult-onset deletion of the Prader-Willi syndrome susceptibility gene Snord116 in mice results in reduced feeding and increased fat mass. *Transl. Pediatr.* 6, 88–97. doi: 10.21037/tp.2017.03.06

Qi, Y., Purtell, L., Fu, M., Lee, N. J., Aepler, J., Zhang, L., et al. (2016). Snord116 is critical in the regulation of food intake and body weight. *Sci. Rep.* 6:18614. doi: 10.1038/srep18614

Raabe, C. A., Voss, R., Kummerfeld, D. M., Brosius, J., Galiveti, C. R., Wolters, A., et al. (2019). Ectopic expression of Snord115 in choroid plexus interferes with editing but not splicing of 5-Ht2c receptor pre-MRNA in mice. *Sci. Rep.* 9, 1–9. doi: 10.1038/s41598-019-39940-6

Rapkins, R. W., Hore, T., Smithwick, M., Ager, E., Pask, A. J., Renfree, M. B., et al. (2006). Recent assembly of an imprinted domain from non-imprinted components. *PLoS Genet.* 2:e182. doi: 10.1371/journal.pgen.0020182

Relkovic, D., and Isles, A. R. (2013). Behavioural and cognitive profiles of mouse models for Prader-Willi syndrome. *Brain Res. Bull.* 92, 41–48. doi: 10.1016/j.brainresbull.2011.09.009

Ren, J., Lee, S., Pagliardini, S., Gérard, M., Stewart, C. L., Greer, J. J., et al. (2003). Absence of Ndn, encoding the Prader-Willi syndrome-deleted gene necdin, results in congenital deficiency of central respiratory drive in neonatal mice. *J. Neurosci.* 23, 1569–1573. doi: 10.1523/JNEUROSCI.23-05-01569.2003

Ribarska, T., Goering, W., Droop, J., Bastian, K. M., Ingenwerth, M., and Schulz, W. A. (2014). Deregulation of an imprinted gene network in prostate cancer. *Epigenetics* 9, 704–717. doi: 10.4161/epi.28006

Rice, L. J., Lagopoulos, J., Brammer, M., and Einfeld, S. L. (2016). Reduced gamma-aminobutyric acid is associated with emotional and behavioral problems in Prader-Willi syndrome. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 171, 1041–1048. doi: 10.1002/ajmg.b.32472

Rittinger, O. (2001). Clinical features and genetic analysis of Prader-Willi syndrome. *Klinische Padiatrie* 4, 387–392. doi: 10.1055/s-2001-15857.

Rozhdestvensky, T. S., Robeck, T., Galiveti, C. R., Raabe, C. A., Seeger, B., Wolters, A., et al. (2016). Maternal transcription of non-protein coding RNAs from the PWS-critical region rescues growth retardation in mice. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep20398

Runte, M., Hüttenhofer, A., Groß, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.* 10, 2687–2700. doi: 10.1093/hmg/10.23.2687

Runte, M., Varon, R., Horn, D., Horsthemke, B., and Buiting, K. (2005). Exclusion of the C/D box SnoRNA gene cluster HBII-52 from a major role in Prader-Willi syndrome. *Hum. Genet.* 116, 228–230. doi: 10.1007/s00439-004-1219-2

Sahoo, T., del Gaudio, D., German, J. R., Shinawi, M., Peters, S. U., Person, R. E., et al. 2008. "Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster." *Nat. Genet.* 40, 719–721. doi: 10.1038/ng.158

Salles, J., Lacassagne, E., Eddiry, S., Franchitto, N., Salles, J. -P., and Tauber, M. (2020). What can we learn from PWS and SNORD116 genes about the pathophysiology of addictive disorders? *Mol. Psychiatry* 26, 51–59. doi: 10.1038/s41380-020-00917-x

Samaco, R. C., Hogart, A., and LaSalle, J. M. (2005). Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. *Hum. Mol. Genet.* 14, 483–492. doi: 10.1093/hmg/ddi045

Schaaf, C. P., Gonzalez-Garay, M. L., Xia, F., Potocki, L., Gripp, K. W., Zhang, B., et al. (2013). Truncating mutations of MAGEL2 cause Prader-Willi phenotypes and autism. *Nat. Genet.* 45, 1405–1408. doi: 10.1038/ng.2776

Schanen, N. C. (2006). Epigenetics of autism spectrum disorders. *Hum. Mol. Genet.* 15(Suppl. 2), R138–R150. doi: 10.1093/hmg/ddl213

Scoles, H. A., Urraca, N., Chadwick, S. W., Reiter, L. T., and Lasalle, J. M. (2011). Increased copy number for methylated maternal 15q duplications leads to changes in gene and protein expression in human cortical samples. *Mol. Autism.* 2:19. doi: 10.1186/2040-2392-2-19

Siegel, J. M. (2005). Clues to the functions of mammalian sleep. *Nature* 437, 1264–1271. doi: 10.1038/nature04285

Skryabin, B. V., Gubar, L. V., Seeger, B., Pfeiffer, J., Handel, S., Robeck, T., et al. (2007). Deletion of the MBII-85 SnoRNA gene cluster in mice results in postnatal growth retardation." Edited by A. C. Ferguson-Smith. *PLoS Genet.* 3:e235. doi: 10.1371/journal.pgen.0030235

Smith, E. Y., Futtner, C. R., Chamberlain, S. J., Johnstone, K. A., and Resnick, J. L. (2011). Transcription is required to establish maternal imprinting at the Prader-Willi syndrome and Angelman syndrome locus. *PLoS Genet.* 7:e1002422. doi: 10.1371/journal.pgen.1002422

Stanurova, J., Neureiter, A., Hiber, M., de Oliveira Kessler, H., Stolp, K., Goetzke, R., et al. (2018). Corrigendum: angelman syndrome-derived neurons display late onset of paternal UBE3A silencing. *Sci. Rep.* 8:46952. doi: 10.1038/srep46952

Stawerska, R., Kolasa-Kicińska, M., Łupińska, A., Hilczer, M., and Lewiński, A. (2020). Comparison of nocturnal and morning ghrelin concentration in children with growth hormone deficiency and with idiopathic short stature. *Chronobiol. Int.* 37, 1629–1635. doi: 10.1080/07420528.2020.1797765

Stelzer, Y., Sagi, I., Yanuka, O., Eiges, R., and Benvenisty, N. (2014). The noncoding RNA IPW regulates the imprinted DLK1-DIO3 locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nat. Genet.* 46, 551–557. doi: 10.1038/ng.2968

Stevenson, T. J., and Prendergast, B. J. (2013). Reversible DNA methylation regulates seasonal photoperiodic time measurement. *Proc. Natl. Acad. Sci. U. S. A.* 110, 16651–16656. doi: 10.1073/pnas.1310643110

Sutcliffe, J. S., Nakao, M., Christian, S., Örstavik, K. H., Tommerup, N., Ledbetter, D. H., et al. (1994). Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nat. Genet.* 8, 52–58. doi: 10.1038/ng0994-52

Tacer, K. F., and Potts, P. R. (2017). Cellular and disease functions of the Prader-Willi syndrome gene Magel2. *Biochem. J.* 474, 2177–2190. doi: 10.1042/BCJ20160616

Takahashi, J. S. (2017). Transcriptional architecture of the mammalian circadian clock. *Nat. Rev. Genet.* 18, 164–179. doi: 10.1038/nrg.2016.150

Tauber, M., Coupaye, M., Diene, G., Molinas, C., Valette, M., and Beauloye, V. (2019). Prader-Willi syndrome: a model for understanding the ghrelin system. *J. Neuroendocrinol.* 31:e12728. doi: 10.1111/jne.12728

Temple, I. K., Cockwell, A., Hassold, T., Pettay, D., and Jacobs, P. (1991). Maternal uniparental disomy for chromosome 14. *J. Med. Genet.* 8, 131–138. doi: 10.1136/jmg.28.8.511

Tennese, A. A., and Wevrick, R. (2011). Impaired hypothalamic regulation of endocrine function and delayed counterregulatory response to hypoglycemia in Magel2-null mice. *Endocrinology* 152, 967–978. doi: 10.1210/en.2010-0709

Tierling, S., Dalbert, S., Schoppenhorst, S., Tsai, C. E., Oliger, S., Ferguson-Smith, A. C., et al. (2006). High-resolution map and imprinting analysis of the gtl2-dnchc1 domain on mouse chromosome 12. *Genomics* 87, 225–235. doi: 10.1016/j.ygeno.2005.09.018

Torrado, M., Araoz, V., Baialardo, E., Abraldes, K., Mazza, C., Krochik, G., et al. (2007). Clinical-etiologic correlation in children with Prader-Willi syndrome (PWS): an interdisciplinary study. *Am. J. Med. Genet. A* 143, 460–468. doi: 10.1002/ajmg.a.31520

Tsai, T. F., Jiang, Y. H., Bressler, J., Armstrong, D., and Beaudet, A. L. (1999). Paternal deletion from Snrpn to Ube3a in the mouse causes hypotonia, growth retardation and partial lethality and provides evidence for a gene contributing to Prader-Willi syndrome. *Hum. Mol. Genet.* 8, 1357–1364. doi: 10.1093/hmg/8.8.1357

Vanessa Carias, K., Zoeteman, M., Seewald, A., Sanderson, M. R., Bischof, J. M., and Wevrick, R. (2020). A MAGEL2-deubiquitinase complex modulates the ubiquitination of circadian rhythm protein CRY1. *PLoS One* 15:e0230874. doi: 10.1371/journal.pone.0230874

Vgontzas, A. N., Bixler, E. O., Kales, A., Centurione, A., Rogan, P. K., Mascari, M., et al. (1996). Daytime Sleepines and rem abrormalities in Prader-Willi syndrome: evidence of generalized hypoarousal. *Int. J. Neurosci.* 87, 127–139. doi: 10.3109/00207459609070832

Vincent, R. N., Gooding, L. D., Louie, K., Wong, E. C., and Ma, S. (2016). Altered DNA methylation and expression of PLAGL1 in cord blood from assisted reproductive technology pregnancies compared with natural conceptions. *Fertil. Steril.* 106, 739.e3–748.e3. doi: 10.1016/j.fertnstert.2016.04.036

Vitali, P., Royo, H., Marty, V., Bortolin-Cavaillé, M. L., and Cavaillé, J. (2010). Long nuclear-retained non-coding RNAs and allele-specific higher-order chromatin organization at imprinted SnoRNA gene arrays. *J. Cell Sci.* 123, 70–83. doi: 10.1242/jcs.054957

Winter, J. (2015). Micrornas of the Mir379–410 cluster: new players in embryonic neurogenesis and regulators of neuronal function. *Neurogenesis* 2:e1004970. doi: 10.1080/23262133.2015.1004970

Wright, K. P., McHill, A. W., Birks, B. R., Griffin, B. R., Rusterholz, T., and Chinoy, E. D. (2013). Entrainment of the human circadian clock to the natural light-dark cycle. *Curr. Biol.* 23, 1554–1558. doi: 10.1016/j.cub.2013.06.039

Wu, H., Ng, C., Villegas, V., Chamberlain, S., Cacace, A., and Wallace, O. (2019). Small molecule inhibitors of G9a reactivate the maternal PWS genes in Prader-Willi-syndrome patient derived neural stem cells and differentiated neurons. BioRxiv [preprint server]. doi: 10.1101/640938

Yan, J., Wang, H., Liu, Y., and Shao, C. (2008). Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput. Biol.* 4:e1000193. doi: 10.1371/journal.pcbi.1000193

Yang, T., Adamson, T. E., Resnick, J. L., Leff, S., Wevrick, R., Francke, U., et al. (1998). A mouse model for Prader-Willi syndrome imprinting-centre mutations. *Nat. Genet.* 19, 25–31. doi: 10.1038/ng0598-25

Yasui, D. H., Scoles, H. A., Horike, S. -I., Meguro-Horike, M., Dunaway, K. W., Schroeder, D. I., et al. (2011). 15q11.2-13.3 chromatin analysis reveals epigenetic regulation of CHRNA7 with deficiencies in rett and autism brain. *Hum. Mol. Genet.* 20, 4311–4323. doi: 10.1093/hmg/ddr357

Zhang, Q., Bouma, G. J., McClellan, K., and Tobet, S. (2012). Hypothalamic expression of SnoRNA Snord116 is consistent with a link to the hyperphagia and obesity symptoms of Prader-Willi syndrome. *Int. J. Dev. Neurosci.* 30, 479–478. doi: 10.1016/j.ijdevneu.2012.05.005

Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E., and Hogenesch, J. B. (2014). A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl. Acad. Sci. U. S. A.* 111, 16219–16224. doi: 10.1073/pnas.1408886111

Zhang, Y. J., Yang, J. H., Shi, Q. S., Zheng, L. L., Liu, J., Zhou, H., et al. (2014). Rapid birth-and-death evolution of imprinted SnoRNAs in the Prader-Willi syndrome locus: implications for neural development in euarchontoglires. *PLoS One* 9:e100329. doi: 10.1371/journal.pone.0100329

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 21 April 2021
doi: 10.3389/fcell.2021.612518

A Chemo-Genomic Approach Identifies Diverse Epigenetic Therapeutic Vulnerabilities in MYCN-Amplified Neuroblastoma

Aleksandar Krstic[1], Anja Konietzny[1,2], Melinda Halasz[1], Peter Cain[3,4], Udo Oppermann[3,4], Walter Kolch[1,5] and David J. Duffy[1,6,7]*

[1] Systems Biology Ireland and Precision Oncology Ireland, School of Medicine, University College Dublin, Dublin, Ireland, [2] Centre for Molecular Neurobiology Hamburg (ZMNH), Emmy-Noether Group "Neuronal Protein Transport", University Medical Centre Hamburg-Eppendorf (UKE), Hamburg, Germany, [3] Botnar Research Centre, NIHR Oxford Biomedical Research Unit, Institute of Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom, [4] Centre for Medicines Discovery, University of Oxford, Oxford, United Kingdom, [5] Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin, Ireland, [6] The Whitney Laboratory for Marine Bioscience and Sea Turtle Hospital, University of Florida, St. Augustine, FL, United States, [7] Department of Biology, University of Florida, Gainesville, FL, United States

Although a rare disease, neuroblastoma accounts for the highest proportion of childhood cancer deaths. There is a lack of recurrent somatic mutations in neuroblastoma embryonal tumours, suggesting a possible role for epigenetic alterations in driving this cancer. While an increasing number of reports suggest an association of MYCN with epigenetic machinery, the mechanisms of these interactions are poorly understood in the neuroblastoma setting. Utilising chemo-genomic approaches we revealed global MYCN-epigenetic interactions and identified numerous epigenetic proteins as MYCN targets. The epigenetic regulators HDAC2, CBX8 and CBP (CREBBP) were all MYCN target genes and also putative MYCN interactors. MYCN-related epigenetic genes included SMARCs, HDACs, SMYDs, BRDs and CREBBP. Expression levels of the majority of MYCN-related epigenetic genes showed predictive ability for neuroblastoma patient outcome. Furthermore, a compound library screen targeting epigenetic proteins revealed broad susceptibility of neuroblastoma cells to all classes of epigenetic regulators, belonging to families of bromodomains, HDACs, HATs, histone methyltransferases, DNA methyltransferases and lysin demethylases. Ninety-six percent of the compounds reduced MYCN-amplified neuroblastoma cell viability. We show that the C646 (CBP-bromodomain targeting compound) exhibits switch-like temporal and dose response behaviour and is effective at reducing neuroblastoma viability. Responsiveness correlates with MYCN expression, with MYCN-amplified cells being more susceptible to C646 treatment. Thus, exploiting the broad vulnerability of neuroblastoma cells to epigenetic targeting compounds represents an exciting strategy in neuroblastoma treatment, particularly for high-risk MYCN-amplified tumours.

Keywords: cAMP-response-element-binding [CREB] protein (known as CBP or CREBBP), E1A Binding Protein P300 (known as EP300 or p300), SMARCA, epigenetic regulation, cancer, precision medicine, zebrafish

OPEN ACCESS

Edited by:
Mojgan Rastegar, University of Manitoba, Canada

Reviewed by:
Matthew Wong, Children's Cancer Institute Australia, Australia
Trond Flægstad, UiT The Arctic University of Norway, Norway

****Correspondence:***
Aleksandar Krstic
aleksandar.krstic@ucd.ie

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 September 2020*
Accepted: *25 March 2021*
Published: *21 April 2021*

Citation:
Krstic A, Konietzny A, Halasz M, Cain P, Oppermann U, Kolch W and Duffy DJ (2021) A Chemo-Genomic Approach Identifies Diverse Epigenetic Therapeutic Vulnerabilities in MYCN-Amplified Neuroblastoma. Front. Cell Dev. Biol. 9:612518. doi: 10.3389/fcell.2021.612518

Abbreviations: ChIP-seq, Chromatin immunoprecipitation sequencing; CPMkb, read counts per million adjusted by gene length in kilobases; DE, differentially expressed; hpf, hours post fertilisation; IPA, Ingenuity pathway analysis; ITR, inferred transcriptional regulator; MNA, MYCN amplified; NB, neuroblastoma; TSS, transcription start site.

INTRODUCTION

Epigenetics is defined as non-DNA encoded heritable modifications, which result in altered gene expression levels (Jin et al., 2011). Increasingly, epigenetics modifications have functional roles in human cancers (Sharma et al., 2010; Wainwright and Scaffidi, 2017). Indeed it has been postulated that neuroblastoma tumours, which lack recurrent somatic mutations (Pugh et al., 2013), may be driven by aberrant epigenetic signalling (Domingo-Fernandez et al., 2013; Veschi et al., 2017; Durinck and Speleman, 2018; Ram Kumar and Schor, 2018; Upton et al., 2020). Unlike more traditional adult tumours, which are largely driven by genomic mutations (Khan and Helman, 2016), epigenetically driven childhood cancers have proven resistant to classical therapeutic target identification approaches, directed against mutations in oncogenes (Johnsen et al., 2018). Furthermore, existing therapies for neuroblastoma have severe and sometimes long-term side effects that include an increased risk of second malignancies (Applebaum et al., 2017).

Neuroblastoma begins *in utero* and the disease is predominantly diagnosed in the first year of life. Despite being rare, it accounts for 8–10% of all diagnosed childhood cancers (Stack et al., 2007; Dreidax et al., 2014; Henrich et al., 2016). However, due to its aggressiveness, neuroblastoma is responsible for 14% of all childhood cancer deaths (Stack et al., 2007). Neuroblastoma is divided into risk groups based on criteria which include: the age of the patient at diagnosis, International Neuroblastoma Risk Group (INRG) tumour stage and MYCN copy number status (Cohn et al., 2009; Monclair et al., 2009). Patients with low-risk (stage 1, 2 and 4s) and intermediate (stage 3) neuroblastoma have event free survival (EFS) of up to 90%, contrary to high-risk patients with EFS of less than 50% (Smith and Foster, 2018; Meany, 2019). Amplification of the MYCN oncogene, which occurs in approximately 20% of cases (Huang and Weiss, 2013), is one of the clearest markers for identifying high-risk neuroblastoma patients, regardless of disease stage. MYCN amplification results in increased cellular proliferation and growth, decreased apoptosis, poor differentiation, and increased vascularity of the tumours (Gustafson and Weiss, 2010). Additionally, is also associated with advanced stage disease, an overall poor prognosis, and therapy resistance (Shimada et al., 2001), with MYCN-amplified tumours being resistant to current therapeutic approaches. Multiple studies have shown that MYCN exerts its functions through interactions with the epigenetic machinery (Domingo-Fernandez et al., 2013; He et al., 2013; Puissant et al., 2013; Carter et al., 2015; Duffy et al., 2015; Yang et al., 2015; Duffy et al., 2016; Henrich et al., 2016; Duffy et al., 2017; Felgenhauer et al., 2018). Therefore, deeper understanding of MYCN dependent epigenetic vulnerabilities provides a novel route for targeted therapies in neuroblastoma.

We investigated the MYCN-related epigenetic signalling network, and the potential of epigenetic lead compounds as therapeutic agents for the treatment of high-risk MYCN amplified neuroblastoma. MYCN is a transcription factor that binds to the promoters of genes critically involved in neuroblastoma oncogenesis (Duffy et al., 2015). We performed an unbiased genome-wide MYCN target gene screen using chromatin immunoprecipitation sequencing (ChIP-seq) to identify MYCN-epigenetic cross-talk, and combined it with a phenotypic screen of small molecule epigenetic targeting compounds. These approaches converged on a number of promising hit compounds which were further characterised.

RESULTS

ChIP-seq Identifies Epigenetic Regulators as MYCN Targets

To identify MYCN's epigenetic-related genomic targets, we mined our MYCN ChIP-seq datasets (Duffy et al., 2015) for known epigenetic genes (ArrayExpress, www.ebi.ac.uk/arrayexpress, accession number E-MTAB-4100). These datasets comprise MYCN ChIP-seq data from the patient matched MYCN amplified cell lines KCN (from a primary tumour at diagnosis) and KCNR (from a secondary tumour after relapse), and of a time-course of MYCN overexpression in the MYCN inducible cell line SY5Y-MYCN at the following time-points: Un-induced (0h), 24h and 48h. The MYCN overexpression achieved in this system is similar to overexpression caused by MYCN gene amplification. To identify the epigenetic-related genes bound by MYCN, we intersected the MYCN ChIP-seq targets with the 167 gene dbEM, Database of Epigenetic Modifiers (Singh Nanda et al., 2016)[1]. Forty-two of the dbEM genes (25%) were identified in our MYCN ChIP-seq datasets (**Figure 1A** and **Table 1**), confirming that MYCN protein targets a relatively high proportion of epigenetic-related genes.

To identify epigenetic proteins which formed protein-protein interactions with MYCN we also mined our MYCN interactome data in SY5Y-MYCN cells (Duffy et al., 2015; Duffy et al., 2017). Thirty-four epigenetic proteins physically, directly or as a part of a complex, bound to MYCN protein (**Supplementary Table 1**), including three epigenetic regulators whose coding genes were genomic targets of MYCN (ChIP-seq); CBP (official gene symbol CREBBP), HDAC2 and CBX8. Expression of these genes was assessed for survival correlations in the SEQC dataset (Zhang et al., 2015) of 498 neuroblastoma patients, using the R2: Genomics Analysis and Visualization Platform[2]. All six MYCN target genes (ChIP-seq) which were common to at least two of the three lists (**Supplementary Table 1**), had prognostic value for neuroblastoma patient outcome, when patients were segregated according to expression of these genes (**Figure 1B**).

Analysis of the Other Known Transcriptional Regulators of MYCN's Genomic Targets Reveals Epigenetic Regulation

To further address the significance of epigenetic mechanisms and their potential interplay with MYCN regulatory networks, we next examined MYCN's genomic targets (ChIP-seq) for known

[1]http://crdd.osdd.net/raghava/dbem
[2]http://r2.amc.nl

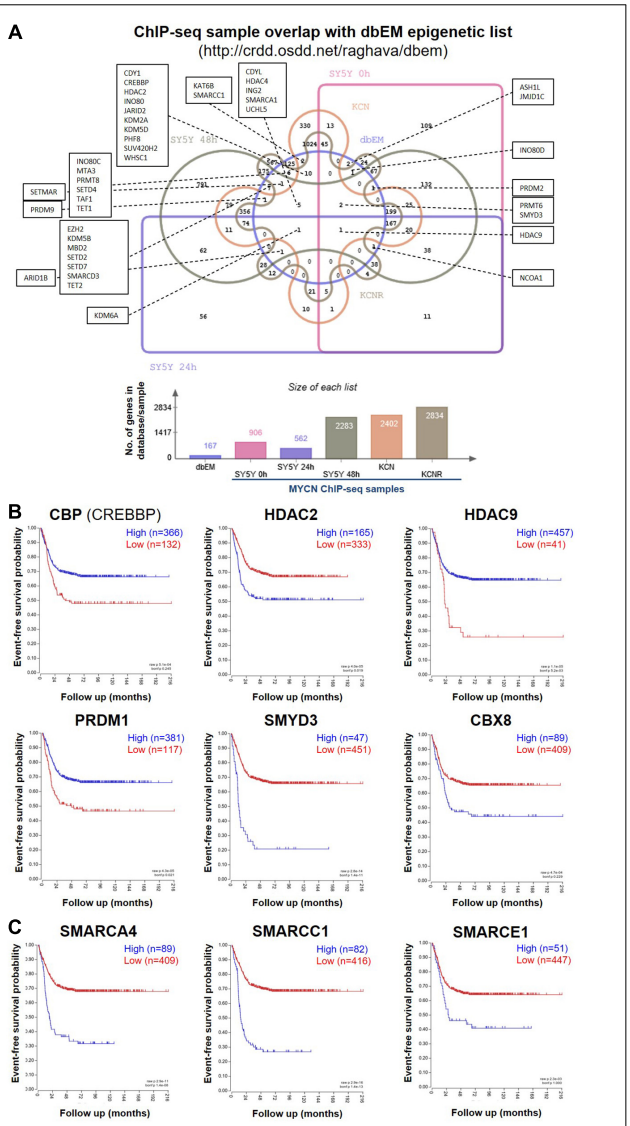

FIGURE 1 | MYCN ChIP-seq epigenetic hits and their survival curves. **(A)** Venn diagram showing the overlap of genes called as genomic MYCN targets (ChIP-seq) in SY5Y-MYCN (un-induced Dox-, 24 h Dox+ and 48 h Dox+), KCN and KCNR cell lines and known epigenetic genes (Singh Nanda et al., 2016) (dbEM, Database of Epigenetic Modifiers, http://crdd.osdd.net/raghava/dbem). The Venn diagram was generated using jvenn (Bardou et al., 2014) (http://jvenn.toulouse.inra.fr/app/index.html). **(B)** Kaplan-Meier survival curves showing the predictive strength of the expression levels of the CBP, HDAC2, HDAC9, PRDM1, SMYD3, and CBX8 mRNAs in neuroblastoma tumours on patient outcome. **(C)** Kaplan-Meier survival curves showing the predictive strength of the expression levels of SMARC genes in neuroblastoma tumours on patient outcome. The SMARCs shown were predicted to be upstream regulators of MYCN's genomic targets (ChIP-seq). SMARCC1 was also a direct MYCN target (ChIP-seq). All survival curves were generated using the SEQC (Zhang et al., 2015) 498 neuroblastoma tumour dataset in the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl).

epigenetic co-regulators, using Ingenuity pathway analysis (IPA). IPA has a manually curated database of transcriptional regulators (Krämer et al., 2014), and certain epigenetic proteins

TABLE 1 | Comparison of ChIP-seq MYCN target gene dataset with a curated epigenetic regulators lists.

MYCN ChIP-seq hits overlapped with the dbEM database of epigenetic modifiers				
CREBBP (CBP)	HDAC9	SUV420H2	JMJD1C	EZH2
HDAC2	HDAC4	SETD2	KDM5B	MBD2
WHSC1	PRMT8	SETD7	CDYL	PRDM2
SMARCA1	SETD4	SETMAR	INO80C	PRDM9
UCHL5	TAF1	MTA3	NCOA1	JARID2
SMARCD3	TET1	ING2	KDM2A	ARID1B
SMARCC1	CDY1	INO80	KDM5D	
ASH1L	PHF8	PRMT6	KDM6A	
TET2	KAT6B	SMYD3	INO80D	
42 / 167 genes				

*Forty-two epigenetic-related genes which were MYCN genomic targets. Gene list obtained by overlapping the 167 gene dbEM, Database of Epigenetic Modifiers (http://crdd.osdd.net/raghava/dbem) with our MYCN ChIP-seq hits (**Figure 1A**; Duffy et al., 2015, 2016; Singh Nanda et al., 2016).*

were consistently recognised as inferred upstream regulators of MYCN's target genes in all ChIP-seq datasets (**Table 2**). The common inferred transcriptional regulators (ITRs) were enriched for key regulators of chromatin structure (**Table 2**), including HDAC and SMARC genes. It has previously been shown that expression levels of 11 HDAC members in primary neuroblastomas are correlated with NB prognosis and stage (Oehme et al., 2009; **Figure 1B**) and that HDAC2 functionally interacts with MYCN (Lodrini et al., 2013; Duffy et al., 2016). The SMARC gene family regulates transcription by altering the chromatin structure around certain genes. This family of proteins still remains poorly understood, although it has been shown that mutations that inactivate their subunits are found in nearly 20% of human cancers (Hohmann and Vakoc, 2014). All six of the SMARC genes, identified as being ITRs of MYCN's genomic targets (ChIP-seq), were able to segregate neuroblastoma patients by outcome, according to the level of SMARC gene expression in tumours (**Figure 1C** and **Supplementary Figure 1A**). For four of these SMARC genes (SMARCA4, SMARCC1, SMARCE1 and SMARCB1) high expression was associated with poor outcome.

Screening an Epigenetic Library in MYCN Amplified Cells Reveals a Broad Vulnerability to Epigenetic Compounds

Having identified novel MYCN-epigenetic links and confirmed that MYCN can interact with epigenetic genes, we next sought to identify small molecules directed against epigenetic regulators with therapeutic potential for treatment of MYCN amplified neuroblastoma. To achieve this, in a non-biased manner, we employed a curated compound library generated by the Structural Genomics Consortium (SGC)[3], which is comprised of 45 epigenetic targeting compounds (**Supplementary Table 2**). The compound library was screened on the IMR32 cell line, and resulting changes to cell phenotype and viability were assessed. IMR32 cells harbour a high level of MYCN amplification with

[3]http://www.thesgc.org/

TABLE 2 | Inferred Transcriptional Regulators (ITRs) for MYCN ChIP-seq samples belong to HDAC and SMARC families of epigenetic regulators.

Inferred Transcriptional Regulator (ITR)	Dataset	*p*-value of overlap
HDAC family	KCNR	2.04E–07
	24h SY5Y-MYCN	5.63E–06
	48h SY5Y-MYCN	5.68E–06
	KCN	1.61E–05
HDAC1	KCNR	3.20E–03
	KCN	6.02E–03
HDAC2	KCN	4.06E–05
	KCNR	4.10E–05
	48h SY5Y-MYCN	1.83E–02
	0h SY5Y-MYCN	4.51E–02
HDAC4	48h	1.20E–07
	KCN	6.20E–05
	KCNR	1.30E–04
	0h SY5Y-MYCN	1.18E–03
	24h SY5Y-MYCN	2.70E–03
HDAC5	48h	2.98E–02
HDAC6	0h SY5Y-MYCN	1.36E–02
SMARCA1	KCN	4.04E–02
SMARCA4	0h SY5Y-MYCN	2.49E–02
	KCNR	2.85E–02
SMARCB1	KCN	1.60E–02
SMARCC1	0h SY5Y-MYCN	4.96E–03
	KCN	3.07E–02
	48h SY5Y-MYCN	4.56E–02
SMARCD3	24h SY5Y-MYCN	3.83E–04
	0h SY5Y-MYCN	9.48E–04
	48h SY5Y-MYCN	1.61E–02
	KCNR	2.65E–02
SMARCE1	0h SY5Y-MYCN	6.08E–03
CREB1	0h SY5Y-MYCN	4.46E–08
(epigenetic related)	24h SY5Y-MYCN	1.28E–07
	48h SY5Y-MYCN	2.99E-16
	KCN	1.56E–08
	KCNR	1.90E–08

MYCN ChIP-seq datasets are from MYCN-amplified KCN and KCNR cell lines and a MYCN overexpression time-course (un-induced [0 h], 24 h, and 48 h induction) in SY5Y-MYCN cells.

a consequent increase in MYCN expression levels (Tumilowicz et al., 1970; Duffy et al., 2014, 2015) (**Supplementary Figure 1B**). Conversely, c-MYC levels are almost undetectable in IMR32 cells (Duffy et al., 2014, 2015; **Supplementary Figure 1B**), qualifying this cell line as a model where MYCN-dependent epigenetic regulatory networks are not obscured by interplay with often overlapping c-MYC-dependent regulatory networks.

We analysed the phenotypic response of IMR32 cells to treatments with library compounds at concentrations that showed physiological/cellular activity in other *in vitro* model systems. Since properties of the leads greatly varied (molecular weight, solubility, IC_{50}), treatments were performed in serial dilutions (1x, 0.1x and 0.01x) of the library working stock concentrations (**Supplementary Table 2**), corresponding to concentration ranges from 0.05 mM for potent histone methyltransferase inhibitors JIB-04 (Wang et al., 2013) and chaetocin (Cherblanc et al., 2013) to 1M for the weak aliphatic HDAC inhibitor valproic acid (Göttlicher et al., 2001; Dokmanovic et al., 2007). The 1x dilution is the suggested SGC working concentration for each compound. Cell morphology was assessed after 48 h (**Supplementary Figure 2B**). The 1x compound dilutions resulted in dramatic changes of cell morphology (**Supplementary Figure 2B**) or almost complete wipe-out of cell numbers in many cases (**Figure 2**, **Supplementary Figure 1C**), except the DNA-methyltransferase inhibitor, Decitabine. However, Decitabine did produce similar cellular responses to the other compounds upon prolonged 72 hour treatment, namely cell debris, reduced cell surface, lost axonal protrusions and clustered cells with highly heterochromatic nuclei (**Supplementary Figure 2C**). Interestingly, treatments with JQ-1 (BET-bromodomain inhibitor) and LLY-507 (inhibitor of SMYD2 protein lysine methyltransferase activity) led to a reduction of contact inhibition and clustering of the cells. SMYD2 expression in neuroblastoma tumours was also predictive of patient outcome (**Supplementary Figure 2D**).

Five compounds reduced IMR32 relative cell viability by at least 50% at nanomolar concentrations (0.01x dilution), i.e., UNC0638 (10 nm), SAHA (10 nm), entinostat (5 nm), TSA (5 nM) and rucaparib (100 nM) (**Figure 2**). These compounds are characterised as a H3K9me2 methyltransferase inhibitor (Vedadi et al., 2011), pan-HDAC inhibitor (Marks, 2007), HDAC1/HDA3 inhibitor (Saletta et al., 2014), pan-HDAC inhibitor (Hřebačková et al., 2009) and poly-ADP-ribosylation inhibitor of histone H1 (D'Amours et al., 1999), respectively. Interestingly, these potent compounds fall across the major epigenetic regulation groups (writers, readers and erasers), again implying a complex interplay of epigenetic regulatory mechanisms with neuroblastoma biology.

A number of the identified MYCN-related epigenetic hits (**Figures 1A–C** and **Tables 1**, **2** and **Supplementary Table 1**) were targets of compounds within the library and where this was the case, compound treatment induced a pronounced loss of cell viability (**Figure 2**, **Supplementary Figure 1C**). Specifically, SMARCA (compound 20), CBP (compounds 13, 40 and 43), SMYD (compound 44) and HDACs (compounds 6, 18, 28, 29, 30, 35, 36, and 37). This observation suggests that epigenetic therapeutics could be used to target MYCN oncogenic networks.

Since epigenetic compounds can generate long term-effects based on modulating epigenetic marks that may persist beyond the treatment period, we employed a washout strategy to assess whether the cells were able to recuperate after each treatment, or if the compounds had sustained effects. Alamar Blue viability assays are non-toxic enabling multi-point monitoring (Rampersad, 2012). After 48 h compound treatments, media containing the drug was removed and the cells were allowed to grow in standard media for a further 72 h. Crucially, even after the additional 72 h washout period, and in the 1x dilution condition the vast majority of compounds resulted in cell viability remaining at or below the post-48h-treatment level (**Figure 2**, **Supplementary Figure 1C**). The epigenetic

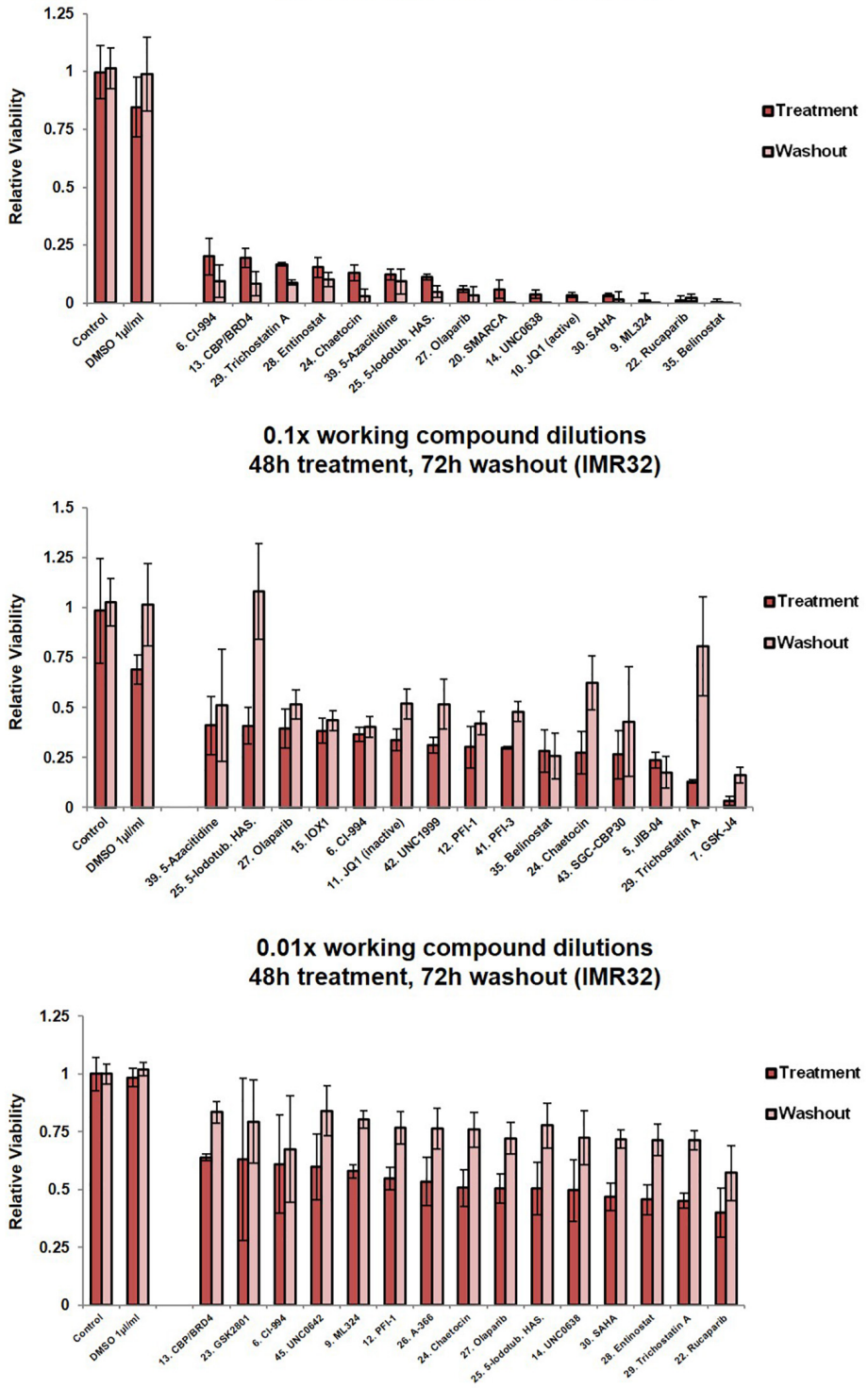

FIGURE 2 | Cell viability assessment for IMR32 cells upon treatment with the top 15 lead SGC library compounds. Compound dilutions are indicated above histograms (1x, 0.1x or 0.01x). Treatment bars – viability upon 48h treatment. Washout bars – viability 72 h upon washout, relative to corresponding untreated control at the same time point. For results from all 45 compounds in the library please see **Supplementary Figure 1C**. Error bars indicate the mean ± SD.

compounds produced long lasting effects, with cells not able to recommence proliferation. For almost half of the compounds (1x dilution condition), a further decline in the cell viability could be observed after the washout period (**Figure 2**, **Supplementary Figure 1C**).

Selective CBP/p300 Inhibitors Strongly Reduce the Viability of MYCN-Amplified Cells

From the library's lead compounds, we further evaluated CBP/p300 bromodomain inhibitors as potential MYCN amplified neuroblastoma therapeutics, since CBP (CREBBP) was prominent in our MYCN-related -omic and neuroblastoma patient outcome analysis. CBP was identified as a MYCN target gene in all MYCN ChIP-seq datasets (**Figure 3A** and **Table 1**, **Supplementary Table 1**), and CBP was a protein-protein interactor of MYCN (Duffy et al., 2015, 2016) in SY5Y-MYCN cells (**Supplementary Table 1**, **Supplementary Figure 3A**). CBP's co-factor p300 (EP300) was also identified as being MYCN bound (**Supplementary Table 1** and **Supplementary Figure 3A**). Similarly, CREB1, which acts in concert with CBP, is also a member of MYCN's transcriptional regulatory network, as revealed by IPA's ITR analysis of the ChIP-seq datasets (**Table 2**). In line with this observation on CBP, we have previously shown that inhibiting β-catenin binding to CBP in neuroblastoma cells alters their proliferative potential (Duffy et al., 2016).

CBP mRNA was consistently expressed in the cell lines utilised in this study (Duffy et al., 2015; **Figure 3B**), and more broadly across a panel of 39 neuroblastoma cell lines (Harenza et al., 2017) (**Supplementary Figure 2A**). Conversely, MYCN expression levels varied across the cell lines selected for this study (Duffy et al., 2014, 2015, 2016, 2017; Schwarzl et al., 2015; **Figure 3B**). CBP, CREB1 and p300 expression were individually prognostic for neuroblastoma patient outcome, with patients with low CBP, CREB1 or p300 mRNA expression only having an event-free survival rate of approximately 50% (**Figure 1B** and **Supplementary Figure 3B**). The library screen revealed that IMR32 cell viability is reduced upon CBP inhibition (**Figures 2**, **3C**).

Given the well-established link between the MYCN oncogene and neuroblastoma patient outcome, the response of neuroblastoma cell lines with varying MYCN status to CBP/p300 inhibition was assessed. The effects of three inhibitors: C646, CBP30 and I-CBP112 (**Figure 2**, **Supplementary Figure 1C**, epigenetic library compound 33, 40 and 43) (Bowers et al., 2010; Hay et al., 2014; SGC, 2020), on relative cell viability were assessed in a neuroblastoma cell line panel. The cell line panel covers the full spectrum of neuroblastoma's MYCN genetic backgrounds, ranging from non-amplified to MYCN amplified cells and from almost undetectable expression of MYCN to highly expressing cells (Duffy et al., 2014, 2015). While I-CBP112 did not exert a strong effect, both C646 and CBP30 reduced cell viability, with the cells being most susceptible to C646 treatment (**Figures 3D,E**). Interestingly, non-amplified MYCN cells (PAR-SY5Y, Dox- and Dox+ SY5Y-MYCN) and the MYCN amplified line with the lowest MYCN expression KCN were almost resistant to CBP30 treatment (**Figures 3D,E**). Conversely, the cell lines with high MYCN amplification IMR32 and KCNR responded strongly to 10μM CBP30 with significant reductions of cell viability, 55% and 49% loss of viability, respectively (**Figure 3E**). This pattern was repeated with 10 μM C646, with non-amplified MYCN cell lines showing minimal loss of viability, while MYCN amplified cell lines IMR32 and KCNR had a loss of viability of 95% and 87% respectively.

Responsiveness to 10 μM C646 was MYCN dose-dependent (**Figure 3F**), with the level of MYCN expression across the cell lines positively correlating to loss of viability (Pearson Correlation Coefficient, $R^2 = 0.469$). While parental (PAR-)SY5Y and un-induced (Dox^-) SY5Y-MYCN cells were practically resistant to 10 μM C646, SY5Y-MYCN cells induced to overexpress MYCN (Dox^+) had cell viability reduced to 45% (**Figures 3D–F**). Within the same genetic background, elevated MYCN expression significantly sensitises the cells to C646 (10 μM C646, Dox+ versus Dox- SY5Y-MYCN, *t*-test, p-value = 0.0190; **Figure 3D**) and even a modest increase in MYCN expression (**Figure 3B**) led to increased C646 sensitivity (10 μM C646, parental SY5Y versus Dox- SY5Y-MYCN, *t*-test, p-value = 0.0470; **Figure 3D**).

As was the case for CBP30, the KCN cell line was only very weakly reactive to C646 treatments (88% relative viability for 10 μM C646). Although KCN cells are MYCN amplified, the extent of amplification (and MYCN expression) is much lower than in IMR32 and KCNR cells (Duffy et al., 2015). KCN cells are derived from the primary tumour of a 1 month old infant, representing a less aggressive form of neuroblastoma than the metastatic tumour derived KCNR and IMR32 cells. KCNR is a patient matched cell line to KCN, being derived after relapse from a secondary tumour from the same individual (one year later). Therefore, the differential response to CBP/p300 inhibition between KCN and KCNR are not due to the initial mutational/epigenetic spectrum that gave rise to neuroblastoma in this patient, but likely from subsequent alterations which occurred between the primary and secondary tumours. One such characterised alteration is an over ten-fold increase in MYCN mRNA expression (Duffy et al., 2015; **Figure 3F**). Due to the lack of cytotoxic effects in the analysed neuroblastoma cell panel, compound I-CBP112 was excluded from further characterisation, while CBP30 and C646 were characterised in more detail. CBP/p300 inhibition shows potential as a therapeutic approach for advanced drug resistant MYCN amplified metastatic neuroblastoma tumours.

Temporal and IC_{50} Characterisation of CBP30 and C646 in IMR32 Cells

We next determined the temporal profile of the activity for 10 μM CBP30 and C646 by treating IMR32 cells from 6 h to 72 h (**Figure 3G**). We observed no significant effect for both leads during short treatments (6 h and 24 h), while prolonged treatments (48 h and 72 h) had a significant and pronounced effect on cell viability (**Figure 3G**). C646 exerted switch-like

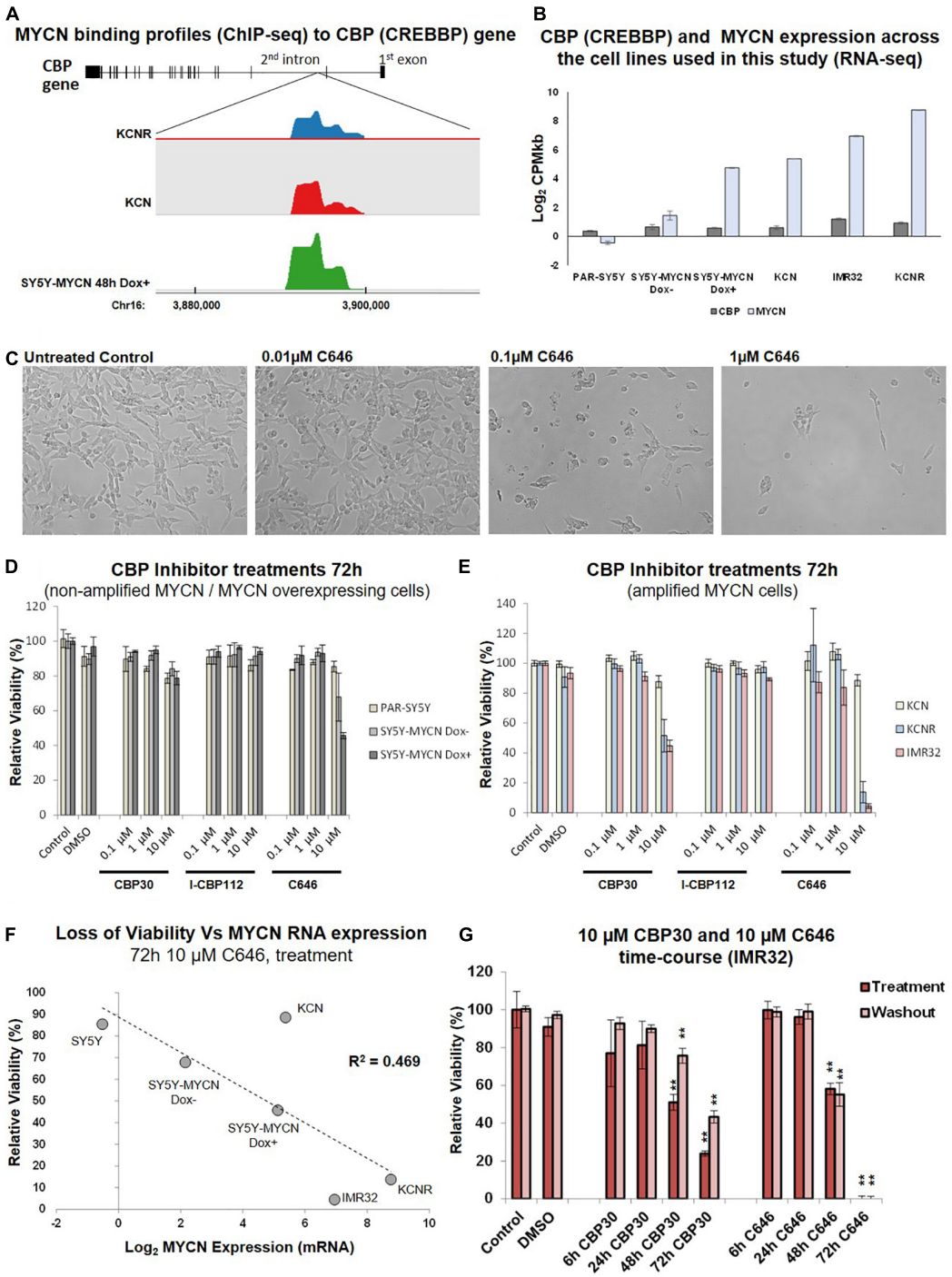

FIGURE 3 | CBP inhibition across a neuroblastoma cell line panel. **(A)** MYCN binding enrichment within the second intron of the CBP gene. Top – CBP gene schematic representation. Middle – Probe enrichment visualisation for KCNR, KCN, and SY5Y-MYCN 48 h Dox induced representative datasets using the SeqMonk tool. Bottom -genomic coordinates. **(B)** CBP and MYCN transcript expression levels across the cell line panel. Read counts per million adjusted by gene length in kilobases (CPMkb), obtained from our previously published RNA-seq of these cell lines (Duffy et al., 2015, 2017). **(C)** Images of IMR32 cells after 48 h treatment with C646. Images were taken with 40x magnification. **(D)** Relative viability (Alamar Blue assay) upon CBP inhibitor treatment (CBP30, I-CBP112 and C646) of MYCN single copy cell lines, including doxycycline inducible MYCN cell line SY5Y-MYCN. **(E)** Relative viability (Alamar Blue assay) upon CBP inhibitor treatment (CBP30, I-CBP112 and C646) of MYCN amplified cell lines, including doxycycline inducible MYCN cell line SY5Y-MYCN. **(F)** Correlation of relative viability loss upon C646 treatment (Alamar Blue assay) and MYCN mRNA expression (Log_2 fold change) obtained from RNA-seq (Duffy et al., 2015). R^2-value reported on the panel is from a Pearson Correlation Coefficient analysis. **(G)** Relative viability (Alamar Blue assay) of C646 time-course (6 h, 24 h, 48 h, and 72 h) treatment of MYCN amplified IMR32 cells. Viability results immediately after treatment (Treatment) and after an additional 48 h in the absence of C646 after treatment ended (Washout) are shown. Paired sample t-test values are designated with asterisks above bars. $P \leq 0.05$ - *; $P \leq 0.01$ - **. Error bars indicate the mean ± SD.

temporal behaviour with a difference in relative viability between 48 h and 72 h treatments. Washout (48 h) experiments did not reveal any significant change in the cell viability when compared to their corresponding treatment, showing that cells are unable to recover from treatment even after the removal of the inhibitor from the growth media (**Figure 3G**).

An expanded dose-dependent response analysis was performed by treating IMR32 cells with 1 – 20 μM CBP30 or C646 for 72 h (**Figure 4A**). Washout relative cell viability was also determined 48 h after the 72 h treatment (**Figure 4A**). For both leads, concentrations above 5 μM were required to induce major cytotoxic effect in IMR32 cells (**Figure 4A**). Even after washout the reduction in viability in the high-dose treatments was maintained with the reduction continuing to intensify even in the absence of the inhibitors (**Figure 4A**). The response curves and IC_{50} values of each compound were estimated using the IC50 Toolkit[4]. The estimated IC_{50} values for C646 and CBP30 were 8.5 μM and 15 μM, respectively (**Figure 4B**), confirming that C646 is the more potent lead. C646 had a sigmoid dose response curve, contrary to CBP30's linear dose response curve (**Figure 4B**). Viability responses to C646 exhibited switch-like behaviour, both in terms of its temporal response and dose response profiles.

CBP and BRD4 Inhibitor Combination Treatments

We next assessed whether targeting multiple epigenetic vulnerabilities simultaneously would have additive or synergistic effects on neuroblastoma cell viability. The compound library results showed that BET-bromodomain BRD4 inhibition strongly reduced IMR32 cell viability (compounds 10, 21, 34 and 41, **Figure 2**, **Supplementary Figure 1C**). Furthermore, a compound [CBP/BRD4 (0383)] co-targeting BRD4 (reader) and the related CBP protein (writer) also resulted in a strong loss of cell viability (compound 13, **Figure 2**). Like CBP and its CREB and p300 co-factors, BRD4 expression level is predictive of neuroblastoma patient outcome (**Figure 1B** and **Supplementary Figure 3B**). We therefore tested BRD4 and CBP combination treatment in IMR32 cells. First, to confirm the effects of BRD4 inhibition in neuroblastoma cells, a dose course of the BRD4 inhibitors JQ1 and I-BET were assessed in the cell line panel (**Figures 4C,D**). Mirroring the library screening (**Figure 2**, **Supplementary Figure 1C**) both inhibitors strongly reduced IMR32 cell viability, with JQ1 being the more potent (**Figure 4D**). Similar to CBP inhibition, BRD4 inhibition tended to reduce viability to a greater degree in high MYCN amplified cells (KCNR and IMR32). Although, non-amplified MYCN cells, MYCN overexpressing cells (SY5Y-MYCN Dox+) and early stage primary MCYN cells (KCN) were more responsive to JQ1 inhibition than they were to CBP inhibition (**Figures 3C,D**, **4C,D**).

Having confirmed the BRD4 inhibitors effect across the cell line panel, we next evaluated potential synergy between these compounds. IMR32 cells were treated with single agents at their near IC_{50} concentrations, C646 (10 μM and 5 μM) and JQ1 (1 μM and 0.5 μM), as well as with their combinations. Combining JQ1 and C646 resulted in slight additive effects at 24 h and 48 h (**Figures 4E,F**), although at 48hr this was non-significant, p-value = 0.1301 (paired sample t-test). Combination treatments were also assessed in the MYCN non-amplified cell line SY5Y-MYCN with Dox inducible MYCN expression. Once again combination treatment only produced a mild additive effect (**Supplementary Figure 3C**).

Zebrafish Developmental Screens and Lack of Cytotoxicity of C646 in Differentiated Neurons

To further asses the lead compounds, we examined how well tolerated they were in zebrafish developmental phenotypic screens. Due to its dynamic nature, embryonic development tends to be more sensitive to pharmaceutical perturbations than non-developmental stages, with responses in rapidly dividing embryonic cells often recapitulating those observed in rapidly proliferating cancer cells. Twenty-four hour treatments with IBET-151, CBP30 and I-CBP112 (concentrations 1, 5, and 10 μM) were well tolerated and resulted in no obvious phenotype changes (**Supplementary Figure 4A**). While 1 μM C646 was well tolerated, 5 μM C646 treated embryos had an overall distorted appearance and an observable increase in heart rate (**Figure 4G**). Furthermore, there was 100% lethality in embryos treated with 10 μM C646. While there were no obvious gross phenotypic effects of 1 μM and 5 μM JQ1 treatment, 10 μM JQ1 resulted in a reduced chromatophores population by 52 h post fertilisation (hpf), resulting in significant embryo depigmentation (**Figure 4G**). Like neuroblastoma precursor cells (neuroblasts), melanocytes (chromatophore cell type) also originate from the transient embryonal neural crest tissue (Kuriyama and Mayor, 2008). Interestingly, inhibition of Wnt signalling similarly affected chromatophore (melanocyte) development in a similar screen (Duffy et al., 2016). While embryonic development is known to be more sensitive to perturbation than post-developmental stages, the potency of C646 and JQ1 should be carefully considered when assessing their utility as therapeutics for childhood cancers.

Cancer therapy, including chemotherapeutic agents, may affect the nervous system in a deleterious manner. In patients this can lead to numerous sequelae, such as chemotherapy-related encephalopathy, meningitis, chronic pain syndrome, psychiatric disorders, neuropathy (Chamberlain, 2010). Often, these neurological problems are associated with neurotoxicity. Laboratory differentiated neuroblastoma cell lines have been used extensively to screen novel compounds for neurotoxic properties and associated mechanisms. Importantly, the response of neuroblastoma cells to compounds and drugs exposure may differ from that of neurons. The differentiation potential of neuroblastoma cell lines into mature neurons, enables the pharmacological and functional differences between neurons and their blast cell counterparts to be assessed (LePage et al., 2005; Duffy et al., 2016, 2017).

To assess the likelihood of unintended neurotoxicity from the clinical use of C646 and JQ1, we analysed the effect on viability of long-term differentiated SY5Y-MYCN cells

[4]http://ic50.tk/index.html

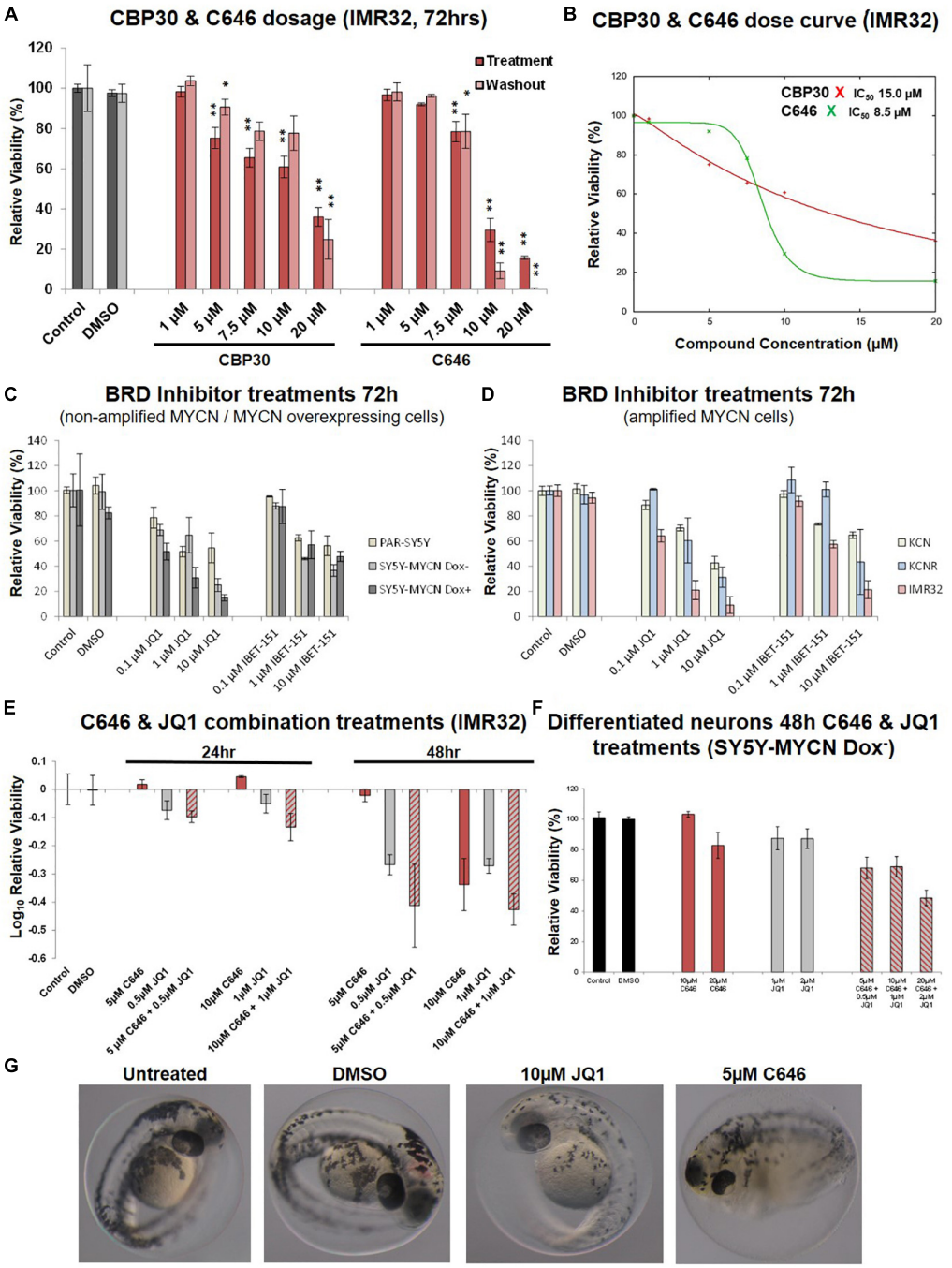

FIGURE 4 | Dose curves, combination treatments, neuronal and phenotypic screens. **(A)** Expanded dose response assessment of the inhibitors C646 and CBP30 in MYCN amplified IMR32 cells. Relative cell viability was determined using Alamar Blue assays. Viability results immediately after the 72 h treatment (Treatment) and after an additional 48 h in the absence of inhibitors after treatment ended (Washout) are shown. Paired sample *t*-test values are designated with asterisks above bars. $P \leq 0.05$ - *; $P \leq 0.01$ - **. Error bars indicate the mean ± SD. **(B)** IC_{50} estimate for CBP30 and C646 for the 48h treatment of IMR32 cells. Asterisks correspond to X axis drug concentrations (μM) and Y axis relative cell viability (%). Curve fitting performed with IC Toolkit (http://ic50.tk/index.html). **(C)** Relative viability (Alamar Blue assay) upon BRD4 inhibitor treatment (JQ1 and I-BET151) of MYCN single copy cell lines, including doxycycline inducible MYCN cell line SY5Y-MYCN. **(D)** Relative viability (Alamar Blue assay) upon BRD4 inhibitor treatment (JQ1 and I-BET151) of MYCN amplified cell lines, including doxycycline inducible MYCN cell line SY5Y-MYCN. **(E)** Relative viability (Alamar Blue assay) upon combination CBP and BRD4 inhibitor treatment (C646 and JQ1) of MYCN amplified IMR32 cells. **(F)** Single agent and combination treatment (C646 and JQ1) of differentiated neurons (pre-inhibitor treatment, 11 day RA treated un-induced SY5Y-MYCN cells). Log_{10} relative viability (Alamar Blue assay) upon 48 h treatments combination CBP and BRD4 inhibitor treatment (C646 and JQ1) of MYCN amplified IMR32 cells. Error bars indicate the mean ± SD. **(G)** Zebrafish embryos treated for 24 h (28 h post fertilisation [hpf]- 52hpf) with epigenetic targeting compounds (JQ1 and C646). Magnification is 10x, images from 52 hpf. Melanocytes appear black due to their endogenous melanin. Otic vesicles (large black circles) mark the anterior pole (head) of the embryos.

upon treatment with these inhibitors as single agents or in combinations. Ahead of the epigenetic drug treatment, SY5Y-MYCN cells were treated with 1 μM all-trans retinoic acid (RA) for 11 days. The differentiation capacity of RA has been well established, with extensive outgrowth of neurites and expression of neuron-specific markers (Encinas et al., 2000; Cheung et al., 2009; Duffy et al., 2016, 2017), and we have previously shown that the SY5Y-MYCN cell line also undergoes neuronal differentiation in response to RA (Duffy et al., 2014, 2017). Differentiated SY5Y-MYCN neurons were fully resistant to 10 μM C646, and only weakly responsive to 20 μM C646 concentration (**Figure 4F**). Similarly, the differentiated neurons were more resistant to 1 μM JQ1 treatment (**Figure 4F**) than their undifferentiated counterparts (**Supplementary Figure 3C**). Forty-eight hour 10 μM C646 and 1 μM JQ1 combination treatment reduced cell viability to approximately 25% in both Dox^- and Dox^+ SY5Y-MYCN cells (**Supplementary Figure 3C**), whereas differentiated neurons were more tolerant of the combination treatment with approximately 70% remaining viable (**Figure 4F**). Taken together, these results demonstrate that the C646 and JQ1 epigenetic targeting compounds are better tolerated by differentiated neurons than their more stem-like neuroblastoma cell counterparts. Furthermore, C646 is better tolerated by neurons than JQ1, with 10 μM C646 having no effect on neuronal cell viability.

DISCUSSION

Epigenetic alterations are strongly associated with the development of cancer, and compounds targeting epigenetic modifier proteins are increasingly being assessed as anti-cancer therapeutic agents (Juergens et al., 2011). Integration of multi-omic MYCN datasets revealed strong support for the hypothesis that targeting epigenetic regulatory mechanisms in MYCN amplified neuroblastoma will likely provide beneficial future therapeutic strategies. Functional support for the applicability of epigenetic targeting compounds was provided by the fact that 43 of the 45 epigenetic library compounds reduced the viability of MYCN amplified neuroblastoma cells. The high responsiveness rate (96%) across a range of epigenetic modulators, confirms a prominent role for epigenetic alterations as drivers of neuroblastoma oncogenesis and progression. Viability assays revealed that modulation of epigenetic regulatory mechanism takes time to exert its effect, but at 1x concentration given sufficient time, dramatic reductions in NB cell viability can be achieved.

In addition to signalling previously reported cross-talk between MYCN and HDAC (Lodrini et al., 2013; Duffy et al., 2016; Fabian et al., 2016; Phimmachanh et al., 2020), we revealed novel MYCN-epigenetic interactions, including those with SMARC genes. SMARC genes were identified as inferred transcriptional regulators of MYCN-bound genes (ChIP-seq), their expression showed strong correlation to neuroblastoma outcome, and SMARCA inhibition strongly reduced IMR32 cell viability.

The SMARC gene family comprises a family of proteins, that display both helicase and ATPase activities, and which play roles in regulation of transcription of certain genes by altering the chromatin structure around those genes (Zofall et al., 2006). Mutations that inactivate SMARC subunits are found in nearly 20% of human cancers, driving aberrant growth (Hohmann and Vakoc, 2014). Certain acute leukaemias and small cell lung cancers, which lack SMARC mutations (similar to neuroblastoma) can be vulnerable to inhibition of SMARCA4 (Hohmann and Vakoc, 2014).

MYCN amplified neuroblastoma cells are resistant to retinoid-based differentiation therapy (Duffy et al., 2017). Interestingly, cell lines lacking SMARCA4 from a variety of cancers do not respond to retinoid therapy, while restoration of SMARCA4 expression restores retinoid sensitivity (Romero et al., 2012). Restoration of SMARCA1 *in vivo* significantly reduced lung cancer invasiveness and c-MYC expression, suggesting that inactivated SMARCA1 keeps cancer cells in an undifferentiated state and prevents its response to developmental and environmental stimuli (Romero et al., 2012). Conversely, in SH-SY5Y neuroblastoma cells, a commonly used model system of retinoid-dependent differentiation, attenuated expression of SMARCA1 prevented response to retinoids (Romero et al., 2012). Here, we show that elevated SMARCA4 expression corresponds strongly to poor outcomes for neuroblastoma patients, whereas elevated expression of SMARCA1 correlates to good outcome, suggesting antagonist roles in neuroblastoma for members of this gene family. SMARCA4 was also one of the genes identified as having protein-protein interactions with MYCN (mass spectrometry Co-IP). The NCI paediatric MATCH trial (NCT03155620) has recently detected SMARC mutations in relapsed neuroblastoma (Kumar et al., 2019), in keeping with the effectiveness of SMARCA inhibitors in strongly reducing the viability of the IMR32 relapsed cell line (**Figure 2**, compound 20). *In vitro* and *in vivo* loss-of-function experiments showed that SMARCA4 is essential for the proliferation of neuroblastoma cells (Jubierre et al., 2016). Taken together, SMARC inhibition provides a promising future therapeutic direction for relapsed MYCN amplified neuroblastoma.

Since few epigenetics drugs, are in clinical use for neuroblastoma, we pursued *in vitro* analyses to define novel potential therapeutics. IMR32, MYCN amplified neuroblastoma cells, showed broad susceptibility to epigenetic targeting compounds. Currently, 14 compounds within the SGC library are in clinical use or undergoing clinical trials (**Supplementary Table 2**), with five - Decitabine, Vorinostat Azacitidine, Olaparib and valproic acid, being tested in neuroblastoma. Given their advanced stage of testing elsewhere, we excluded these molecules from further characterisation. However, these agents are pan-antagonists that target DNMTs (Decitabine, Azacitidine), PARP (Olaparib) and HDACs (vorinostat and valproic acid), which usually results in numerous off-target effects, high toxicity and acquired resistance with prolonged drug exposure. Therefore, leads that have the potential to directly or indirectly modulate the activity of the MYCN oncogene are desirable. To identify such leads, we examined

our omic datasets for genes targeted by SGC library compounds which significantly reduced IMR32 cell viability, which led to a primary focus on the CBP/p300 inhibitors: C646, I-CBP112 and CBP30.

The CBP and E1A Binding Protein P300 (known as EP300 or p300) proteins are closely related histone acetyltransferases (HATs) that act as transcriptional coactivators (Bannister and Kouzarides, 1996; Ogryzko et al., 1996). CBP/p300 inhibition can facilitate cellular reprogramming (Ebrahimi et al., 2019). Histone acetyltransferase inhibitors block SK-N-SH neuroblastoma cell growth *in vivo*, partly through CBP and p300 interactions (Gajer et al., 2015). Importantly, CBP/p300 also functions as a scaffold for various members of the transcriptional machinery and/or transcription factors (Barbieri et al., 2013). Among numerous partners forming transcriptional complexes, c-Myc binds to TATA-binding protein (TBP) and CBP, one half of the CBP/p300 coactivator complex, which has HAT activity and scaffolding functions (Patel et al., 2004). Thus, bound c-MYC together with CBP/p300 is localised to acetylated chromatin becoming a part of the transcriptional machinery to potentiate/modulate gene expression (He et al., 2013). CBP is also able to bind acetylated lysines on non-histone proteins, such as the p53 oncogene, and this interaction is required for the activation of the cyclin-dependent kinase inhibitor p21 (Mujtaba et al., 2004). It has also been shown that p300 is able to associate with c-Myc in mammalian cells through direct interactions with transactivation residues in c-MYC (N-terminal amino acids 1 to 110) and acetylates c-Myc protein. The resulting acetylation modulates the activity of c-MYC as well as the turnover of the protein (Faiola et al., 2005). Although highly similar in structure to c-MYC, there has been no study to reveal whether CBP/p300 co-factors are also able to modulate transcriptional activity and stability of the MYCN protein. Since CBP/p300 act as co-factors that facilitate localisation of transcription factors, among them MYC family members, modulating their activity represents a valuable and still poorly understood potential therapeutic approach.

Here, we show that MYCN amplified neuroblastoma cells are susceptible to C646 treatment, and that CBP targeting compounds should be further investigated as a single modality or combination therapeutic approach for high-risk neuroblastoma. C646 has been shown to increase survival in a mouse model of leukaemia (Wang et al., 2011), suggesting therapeutic utility across a range of cancer types. Since single agent treatments have thus far not provided the desired outcome benefit for high-risk neuroblastoma patients and since they can lead to acquired drug resistance, we also assessed the utility of a BRD4 inhibitor (JQ1) and CBP/p300 inhibitor (C646 and I-CBP112) combination. The combination treatments demonstrated that BRD and CBP/p300 inhibitors do not exert a synergistic effect, but rather a mild additive one. These compounds should continue to be assessed as single agents until more effective combination partners can be found for each. C646 was well tolerated in our differentiated neuron assay, Interestingly, the use of C646 has been reported in a study to impair memory formation during fear conditioning in mice, without reported adverse neurotoxic effects (Maddox et al., 2013).

CONCLUSION

We reveal that MYCN amplified neuroblastoma cells are exquisitely sensitive to pharmaceutically induced epigenetic perturbations, with inhibition of numerous epigenetic mechanisms resulting in extensive loss of cell viability. We also show that MYCN interacts with the epigenetic machinery, both by protein-protein interactions, as well as, by directly regulating the expression of epigenetic modifiers. Thus, targeting of MYCN's epigenetic network may prove an effective therapeutic avenue for high-risk neuroblastoma. In particular, the CBP/p300 inhibitor C646 shows promise as a targeted MYCN amplified neuroblastoma therapeutic.

MATERIALS AND METHODS

Cell Culture and Treatments

For cell culture conditions and cell sources see Duffy et al. (2014). Briefly, the six neuroblastoma cell lines used were received as generous gifts from Dr. Frank Westermann (DKFZ, Heidelberg University, Germany) and Dr. Johannes Schulte (University Children's Hospital Essen, Germany). The following lines were used: IMR-32 (IMR32) – ATCC code CCL-127, SH-SY5Y parental cell lines (PAR-SY5Y) (ATCC code CRL-2266), SMS-KCN (KCN)[5], SMS-KCNR (KCNR)[6], SH-SY5Y/6TR(EU)/pTrex-Dest-30/MYCN (SY5Y-MYCN) [36,71] – (generated in the Westermann Laboratory, DKFZ, Heidelberg). All NB cell lines were maintained in RPMI 1640 media (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 1% Penicillin-Streptomycin solution (Gibco). Transgenic cell line SY5Y-MYCN was maintained in the same media as for other cell lines and supplemented with G418 (0.2 mg/ml) (Sigma-Aldrich) and Blasticitidin (7.5 μg/ml) (Invitrogen). Induction of MYCN overexpression in SY5Y-MYCN cells was performed by adding Doxycycline hyclate (Sigma-Aldrich) at the final concentration of 1 μg/ml. Cells were cultivated in cell culture incubator (Thermo Scientific) at 37°C and 5% CO2. For differentiation assays and generation of neurons from NB cells, all-*trans* retinoic acid (RA) was added to growth media at the final concentration of 1 μM.

The SGC chemical library (Plate 9C – 45 compounds, **Supplementary Table 2**) was used for compound screening. Additional small molecules beyond the SGC library stocks used were from the following sources: C646 (SML0002, Sigma Aldrich), CBP30 (#4889, Tocris Bioscience), I-CBP112 (#4891, Tocris Bioscience), (+) -JQ1 (#87110, Selleck Chemicals), I-BET151 (ML0666, Sigma Aldrich) and all-*trans* Retinoic Acid (#R2625, Sigma Aldrich). Stock solutions were dissolved in DMSO. Compounds were replenished every 24 h for any treatment longer than a 24 h duration.

[5]http://www.cogcell.org/dl/NB_Data_Sheets/SMS-KCN_Cell_Line_Data_Sheet_COGcell_org.pdf

[6]http://www.cogcell.org/dl/NB_Data_Sheets/SMS-KCNR_Cell_Line_Data_Sheet_COGcell_org.pdf

Cell Treatments

Doxycycline induction - SY5Y-MYCN cells were induced for MYCN expression by adding 1 μg/ml (final concentration) of Doxycycline to growth media. As the half-life of this drug in the culture media is approximately 48 h, growth media was replaced every day.

All-*trans* retinoic RA induction - For the generation of neurons from SH-SY5Y-MYCN cells, all-*trans* retinoic acid (RA) was added to growth media at the final concentration of 1 μM. RA was prepared in storage stocks (25 mg/ml) by dissolving the compound in DMSO. These stocks were stored at −80°C no longer than 6 months past the date of the stock preparation. Working stock (1 mM; 0.300044 mg/ml) was also kept at −80°C. Upon treatment, RA reagent and tissue culture dishes were kept protected from light.

Cell treatments with small molecule compounds - Cells were seeded in 96-well plates with 10,000 cells/well, with at least technical triplicates per treatment. Compounds were used in various concentrations, in accordance with experimental setup. In order to minimize the effect of compound degradation in the growth media, all compounds were replenished every 24 h for any treatment longer than the 24 h duration.

Alamar Blue Cell Viability Assays

Alamar Blue assay (Resazurin sodium salt R7017, Sigma) provides the ability to monitor the combined effects of cellular health, apoptosis, proliferation, cell cycle function and control in a single assay. The active ingredient for Alamar Blue assay is resazurin which is water-soluble and stable in culture medium. Compared to other cell viability reagents it's non-toxic and easily permeable through cell membranes. Measurements were performed in at least triplicate. Treatments were performed in 96-well plates. Upon the conclusion of treatments culture media was discarded and 0.1 ml of fresh RPMI-1640 medium with fresh Alamar Blue 0.05% stock (1:10 Alamar Blue to medium) was added directly to the wells. Cells were incubated for at least 3 h at 37°C and 5% CO2. After this time, absorbance was measured (Spectramax Plus384 Plate Reader, Molecular Devices accompanied with SoftMaxPro software) at excitation 570 nm and emission 600 nm. Calculations were performed in accordance with the manufacturer's protocol[7]. The results are shown as the mean inhibition indices calculated by dividing each experimental result by the mean of the respective control values. For the washout experiments, media containing Alamar Blue was removed upon measurement and replaced with fresh media in order to facilitate the recovery of cells from the drug treatment.

Statistical Analysis

Statistical testing was performed in MS Excel. Data was tested for normal distribution (Shapiro-Wilk Test) using the online tool; https://www.statskingdom.com/320ShapiroWilk.html. Two sided t-tests with equal variance were performed by using the online tool; http://studentsttest.com/. Error bars are represented as plus/minus one standard deviation. Linear regression analysis was also performed in MS Excel (Pearson Correlation Coefficient test). IC_{50} calculation and dose response curve fitting were performed using the online tool, IC50 Toolkit[8].

Zebrafish Treatments

Zebrafish larvae (*Danio rerio*, AB and Tg[Fli1:EGFP] strains) were maintained on a 14h light/10h dark lighting cycle at 28.5°C. Drug treatments were conducted for 24 h (28 h post fertilisation [hpf]- 52 hpf), by waterborne exposure i.e., incubating the embryos in water containing epigenetic targeting compounds; 10 μM IBET-151, 5 μM C646, 10 μM CB112, 10 μM JQ1 and 10 μM CBP30. Embryos studies were approved by the UCD Animal Research Ethics Committee.

Omic Datasets and Bioinformatics Tools

Ingenuity Pathway Analysis (IPA) software was also used for the ITR and pathway analysis. All survival curves were generated using the SEQC (Zhang et al., 2015) 498 neuroblastoma tumour dataset in the R2: Genomics Analysis and Visualization Platform[2]. Kaplan Meier curves were generated using the KaplanScan function that segregates a patient cohort in 2 groups on the basis of gene expression. The scanning function of this tool yields a cut-off where the difference in survival is most significant. KaplanScan was run with the following parameters: Type of Survival - event free; minimal groups size = 8. Multi-omic datasets (ChIP-seq, RNA-seq and interactome) were mined from published studies (Duffy et al., 2015, 2016; Harenza et al., 2017). Peaks were visualized using the SeqMonk analysis tool[9] using default parameter settings. Briefly, using input ChIP-seq BAM files, sequencing probes were generated and peaks were called with MACS2, followed by annotation for the nearest gene. Distinct peaks were quantified and normalized against the largest dataset, while smoothing was performed using the pipeline for adjacent probes. Venn diagrams were generated using jvenn (Bardou et al., 2014)[10]. Epigenetic genes78 The dbEM, Database of Epigenetic Modifiers (Singh Nanda et al., 2016), was used for identifying MYCN's epigenetic related targets[1]. Protein interaction networks were generated using the String database[11].

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: ArrayExpress, www.ebi.ac.uk/arrayexpress, accession number E-MTAB-4100 (previously published, was mined for the current study).

ETHICS STATEMENT

The animal study was reviewed and approved by The UCD Animal Research Ethics Committee.

[7]http://tools.lifetechnologies.com/content/sfs/manuals/PI-DAL1025-1100_TI%20alamarBlue%20Rev%201.1.pdf

[8]http://ic50.tk/

[9]https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/

[10]http://jvenn.toulouse.inra.fr/app/index.html

[11]https://string-db.org/

AUTHOR CONTRIBUTIONS

AKr, WK, and DD designed and supervised the project. AKr, AKo, DD, and MH generated the data and performed data and bioinformatics analysis. UO and PC generated, designed and curated the epigenetic library. AKr and DD wrote the manuscript, with all authors (AKr, AKo, MH, UO, PC, WK, and DD) contributing to reading and editing the manuscript.

FUNDING

The research leading to these results received funding from the European Union Seventh Framework Programme (FP7/2007- 2013) ASSET project under grant agreement number FP7-HEALTH-2010-259348-2, Cancer Research United Kingdom (A23900), the LEAN program grant of the Leducq Foundation, and with the financial support provided by the National Children's Research Centre and Science Foundation Ireland under the Precision Oncology Ireland Grant number 18/SPP/3522.

ACKNOWLEDGMENTS

Warmest thanks to Breandán Kennedy for the generous use of his zebrafish facilities, and all the SBI support staff (Amaya Garcia Munoz and Ruth Pilkington). Thanks are also due to Frank Westermann, Johannes Schulte, Sven Lidner and Andrea Odersky for the generous gifting of cell lines.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.612518/full#supplementary-material

Supplementary Figure 1 | Additional SMARC survival curves, MYC family gene expression and SGC library compound treatment cell viability results. **(A)** Kaplan-Meier survival curves showing the predictive strength of the expression levels of SMARC genes in neuroblastoma tumours on patient outcome. The SMARCs shown were predicted to be upstream regulators of MYCN's genomic targets (ChIP-seq). SMARCA1and SMARCD3 was also a direct MYCN target (ChIP-seq). All survival curves were generated using the SEQC (Zhang et al., 2015) 498 neuroblastoma tumour dataset in the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl). **(B)** Gene expression levels of MYC family gene members in the IMR32 cell line, as assessed by RNA-seq (Duffy et al., 2015) **(C)** Cell viability assessment for IMR32 cells upon treatment with SGC library compounds. Compound dilutions are indicated above histograms (1x, 0.1x or 0.01x). Treatment bars – viability upon 48 h treatment. Washout bars – viability 72h upon washout, relative to corresponding untreated control at the same time point. Error bars indicate the mean ± SD.

Supplementary Figure 2 | Additional expression and post-treatment imaging data. **(A)** CBP (CREBBP) transcript expression across the 39 neuroblastoma cell lines and two control non-neuroblastoma samples (human foetal brain tissue and a retinal pigment epithelial cell line, RPE1 cells) profiled by RNA-seq by Harenza et al. (2017), expression measured in FPKMBs. **(B)** Images of IMR32 cells after 48 h treatment with 1x dilution of epigenetic targeting compounds. Images were taken with 40x magnification. **(C)** Images of IMR32 cells after 72 h treatment with decitabine (5-aza-dC). **(D)** Kaplan-Meier survival curves showing the predictive strength of the expression level of SMYD2 in neuroblastoma tumours on patient outcome. Survival curve was generated using the SEQC (Zhang et al., 2015) 498 neuroblastoma tumour dataset in the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl).

Supplementary Figure 3 | Additional network, outcome and viability data. **(A)** Protein-protein interaction network of the epigenetic MYCN interactome hits, and the MYCN protein, as generated by STRING (https://string-db.org/). Epigenetic hits were generated by overlapping our previously published MYCN mass spectrometry protein-protein interactome (Duffy et al., 2015, 2016) with epigenetic gene lists from Miremadi et al. (2007) and EpiDBase (Loharch et al., 2015) (**Supplementary Table 1**). **(B)** Kaplan-Meier survival curves showing the predictive strength of the expression levels of the CREB1, EP300 and BRD4 genes in neuroblastoma tumours on patient outcome. All survival curves were generated using the SEQC (Zhang et al., 2015) 498 neuroblastoma tumour dataset in the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl). **(C)** Log_{10} relative viability (Alamar Blue assay) upon 48 h combination CBP and BRD4 inhibitor treatment (C646 and JQ1) of MYCN amplified induced (Dox^+) and non-induced (Dox^-) SY5Y-MYCN cells. Error bars indicate the mean ± SD.

Supplementary Figure 4 | Additional phenotypic screen data. **(A)** Zebrafish embryos treated for 24 h (28 h post fertilisation [hpf]- 52 hpf) with epigenetic targeting compounds (IBET-151, I-CB112 and CBP30). Magnification is 10x, images from 52 hpf. Melanocytes appear black due to their endogenous melanin. Otic vesicles (large black circles) mark the anterior pole (head) of the embryos.

REFERENCES

Applebaum, M. A., Vaksman, Z., Lee, S. M., Hungate, E. A., Henderson, T. O., London, W. B., et al. (2017). Neuroblastoma survivors are at increased risk for second malignancies: a report from the international neuroblastoma risk group project. *Eur. J. Cancer* 72, 177–185. doi: 10.1016/j.ejca.2016.11.022

Bannister, A. J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature* 384, 641–643. doi: 10.1038/384641a0

Barbieri, I., Cannizzaro, E., and Dawson, M. A. (2013). Bromodomains as therapeutic targets in cancer. *Brief. Funct. Genomics* 12, 219–230. doi: 10.1093/bfgp/elt007

Bardou, P., Mariette, J., Escudié, F., Djemiel, C., and Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* 15:293. doi: 10.1186/1471-2105-15-293

Bowers, E. M., Yan, G., Mukherjee, C., Orry, A., Wang, L., Holbert, M. A., et al. (2010). Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem. Biol.* 17, 471–482. doi: 10.1016/j.chembiol.2010.03.006

Carter, D. R., Murray, J., Cheung, B. B., Gamble, L., Koach, J., Tsang, J., et al. (2015). Therapeutic targeting of the MYC signal by inhibition of histone chaperone FACT in neuroblastoma. *Sci. Transl. Med.* 7:312ra176. doi: 10.1126/scitranslmed.aab1803

Chamberlain, M. C. (2010). Neurotoxicity of cancer treatment. *Curr. Oncol. Rep.* 12, 60–67. doi: 10.1007/s11912-009-0072-9

Cherblanc, F. L., Chapman, K. L., Brown, R., and Fuchter, M. J. (2013). Chaetocin is a nonspecific inhibitor of histone lysine methyltransferases. *Nat. Chem. Biol.* 9, 136–137. doi: 10.1038/nchembio.1187

Cheung, Y.-T., Lau, W. K.-W., Yu, M.-S., Lai, C. S.-W., Yeung, S.-C., So, K.-F., et al. (2009). Effects of all-trans-retinoic acid on human SH-SY5Y neuroblastoma as in vitro model in neurotoxicity research. *Neurotoxicology* 30, 127–135. doi: 10.1016/j.neuro.2008.11.001

Cohn, S. L., Pearson, A. D. J., London, W. B., Monclair, T., Ambros, P. F., Brodeur, G. M., et al. (2009). The international neuroblastoma risk group (INRG) classification system: an INRG task force report. *J. Clin. Oncol.* 27, 289–297. doi: 10.1200/jco.2008.16.6785

D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G. G. (1999). Poly (ADP-ribosyl) ation reactions in the regulation of nuclear functions. *Biochem. J.* 342, 249–268. doi: 10.1042/0264-6021:3420249

Dokmanovic, M., Clarke, C., and Marks, P. A. (2007). Histone deacetylase inhibitors: overview and perspectives. *Mol. Cancer Res.* 5, 981–989. doi: 10.1158/1541-7786.mcr-07-0324

Domingo-Fernandez, R., Watters, K., Piskareva, O., Stallings, R. L., and Bray, I. (2013). The role of genetic and epigenetic alterations in neuroblastoma disease pathogenesis. *Pediatr. Surg. Int.* 29, 101–119. doi: 10.1007/s00383-012-3239-7

Dreidax, D., Bannert, S., Henrich, K.-O., Schröder, C., Bender, S., Oakes, C. C., et al. (2014). p19-INK4d inhibits neuroblastoma cell growth, induces differentiation and is hypermethylated and downregulated in MYCN-amplified neuroblastomas. *Hum. Mol. Genet.* 23, 6826–6837. doi: 10.1093/hmg/ddu406

Duffy, D. J., Krstic, A., Halasz, M., Schwarzl, T., Fey, D., Iljin, K., et al. (2015). Integrative omics reveals MYCN as a global suppressor of cellular signalling and enables network-based therapeutic target discovery in neuroblastoma. *Oncotarget* 6, 43182–43201. doi: 10.18632/oncotarget.6568

Duffy, D. J., Krstic, A., Halasz, M., Schwarzl, T., Konietzny, A., Iljin, K., et al. (2017). Retinoic acid and TGF-β signalling cooperate to overcome MYCN-induced retinoid resistance. *Genome Med.* 9:15. doi: 10.1186/s13073-017-0407-3

Duffy, D. J., Krstic, A., Schwarzl, T., Halasz, M., Iljin, K., Fey, D., et al. (2016). Wnt signalling is a bi-directional vulnerability of cancer cells. *Oncotarget* 7, 60310–60331. doi: 10.18632/oncotarget.11203

Duffy, D. J., Krstic, A., Schwarzl, T., Higgins, D. G., and Kolch, W. (2014). GSK3 inhibitors regulate MYCN mRNA levels and reduce neuroblastoma cell viability through multiple mechanisms, including p53 and wnt signaling. *Mol, Cancer Ther.* 13, 454–467. doi: 10.1158/1535-7163.mct-13-0560-t

Durinck, K., and Speleman, F. (2018). Epigenetic regulation of neuroblastoma development. *Cell Tissue Res.* 372, 309–324. doi: 10.1007/s00441-017-2773-y

Ebrahimi, A., Sevinç, K., Gürhan Sevinç, G., Cribbs, A. P., Philpott, M., Uyulur, F., et al. (2019). Bromodomain inhibition of the coactivators CBP/EP300 facilitate cellular reprogramming. *Nat. Chem. Biol.* 15, 519–528. doi: 10.1038/s41589-019-0264-z

Encinas, M., Iglesias, M., Liu, Y., Wang, H., Muhaisen, A., Cena, V., et al. (2000). Sequential treatment of SH−SY5Y cells with retinoic acid and brain−derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor−dependent, human neuron−like cells. *J. Neurochem.* 75, 991–1003. doi: 10.1046/j.1471-4159.2000.0750991.x

Fabian, J., Opitz, D., Althoff, K., Lodrini, M., Hero, B., Volland, R., et al. (2016). MYCN and HDAC5 transcriptionally repress CD9 to trigger invasion and metastasis in neuroblastoma. *Oncotarget* 7, 66344–66359. doi: 10.18632/oncotarget.11662

Faiola, F., Liu, X., Lo, S., Pan, S., Zhang, K., Lymar, E., et al. (2005). Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription. *Mol. Cell. Biol.* 25, 10220–10234. doi: 10.1128/mcb.25.23.10220-10234.2005

Felgenhauer, J., Tomino, L., Selich-Anderson, J., Bopp, E., and Shah, N. (2018). Dual BRD4 and AURKA inhibition is synergistic against MYCN-Amplified and nonamplified neuroblastoma. *Neoplasia* 20, 965–974. doi: 10.1016/j.neo.2018.08.002

Gajer, J. M., Furdas, S. D., Gründer, A., Gothwal, M., Heinicke, U., Keller, K., et al. (2015). Histone acetyltransferase inhibitors block neuroblastoma cell growth in vivo. *Oncogenesis* 4:e137. doi: 10.1038/oncsis.2014.51

Göttlicher, M., Minucci, S., Zhu, P., Krämer, O. H., Schimpf, A., Giavara, S., et al. (2001). Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *EMBO J.* 20, 6969–6978. doi: 10.1093/emboj/20.24.6969

Gustafson, W. C., and Weiss, W. A. (2010). Myc proteins as therapeutic targets. *Oncogene* 29, 1249–1259. doi: 10.1038/onc.2009.512

Harenza, J. L., Diamond, M. A., Adams, R. N., Song, M. M., Davidson, H. L., Hart, L. S., et al. (2017). Transcriptomic profiling of 39 commonly-used neuroblastoma cell lines. *Sci. Data* 4:170033. doi: 10.1038/sdata.2017.33

Hay, D. A., Fedorov, O., Martin, S., Singleton, D. C., Tallant, C., Wells, C., et al. (2014). Discovery and optimization of small-molecule ligands for the CBP/p300 bromodomains. *J. Am. Chem. Soc.* 136, 9308–9319. doi: 10.1021/ja412434f

He, S., Liu, Z., Oh, D.-Y., and Thiele, C. J. (2013). MYCN and the Epigenome. *Front. Oncol.* 3:1. doi: 10.3389/fonc.2013.00001

Henrich, K.-O., Bender, S., Saadati, M., Dreidax, D., Gartlgruber, M., Shao, C., et al. (2016). Integrative genome-scale analysis identifies epigenetic mechanisms of transcriptional deregulation in unfavorable neuroblastomas. *Cancer Res.* 76, 5523–5537. doi: 10.1158/0008-5472.can-15-2507

Hohmann, A. F., and Vakoc, C. R. (2014). A rationale to target the SWI/SNF complex for cancer therapy. *Trends Genet.* 30, 356–363. doi: 10.1016/j.tig.2014.05.001

Hřebačková, J., Poljaková, J., Eckschlager, T., Hraběta, J., Procházka, P., Smutný, S., et al. (2009). Histone deacetylase inhibitors valproate and trichostatin A are toxic to neuroblastoma cells and modulate cytochrome P450 1A1, 1B1 and 3A4 expression in these cells. *Interdiscip. Toxicol.* 2, 205–210. doi: 10.2478/v10102-009-0019-x

Huang, M., and Weiss, W. A. (2013). Neuroblastoma and MYCN. *Cold Spring Harbor Perspect. Med.* 3:a014415. doi: 10.1101/cshperspect.a014415

Jin, B., Li, Y., and Robertson, K. D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* 2, 607–617. doi: 10.1177/1947601910393957

Johnsen, J. I., Dyberg, C., Fransson, S., and Wickström, M. (2018). Molecular mechanisms and therapeutic targets in neuroblastoma. *Pharmacol. Res.* 131, 164–176. doi: 10.1016/j.phrs.2018.02.023

Jubierre, L., Soriano, A., Planells-Ferrer, L., París-Coderch, L., Tenbaum, S. P., Romero, O. A., et al. (2016). BRG1/SMARCA4 is essential for neuroblastoma cell viability through modulation of cell death and survival pathways. *Oncogene* 35, 5179–5190. doi: 10.1038/onc.2016.50

Juergens, R. A., Wrangle, J., Vendetti, F. P., Murphy, S. C., Zhao, M., Coleman, B., et al. (2011). Combination epigenetic therapy has efficacy in patients with refractory advanced non–small cell lung cancer. *Cancer Discov.* 1, 598–607. doi: 10.1158/2159-8290.cd-11-0214

Khan, J., and Helman, L. J. (2016). Precision therapy for pediatric cancers. *JAMA Oncol.* 2, 575–577. doi: 10.1001/jamaoncol.2015.5685

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 30, 523–530. doi: 10.1093/bioinformatics/btt703

Kumar, P., Matthay, K. K., and Gustafson, W. C. (2019). "Chapter 14 - novel therapeutic targets in neuroblastoma," in *Neuroblastoma: Molecular Mechanisms and Therapeutic Interventions*, ed. S. K. Ray (Cambridge, MA: Academic Press), 231–261. doi: 10.1016/b978-0-12-812005-7.00014-x

Kuriyama, S., and Mayor, R. (2008). Molecular analysis of neural crest migration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1349–1362.

LePage, K. T., Dickey, R. W., Gerwick, W. H., Jester, E. L., and Murray, T. F. (2005). On the use of neuro-2a neuroblastoma cells versus intact neurons in primary culture for neurotoxicity studies. *Crit. Rev. Neurobiol.* 17, 27–50. doi: 10.1615/critrevneurobiol.v17.i1.20

Lodrini, M., Oehme, I., Schroeder, C., Milde, T., Schier, M. C., Kopp-Schneider, A., et al. (2013). MYCN and HDAC2 cooperate to repress miR-183 signaling in neuroblastoma. *Nucleic Acids Res.* 41, 6018–6033. doi: 10.1093/nar/gkt346

Loharch, S., Bhutani, I., Jain, K., Gupta, P., Sahoo, D. K., and Parkesh, R. (2015). EpiDBase: a manually curated database for small molecule modulators of epigenetic landscape. *Database* 2015:bav013.

Maddox, S. A., Watts, C. S., and Schafe, G. E. (2013). p300/CBP histone acetyltransferase activity is required for newly acquired and reactivated fear memories in the lateral amygdala. *Learn. Mem.* 20, 109–119. doi: 10.1101/lm.029157.112

Marks, P. (2007). Discovery and development of SAHA as an anticancer agent. *Oncogene* 26, 1351–1356. doi: 10.1038/sj.onc.1210204

Meany, H. J. (2019). Non-high-risk neuroblastoma: classification and achievements in therapy. *Children* 6:5. doi: 10.3390/children6010005

Miremadi, A., Oestergaard, M. Z., Pharoah, P. D., and Caldas, C. (2007). Cancer genetics of epigenetic genes. *Hum. Mol. Genet.* 16, R28–R49.

Monclair, T., Brodeur, G. M., Ambros, P. F., Brisse, H. J., Cecchetto, G., Holmes, K., et al. (2009). The international neuroblastoma risk group (INRG) staging system: an INRG task force report. *J. Clin. Oncol.* 27, 298–303. doi: 10.1200/jco.2008.16.6876

Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sanchez, R., et al. (2004). Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol. Cell* 13, 251–263. doi: 10.1016/s1097-2765(03)00528-8

Oehme, I., Deubzer, H. E., Wegener, D., Pickert, D., Linke, J.-P., Hero, B., et al. (2009). Histone deacetylase 8 in neuroblastoma tumorigenesis. *Clin. Cancer Res.* 15, 91–99. doi: 10.1158/1078-0432.ccr-08-0684

Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and cbp are histone acetyltransferases. *Cell* 87, 953–959. doi: 10.1016/S0092-8674(00)82001-2

Patel, J. H., Du, Y., Ard, P. G., Phillips, C., Carella, B., Chen, C.-J., et al. (2004). The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol. Cell. Biol.* 24, 10826–10834. doi: 10.1128/mcb.24.24.10826-10834.2004

Phimmachanh, M., Han, J. Z., O'Donnell, Y. E., Latham, S. L., and Croucher, D. R. (2020). Histone deacetylases and histone deacetylase inhibitors in neuroblastoma. *Front. Cell Dev. Biol.* 8:578770. doi: 10.3389/fcell.2020.578770

Pugh, T. J., Morozova, O., Attiyeh, E. F., Asgharzadeh, S., Wei, J. S., Auclair, D., et al. (2013). The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* 45, 279–284.

Puissant, A., Frumm, S. M., Alexe, G., Bassil, C. F., Qi, J., Chanthery, Y. H., et al. (2013). Targeting MYCN in Neuroblastoma by BET Bromodomain Inhibition. *Cancer Discov.* 3, 308–323. doi: 10.1158/2159-8290.cd-12-0418

Ram Kumar, R. M., and Schor, N. F. (2018). Methylation of DNA and chromatin as a mechanism of oncogenesis and therapeutic target in neuroblastoma. *Oncotarget* 9, 22184–22193. doi: 10.18632/oncotarget.25084

Rampersad, S. N. (2012). Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays. *Sensors* 12, 12347–12360. doi: 10.3390/s120912347

Romero, O. A., Setien, F., John, S., Gimenez−Xavier, P., Gómez−López, G., Pisano, D., et al. (2012). The tumour suppressor and chromatin−remodelling factor BRG1 antagonizes Myc activity and promotes cell differentiation in human cancer. *EMBO Mol. Med.* 4, 603–616. doi: 10.1002/emmm.201200236

Saletta, F., Wadham, C., Ziegler, D. S., Marshall, G. M., Haber, M., McCowage, G., et al. (2014). Molecular profiling of childhood cancer: biomarkers and novel therapies. *BBA Clin.* 1, 59–77. doi: 10.1016/j.bbacli.2014.06.003

Schwarzl, T., Higgins, D. G., Kolch, W., and Duffy, D. J. (2015). Measuring transcription rate changes via time-course 4-thiouridine pulse-labelling improves transcriptional target identification. *J. Mol. Biol.* 427, 3368–3374. doi: 10.1016/j.jmb.2015.09.006

SGC (2020). *Consortium Database. I-CBP112 - a CREBBP/EP300-Selective Chemical Probe*. Avaliable at: https://www.thesgc.org/chemical-probes/I-CBP112 (accessed March 15, 2021).

Sharma, S., Kelly, T. K., and Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis* 31, 27–36.

Shimada, H., Umehara, S., Monobe, Y., Hachitanda, Y., Nakagawa, A., Goto, S., et al. (2001). International neuroblastoma pathology classification for prognostic evaluation of patients with peripheral neuroblastic tumors: a report from the Children;s Cancer Group. *Cancer* 92, 2451–2461. doi: 10.1002/1097-0142(20011101)92:9<2451::aid-cncr1595>3.0.co;2-s

Singh Nanda, J., Kumar, R., and Raghava, G. P. S. (2016). dbEM: a database of epigenetic modifiers curated from cancerous and normal genomes. *Sci. Rep.* 6:19340. doi: 10.1038/srep19340

Smith, V., and Foster, J. (2018). High-risk neuroblastoma treatment review. *Children* 5:114. doi: 10.3390/children5090114

Stack, M., Walsh, P. M., Comber, H., Ryan, C. A., and O'Lorcain, P. (2007). Childhood cancer in Ireland: a population-based study. *Archiv. Dis. Childh.* 92, 890–897. doi: 10.1136/adc.2005.087544

Tumilowicz, J. J., Nichols, W. W., Cholon, J. J., and Greene, A. E. (1970). Definition of a continuous human cell line derived from neuroblastoma. *Cancer Res.* 30, 2110–2118.

Upton, K., Modi, A., Patel, K., Kendsersky, N. M., Conkrite, K. L., Sussman, R. T., et al. (2020). Epigenomic profiling of neuroblastoma cell lines. *Sci. Data* 7:116. doi: 10.1038/s41597-020-0458-y

Vedadi, M., Barsyte-Lovejoy, D., Liu, F., Rival-Gervier, S., Allali-Hassani, A., Labrie, V., et al. (2011). A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nat. Chem. Biol.* 7, 566–574. doi: 10.1038/nchembio.599

Veschi, V., Liu, Z., Voss, T. C., Ozbun, L., Gryder, B., Yan, C., et al. (2017). Epigenetic siRNA and chemical screens identify SETD8 inhibition as a therapeutic strategy for p53 activation in high-risk neuroblastoma. *Cancer Cell* 31, 50–63. doi: 10.1016/j.ccell.2016.12.002

Wainwright, E. N., and Scaffidi, P. (2017). Epigenetics and cancer stem cells: unleashing, hijacking, and restricting cellular plasticity. *Trends Cancer* 3, 372–386. doi: 10.1016/j.trecan.2017.04.004

Wang, L., Chang, J., Varghese, D., Dellinger, M., Kumar, S., Best, A. M., et al. (2013). A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth. *Nat. Commun.* 4, 1–13. doi: 10.1002/9783527678679.dg06526

Wang, L., Gural, A., Sun, X.-J., Zhao, X., Perna, F., Huang, G., et al. (2011). The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science* 333, 765–769. doi: 10.1126/science.1201662

Yang, J., AlTahan, A. M., Hu, D., Wang, Y., Cheng, P.-H., Morton, C. L., et al. (2015). The role of histone demethylase KDM4B in Myc signaling in neuroblastoma. *J. Natl. Cancer Inst.* 107:djv080. doi: 10.1093/jnci/djv080

Zhang, W., Yu, Y., Hertwig, F., Thierry-Mieg, J., Zhang, W., Thierry-Mieg, D., et al. (2015). Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biol.* 16:133. doi: 10.1186/s13059-015-0694-1

Zofall, M., Persinger, J., Kassabov, S. R., and Bartholomew, B. (2006). Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat. Struct. Mol. Biol.* 13, 339–346. doi: 10.1038/nsmb1071

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 23 April 2021
doi: 10.3389/fgene.2021.624290

The Molecular Functions of MeCP2 in Rett Syndrome Pathology

*Osman Sharifi and Dag H. Yasui**

LaSalle Laboratory, Department of Medical Microbiology and Immunology, UC Davis School of Medicine, Davis, CA, United States

MeCP2 protein, encoded by the *MECP2* gene, binds to DNA and affects transcription. Outside of this activity the true range of MeCP2 function is still not entirely clear. As *MECP2* gene mutations cause the neurodevelopmental disorder Rett syndrome in 1 in 10,000 female births, much of what is known about the biologic function of MeCP2 comes from studying human cell culture models and rodent models with *Mecp2* gene mutations. In this review, the full scope of MeCP2 research available in the NIH Pubmed (https://pubmed.ncbi.nlm.nih.gov/) data base to date is considered. While not all original research can be mentioned due to space limitations, the main aspects of MeCP2 and Rett syndrome research are discussed while highlighting the work of individual researchers and research groups. First, the primary functions of MeCP2 relevant to Rett syndrome are summarized and explored. Second, the conflicting evidence and controversies surrounding emerging aspects of MeCP2 biology are examined. Next, the most obvious gaps in MeCP2 research studies are noted. Finally, the most recent discoveries in MeCP2 and Rett syndrome research are explored with a focus on the potential and pitfalls of novel treatments and therapies.

Keywords: Rett syndrome, epigenetic, DNA, chromatin, gene expression, MeCP2

OPEN ACCESS

Edited by:
Andrei V. Chernov, University of California, San Diego, United States

Reviewed by:
Nicoletta Landsberger, University of Milan, Italy
Wei Li, University of Alabama at Birmingham, United States
Alessandra Pecorelli, North Carolina State University, United States

***Correspondence:**
Dag H. Yasui
dhyasui@ucdavis.edu

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: *31 October 2020*
Accepted: *08 February 2021*
Published: *23 April 2021*

Citation:
Sharifi O and Yasui DH (2021) The Molecular Functions of MeCP2 in Rett Syndrome Pathology. Front. Genet. 12:624290. doi: 10.3389/fgene.2021.624290

INTRODUCTION

MeCP2 research efforts have been extensive and have therefore greatly advanced the fields of epigenetics, neuroscience and chromatin research. MeCP2, encoded by the *MECP2* gene in humans and the *Mecp2* gene in rodents, was first characterized as a methyl CpG DNA binding protein in 1992, thereby establishing it as an epigenetic reader of DNA methylation (Meehan et al., 1992). It was not until the 1999 discovery that mutations in *MECP2* contribute to the pathology of the rare disease, Rett syndrome (RTT, OMIM #312750), that the relevance of MeCP2 to normal human development was established (Amir et al., 1999). Virtually all Rett patients are female (Reichow et al., 2015) and mosaic for cells with *MECP2* expression as point mutations on the X chromosome are almost always paternally inherited (Trappe et al., 2001). Despite this fact, the function of MeCP2 has been predominately studied in the nervous system of *Mecp2* null male mice as ablation of MeCP2 in all brain cells accounts for the most severe disease phenotypes (Guy et al., 2001). Since the initial characterization of MeCP2 as a DNA binding protein in 1992 (Meehan et al., 1992) a whole field of biochemical research has emerged, culminating in the discovery that MeCP2 can organize chromatin by liquid phase separation (Fan et al., 2020; Li et al., 2020; Wang et al., 2020). Since 1999 and the discovery that mutations in *MECP2* contribute to Rett syndrome, at

least 3,194 papers mentioning MeCP2 have accumulated in the NIH Pubmed database. In this review we will examine the biologic function of MeCP2 in mammals, highlight controversial aspects of MeCP2 research, point out significant gaps in knowledge, and report on paradigm shifting advances in the Rett syndrome field.

Fundamental Aspects of MeCP2 Function in Mammals

MECP2/*Mecp2* Mutations and Resulting MeCP2 Deficiencies in Brain Underlie the Majority of Rett Syndrome Like Phenotypes

The key to defining the functions of MeCP2 is to understand the effects that hypomorphic *MECP2* has at the whole organism level. Prior to the identification of *MECP2* mutations in Rett patients, a general model of *MECP2*/*MeCP2* function had been developed from early studies on the *Mecp2* gene and MeCP2 protein in mice (Lewis et al., 1992). Evidence from *in vitro* studies concluded that MeCP2 acted as a transcriptional repressor of genes in *cis* (Nan et al., 1997). In 1999 a team led by Huda Zoghbi at Baylor College of Medicine and Uta Francke at Stanford University identified the common mutations for Rett syndrome and mapped them to the *MECP2* gene located on the X chromosome (Amir et al., 1999). A Rett syndrome model with a germline *Mecp2* exon 3 and 4 deletion allele produced healthy mice at birth but male pups ($Mecp2^{-/y}$ null) soon displayed impaired motor skills and premature lethality, while female mice ($Mecp2^{-/+}$ deficient heterozygotes) became hypoactive and exhibited motor and breathing defects after 3 months of age Post mortem examination of male $Mecp2^{-/y}$ null mice revealed reduced brain and neuronal size (Guy et al., 2001). These *Mecp2* exon 3 and 4 deletion or "Bird *Mecp2* deletion" mice have been the primary animal model of Rett syndrome research. A simultaneous germline *Mecp2* exon 3 deletion mouse which will be henceforth referred to as the "Jaenisch *Mecp2* deletion" model displayed very similar motor defects and premature death as the Bird *Mecp2* null mice in males and motor, hypoactivity and respiratory phenotypes in females (Chen et al., 2001). To test the hypothesis that *Mecp2* expression is necessary for normal brain function, deletion of *Mecp2* in *nestin* expressing cells, encompassing neurons and glia, also displayed premature death and motor defects in Bird $Mecp2^{-/y}$ null deletion males and delayed motor defects in Bird $Mecp2^{-/+}$ deficient heterozygous female mice (Guy et al., 2001). Simultaneous histological examination of Jaenisch $Mecp2^{-/y}$ null deletion mice revealed reduced neuronal size along with reduced brain weight, suggestive of brain restricted defects underlying Rett like phenotypes (Chen et al., 2001). To better examine the requirement for *Mecp2* expression during brain development, the Jaenisch team also generated mice conditional for *Mecp2* deletion in *nestin* expressing neurons and glia. These mice exhibited the same phenotypes as seen in *Mecp2* germline deletion mice as did mice conditional for *Mecp2* deletion in *CamK* expressing, post-mitotic neurons that presented with delayed neurological defects in both $Mecp2^{-/y}$ null and $Mecp2^{-/+}$ deficient female mice and premature death in male mice (Chen et al., 2001). Together, these Bird and Jaenisch germline and conditional *Mecp2* deletion mouse models led to the conclusion that defects in *Mecp2* expression of MeCP2 protein during development is necessary for normal central nervous system function and a normal lifespan. Interestingly, subsequent deletion of *Mecp2* in specific neuronal subtypes revealed specific network defects but not reduced life span (Fyffe et al., 2008; Samaco et al., 2009; Chao et al., 2010) except for male mice with *Mecp2* deletion in somatostatin and parvalbumin neurons (Ito-Ishida et al., 2015). However, *MECP2* exon deletions homologous to *Mecp2* exon deletion model mice are rarely found in Rett patients and thus represent somewhat artificial genetic constructs of the disease.

The requirement for MeCP2 activity during murine development was examined by leveraging a system in which a stop codon in *Mecp2* could be removed using a systemically administered, tamoxifen induced, cre-lox deletion. Using this tool, a landmark series of experiments performed by teams from the Bird and Jaenisch labs indicated that the expression of *Mecp2* in neurons during early adult development was able to rescue motor defects, hypo-activity and premature death in $Mecp2^{-/y}$ null male mice bearing Bird and Jaenisch *Mecp2* deletion alleles (Giacometti et al., 2007; Guy et al., 2007). To further test the hypothesis that *Mecp2* expression is necessary for neurologic function after development in a whole animal, a team led by Huda Zoghbi examined the effects of the Bird *Mecp2* deletion allele in adult mice. The $Mecp2^{-/y}$ null male mice acquired motor defects, learning defects, apraxia and lethality within 15 weeks after *Mecp2* gene deletion (McGraw et al., 2011). From these collective *Mecp2* deletion mouse model studies, it is clear that MeCP2 is necessary for normal neuronal function and overall health throughout the lifespan. One caveat of these and other studies is that Bird $Mecp2^{-/y}$ null deletion males were the primary focus of these studies as they have a robust disease endpoint (death) by 20 weeks of age and present with motor defects at 6 weeks of age (Guy et al., 2001). However, the vast majority of Rett patients are female and are heterozygous for *MECP2* mutations as spontaneous *MECP2* point mutations are passed from the paternal X chromosome in sperm (Trappe et al., 2001). Female $Mecp2^{-/+}$ deficient heterozygous deletion mice are more difficult to study as life span is normal and the motor, altered anxiety and respiratory symptoms common to $Mecp2^{-/y}$ null deletion males arise after 6 months of age and are often subtle in presentation (Chen et al., 2001; Guy et al., 2001). It should be noted that maternally inherited *MECP2* mutations are extremely rare, tend to be gene duplications and present as a different disease in male patients (Del Gaudio et al., 2006).

Since the generation of the initial *Mecp2* deletion mouse models, multiple *Mecp2* gene "knock in" models that are based on actual *MECP2* mutations identified in Rett patients have been developed. The model that may have the greatest relevance to Rett is the T158A model developed by Zhou (Goffin et al., 2012) as mutation of threonine 158 is the most common *MECP2* missense mutation according to the RettBASE database (mecp2.chw.edu.au) and affects the ability of MeCP2 to bind to DNA. The non-sense *MECP2* mutation R168X correlates with severe disease symptoms in Rett. This mutation when recapitulated in *Mecp2* resulted in motor, cognitive and anxiety defects in $Mecp2^{R168X/y}$ males and females, although the time

of onset was delayed in females (Schaevitz et al., 2013). In fact, male T158A mice (*Mecp2*$^{T158A/y}$) recapitulated the motor and learning defects as well as premature mortality observed in male *Mecp2* null mice (Goffin et al., 2012) while female *Mecp2*$^{T158A/+}$ exhibit breathing abnormalities similar to those observed in Rett females (Bissonnette et al., 2014). In 2014 a mouse model based on a human *MECP2* exon 1 mutation showed motor defects similar to Bird *Mecp2* null mice along with altered anxiety and stereotypic behavior in *Mecp2-e1*$^{-/y}$ males (Yasui et al., 2014). Additional *Mecp2* knock in mouse models based on Rett *MECP2* mutations have been developed and exhibit a range of motor and behavioral phenotypes as reviewed in Schmidt and Cardoso (Schmidt et al., 2020). Rat *Mecp2* knock in models may have advantages over mouse models as a *Mecp2* truncation model exhibits neurologic regression in *Mecp2*$^{-/+}$ females and thus more accurately models Rett syndrome than some existing mouse models (Veeraragavan et al., 2015).

Mutations in *MECP2*/*Mecp2* Affect Virtually All Organs and Tissues

Although the most severe disease phenotypes were observed in *Mecp2*$^{-/y}$ null deletion male mice with deletion engineered in brain (Chen et al., 2001; Guy et al., 2001), defects in *MECP2* expression in humans has the potential to affect almost all organs and cell types, as transcripts are detected at significant levels in virtually all adult human tissues (**Figure 1**). At the systems biology level, male mice engineered to have only peripheral deletion of *Mecp2* outside of the nervous system were hypoactive, had reduced exercise capacity and bone defects, but survived beyond 50 weeks after birth (Ross et al., 2016). This study, led by Stuart Cobb, established that *Mecp2* expression in multiple organs and cell types likely accounts for these broad phenotypes thereby underlining the importance of MeCP2 function outside of the nervous system.

These limited studies indicate that while MeCP2 is necessary for normal neurologic function in mice, disease is still present in other tissues. Similarly, loss or diminished MeCP2 activity contributes to Rett syndrome pathology outside of the nervous system. For example, an early report in 1994 described extended QTc intervals in Rett patients (Sekul et al., 1994). Analysis of *Mecp2* deletion mice revealed similar cardiac abnormalities as the patients as well as ventricular tachycardia (McCauley et al., 2011). To further investigate this finding, conditional deletion of *Mecp2* from cholinergic neurons was performed by a team led by Jeff Neul. They found that loss of *Mecp2* from cholinergic, parasympathetic neurons recapitulated the previous cardiac findings in Bird *Mecp2* null mice and symptoms in Rett patients (Herrera et al., 2016). Interestingly, while *MECP2* transcripts have relatively low abundance in the human liver (**Figure 1**) an ENU mutagenesis screen in Bird *Mecp2*$^{-/y}$ null male mice revealed that MeCP2 function is required for normal lipid metabolism as these animals have dysregulated cholesterol synthesis in brain and elevated cholesterol in liver and serum (Buchovecky et al., 2013; Kyle et al., 2016; Kyle et al., 2018). *MECP2* is also expressed at high level in human intestine (**Figure 1**) and 66% of Rett patients report gastrointestinal pain (Symons et al., 2013). A 2016 study found that *Mecp2* null mice had gut hypomotility and reduced nitric oxide synthase expression in enteric neurons (Wahba et al., 2015). To examine the hypothesis that microbial alterations underlie intestinal defects, the overall gut microbiome diversity was examined and appeared to be reduced in Rett patients (Strati et al., 2016; Borghi et al., 2017). As the diets were similar between the Rett and control subjects, these results suggest that altered *MECP2* levels in the digestive tract may alter the intestinal environment and thus microbe growth. In fact, it was recently shown that mice with loss of *Mecp2* expression solely in the intestine have severe colonic epithelial defects (Millar-Büchner et al., 2016). As there is direct signaling from the gut to the brain, extension of these gut/microbiome studies should be extended to Bird *Mecp2*$^{-/+}$ deficient female deletion mice and would thus provide critical insights for female Rett patients. While studies on the role of MeCP2 function in tissues outside of the brain, a recent study found evidence that MeCP2 represses LINE-1 activity in brain, but not in peripheral tissues such as heart and eye (Zhao et al., 2019).

MeCP2 Is a DNA Binding Protein

MeCP2 was first described in the lab of Adrian Bird as a DNA binding protein with a high affinity for CpG methylated DNA (Meehan et al., 1992) via the methyl DNA binding domain (MBD) (Free et al., 2001). Although this report is often overlooked, MeCP2 was later found to be homologous to the nuclear matrix binding protein ARBP, cloned from chicken (Weitzel et al., 1997). This finding is intriguing as the nuclear matrix structures the nucleus in eukaryotic cells and organizes chromatin for replication, DNA repair and transcription among other functions (Wasag and Lenartowski, 2016). These *in vitro* binding results from the Stratling lab suggested that ARBP/MeCP2 was involved in chromatin loop organization and heterochromatin structure (Weitzel et al., 1997). Basic biochemical analyses of MeCP2 by Jeff Hansen and colleagues revealed that MeCP2 is a highly disordered protein outside of the MBD domain (Adams et al., 2007). Significantly, another study by the Hansen and Woodcock team demonstrated that a common Rett *MECP2* mutation, R106W in the MBD altered MeCP2 interaction with nucleosomal DNA (Nikitina et al., 2007). Consistent with these results, MeCP2 lacking the MBD domain was found to bind with low affinity to chromatin *in vivo* (Stuss et al., 2013). The Hansen group later determined that MeCP2 contains at least two distinct DNA binding domains, the MBD and sequences in the carboxyl terminus (Ghosh et al., 2010).

While Woodcock and colleagues had identified DNA binding regions in MeCP2 outside of the MBD, these sites were not well characterized at the time (Ghosh et al., 2010). Later however, these non-MBD binding sites were identified as AT hook regions by the Zoghbi lab in 2013 (Baker et al., 2013). The description of AT hook domains in MeCP2 fits well with an 2005 description of high affinity binding sites containing AT runs adjacent to methylated CpG sites (Klose et al., 2005). To summarize these diverse findings, a very recent paper nicely illustrates the *in vivo* activity of MeCP2 in live cells. In this study Nat Heintz and colleagues leveraged MeCP2 fluorescent tagging and real-time

visual mobility analysis to describe the sum effect of stable MBD domain interactions and transient AT hook interactions that slow the diffusion of MeCP2 from DNA in living cells (Piccolo et al., 2019).

Since the early characterization of MeCP2 as a methylated CpG binding protein (Meehan et al., 1992), the range of MeCP2 binding activity has been extended. A discovery by the Heintz lab revealed that 5-hydroxymethylcytosine (5hmc) is highly abundant in mouse brain (Kriaucionis and Heintz, 2009). This finding was confirmed by Anjana Rao and colleagues who also identified TET1 as the enzyme responsible for the conversion of 5-methylcytosine (5mc) to 5hmc in brain (Tahiliani et al., 2009). Later it was determined by the Heintz lab that MeCP2 can bind to 5hmc with similar affinity as 5mc in actively transcribed genes (Mellén et al., 2012). However, *in vitro* results had previously shown that 5hmc inhibits binding of the MBD domain of MeCP2 to DNA (Valinluck et al., 2004). Although the activity of TET1 and related factors TET2 and TET3 can each lead to active CpG demethylation, potentially altering the transcriptional state of genes (He et al., 2011) the biologic implications of MeCP2 binding to 5hmc is still not clear.

While 5mc and 5hmc in CpG dinucleotides account for the majority of MeCP2 binding sites in the mammalian genome, MeCP2 can bind to other motifs. In 2014 the Hongjun Song lab revealed that surprisingly, 25% of CpA, CpC, and CpT (CH) dinucleotides are methylated in the brain compared with 75% of CpG sites (Guo et al., 2014). The Song lab also showed that MeCP2 was able to bind to methylated CpH *in vitro* (Guo et al., 2014). The following year, Harrison Gabel and other members of the Greenberg lab determined that MeCP2 binds to methyl CpA sites within gene bodies, although the effect of this binding is controversial as described in a subsequent section of this review (Gabel et al., 2015). More recently the binding of MeCP2 was correlated with the methylation status of CH sites in the brain (Lavery et al., 2020). These results conflict with data indicating that MeCP2 binding to DNA is mostly due to CpG methylation and that MeCP2 binds to promoters with low CpG density (Baubec et al., 2013). While it is clear that MeCP2 binds to dinucleotides with methylated cytosine, an unbiased statistical analysis of genome-wide MeCP2 ChIP-seq data sets indicated that the GC content of a particular genomic region is the most predictive of MeCP2 binding (Rube et al., 2016). In contrast to these results, recent *in vitro* and *in vivo* studies indicate that MeCP2 has minimal binding to non-methylated GT rich DNA (Connelly et al., 2020). To further define MeCP2 binding, Buchmuller et al. (2020) found that MeCP2 was able to bind to DNA *in vitro* with the asymmetric cytosine modifications C/mC, mC/mC, mC/hmC, and mC/fC, where C is cytosine, mC is 5 methyl cytosine, hmC is 5 hydroxymethylcytosine and fC is 5 formylcytosine. These recent findings appear to account for the ability of MeCP2 to bind virtually anywhere in the genome. It is important to note that the level of MeCP2 protein expression is critical for normal brain function as mice with 50% of normal expression (Kerr et al., 2008; Samaco et al., 2008) and twice normal expression (Collins et al., 2004) exhibit neurological defects. Recent studies suggest that altered neuronal MeCP2 levels correlate with heterochromatin changes and behavioral symptoms (Ito-Ishida et al., 2020).

MeCP2 Is Involved in Higher Order Chromatin Organization

Another fundamental aspect of MeCP2 function that is often overlooked is its ability to regulate large chromosomal domains. For example, some of the earliest *in vitro* structural studies revealed that MeCP2 could bind and condense nucleosomal arrays that incorporate both methylated and unmethylated DNA (Georgel et al., 2003). Interestingly, these studies did not indicate that MeCP2 was able to form dimers, suggesting that one MeCP2 molecule could bridge two independent arrays. A similar chromatin condensation activity was shown in myoblast cell lines by the Cardoso lab (Brero et al., 2005). Co-immunoprecipitation studies by the El-Osta lab found that the SWI/SNF chromatin remodeling factors Brahma and BAF57 can be recruited to DNA by MeCP2 (Harikrishnan et al., 2005). While these studies examined recombinant MeCP2 linking of artificial nucleosomal constructs, a study from the Kohwi-Shigematsu lab identified a functional chromatin looping activity *in vivo*. In this study Shin-Ichi Horike and colleagues found that MeCP2 mediated long range looping and thus transcriptional activity of the *Dlx5* gene (Horike et al., 2005). In a similar study it was found that long range interactions between the Prader-Willi imprinting center (PWS-IC) and *CHRNA7* modulated gene expression in neurons (Yasui et al., 2011). These studies in cell lines were also consistent with *in vitro* studies from the Hansen lab revealing that MeCP2 was able to interact with DNA outside of the MBD domain, thereby accounting for the ability of one MeCP2 molecule to bind to two sites simultaneously in chromatin looping (Nikitina et al., 2007). This 2007 study also found that MeCP2 molecules encoded by frequent Rett *MECP2* mutations were deficient in their ability to compact chromatin (Nikitina et al., 2007). Studies from the LaSalle lab extended this chromatin looping aspect of MeCP2 function further by showing that MeCP2 along with CTCF contributes to both inter and intra chromosomal interactions and gene regulation of 15q11-13 (Meguro-Horike et al., 2011; Yasui et al., 2011). Later studies from the Natalie Berube lab showed that MeCP2 recruits ATRX to imprinted genes to regulate CTCF chromatin looping interactions (Kernohan et al., 2014). Recent atomic force microscopy (AFM) studies reveal that MeCP2 forms loops by bridging distant binding sites on continuous DNA strands (Liu et al., 2020).

MeCP2 Regulates RNA Splicing

While MeCP2 can affect gene transcription directly, one relatively unexplored activity of MeCP2 is its effect on RNA splicing. The first genome wide study of MeCP2 effects on splicing was performed in the Huda Zoghbi lab in 2005. These studies revealed significant splicing defects in *Mecp2*$^{308/y}$ male mice expressing a truncated MeCP2 protein (Young et al., 2005). *In vitro* studies revealed that MeCP2 bound to RNA along with YB-1 resulting in gene construct splicing events in neuronal cell lines (Young et al., 2005). Following this report, few attempts were made to confirm the role of MeCP2 in RNA splicing regulation. It was not until 2013 that the Keji Zhao lab reported

that *Mecp2* ablation correlated with alternatively spliced exon skipping (Maunakea et al., 2013). Research from the LaSalle lab in 2014 confirmed MeCP2 association with YB-1 shown by Young et al. (2005) as well as interaction with the RNA splicing factors MATR3, SFPQ, and SFRS1 (Yasui et al., 2014). The mechanism underlying MeCP2 regulation of splicing was further explored by the Rasko lab. They found that MeCP2 depletion reduces splicing factor recruitment to methylated DNA and increases intron retention events in blood cells (Wong et al., 2017). More recently Nurit Ballas and colleagues reported abnormal intron retention exon skipping in activation induced genes in *Mecp2* null neurons compared to wild-type neurons (Osenberg et al., 2018). The latest investigation of the role of MeCP2 in RNA splicing regulation using a machine learning approach, however, found subtle changes in RNA splicing in cells with varying levels of MeCP2 (Chhatbar et al., 2020). While the results may be not always be direct, or of great magnitude, MeCP2 binding to DNA has an effect on mRNA splicing events. However, one wonders how much current RNA-seq methodologies bias this sort of approach.

Unresolved Details of MeCP2 Function

MeCP2 Functions as a Transcriptional Repressor and Activator

The model of MeCP2 as a direct transcriptional repressor of single copy genes in *cis* derived from *in vitro* studies dating back to 1992. In 1997 this model was formally proposed in a report showing that affinity purified rat MeCP2 showed repressive activity of CpG methylated gene promoter constructs *in vitro* (Nan et al., 1997). Subsequent studies from the Alan Wolffe lab demonstrated that MeCP2 recruitment of histone deacetylase activity was correlated with transcriptional repression (Jones et al., 1998). An example of MeCP2 functioning as a transcriptional repressor is the regulation of viral elements that exist in the genome as multiple, highly CpG methylated elements. Specifically, LINE-1 elements were found to be transcriptionally repressed by MeCP2 binding (Lorincz et al., 2001; Yu et al., 2001). In 2010, a paradigm shifting report published by the Gage lab found that mammalian neuronal progenitor cells have a massive activation of LINE-1 transcription and transposition that is absent in cells from peripheral tissues (Muotri et al., 2010). More recently the Cardoso lab demonstrated in cell lines that MBD proteins including MeCP2 repress TET1 mediated 5hmc conversion and activation of LINE1 elements in human cells (Zhang et al., 2017). A recent analysis of LINE1 insertions revealed in that Rett tissues had fewer cells with exonic insertion, consistent with selection against neurons with multiple LINE1 insertion events and other cell types (Zhao et al., 2019). Thus, *in vitro* and *in vivo* evidence indicates that MeCP2 binds to CpG methylated DNA promoters and represses transcription of genes and viral elements along with co-factors such as HDACs.

The model of MeCP2 as a transcriptional repressor of single copy genes was first challenged by the genome wide correlation of MeCP2 binding with gene transcripts. In 2007, MeCP2 ChIP-seq analysis combined with gene transcript analysis from the LaSalle lab found that MeCP2 bound to active promoters with low levels of DNA methylation (Yasui et al., 2007). This was closely followed by results published in 2008 from the Zoghbi lab showing that loss of MeCP2 correlated with both reduced and elevated levels of gene transcripts in murine hypothalamus (Chahrour et al., 2008). Further analyses of select gene promoters upregulated by loss of MeCP2 activity revealed binding of the transcriptional activator CREB1 along with MeCP2 to expressed genes (Chahrour et al., 2008). In support of these findings, in 2017 it was shown that *MECP2* mutant and *MECP2* null human embryonic stem cell derived forebrain neurons had reduced CREB levels along with reduced dendritic complexity, neurite growth, and mitochondrial function (Bu et al., 2017). Most recently, results from the Harrison Gabel lab suggest that DNA methylation at enhancers may repress genes in trans (Clemens et al., 2020). Clearly these results show that MeCP2 binding and recruitment of co-factors correlates with both activation and repression of transcription. However, the location of MeCP2 binding and co-factor interactions appear to be critical for how gene expression is affected.

MeCP2e1 and MeCP2e2 Protein Isoforms Have Distinct Functions

Prior to 2004, MeCP2 was believed to be the gene product of three exons comprising one protein isoform. In that year a team from The Hospital for Sick Children in Toronto described a novel exon upstream of the known *MECP2* exons that could splice to exons 3 and 4 to produce a novel protein isoform with a distinctly different amino terminus than the originally described MeCP2 isoform (Mnatzakanian et al., 2004). The presence of this novel exon was also shown in mouse *Mecp2* by RT-PCR (Mnatzakanian et al., 2004). Around the same time, the homologous upstream exon in murine *Mecp2* was also reported by the lab of Adrian Bird (Kriaucionis and Bird, 2004). Therefore, the alternative splicing of *MECP2*/*Mecp2* exon 1 to exons 3 and 4 produces the MeCP2e1 isoform while translation of mRNA with all four exons leads to production of MeCP2e2 due to the use of an alternative translational start site (**Figure 2**). In humans and rodents this produces MeCP2 isoforms which have almost identical amino acid sequences but with distinctly different amino termini (**Figure 2**). Furthermore, it was shown that mutations in *MECP2* exon 1 are present in Rett patients (Mnatzakanian et al., 2004) and that *MECP2* exon 1 mutations disrupt translation of the MeCP2e1 but not MeCP2e2 isoforms (Gianakopoulos et al., 2012).

Evidence that MeCP2e1 is the isoform underlying Rett syndrome was provided by two key studies. In the first study, knockout of *Mecp2* exon 2 which only effects production of MeCP2e2 resulted in mice with normal development and function (Itoh et al., 2012). However, *Mecp2* exon 2 deletion did result in placental defects (Itoh et al., 2012). To follow up on these findings the Berge Minassian lab screened a cohort of patients with idiopathic RTT. In one patient they found an *MECP2* exon 1 A to T point mutation affecting the translational start site (Saunders et al., 2009). This exon 1 mutation was predicted to prevent translation of only the MeCP2e1 isoform but not the MeCP2e2 isoform, suggesting that it is the loss of MeCP2e1

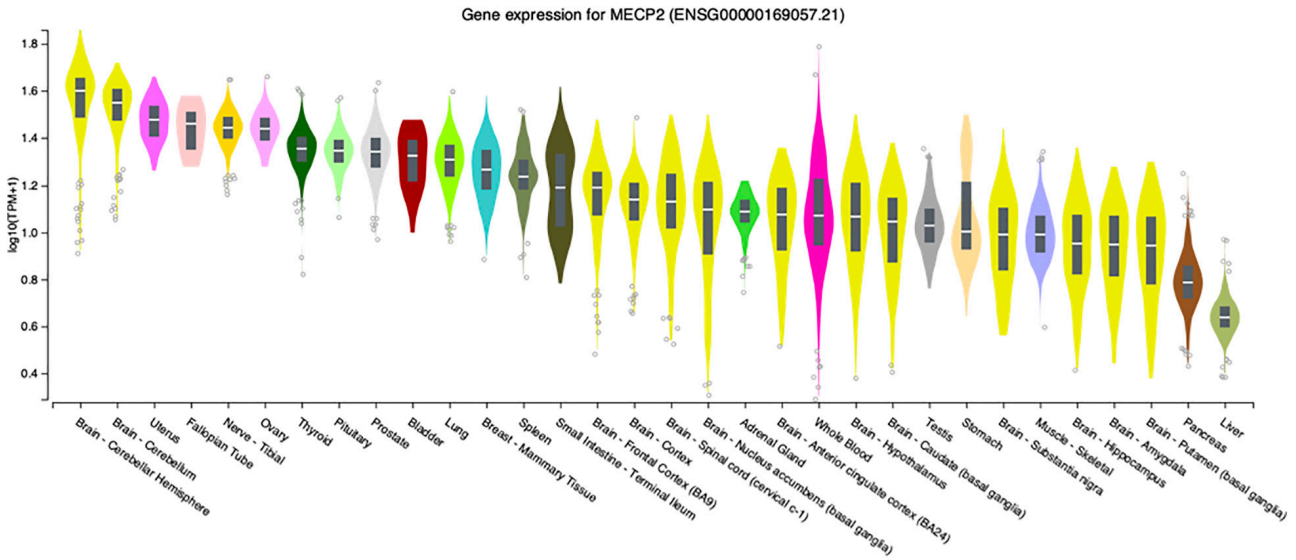

FIGURE 1 | *MECP2* transcript levels in select human tissues. Violin plots from the Genome-Tissue expression (GTEx) project are shown for *MECP2* with log2 values ranked from high to low levels.

activity that results in Rett syndrome. Therefore, the LaSalle lab replicated this mutation in a mouse model which resulted in the complete loss of the MeCP2e1 protein isoform while retaining expression of MeCP2e2 (Yasui et al., 2014). Most importantly the results from Yasui and colleagues found that *Mecp2e1*$^{-/y}$ males had impaired motor function, exhibited apraxia like limb clasping with altered anxiety behavior and premature death (Yasui et al., 2014) similar to Bird *Mecp2* deletion male mice (Guy et al., 2001). These MeCP2e1 deficient mice accurately model Rett syndrome as heterozygous *Mecp2e1*$^{-/+}$ females exhibited significant motor impairment and altered body composition (Yasui et al., 2014; Vogel Ciernia et al., 2017). Interestingly, both *Mecp2e1*$^{-/y}$ males and *Mecp2e1*$^{-/+}$ females had elevated body fat accumulation correlating with reduced energy expenditure in early adulthood (Vogel Ciernia et al., 2018).

Additional evidence that MeCP2e1 has non-overlapping function with MeCP2e2 is the finding by the Rastegar lab that MeCP2e1 is expressed in brain prior to MeCPe2 at higher levels and a more consistent distribution pattern between brain regions (Olson et al., 2014). Consistent with the mouse studies, the Ellis lab found that human MeCP2e1 deficient iPS derived neurons recapitulate synaptic current, soma size, membrane potential and maturation defects common to Rett syndrome further establishing the necessity of MeCP2e1 in neuronal development (Djuric et al., 2015). MeCP2e1 and MeCPe2 were also found by the Ausio lab to have differential DNA binding kinetics, distinct co-factor association and preferred different DNA binding motifs (Martínez De Paz et al., 2019). Perhaps the most compelling evidence that MeCP2e1 has an essential function that differs from MeCP2e2 is the finding that MeCP2e1 is the evolutionarily older isoform, as orthologs have been identified in vertebrates back to bony fish (Osteichthyes) and amphibians (Mnatzakanian et al., 2004).

Glial Cells Are Critical for Rett Syndrome Pathology

Another contentious issue in the MeCP2/Rett field is the contribution of non-neuronal cells to the disease process, despite considerable evidence that *MECP2* transcripts are present at high levels in virtually all tissues (**Figure 1**). Although cell sorting experiments show that MeCP2 is present at near octomer levels in neurons, MeCP2 is present in glia, albeit at much lower levels (Skene et al., 2010). Direct evidence for the role of non-neural cells in Rett pathology was provided by research performed by the Jin lab at UC Davis where it was found that astrocytes from Bird *Mecp2*$^{-/y}$ null male mice were abnormal and produced underdeveloped wild-type neurons compared with wild-type astrocytes in a co-culture system (Maezawa et al., 2009). A subsequent study by the Jin lab revealed that Bird *Mecp2*$^{-/y}$ null male microglia damaged adjacent neurons in a non-cell autonomous manner by releasing glutamate (Maezawa and Jin, 2010). The relevance of cell autonomous vs. non-cell autonomous effects in the *Mecp2* mutant brain is discussed further in the next section.

In support of these studies, a paper by the Mandel lab described that re-expression of *Mecp2* solely in astrocytes was able to reduce motor defects, breathing irregularity as well as prolong life in Bird *Mecp2*$^{-/y}$ null male mice (Lioy et al., 2011). Surprisingly, microglia, which comprise a tiny percentage of brain cells, were found to be important for Rett pathology when it was shown that replacement of *Mecp2* null microglia with wild-type microglia in Bird *Mecp2*$^{-/y}$ null male brain improved life span, normalized breathing and reduced motor impairment (Derecki et al., 2012). Although these results were not independently reproduced, a recent study by the Stevens lab established that microglial engulfment of synapses damages neural circuits in *Mecp2*$^{-/y}$ null male mice (Schafer et al., 2016). To further address the specific contribution of

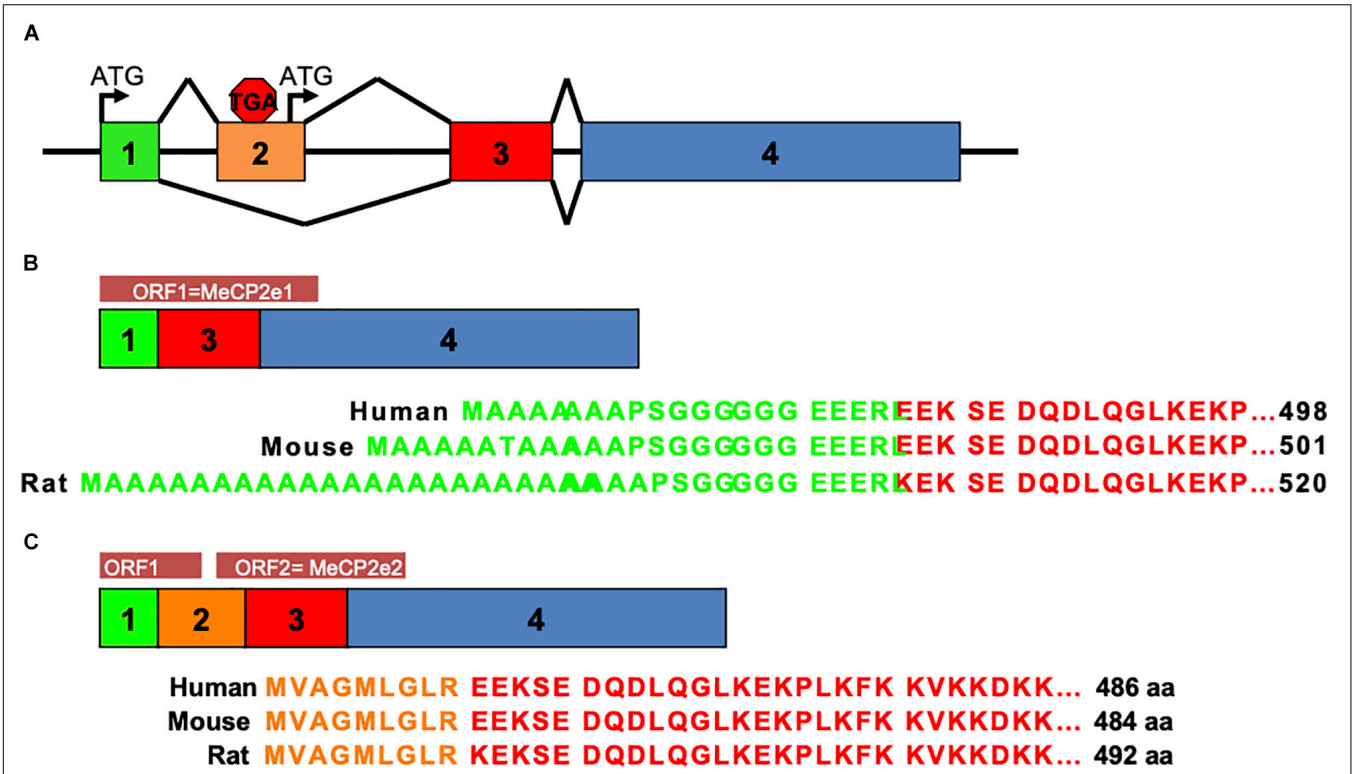

FIGURE 2 | MeCP2 protein isoforms. MeCP2e1 and MeCP2e2 protein isoforms are produced from alternative splicing events. The MeCP2e1 isoform is translated from splicing of exon 1 to exons 3 and 4 with the translational start ATG in exon 1. The MeCP2e2 protein isoform is produced from a transcript from all four exons with a translational start ATG in exon 2. Interestingly, this transcript may encode a small peptide from an ORF in exons 1 and 2. **(A)** *MECP2/Mecp2* exon splicing. **(B)** MeCP2-e1 isoform. **(C)** MeCP2-e2 isoform.

astrocytes to Rett phenotypes, Qiang Chang from the University of Wisconsin Madison derived wild-type and *MECP2* mutant iPS astrocytes and examined their function in culture with neurons. They found that *MECP2* mutant astrocytes have excess calcium stores that correlated with elevated NMDA receptor expression in neighboring neurons and hyper neuronal network excitability (Dong et al., 2018). Therefore, the authors conclude that astrocytes act both cell-autonomously and non-cell autonomously to mediate Rett like pathology (Dong et al., 2018). An extensive review on the function of MeCP2 in glia has recently become available (Kahanovitch et al., 2019).

MECP2/Mecp2 Has Both Cell Autonomous and Non-cell Autonomous Effects

The vast majority of Rett patients are female *MECP2* heterozygotes and are thus cellular mosaics for wild-type and mutant *MECP2* expressing cells. It was first reported by the LaSalle lab that Bird *Mecp2* null neurons affected the expression of adjacent wild-type neurons in *Mecp2*$^{-/+}$ heterozygous female brains (Braunschweig et al., 2004). This finding led to the "bad neighborhood" or non-cell autonomous effects hypothesis of *Mecp2* deficiency where defects in *Mecp2* mutant cells affect the function of local wild-type cells in the brain or other tissues (Braunschweig et al., 2004). This "bad neighborhood" becomes important when examining disease processes in female but not male mice where all cells are deficient in *Mecp2* expression. Subsequent studies by the Gail Mandel lab showed that MeCP2 deficient glia impair neuronal function in a non-cell autonomous manner *in vitro* (Ballas et al., 2009). As mentioned in the previous section, a similar *in vitro* culture system showed that *Mecp2*$^{-/y}$ null astrocytes were able to impair the growth of wild-type neuronal dendrites through soluble factors (Maezawa et al., 2009). To study these effects *in vivo* by transplanting either *Mecp2* mutant or wild-type neuroblasts into wild-type brain Kishi and Macklis found that MeCP2 functions largely cell autonomously in development of the cortex but that non-cell autonomous effects on neuronal function exist (Kishi and Macklis, 2010). In the most definitive study to date, the Zhou lab examined *Mecp2* mutant tagged and wild-type cells from female brains and found that non-cell autonomous effects were greater in magnitude than cell autonomous effects and tended to effect cell to cell signaling and phosphorylation (Johnson et al., 2017). These results were consistent with subsequent single cell RNA-sequencing (sc-RNAseq) analysis of neurons from *Mecp2*$^{-/+}$ deficient female brain suggesting that MeCP2 can affect gene expression in a non-cell autonomous manner in different neuronal cell types, Renthal et al. (2018).

MeCP2 Preferentially Regulates Long Genes Over Short Genes

The hypothesis that MeCP2 preferentially regulates long genes was first proposed by Sacha Nelson in 2014 (Sugino et al., 2014). This report was supported by data published from the Greenberg lab in 2015 (Gabel et al., 2015). This hypothesis was further tested in an important study mentioned previously by the Zhou lab which found a trend toward long genes being more severely affected by MeCP2 deficiency than short genes (Johnson et al., 2017). However, these findings were refuted by a re-analysis of the data and transcriptional analysis using alternative methods which together, suggested that traditional PCR methods bias the RNA-seq results toward long genes (Raman et al., 2018). The controversy over regulation of long genes is discussed further in a review by Connolly and Zhou (2019).

At What Developmental Stage Do MeCP2 Defects Impair Development

A lack of a developmental time course of how MeCP2 defects are manifested impedes Rett research. Although there is an abundance of data about MeCP2 defects in adult mice, the earliest stages of development have not been well characterized, aside from a 2003 study which found MeCP2 positive cells in E14 rat cortex (Jung et al., 2003). To address this gap, the Landsberger lab analyzed neurons from embryonic Bird *Mecp2*$^{-/y}$ null male cortex (Bedogni et al., 2016). They found that Bird *Mecp2*$^{-/y}$ null embryonic neurons have altered gene expression, altered morphology, reduced calcium flux and mobility compared to wild-type neurons (Bedogni et al., 2016). Independent evidence suggests that there are deficiencies in *Mecp2* transcripts and MeCP2 protein in the mouse cortex as early as E14 (Zachariah et al., 2012). However, one limitation of these molecular studies of early MeCP2 expression in mice is that they employed conventional methods in bulk tissue such as qRT-PCR and Western blot and that the investigators studied Bird *Mecp2*$^{-/y}$ null male mice which do accurately model Rett syndrome in *MECP2* heterozygous females. To date there has been only one study of MeCP2 embryonic function at single-cell resolution in disease relevant Bird *Mecp2*$^{-/+}$deficient mouse brain by the Greenberg group. However only adult Bird *Mecp2*$^{-/+}$ deficient deletion female mice and adult human Rett brains were analyzed (Renthal et al., 2018). Yet it is known that Rett syndrome girls experience a postnatal developmental regression at about 6–18 months of age (Zoghbi, 2005). Thus, the molecular changes during this critical early time are still not well defined. While it is clear that MeCP2 function is critical for normal neurological function, the molecular phenotypes in distinct brain cell types have not been investigated over the full-time course of development in mouse models. The Rett field is lacking an extensive, longitudinal study in mosaic *Mecp2*$^{-/+}$ deficient mice investigating the mechanisms and molecular pathways that are impacted by perturbed *Mecp2*/MeCP2 expression throughout disease progression.

MeCP2 Is an RNA Binding Protein

Coding RNAs and non-coding RNA such as long non-coding RNAs, small nucleolar RNAs, and micro RNAs establish a diverse set of functions due to their direct interactions with RNA-binding proteins (RBPs). In one of the few studies on this topic it was found that MeCP2 binds with high affinity to mRNA and siRNA outside of the MBD *in vitro* (Jeffery and Nakielny, 2004). MeCP2 was also found to associate with long non-coding RNAs and imprinted genes in mouse brain extracts (Maxwell et al., 2013). To examine MeCP2 interaction with RNA *in vivo*, MeCP2 RNA-immunoprecipitation of small RNAs was performed in the lab of Assam El-Osta who found that MeCP2 bound to specific RNAs including micro RNAs (Khan et al., 2017). These studies investigating MeCP2 interactions with RNA were performed in a targeted manner, therefore an unbiased, genome wide screen of RNA binding would provide a significant advance in understanding this MeCP2 activity.

Recent Developments in the MeCP2 Field

MeCP2 Enables Liquid Phase Separation Events in Nuclear Compartmentalization

In the last 2 years the biologic functions of MeCP2 have expanded appreciably. One concept that has emerged is the effect that liquid-liquid phase separations (LLPS) have on the formation of sub-nuclear compartments (Alberti et al., 2019). In their review Alberti, Gladfelter, and Mittag explain that the nucleus of eukaryotic cells consists of membrane less structures such as nucleoli and heterochromatin that form by LLPS through the activity of proteins and nucleic acids that condense into a dense phase (heterochromatin) and a loose phase depending on conditions such as molecular concentration, salt concentration and pH (Alberti et al., 2019). Thus, the three near simultaneous reports that MeCP2 is involved with LLPS DNA phase separation in the cell nucleus was of great relevance to the field. In the first report, Wang and colleagues found that MeCP2 was able to condense chromatin constructs *in vitro* and that DNA methylation enhanced this effect (Wang et al., 2020). Furthermore, Wang et al found that mutant MeCP2 proteins from Rett patients were defective in chromatin condensation and LLPS (Wang et al., 2020). Results from Fan and colleagues confirmed that wild-type MeCP2 was able to form condensates with DNA *in vitro* (Fan et al., 2020). Finally, research published by Richard Young and Rudolph Jaenisch at MIT also showed that MeCP2 condensed chromatin *in vitro* and MeCP2 mutant forms were defective in LLPS activity (Li et al., 2020). What set apart the Young and Jaenisch study from the other studies is the finding that the intrinsically disordered regions (IDR) of MeCP2 mediate the phase separation activity in the nucleus (Li et al., 2020). The concept of MeCP2 as a nuclear organizer via LLPS is reminiscent of early reports that an MeCP2 homolog, ARBP, functions as a component of the nuclear matrix as mentioned previously (Weitzel et al., 1997).

MeCP2 Functions With DNMT3A/Dnmt3a to Regulate Gene Expression

MeCP2 and DNMT3A, a *de novo* DNA methyltransferase have long been linked by the fact that the MeCP2 MBD preferentially binds to CpG methylated DNA (mCG). Recently it was found that there is a direct molecular link between the

two molecules as the MeCP2 TRD binds to the DNMT3A ADD domain. This interaction inhibits DNMT3A activity and this inhibition is countered by DNMT3A interaction with H3K4 through the ADD domain *in vitro* (Rajavelu et al., 2018). Similar results were shown by a collaboration between the Joe Ecker and Huda Zoghbi labs that examined the *in vivo* effects of MeCP2 and DNMT3A interaction. It had been previously established by the Ecker lab that high levels of DNA methylation at CH (mCH) sites (H = Adenosine, Cytosine, and Thymidine) in neurons, is due to the activity of DNMT3A (Lister et al., 2009). Therefore, Zoghbi and Ecker examined the hypothesis that as MeCP2 is the sole reader of methyl mCH and that DNMT3A is the sole writer of methyl CH, deletion of either factor in neurons would have similar effects on gene transcription. Surprisingly, loss of DNMT3A in neurons affected transcription of significantly more genes than the loss of MeCP2, highlighting the importance of mCH to neurologic gene regulation and function (Lavery et al., 2020). The relationship between mCH, mCG, and MeCP2 during development is explored further in a review by Lavery and Zoghbi (2019). These advances also underscore the concept that much of MeCP2 function is mediated through co-factor association.

MeCP2 May Function in DNA Repair Processes

The first evidence to suggest that MeCP2 could function in DNA repair processes was first reported in 2005. Valinluck and colleagues found that MeCP2 binds with high affinity to DNA containing halogenated pyrimidines, a modification that can result from inflammation (Valinluck et al., 2005). More recently it was found that neural stem cells from Bird *Mecp2*$^{-/+}$ deficient female mice are prone to early senescence and have elevated rates of cell death when exposed to DNA damaging H_2O_2, UV light and doxorubicin (Alessio et al., 2018). However, the most convincing evidence for the role of MeCP2 in DNA repair comes from an unbiased N-ethyl-N-nitrosourea (ENU) suppressor screen in Bird *Mecp2*$^{-/y}$ null male mice (Enikanolaiye et al., 2020). Out of 2,498 males born to ENU treated wild-type males mated with *Mecp2*$^{-/+}$ females, 96 males had ameliorated neurologic symptoms from which 32 genes which suppressed the *Mecp2* null phenotype were identified. Of these 32 genes, *Tet1*, *Birc6,* and *Spin1* function in the DNA damage response while *Rbbp8, Rad50, Fan1, Brca1,* and *Brca2* encode factors involved in DNA double strand break repair (Enikanolaiye et al., 2020).

Potential Treatments Under Development for Rett Syndrome

The remarkable reversal of Rett like phenotypes in Bird *Mecp2*$^{-/y}$ null male mice provides a basis for clinical therapies that seek to restore expression of *MECP2* (Guy et al., 2007). To this aim, the potential of adeno associated viral (AAV) vectors to deliver MeCP2 to deficient cells in a whole animal model was shown independently in 2013 by two groups. One group led by Stuart Cobb showed that AAV9/*MECP2* delivered intravenously in pre-symptomatic Bird *Mecp2*$^{-/y}$ null male mice extended survival modestly (Gadalla et al., 2013). Gail Mandel led a second study that showed that intravenous AAV9/MeCP2e1 vector reduced motor defects in Bird *Mecp2*$^{-/+}$ deficient female mice, which thus had greater relevance to Rett syndrome as 95% of Rett patients are female (Garg et al., 2013). In 2017, after these early proofs of principle, advances in *AAV9/MECP2* gene therapy were reported by the combined efforts of Stuart Cobb and Steven J. Gray. The authors in two manuscripts reported improvement in cell transduction with a reduced titer of a second generation AAV9/hMeCP2 virus, a reduction in liver toxicity and prolonged survival in Bird *Mecp2*$^{-/y}$ null mice by delivery through the cerebral spinal fluid (CSF) or directly into brain (Gadalla et al., 2017; Sinnett et al., 2017). These studies, while they represent significant improvements in *MECP2* gene therapy tools, should be performed in Bird *Mecp2*$^{-/+}$ deficient heterozygous female deletion mice to be more relevant to Rett. Furthermore, the risks of treating human children with AAV vectors cannot be fully mitigated.

Alternatively, the use of site directed RNA editing is being developed for potential Rett syndrome therapy. RNA editing strategies leverage an Adenosine Deaminase Acting on RNA (ADAR2) RNA editing factors with a guide RNA to repair *Mecp2/MECP2* point mutations in RNA. The first report on *Mecp2* RNA editing was published by the Gail Mandel lab in 2017 where an exogenous, modified ADAR2 and guide was used to successfully edit 72% G to A mutations in RNA from *Mecp*^{R106Q} mouse neurons (Sinnamon et al., 2017). In 2020 Sinnamon and colleagues successfully edited up to 50% of mRNA in neurons in developing brains of *Mecp2* mutant mice using a similar approach (Sinnamon et al., 2020). The potential of RNA editing as a Rett therapy is limited by the fact that although endogenous RNA editing proteins such as ADAR1, exist naturally in the brain, current approaches express a modified RNA editing protein delivered by an AAV virus and the fact that these ADAR2 like proteins primarily edit only G to A mutations (Sinnamon et al., 2017, 2020).

Despite the excitement for *MECP2* gene replacement and RNA editing strategies, other therapeutic methods are in development. One strategy explored by the Zoghbi lab in Rett model mice is deep brain stimulation. Electrode based stimulation of symptomatic *Mecp2*$^{-/+}$ female mice normalized learning and memory defects (Hao et al., 2015; Pohodich et al., 2018). Another promising area for potential Rett therapies, is the repurposing of existing FDA approved compounds. For example, in 2020, it was found that the diabetes drug, metformin, reduces mitochondrial defects and damage from oxidative stress in *Mecp2 308*$^{-/+}$ truncation female mice (Zuliani et al., 2020). Finally, a recently developed compound for Alzheimer's disease reduced motor defects, learning deficits and breathing abnormalities in *Mecp2*$^{-/+}$ female mice with minimal side effects (Kaufmann et al., 2019). Until effective and safe *MECP2* gene replacement and/or RNA editing strategies are available, existing or novel chemical compounds could be used to reduce the most severe symptoms in Rett patients.

DISCUSSION

Currently, there is debate as to whether patients with overlapping phenotypes have Rett syndrome or some other disease. For example, patients with mutations in *CDKL5* were considered to have a severe variant of Rett, however, this was later categorized as a separate disease (Fehr et al., 2013). Patients with *FOXG1* mutations are currently classified with Rett. However, as these rare *FOXG1* mutation patients present with disease from birth and have other symptoms that do not overlap with Rett, a proposal to classify these patients as a separate syndrome was recently submitted (Cutri-French et al., 2020). Therefore, with few exceptions, any discussion about Rett syndrome phenotypes is about how defective *MECP2/Mecp2* expression presents in patients and animal models and manifests as a neurodevelopmental disease.

At this time it is still unclear how *MECP2* expression defects in non-neural tissues contribute to Rett phenotypes. The answer may be complex as illustrated by a report from the Landsberger lab showing that male Bird *Mecp2* deletion mice have abnormal muscle tissue while deletion of the Bird *Mecp2* allele only in muscle resulted in normal muscle tissue (Conti et al., 2015). The conclusion from this study is that non-cell-autonomous signaling to the muscle by defective neurons in the Bird *Mecp2* deletion mice contributes to the muscle defect (Conti et al., 2015). So, there is the possibility that although *MECP2* transcripts are abundant in all tissues and cell types, extracellular neuronal signaling could be responsible for some, if not all of the defects.

The recent identification of MeCP2 binding beyond methylated CpG sites has greatly expanded the range of potential functions of the protein. These advances have been driven by the discovery of additional nucleotide modifications in mammalian brain. Yet many questions remain. For example, does MeCP2 binding to 5-hmc (Kriaucionis and Heintz, 2009) in cerebellum have a significant neurologic function (Mellén et al., 2012)? Do MeCP2 isoforms have different 5-hmc binding activity? Some data suggests that MeCP2 protects 5mc from TET1 conversion to 5 hmc (Szulwach et al., 2011) but more study is needed to resolve this important question. The 2014 discovery of DNMT3A dependent CH methylation in mouse brain by the Song lab (Guo et al., 2014) and the description of MeCP2 binding to methyl CpA sites in brain *in vivo* (Gabel et al., 2015) has been suggested to be key to the pathogenesis of Rett (Lavery et al., 2020). Here again additional research is needed to clearly establish the effect of MeCP2 binding outside of methyl CpG sites.

Another evolving area of research is how MeCP2 functions as a transcriptional repressor and activator. The model of MeCP2 as a transcriptional repressor of single copy genes was developed over twenty years ago (Nan et al., 1997; Jones et al., 1998). Since that time it was found that MeCP2 binds to partially CpG methylated promoters of active genes in neurons (Yasui et al., 2007) and can act as a transcriptional activator by recruitment of CREB in neurons (Chahrour et al., 2008). The latest evidence suggests that MeCP2 modulates gene transcription depending on co-factor interaction. Future gene expression analyses should focus on the differential activity of the MeCP2-e1 and MeCP2-e2 isoforms.

The role of glial cell defects in Rett syndrome pathology is still being described. However, the very fact that glial cells express *MECP2/Mecp2* and are defective in Rett patients and *Mecp2* null and deficient mice is relevant to neuronal disease phenotypes. As Rett females and *Mecp2*$^{-/+}$ deficient heterozygous female mice are mosaic for normal and mutant cells, non-cell autonomous effects are clearly contributing to neuronal dysfunction (Johnson et al., 2017). Analysis of neurons from mosaic Bird *Mecp2*$^{-/+}$ deficient female mouse brain and human Rett brain suggest that MeCP2 preferentially represses methylated long genes, Sugino et al. (2014), in a cell autonomous manner in neurons, although non-cell autonomous effects on short genes cannot be excluded (Renthal et al., 2018). As recent study concluded that MeCP2 bias toward long gene regulation may be an artifact of PCR based transcript detection (Raman et al., 2018), single cell analyses using alternatives to current RNA-seq methods in brain cell types will be necessary to resolve the role of non-cell autonomous effects in brains mosaic for wild-type and *Mecp2* mutant neurons and glia.

Although it is widely assumed that Rett patients are phenotypically normal at birth this conclusion is based on limited behavioral observation. As there is little incentive and ethical limitations preclude studying apparently normal human infants, a full developmental time course of the disease is needed in Rett model *Mecp2*$^{-/+}$ heterozygous mice that more accurately model Rett *MECP2* mutations such *Mecp2-e1* mutant mice (Yasui et al., 2014). Surprisingly, apart from a study that revealed defects in Bird *Mecp2*$^{-/y}$ null male embryonic cells, little analysis has been done of early developmental time points (Bedogni et al., 2016). Clearly, a developmental time analyses of both *Mecp2*$^{-/+}$ females and *Mecp2*$^{-/y}$ males at embryonic, pre-disease, early disease, and late disease is needed to identify pathological mechanisms.

There are aspects of MeCP2 function that are oddly neglected. For example, surprisingly, few studies have directly examined the ability of MeCP2 to bind to RNA. The studies that do show a strong *in vitro* interaction of MeCP2 with mRNA and siRNA via the RG domain between the MBD and TRD (Jeffery and Nakielny, 2004) and *in vivo* interaction with miRNAs (Khan et al., 2017). A logical progression of these studies would be to examine RNA binding ability of common Rett mutant MeCP2 proteins such as T158M *in vitro* and *in vivo*. The Rett field is lacking studies using more unbiased high-through methods such as targets of RNA-binding proteins identified by editing (TRIBE) (McMahon et al., 2016) which fuses MeCP2 to ADAR, an enzyme that modifies RNA where MeCP2 binds. An additional advantage of this approach is that it can be used for *in vivo* studies.

Finally, unexpected discoveries continue to revitalize the study of MeCP2. Recent cell biology studies reveal that compartmentalization of the nucleus by liquid-liquid phase separation forms critical regulatory compartments such as the nucleolus and heterochromatin (reviewed in Alberti et al., 2019). Now it appears that MeCP2 plays a key role in the formation of nuclear heterochromatin domains according to new *in vitro*

studies (Fan et al., 2020; Wang et al., 2020) and that the MeCP2 intrinsically disorder region plays a key role in this process (Li et al., 2020). These results build on the extensive study of how MeCP2 organizes chromatin in neuronal cells and may provide key insights into Rett syndrome defects.

The most interesting advances in MeCP2 research concern Rett therapies, as restoration of MeCP2 in mouse brain led to disease reversal in a mouse model (Guy et al., 2007). Current *MECP2* gene therapy for Rett patients has great potential and some risks. The latest gene therapy treatment of Bird *Mecp2*$^{-/+}$ heterozygous mice with AAV9/*MECP2* was able to normalize breathing but high viral doses produced severe liver toxicity and death in some animals (Matagne et al., 2020). While RNA editing strategies for repairing mutant MECP2 have great potential as a therapy, delivery of editing proteins to the brain is also as problematic as delivering functional MeCP2. Therefore, considerable time and effort will be needed to refine *MECP2* gene replacement and RNA editing based therapies. Until then, alternative treatments are needed to treat Rett.

One potential therapy for Rett is deep brain stimulation (DSB) which has been approved by the FDA for epilepsy. Nonetheless, DSB also has its risks as electrodes are implanted into the brain. Existing drugs approved by the FDA for other diseases are being re-directed to treat Rett symptoms. Clinical trials of IGF-1 and IGF-1 analogs that have been approved for short stature are ongoing for Rett and offer some hope of reducing symptoms (Khwaja et al., 2014). Another compound, ANAVEX 2-73 which targets the sigma-1 receptor that affects learning, memory, and neuronal development and is in phase II clinical trials for Alzheimer's disease is also scheduled for Phase I trials in Rett girls (Kaufmann et al., 2019). In some cases, effective treatment for a specific disease have been employed without fully understanding the mechanisms of the disease process. While this may be the case with Rett syndrome, that should not preclude the continuing research on the evolving molecular function of MeCP2 in neurons and other cell types.

AUTHOR CONTRIBUTIONS

DY and OS conceived of, planned and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the NIH NIAA Grant 1RO1AA027075 Neuroimmune interactions in Rett syndrome to Janine M. LaSalle.

REFERENCES

Adams, V. H., McBryant, S. J., Wade, P. A., Woodcock, C. L., and Hansen, J. C. (2007). Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *J. Biol. Chem.* 282, 15057–15064. doi: 10.1074/jbc.M700855200

Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* 176, 419–434. doi: 10.1016/j.cell.2018.12.035

Alessio, N., Riccitiello, F., Squillaro, T., Capasso, S., Del Gaudio, S., Di Bernardo, G., et al. (2018). Neural stem cells from a mouse model of Rett syndrome are prone to senescence, show reduced capacity to cope with genotoxic stress, and are impaired in the differentiation process. *Exper. Mol. Med.* 50:1. doi: 10.1038/s12276-017-0005-x

Amir, R. E., Van Den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. *Nat. Genet.* 23, 185–188. doi: 10.1038/13810

Baker, S. A., Chen, L., Wilkins, A. D., Yu, P., Lichtarge, O., and Zoghbi, H. Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell* 152, 984–996. doi: 10.1016/j.cell.2013.01.038

Ballas, N., Lioy, D. T., Grunseich, C., and Mandel, G. (2009). Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. *Nat. Neurosci.* 12, 311–317. doi: 10.1038/nn.2275

Baubec, T., Ivánek, R., Lienert, F., and Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the mbd protein family. *Cell* 153, 480–492. doi: 10.1016/j.cell.2013.03.011

Bedogni, F., Gigli, C. C., Pozzi, D., Rossi, R. L., Scaramuzza, L., Rossetti, G., et al. (2016). Defects during Mecp2 null embryonic cortex development precede the onset of overt neurological symptoms. *Cereb. Cortex* 26, 2517–2529. doi: 10.1093/cercor/bhv078

Bissonnette, J. M., Schaevitz, L. R., Knopp, S. J., and Zhou, Z. (2014). Respiratory phenotypes are distinctly affected in mice with common Rett syndrome mutations MECP2 T158A and R168X. *Neuroscience* 267, 166–176. doi: 10.1016/j.neuroscience.2014.02.043

Borghi, E., Borgo, F., Severgnini, M., Savini, M. N., Casiraghi, M. C., and Vignoli, A. (2017). Rett syndrome: a focus on gut microbiota. *Intern. J. Mol. Sci.* 18:344. doi: 10.3390/ijms18020344

Braunschweig, D., Simcox, T., Samaco, R. C., and LaSalle, J. M. (2004). X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and Mecp2-/+ mouse brain. *Hum. Mol. Genet.* 13, 1275–1286. doi: 10.1093/hmg/ddh142

Brero, A., Easwaran, H. P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., et al. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J. Cell Biol.* 169, 733–743. doi: 10.1083/jcb.200502062

Bu, Q., Wang, A., Hamzah, H., Waldman, A., Jiang, K., Dong, Q., et al. (2017). CREB signaling is involved in rett syndrome pathogenesis. *J. Neurosci.* 37, 3671–3685. doi: 10.1523/JNEUROSCI.3735-16.2017

Buchmuller, B. C., Kosel, B., and Summerer, D. (2020). Complete profiling of Methyl-CpG-binding domains for combinations of cytosine modifications at CpG dinucleotides reveals differential read-out in normal and Rett-associated states. *Sci. Rep.* 10:4053. doi: 10.1038/s41598-020-61030-1

Buchovecky, C. M., Turley, S. D., Brown, H. M., Kyle, S. M., McDonald, J. G., Liu, B., et al. (2013). A suppressor screen in Mecp2 mutant mice implicates cholesterol metabolism in Rett syndrome. *Nat. Genet.* 45, 1013–1020. doi: 10.1038/ng.2714

Chahrour, M., Sung, Y. J., Shaw, C., Zhou, X., Wong, S. T. C., Qin, J., et al. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* 320, 1224–1229. doi: 10.1126/science.1153252

Chao, H. T., Chen, H., Samaco, R. C., Xue, M., Chahrour, M., Yoo, J., et al. (2010). Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* 468, 263–269. doi: 10.1038/nature09582

Chen, R. Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* 27, 327–331. doi: 10.1038/85906

Chhatbar, K., Cholewa-Waclaw, J., Shah, R., Bird, A., and Sanguinetti, G. (2020). Quantitative analysis questions the role of MeCP2 as a global regulator of alternative splicing. *PLoS Genet.* 16:1009087. doi: 10.1371/journal.pgen.1009087

Clemens, A. W., Wu, D. Y., Moore, J. R., Christian, D. L., Zhao, G., and Gabel, H. W. (2020). MeCP2 represses enhancers through chromosome topology-associated DNA methylation. *Mol. Cell* 77, 279–293.e8. doi: 10.1016/j.molcel.2019.10.033

Collins, A. L., Levenson, J. M., Vilaythong, A. P., Richman, A., Armstrong, A. L., Noebels, J. L., et al. (2004). Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* 13, 2679–2689. doi: 10.1093/hmg/ddh282

Connelly, J. C., Cholewa, J., Webb, S., Steccanella, V., Waclaw, B., and Bird, A. (2020). Absence of MeCP2 binding to non-methylated GT sequences *in vivo*. *Nucleic Acids Research* 48, 3542–3552. doi: 10.1093/nar/gkaa102

Connolly, D. R., and Zhou, Z. (2019). Genomic insights into MeCP2 function: a role for the maintenance of chromatin architecture. *Curr. Opin. Neurobiol.* 59, 174–179. doi: 10.1016/j.conb.2019.07.002

Conti, V., Gandaglia, A., Galli, F., Tirone, M., Bellini, E., Campana, L., et al. (2015). MeCP2 affects skeletal muscle growth and morphology through non cell-autonomous mechanisms. *PLoS One* 10:e0130183. doi: 10.1371/journal.pone.0130183

Cutri-French, C., Armstrong, D., Saby, J., Gorman, C., Lane, J., Fu, C., et al. (2020). Comparison of core features in four developmental encephalopathies in the Rett natural history study. *Ann. Neurol.* 88, 396–406. doi: 10.1002/ana.25797

Del Gaudio, D., Fang, P., Scaglia, F., Ward, P. A., Craigen, W. J., Glaze, D. G., et al. (2006). Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genet. Med.* 8, 784–792. doi: 10.1097/01.gim.0000250502.28516.3c

Derecki, N. C., Cronk, J. C., Lu, Z., Xu, E., Abbott, S. B. G., Guyenet, P. G., et al. (2012). Wild-type microglia arrest pathology in a mouse model of Rett syndrome. *Nature* 484, 105–109. doi: 10.1038/nature10907

Djuric, U., Cheung, A. Y. L., Zhang, W., Mok, R. S., Lai, W., Piekna, A., et al. (2015). MECP2e1 isoform mutation affects the form and function of neurons derived from Rett syndrome patient iPS cells. *Neurobiol. Dis.* 76, 37–45. doi: 10.1016/j.nbd.2015.01.001

Dong, Q., Liu, Q., Li, R., Wang, A., Bu, Q., Wang, K. H., et al. (2018). Mechanism and consequence of abnormal calcium homeostasis in rett syndrome astrocytes. *eLife* 7:33417. doi: 10.7554/eLife.33417

Enikanolaiye, A., Ruston, J., Zeng, R., Taylor, C., Schrock, M., Buchovecky, C. M., et al. (2020). Suppressor mutations in Mecp2-null mice implicate the DNA damage response in Rett syndrome pathology. *Genome Res.* 30, 540–552. doi: 10.1101/gr.258400.119

Fan, C., Zhang, H., Fu, L., Li, Y., Du, Y., Qiu, Z., et al. (2020). Rett mutations attenuate phase separation of MeCP2. *Cell Discov.* 6:38. doi: 10.1038/s41421-020-0172-0

Fehr, S., Wilson, M., Downs, J., Williams, S., Murgia, A., Sartori, S., et al. (2013). The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *Eur. J. Hum. Genet.* 21, 266–273. doi: 10.1038/ejhg.2012.156

Free, A., Wakefield, R. I. D., Smith, B. O., Drydenll, D. T. F., Barlow, P. N., and Bird, A. P. (2001). DNA Recognition by the Methyl-CpG binding domain of MeCP2. *J. Biol. Chem.* 276, 3353–3360. doi: 10.1074/jbc.M007224200

Fyffe, S. L., Neul, J. L., Samaco, R. C., Chao, H. T., Ben-Shachar, S., Moretti, P., et al. (2008). Deletion of Mecp2 in Sim1-expressing neurons reveals a critical role for MeCP2 in feeding behavior, aggression, and the response to stress. *Neuron* 59, 947–958. doi: 10.1016/j.neuron.2008.07.030

Gabel, H. W., Kinde, B., Stroud, H., Gilbert, C. S., Harmin, D. A., Kastan, N. R., et al. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* 522, 89–93. doi: 10.1038/nature14319

Gadalla, K. K., Bailey, M. E., Spike, R. C., Ross, P. D., Woodard, K. T., Kalburgi, S. N., et al. (2013). Improved survival and reduced phenotypic severity following AAV9/MECP2 gene transfer to neonatal and juvenile male Mecp2 knockout mice. *Mol. Therapy* 21, 18–30. doi: 10.1038/mt.2012.200

Gadalla, K. K. E., Vudhironarit, T., Hector, R. D., Sinnett, S., Bahey, N. G., Bailey, M. E. S., et al. (2017). Development of a Novel AAV Gene therapy cassette with improved safety features and efficacy in a mouse model of Rett syndrome. *Mol. Ther. Methods Clin. Dev.* 5, 180–190. doi: 10.1016/j.omtm.2017.04.007

Garg, S. K., Lioy, D. T., Cheval, H., McGann, J. C., Bissonnette, J. M., Murtha, M. J., et al. (2013). Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *J. Neurosci.* 33, 13612–13620. doi: 10.1523/JNEUROSCI.1854-13.2013

Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A., and Hansen, J. C. (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.* 278, 32181–32188. doi: 10.1074/jbc.M305308200

Ghosh, R. P., Nikitina, T., Horowitz-Scherer, R. A., Gierasch, L. M., Uversky, V. N., Hite, K., et al. (2010). Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochemistry* 49, 4395–4410. doi: 10.1021/bi9019753

Giacometti, E., Luikenhuis, S., Beard, C., and Jaenisch, R. (2007). Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1931–1936. doi: 10.1073/pnas.0610593104

Gianakopoulos, P. J., Zhang, Y., Pencea, N., Orlic-Milacic, M., Mittal, K., Windpassinger, C., et al. (2012). Mutations in *MECP2* exon 1 in classical rett patients disrupt *MECP2_e1* transcription, but not transcription of *MECP2_e2*. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 159B, 210–216. doi: 10.1002/ajmg.b.32015

Goffin, D., Allen, M., Zhang, L., Amorim, M., Wang, I. T. J., Reyes, A. R. S., et al. (2012). Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. *Nat. Neurosci.* 15, 274–283. doi: 10.1038/nn.2997

Guo, J. U., Su, Y., Shin, J. H., Shin, J., Li, H., Xie, B., et al. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat. Neurosci.* 17, 215–222. doi: 10.1038/nn.3607

Guy, J., Gan, J., Selfridge, J., Cobb, S., and Bird, A. (2007). Reversal of neurological defects in a mouse model of Rett syndrome. *Science* 315, 1143–1147. doi: 10.1126/science.1138389

Guy, J., Hendrich, B., Holmes, M., Martin, J. E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic rett syndrome. *Nat. Genet.* 27, 322–326. doi: 10.1038/85899

Hao, S., Tang, B., Wu, Z., Ure, K., Sun, Y., Tao, H., et al. (2015). Forniceal deep brain stimulation rescues hippocampal memory in Rett syndrome mice. *Nature* 526, 430–434. doi: 10.1038/nature15694

Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., et al. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat. Genet.* 37, 254–264. doi: 10.1038/ng1516

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307. doi: 10.1126/science.1210944

Herrera, J. A., Ward, C. S., Wehrens, X. H. T., and Neul, J. L. (2016). Methyl-CpG binding-protein 2 function in cholinergic neurons mediates cardiac arrhythmogenesis. *Hum. Mol. Genet.* 25, 4983–4995. doi: 10.1093/hmg/ddw326

Horike, S. I., Cai, S., Miyano, M., Cheng, J. F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* 37, 31–40. doi: 10.1038/ng1491

Itoh, M., Tahimic, C. G. T., Ide, S., Otsuki, A., Sasaoka, T., Noguchi, S., et al. (2012). Methyl CpG-binding protein isoform MeCP2-e2 is dispensable for rett syndrome phenotypes but essential for embryo viability and placenta development. *J. Biol. Chem.* 287, 13859–13867. doi: 10.1074/jbc.M111.309864

Ito-Ishida, A., Baker, S. A., Sillitoe, R. V., Sun, Y., Zhou, J., Ono, Y., et al. (2020). MeCP2 levels regulate the 3d structure of heterochromatic foci in mouse neurons. *J. Neurosci.* 40, 8746–8766. doi: 10.1523/JNEUROSCI.1281-19.2020

Ito-Ishida, A., Ure, K., Chen, H., Swann, J. W., and Zoghbi, H. Y. (2015). Loss of MeCP2 in Parvalbumin-and Somatostatin-expressing neurons in mice leads to distinct rett syndrome-like phenotypes. *Neuron* 88, 651–658. doi: 10.1016/j.neuron.2015.10.029

Jeffery, L., and Nakielny, S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. *J. Biol. Chem.* 279, 49479–49487. doi: 10.1074/jbc.M409070200

Johnson, B. S., Zhao, Y. T., Fasolino, M., Lamonica, J. M., Kim, Y. J., Georgakilas, G., et al. (2017). Biotin tagging of MeCP2 in mice reveals contextual insights into the Rett syndrome transcriptome. *Nat. Med.* 23, 1203–1214. doi: 10.1038/nm.4406

Jones, P. L., Veenstra, G. J. C., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., et al. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* 19, 187–191. doi: 10.1038/561

Jung, B. P., Jugloff, D. G. M., Zhang, G., Logan, R., Brown, S., and Eubanks, J. H. (2003). The expression of methyl CpG binding factor MeCP2 correlates

with cellular differentiation in the developing rat brain and in cultured cells. *J. Neurobiol.* 55, 86–96. doi: 10.1002/neu.10201

Kahanovitch, U., Patterson, K. C., Hernandez, R., and Olsen, M. L. (2019). Glial dysfunction in meCP2 deficiency models: implications for rett syndrome. *Intern. J. Mol. Sci.* 20:3813. doi: 10.3390/ijms20153813

Kaufmann, W. E., Sprouse, J., Rebowe, N., Hanania, T., Klamer, D., and Missling, C. U. (2019). ANAVEX® 2-73 (blarcamesine), a Sigma-1 receptor agonist, ameliorates neurologic impairments in a mouse model of Rett syndrome. *Pharmacol. Biochem. Behav.* 187:172796. doi: 10.1016/j.pbb.2019.172796

Kernohan, K. D., Vernimmen, D., Gloor, G. B., and Bérubé, N. G. (2014). Analysis of neonatal brain lacking ATRX or *MeCP2* reveals changes in nucleosome density, CTCF binding and chromatin looping. *Nucleic Acids Res.* 42, 8356–8368. doi: 10.1093/nar/gku564

Kerr, B., Alvarez-saavedra, M., Sáez, M. A., Saona, A., and Young, J. I. (2008). Defective body-weight regulation, motor control and abnormal social interactions in Mecp2 hypomorphic mice. *Hum. Mol. Genet.* 17, 1707–1717. doi: 10.1093/hmg/ddn061

Khan, A. W., Ziemann, M., Rafehi, H., Maxwell, S., Ciccotosto, G. D., and El-Osta, A. (2017). MeCP2 interacts with chromosomal microRNAs in brain. *Epigenetics* 12, 1028–1037. doi: 10.1080/15592294.2017.1391429

Khwaja, O. S., Ho, E., Barnes, K. V., O'Leary, H. M., Pereira, L. M., Finkelstein, Y., et al. (2014). Safety, pharmacokinetics, and preliminary assessment of efficacy of mecasermin (recombinant human IGF-1) for the treatment of Rett syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4596–4601. doi: 10.1073/pnas.1311141111

Kishi, N., and Macklis, J. D. (2010). MeCP2 functions largely cell-autonomously, but also non-cell-autonomously, in neuronal maturation and dendritic arborization of cortical pyramidal neurons. *Exper. Neurol.* 222, 51–58. doi: 10.1016/j.expneurol.2009.12.007

Klose, R. J., Sarraf, S. A., Schmiedeberg, L., McDermott, S. M., Stancheva, I., and Bird, A. P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell* 19, 667–678. doi: 10.1016/j.molcel.2005.07.021

Kriaucionis, S., and Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res.* 32, 1818–1823. doi: 10.1093/nar/gkh349

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science* 324, 929–930. doi: 10.1126/science.1169786

Kyle, S. M., Saha, P. K., Brown, H. M., Chan, L. C., and Justice, M. J. (2016). MeCP2 co-ordinates liver lipid metabolism with the NCoR1/HDAC3 corepressor complex. *Hum. Mol. Genet.* 25, 3029–3041. doi: 10.1093/hmg/ddw156

Kyle, S. M., Vashi, N., and Justice, M. J. (2018). Rett syndrome: a neurological disorder with metabolic components. *Open Biol.* 8:170216. doi: 10.1098/rsob.170216

Lavery, L. A., Ure, K., Wan, Y. W., Luo, C., Trostle, A. J., Wang, W., et al. (2020). Losing dnmt3a dependent methylation in inhibitory neurons impairs neural function by a mechanism impacting rett syndrome. *eLife* 9:52981. doi: 10.7554/eLife.52981

Lavery, L. A., and Zoghbi, H. Y. (2019). The distinct methylation landscape of maturing neurons and its role in Rett syndrome pathogenesis. *Curr. Opin. Neurobiol.* 59, 180–188. doi: 10.1016/j.conb.2019.08.001

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., et al. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to Methylated DNA. *Cell* 69, 905–914. doi: 10.1016/0092-8674(92)90610-O

Li, C. H., Coffey, E. L., Dall'Agnese, A., Hannett, N. M., Tang, X., Henninger, J. E., et al. (2020). MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature* 586, 440–444. doi: 10.1038/s41586-020-2574-4

Lioy, D. T., Garg, S. K., Monaghan, C. E., Raber, J., Foust, K. D., Kaspar, B. K., et al. (2011). A role for glia in the progression of Rett-syndrome. *Nature* 475, 497–500. doi: 10.1038/nature10214

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322. doi: 10.1038/nature08514

Liu, M., Movahed, S., Dangi, S., Pan, H., Kaur, P., Bilinovich, S. M., et al. (2020). DNA looping by two 5-methylcytosine-binding proteins quantified using nanofluidic devices. *Epigenet. Chrom.* 13:18. doi: 10.1186/s13072-020-00339-7

Lorincz, M. C., Schübeler, D., and Groudine, M. (2001). Methylation-mediated proviral silencing is associated with MeCP2 recruitment and localized Histone H3 deacetylation. *Mol. Cell. Biol.* 21, 7913–7922. doi: 10.1128/mcb.21.23.7913-7922.2001

Maezawa, I., and Jin, L. W. (2010). Rett syndrome microglia damage dendrites and synapses by the elevated release of glutamate. *J. Neurosci.* 30, 5346–5356. doi: 10.1523/JNEUROSCI.5966-09.2010

Maezawa, I., Swanberg, S., Harvey, D., LaSalle, J. M., and Jin, L. W. (2009). Rett syndrome astrocytes are abnormal and spread MeCP2 deficiency through gap junctions. *J. Neurosci.* 29, 5051–5061. doi: 10.1523/JNEUROSCI.0324-09.2009

Martínez De Paz, A., Khajavi, L., Martin, H., Claveria-Gimeno, R., Tom Dieck, S., Cheema, M. S., et al. (2019). MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. *Epigenet. Chrom.* 12:63. doi: 10.1186/s13072-019-0298-1

Matagne, V., Borloz, E., Ehinger, Y., Saidi, L., Villard, L., and Roux, J.-C. (2020). Severe offtarget effects following intravenous delivery of AAV9-MECP2 in a female mouse model of Rett syndrome. *Neurobiol. Dis.* 149:105235. doi: 10.1016/j.nbd.2020.105235

Maunakea, A. K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* 23, 1256–1269. doi: 10.1038/cr.2013.110

Maxwell, S. S., Pelka, G. J., Tam, P. P., and El-Osta, A. (2013). Chromatin context and ncRNA highlight targets of MeCP2 in brain. *RNA Bio.* 10, 1741–1757. doi: 10.4161/rna.26921

McCauley, M. D., Wang, T., Mike, E., Herrera, J., Beavers, D. L., Huang, T. W., et al. (2011). Rett syndrome: pathogenesis of lethal cardiac arrhythmias in Mecp2 mutant mice: implication for therapy in Rett syndrome. *Sci. Transl. Med.* 3:2982. doi: 10.1126/scitranslmed.3002982

McGraw, C. M., Samaco, R. C., and Zoghbi, H. Y. (2011). Adult neural function requires MeCP2. *Science* 333:186. doi: 10.1126/science.1206593

McMahon, A. C., Rahman, R., Jin, H., Shen, J. L., Fieldsend, A., Luo, W., et al. (2016). TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell* 165, 742–753. doi: 10.1016/j.cell.2016.03.007

Meehan, R., Lewis, J. D., and Bird, A. P. (1992). Characterization of MECP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.* 20, 5085–5092. doi: 10.1093/nar/20.19.5085

Meguro-Horike, M., Yasui, D. H. D. H., Powell, W., Schroeder, D. I., Oshimura, M., LaSalle, J. M., et al. (2011). Neuron-specific impairment of inter-chromosomal pairing and transcription in a novel model of human 15q-duplication syndrome. *Hum. Mol. Genet.* 20, 3798–3810. doi: 10.1093/hmg/ddr298

Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* 151, 1417–1430. doi: 10.1016/j.cell.2012.11.022

Millar-Büchner, P., Philp, A. R., Gutierrez, N., Villanueva, S., Kerr, B., and Flores, C. A. (2016). Severe changes in colon epithelium in the Mecp2-null mouse model of Rett syndrome. *Mol. Cell. Pediatr.* 3:37. doi: 10.1186/s40348-016-0065-3

Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J. M., et al. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.* 36, 339–341. doi: 10.1038/ng1327

Muotri, A. R., Marchetto, M. C. N., Coufal, N. G., Oefner, R., Yeo, G., Nakashima, K., et al. (2010). L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468, 443–446. doi: 10.1038/nature09544

Nan, X., Campoy, F. J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88, 471–481. doi: 10.1016/S0092-8674(00)81887-5

Nikitina, T., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., Grigoryev, S. A., and Woodcock, C. L. (2007). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. *J. Biol. Chem.* 282, 28237–28245. doi: 10.1074/jbc.M704304200

Olson, C. O., Zachariah, R. M., Ezeonwuka, C. D., Liyanage, V. R. B., and Rastegar, M. (2014). Brain region-specific expression of MeCP2 isoforms correlates with

DNA methylation within Mecp2 regulatory elements. *PLoS One* 9:e0090645. doi: 10.1371/journal.pone.0090645

Osenberg, S., Karten, A., Sun, J., Li, J., Charkowick, S., Felice, C. A., et al. (2018). Activity-dependent aberrations in gene expression and alternative splicing in a mouse model of Rett syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 115, E5363–E5372. doi: 10.1073/pnas.1722546115

Piccolo, F. M., Liu, Z., Dong, P., Hsu, C. L., Stoyanova, E. I., Rao, A., et al. (2019). MeCP2 nuclear dynamics in live neurons results from low and high affinity chromatin interactions. *eLife* 8:e051449. doi: 10.7554/eLife.51449

Pohodich, A. E., Yalamanchili, H., Raman, A. T., Wan, Y. W., Gundry, M., Hao, S., et al. (2018). Forniceal deep brain stimulation induces gene expression and splicing changes that promote neurogenesis and plasticity. *eLife* 7:e034021. doi: 10.7554/eLife.34031

Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., Bröhm, A., Broche, J., et al. (2018). Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Nucleic Acids Res.* 46, 9044–9056. doi: 10.1093/nar/gky715

Raman, A. T., Pohodich, A. E., Wan, Y. W., Yalamanchili, H. K., Lowry, W. E., Zoghbi, H. Y., et al. (2018). Apparent bias toward long gene misregulation in MeCP2 syndromes disappears after controlling for baseline variations. *Nat. Commun.* 9:3225. doi: 10.1038/s41467-018-05627-1

Reichow, B., George-Puskar, A., Lutz, T., Smith, I. C., and Volkmar, F. R. (2015). Brief report: systematic review of Rett syndrome in males. *J. Autism Dev. Disord.* 45, 3377–3383. doi: 10.1007/s10803-015-2519-1

Renthal, W., Boxer, L. D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M. A., et al. (2018). Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nat. Neurosci.* 21, 1670–1679. doi: 10.1038/s41593-018-0270-6

Ross, P. D., Guy, J., Selfridge, J., Kamal, B., Bahey, N., Elizabeth Tanner, K., et al. (2016). Exclusive expression of MeCP2 in the nervous system distinguishes between brain and peripheral Rett syndrome-like phenotypes. *Hum. Mole. Genet.* 25, 4389–4404. doi: 10.1093/hmg/ddw269

Rube, H. T. T., Lee, W., Hejna, M., Chen, H., Yasui, D. H. D. H., Hess, J. F. J. F., et al. (2016). Sequence features accurately predict genome-wide MeCP2 binding in vivo. *Nat. Commun.* 7:11025. doi: 10.1038/ncomms11025

Samaco, R. C., Fryer, J. D., Ren, J., Fyffe, S., Chao, H. T., Sun, Y., et al. (2008). A partial loss of function allele of Methyl-CpG-binding protein 2 predicts a human neurodevelopmental syndrome. *Hum. Mol. Genet.* 17, 1718–1727. doi: 10.1093/hmg/ddn062

Samaco, R. C., Mandel-Brehm, C., Chao, H. T., Ward, C. S., Fyffe-Maricich, S. L., Ren, J., et al. (2009). Loss of MeCP2 in aminergic neurons causes cell-autonomous defects in neurotransmitter synthesis and specific behavioral abnormalities. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21966–21971. doi: 10.1073/pnas.0912257106

Saunders, C. J., Minassian, B. E., Chow, E. W. C., Zhao, W., and Vincent, J. B. (2009). Novel exon 1 mutations in MECP2 implicate isoform MeCP2-e1 in classical rett syndrome. *Am. J. Med. Genet. Part A* 149, 1019–1023. doi: 10.1002/ajmg.a.32776

Schaevitz, L. R., Gómez, N. B., Zhen, D. P., and Berger-Sweeney, J. E. (2013). MeCP2 R168X male and female mutant mice exhibit Rett-like behavioral deficits. *Genes Brain Behav.* 12, 732–740. doi: 10.1111/gbb.12070

Schafer, D. P., Heller, C. T., Gunner, G., Heller, M., Gordon, C., Hammond, T., et al. (2016). Microglia contribute to circuit defects in Mecp2 null mice independent of microglia-specific loss of Mecp2 expression. *eLife* 5:e015224. doi: 10.7554/eLife.15224

Schmidt, A., Zhang, H., and Cardoso, M. C. (2020). MeCP2 and chromatin compartmentalization. *Cells* 9:878. doi: 10.3390/cells9040878

Sekul, E. A., Moak, J. P., Schultz, R. J., Glaze, D. G., Dunn, J. K., and Percy, A. K. (1994). Electrocardiographic findings in Rett syndrome: an explanation for sudden death? *J. Pediatr.* 125, 80–82. doi: 10.1016/S0022-3476(94)70128-8

Sinnamon, J. R., Kim, S. Y., Corson, G. M., Song, Z., Nakai, H., Adelman, J. P., et al. (2017). Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9395–E9402. doi: 10.1073/pnas.1715320114

Sinnamon, J. R., Kim, S. Y., Fisk, J. R., Song, Z., Nakai, H., Jeng, S., et al. (2020). In vivo repair of a protein underlying a neurological disorder by programmable RNA editing. *Cell Rep.* 32:107878. doi: 10.1016/j.celrep.2020.107878

Sinnett, S. E., Hector, R. D., Gadalla, K. K. E., Heindel, C., Chen, D., Zaric, V., et al. (2017). Improved MECP2 gene therapy extends the survival of MeCP2-null mice without apparent toxicity after intracisternal delivery. *Mol. Therapy Methods Clin. Dev.* 5, 106–115. doi: 10.1016/j.omtm.2017.04.006

Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R. W., James, K. D., Turner, D. J., et al. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell* 37, 457–468. doi: 10.1016/j.molcel.2010.01.030

Strati, F., Cavalieri, D., Albanese, D., De Felice, C., Donati, C., Hayek, J., et al. (2016). Altered gut microbiota in Rett syndrome. *Microbiome* 4:41. doi: 10.1186/s40168-016-0185-y

Stuss, D. P., Cheema, M., Ng, M. K., Martinez De Paz, A., Williamson, B., Missiaen, K., et al. (2013). Impaired in vivo binding of MeCP2 to chromatin in the absence of its DNA methyl-binding domain. *Nucleic Acids Res.* 41, 4888–4900. doi: 10.1093/nar/gkt213

Sugino, K., Hempel, C. M., Okaty, B. W., Arnson, H. A., Kato, S., Dani, V. S., et al. (2014). Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. *J. Neurosci.* 34, 12877–12883. doi: 10.1523/JNEUROSCI.2674-14.2014

Symons, F. J., Byiers, B., Tervo, R. C., and Beisang, A. (2013). Parent-reported pain in Rett syndrome. *Clin. J. Pain* 29, 744–746. doi: 10.1097/AJP.0b013e318274b6bd

Szulwach, K. E., Li, X., Li, Y., Song, C. X., Wu, H., Dai, Q., et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.* 14, 1607–1616. doi: 10.1038/nn.2959

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935. doi: 10.1126/science.1170116

Trappe, R., Laccone, F., Cobilanschi, J., Meins, M., Huppke, P., Hanefeld, F., et al. (2001). MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am. J. Hum. Genet.* 68, 1093–1101. doi: 10.1086/320109

Valinluck, V., Liu, P., Kang, J. I., Burdzy, A., and Sowers, L. C. (2005). 5-Halogenated pyrimidine lesions within a CpG sequence context mimic 5-methylcytosine by enhancing the binding of the methyl-CpG-binding domain of methyl-CpG-binding protein 2 (MeCP2). *Nucleic Acids Res.* 33, 3057–3064. doi: 10.1093/nar/gki612

Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A., and Sowers, L. C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* 32, 4100–4108. doi: 10.1093/nar/gkh739

Veeraragavan, S., Wan, Y.-W., Connolly, D. R., Hamilton, S. M., Ward, C. S., Soriano, S., et al. (2015). Loss of MeCP2 in the rat models regression, impaired sociability and transcriptional deficits of Rett syndrome. *Hum. Mol. Genet.* 25:ddw178. doi: 10.1093/hmg/ddw178

Vogel Ciernia, A., Pride, M. C., Durbin-Johnson, B., Noronha, A., Chang, A., Yasui, D. H. D. H., et al. (2017). Early motor phenotype detection in a female mouse model of Rett syndrome is improved by cross-fostering. *Hum. Mol. Genet.* 26, 1839–1854. doi: 10.1093/hmg/ddx087

Vogel Ciernia, A., Yasui, D. H., Pride, M. C., Durbin-Johnson, B., Noronha, A. B., Chang, A., et al. (2018). MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Hum. Mol. Genet.* 27, 4077–4093. doi: 10.1093/hmg/ddy301

Wahba, G., Schock, S. C., Claridge, E., Bettolli, M., Grynspan, D., Humphreys, P., et al. (2015). MeCP2 in the enteric nervous system. *Neurogastroenterol. Motil.* 27, 1156–1161. doi: 10.1111/nmo.12605

Wang, L., Hu, M., Zuo, M. Q., Zhao, J., Wu, D., Huang, L., et al. (2020). Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Res.* 30, 393–407. doi: 10.1038/s41422-020-0288-7

Wasag, P., and Lenartowski, R. (2016). Nuclear matrix - structure, function and pathogenesis. *Postepy Hog. Med. Dosw.* 70, 1206–1219.

Weitzel, J. M., Buhrmester, H., and Strätling, W. H. (1997). Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. *Mol. Cell. Biol.* 17, 5656–5666. doi: 10.1128/mcb.17.9.5656

Wong, J. J. L., Gao, D., Nguyen, T. V., Kwok, C. T., Van Geldermalsen, M., Middleton, R., et al. (2017). Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. *Nat. Commun.* 8:134. doi: 10.1038/ncomms15134

Yasui, D. H., Gonzales, M. L., Aflatooni, J. O., Crary, F. K., Hu, D. J., Gavino, B. J., et al. (2014). Mice with an isoform-ablating Mecp2exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum. Mole. Genet.* 23, 2447–2458. doi: 10.1093/hmg/ddt640

Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., et al. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9416–9421. doi: 10.1073/pnas.0707442104

Yasui, D. H., Scoles, H. A., Horike, S.-I., Meguro-Horike, M., Dunaway, K. W., Schroeder, D. I., et al. (2011). 15q11.2-13.3 chromatin analysis reveals epigenetic regulation of CHRNA7 with deficiencies in Rett and autism brain. *Hum. Mol. Genet.* 20, 4311–4323. doi: 10.1093/hmg/ddr357

Young, J. I., Hong, E. P., Castle, J. C., Crespo-Barreto, J., Bowman, A. B., Rose, M. F., et al. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17551–17558. doi: 10.1073/pnas.0507856102

Yu, F., Zingler, N., Schumann, G., and Strätling, W. H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Res.* 29, 4493–4501. doi: 10.1093/nar/29.21.4493

Zachariah, R. M., Olson, C. O., Ezeonwuka, C., and Rastegar, M. (2012). Novel MeCP2 isoform-specific antibody reveals the endogenous MeCP2E1 expression in murine brain, primary neurons and Astrocytes. *PLoS One* 7:e0049763. doi: 10.1371/journal.pone.0049763

Zhang, P., Ludwig, A. K., Hastert, F. D., Rausch, C., Lehmkuhl, A., Hellmann, I., et al. (2017). L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins. *Nucleus* 8, 548–562.

Zhao, B., Wu, Q., Ye, A. Y., Guo, J., Zheng, X., Yang, X., et al. (2019). Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. *PLoS Genet.* 15:1008043. doi: 10.1371/journal.pgen.1008043

Zoghbi, H. Y. (2005). MeCP2 dysfunction in humans and mice. *J. Child Neurol.* 20, 736–740. doi: 10.1177/08830738050200090701

Zuliani, I., Urbinati, C., Valenti, D., Quattrini, M. C., Medici, V., Cosentino, L., et al. (2020). The anti-diabetic drug metformin rescues aberrant mitochondrial activity and restrains oxidative stress in a female mouse model of Rett syndrome. *J. Clin. Med.* 9:1669. doi: 10.3390/jcm9061669

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 06 August 2021
doi: 10.3389/fcell.2021.655703

A Comprehensive Genomic Analysis Constructs miRNA–mRNA Interaction Network in Hepatoblastoma

Tong Chen[1†], Linlin Tian[2†], Jianglong Chen[1], Xiuhao Zhao[1], Jing Zhou[1], Ting Guo[1], Qingfeng Sheng[1], Linlin Zhu[1], Jiangbin Liu[1] and Zhibao Lv[1*]*

[1] Department of General Surgery, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China,
[2] Department of Microbiology, Faculty of Basic Medical Sciences, Guilin Medical University, Guilin, China

OPEN ACCESS

Edited by:
Mojgan Rastegar, University of Manitoba, Canada

Reviewed by:
Amancio Carnero, Seville University, Spain
Nicholas B. Larson, Mayo Clinic, United States

***Correspondence:**
Zhibao Lv
zhibaolyu@sina.cn
Jiangbin Liu
liujb@shchildren.com.cn

† These authors have contributed equally to this study

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Cell and Developmental Biology

Received: *19 January 2021*
Accepted: *13 July 2021*
Published: *06 August 2021*

Citation:
Chen T, Tian L, Chen J, Zhao X, Zhou J, Guo T, Sheng Q, Zhu L, Liu J and Lv Z (2021) A Comprehensive Genomic Analysis Constructs miRNA–mRNA Interaction Network in Hepatoblastoma. Front. Cell Dev. Biol. 9:655703. doi: 10.3389/fcell.2021.655703

Hepatoblastoma (HB) is a rare disease but nevertheless the most common hepatic tumor in the pediatric population. For patients with advanced HB, the prognosis is dismal and there are limited therapeutic options. Multiple microRNAs (miRNAs) were reported to be involved in HB development, but the miRNA–mRNA interaction network in HB remains elusive. Through a comparison between HB and normal liver samples in the GSE131329 dataset, we detected 580 upregulated differentially expressed mRNAs (DE-mRNAs) and 790 downregulated DE-mRNAs. As for the GSE153089 dataset, the first cluster of differentially expressed miRNAs (DE-miRNAs) were detected between fetal-type tumor and normal liver groups, while the second cluster of DE-miRNAs were detected between embryonal-type tumor and normal liver groups. Through the intersection of these two clusters of DE-miRNAs, 33 upregulated hub miRNAs, and 12 downregulated hub miRNAs were obtained. Based on the respective hub miRNAs, the upstream transcription factors (TFs) were detected via TransmiR v2.0, while the downstream target genes were predicted via miRNet database. The intersection of target genes of respective hub miRNAs and corresponding DE-mRNAs contributed to 250 downregulated candidate genes and 202 upregulated candidate genes. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses demonstrated the upregulated candidate genes mainly enriched in the terms and pathways relating to the cell cycle. We constructed protein–protein interaction (PPI) network, and obtained 211 node pairs for the downregulated candidate genes and 157 node pairs for the upregulated candidate genes. Cytoscape software was applied for visualizing the PPI network and respective top 10 hub genes were identified using CytoHubba. The expression values of hub genes in the PPI network were subsequently validated through Oncopression database followed by quantitative real-time polymerase chain reaction (qRT-PCR) in HB and matched normal liver tissues, resulting in six

significant downregulated genes and seven significant upregulated genes. The miRNA–mRNA interaction network was finally constructed. In conclusion, we uncover various miRNAs, TFs, and hub genes as potential regulators in HB pathogenesis. Additionally, the miRNA–mRNA interaction network, PPI modules, and pathways may provide potential biomarkers for future HB theranostics.

Keywords: hepatoblastoma, miRNA, mRNA, PPI, TF

INTRODUCTION

Hepatoblastoma (HB) is a rare disease with an annual incidence of 1.5 cases per million children per year (Spector and Birch, 2012). Nevertheless, it is the predominant hepatic tumor in the pediatric population (Schnater et al., 2003). The past three decades have witnessed a consistently increasing incidence of HB (Linabery and Ross, 2008). Surgical resection and chemotherapy have dramatically improved the prognosis for HB children, with the 3-years event-free survival (EFS) > 80% (Aronson et al., 2014). However, there are limited therapeutic strategies for advanced HB children, with the 3-years EFS of only 34% (Semeraro et al., 2013). In addition, patient survivors may suffer severe side effects of chemotherapeutic or immunosuppressive agents. Therefore, there is an urgent need to unveil the molecular mechanisms underlying this rare tumor in order to identify novel biomarkers for therapeutic tailoring.

MicroRNAs (miRNAs) are ∼22 nucleotide non-coding RNAs that post-transcriptionally suppress messenger RNAs (mRNAs) expression (Lou et al., 2019). Through base-pairing interactions with mRNAs, miRNAs play crucial roles in proliferation (Roy et al., 2017), apoptosis (Liu et al., 2020), epithelial-mesenchymal transition (Weng et al., 2019), and autophagy (Kuang et al., 2020) of human cells. Moreover, the dysregulated expression of miRNA is associated with the pathogenesis of various human tumors, including HB (Cui et al., 2019b). In the context of HB, miR-193a-5p promotes proliferative, migrative, and invasive properties of HB through targeting DPEP1 and augmenting PI3K/AKT/mTOR signaling pathway (Cui et al., 2019a); miR-492 serves as an endogenous tumor-promoting factor to induce proliferation, anchorage-independent growth, migrative and invasive properties of HB cells by targeting CD44, and high level of miR-492 expression is correlated with high-risk or aggressive HB (von Frowein et al., 2018); miR-21 enhances apoptosis in HB cells through targeting ASPP2 and augmenting ASPP2/p38 signaling pathway (Liu et al., 2019). In other words, the intimate relationship between altered expression of certain miRNA and its target gene has been uncovered in HB. Transcription factors (TFs) are endogenous proteins that regulate the transcription process of mRNAs or miRNAs. The function of TFs can be either oncogenic or tumor suppressive depending on context (Lambert et al., 2018). Recently, multiple TFs have been demonstrated to modulate the aggressive phenotype and cellular process in HB development (Zhang et al., 2019; Nakra et al., 2020; Wagner et al., 2020).

In recent years, high-throughput technologies have enabled us to identify the key genes, miRNAs, and TFs in the initiation and progression of human tumors. To date, there has been a scarce number of integrated genome-wide studies on HB (Zhang L. et al., 2018; Aghajanzadeh et al., 2020) via research on several cases or one dataset. To gain a better understanding of the underlying mechanisms behind HB, this study aimed to explore the miRNA–mRNA interaction network, TFs, and biological pathways involved in HB through comprehensive bioinformatic approaches.

MATERIALS AND METHODS

Data Retrieval and Extraction

HB-related data were obtained from the Gene Expression Omnibus (GEO[1]) database portal via the keyword "hepatoblastoma." The dataset was included when all four items of the following criteria were met: (1) there were both HB and normal liver samples; (2) the dataset had miRNA or mRNA transcriptome data; (3) data for all samples were completely presented; (4) HB and normal liver samples could be clearly distinguished using principal component analysis (PCA). After screening, we chose one mRNA dataset (accession number: GSE131329) and one miRNA dataset (accession number: GSE153089) for further analysis. GSE131329 (Hiyama et al., 2019), consisting of 53 HB samples and 14 normal liver samples, was analyzed via GPL6244 platform (Affymetrix Human Gene 1.0 ST Array). GSE153089 (Honda et al., 2020), comprising of 30 HB samples and 14 normal liver samples, was analyzed via GPL21572 platform (Affymetrix Multispecies miRNA-4 Array). General information of the two datasets used for the present study is shown in **Supplementary Table 1**.

The GSE153089 dataset included nine specimens from metastatic tumor, 21 specimens from primary tumor (11 fetal subtypes and 10 embryonal subtypes), and 14 specimens from surrounding normal liver (Honda et al., 2020). Due to the lack of metastatic tumor samples in the GSE131329 dataset (Hiyama et al., 2019), we excluded all specimens from metastatic tumor in the GSE153089 dataset before further analysis. The remaining specimens in the GSE153089 dataset were subsequently divided into three groups, namely, normal surrounding liver, fetal-type tumor, and embryonal-type tumor groups. Each patient in the GSE153089 dataset possessed no more than one specimen from the same group except for patient 7. There were two fetal-type tumor specimens for patient 7 (Sample ID: 25F-1 and 25F-2), one (Sample ID: 25F-2) of which was randomly excluded for further analysis. Detailed information of samples

[1]https://www.ncbi.nlm.nih.gov/geo/

in the GSE153089 dataset used for the present study is listed in **Supplementary Table 2**.

Screening of Differentially Expressed miRNAs and Differentially Expressed mRNAs

Raw data files (*.CEL) of GSE153089 and GSE131329 were imported using the *oligo* (Carvalho and Irizarry, 2010) R package. The data were sequentially filtered, background corrected, log base 2 transformed, and normalized. Based on the platform annotation information, gene symbol was obtained via conversion of the probe. If one gene symbol corresponded to two or more probes, the mean expression level of these corresponding mRNAs or miRNAs was treated as the final expression value. Before and after clustering and removing outliers, we detected the distribution patterns of HB and normal liver samples via PCA. DE-mRNAs and DE-miRNAs were then detected using the *limma* R package (17). An adjusted $P < 0.05$ and $|\log2FC| > 1$ indicated statistical significance. Benjamini–Hochberg (BH) method was used to adjust the *P* value. Regarding the GSE131329 dataset, DE-mRNAs were obtained based on the comparison between HB and normal liver samples. As for the GSE153089 dataset, the first cluster of DE-miRNAs were detected between fetal-type tumor and normal liver groups, while the second cluster of DE-miRNAs were detected between embryonal-type tumor and normal liver groups. Through the intersection of these two clusters of DE-miRNAs, the upregulated or downregulated hub miRNAs were obtained.

Prediction of Potential TFs and Target Genes of Hub miRNAs

Based on the hub miRNAs, we predicted the upstream TFs via TransmiR v2.0 (Tong et al., 2019), an easy-accessible public tool integrating experimentally verified TF-miRNA regulatory relationships from the publications. The Cytoscape software was subsequently utilized to visualize TF-miRNA regulatory relationships (Shannon et al., 2003). In addition, miRNet database was used for the prediction of the downstream target genes of hub miRNAs (Fan et al., 2016).

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Analyses

To further explore functional annotation of the candidate genes, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses via the *clusterProfiler* R package (Yu et al., 2012). GO terms consisted of biological process (BP), cellular component (CC), and molecular function (MF). An adjusted $P < 0.05$ was considered significantly enriched, and BH method was used to adjust the *P* value.

Protein–Protein Interaction Network

To unveil the relationships between the candidate genes, we established the PPI network via the STRING database (Szklarczyk et al., 2015). PPI pairs were considered significant with a combined score ≥ 0.4. Cytoscape software was subsequently applied to visualize the network (Shannon et al., 2003). On the basis of the degree obtained through Cytoscape plugin CytoHubba (Chin et al., 2014), top 10 hub genes were detected in the PPI network.

Hub Genes Verification Through Oncopression Database

We applied Oncopression database[2] to validate expression levels of top 10 up-regulated hub genes and top 10 down-regulated

[2]http://oncopression.com/

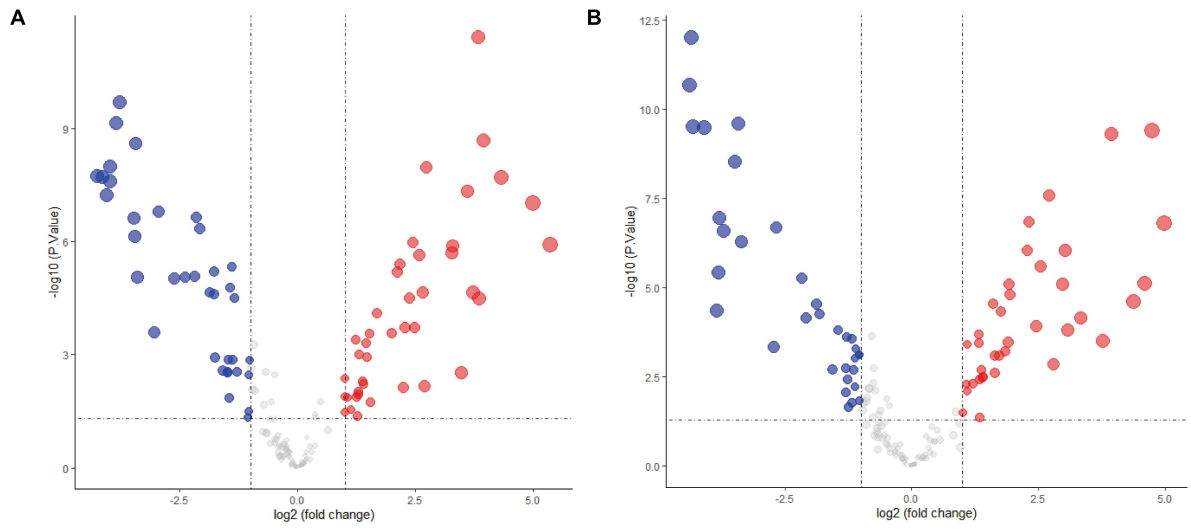

FIGURE 1 | Differentially expressed microRNAs analysis of the GSE153089 dataset. **(A)** DE-miRNAs between fetal-type tumor and normal liver samples were visualized via volcano plot. **(B)** DE-miRNAs between embryonal-type tumor and normal liver samples were visualized via volcano plot. Red points representing up-regulation; blue points indicating down-regulation; gray points representing normal expression. DE-miRNAs, differentially expressed microRNAs.

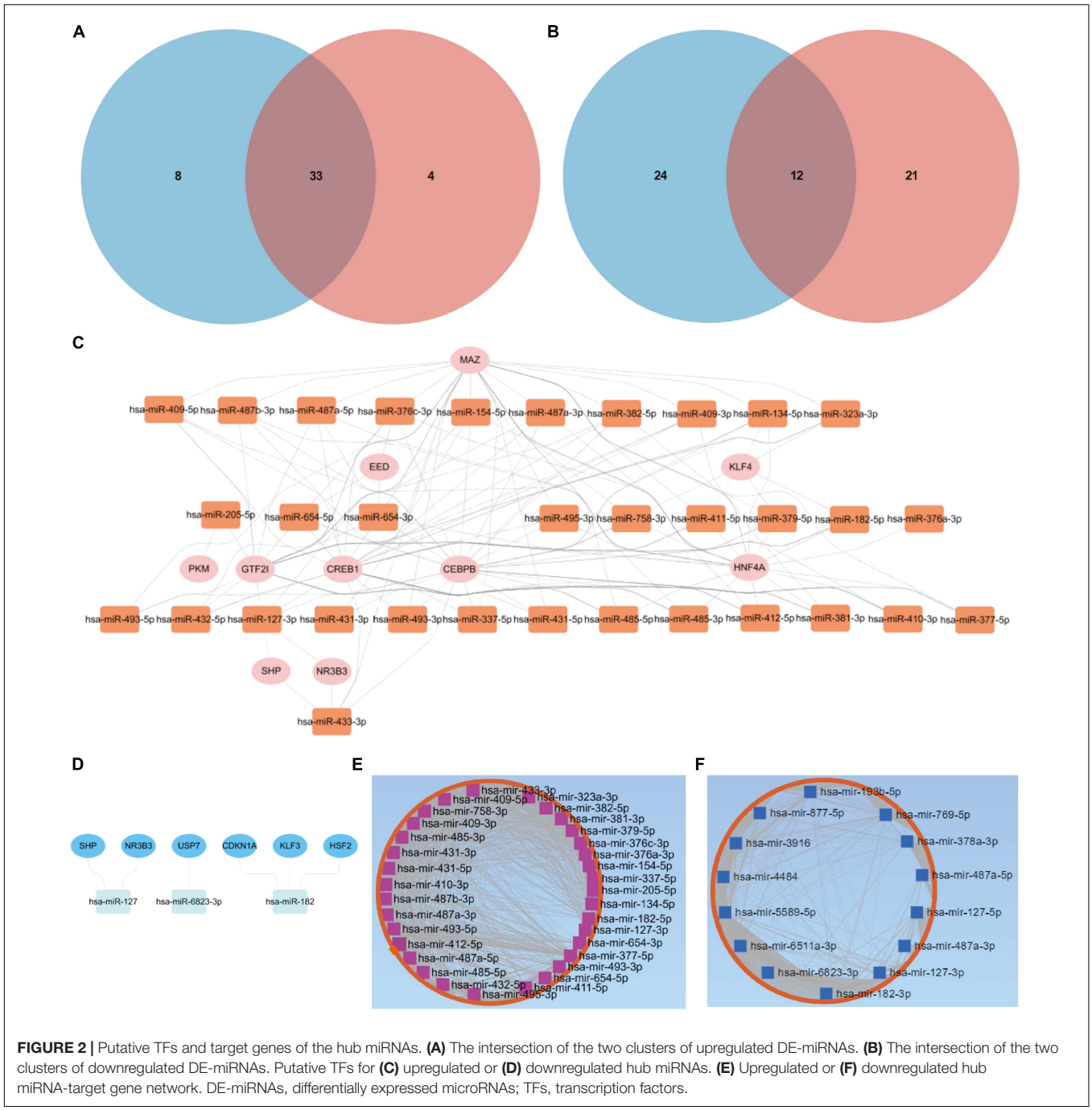

FIGURE 2 | Putative TFs and target genes of the hub miRNAs. **(A)** The intersection of the two clusters of upregulated DE-miRNAs. **(B)** The intersection of the two clusters of downregulated DE-miRNAs. Putative TFs for **(C)** upregulated or **(D)** downregulated hub miRNAs. **(E)** Upregulated or **(F)** downregulated hub miRNA-target gene network. DE-miRNAs, differentially expressed microRNAs; TFs, transcription factors.

hub genes. Oncopression is a web-based integrated gene expression profile using single sample normalization method UPC (Lee and Choi, 2017).

Tissue Samples

Hepatoblastoma and matched normal liver tissue samples from eight children undergoing surgical excision for primary HB were obtained from our hospital between 2014 and 2019. None of the patients received adjuvant radiotherapy or chemotherapy prior to surgery. Tissues were stored at −80°C immediately after harvest until further use. The pathological diagnosis of the tissue adjacent to each frozen tissue specimen was confirmed by at least two independent pathologists.

RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction

Total RNAs were isolated from the tissues using TRIzol reagent (Life Technologies, Carlsbad, CA, United States). Total mRNA was subsequently reverse-transcribed to produce complementary DNA (cDNA) using TaKaRa reverse transcription kit (TaKaRa

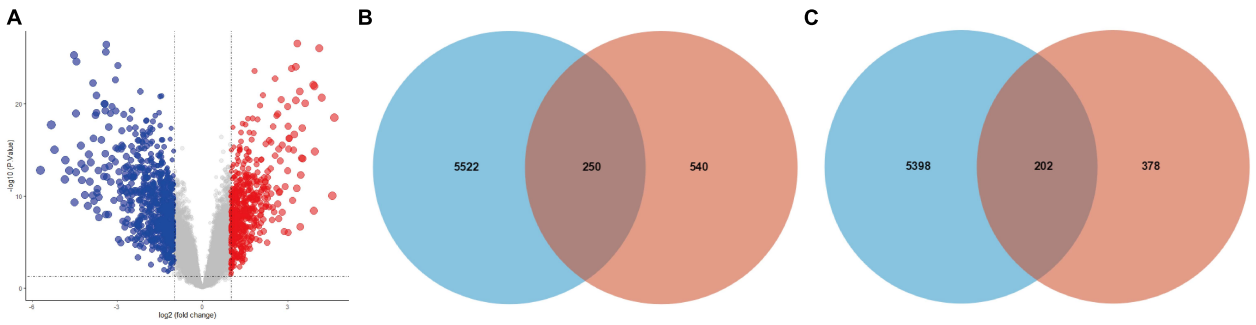

FIGURE 3 | Intersection of target genes of hub miRNAs and corresponding DE-mRNAs. **(A)** DE-mRNAs between HB and normal liver samples in the GSE131329 dataset were visualized using volcano plot. Red points representing up-regulation; blue points indicating down-regulation; gray points representing normal expression. The intersection of target genes of **(B)** upregulated or **(C)** downregulated hub miRNAs and corresponding DE-mRNAs. DE-mRNAs, differentially expressed mRNAs.

Bio, Shiga, Japan). The SYBR Green fluorescence system (Roche, IN, United States) was used, and mRNA qRT-PCR was performed using a quantitative mRNA kit (TaKaRa Bio, Shiga, Japan). Based on the $2^{-\Delta\Delta Ct}$ method, the relative mRNA levels were normalized to GAPDH mRNA levels. All primers were synthesized by Sangon (Shanghai, China). The sequence of primers is summarized in **Supplementary Table 3**.

Statistical Analysis

We conducted data analysis and visualization using R software (version 3.6.3) and GraphPad Prism (version 8.0.1). The expression levels of mRNAs or miRNAs between groups in the datasets were compared via a moderated *t*-test. For differential expression analysis of mRNAs or miRNAs in the datasets, a *P* value < 0.05 and $|\log2FC| > 1$ were considered statistically significant. The mRNA expression levels of hub genes in HB and matched normal liver tissues from our hospital were statistically analyzed by a paired Student *t* test, and *P* values below 0.05 were considered significant.

RESULTS

Hub miRNAs Identification

The expression levels of fetal-type tumor and normal liver samples in the GSE153089 dataset prior to and after normalization are shown (**Supplementary Figures 1A,B**). PCA results before and after removing outliers (GSM4633970, GSM4633988, and GSM4633998) are also presented (**Supplementary Figures 1C,D**). Based on the differential expression analysis, we detected 41 upregulated DE-miRNAs and 36 downregulated DE-miRNAs, which are presented via volcano plot in **Figure 1A**. In addition, these DE-miRNAs between fetal-type tumor and normal liver samples were regarded as the first cluster of DE-miRNAs.

The expression values of embryonal-type tumor and normal liver samples in the GSE153089 dataset prior to and after normalization are shown (**Supplementary Figures 2A,B**). PCA results before and after removing outliers (GSM4633970 and GSM4633988) are also presented (**Supplementary Figures 2C,D**). Through the differential expression analysis, we detected 37 upregulated DE-miRNAs and 33 downregulated DE-miRNAs, which are presented via volcano plot in **Figure 1B**. Additionally, these DE-miRNAs between embryonal-type tumor and normal liver samples were regarded as the second cluster of DE-miRNAs.

Through the intersection of the aforementioned two clusters of DE-miRNAs, a total of 33 upregulated DE-miRNAs and 12 downregulated hub miRNAs were obtained (**Figures 2A,B**). Detailed information of respective hub miRNAs is also listed (**Supplementary Tables 4,5**).

TFs and Target Genes Predicted by Hub miRNAs

As for the upregulated hub miRNAs, the predicted TFs included HNF4A, GTF2I, CEBPB, CREB1, MAZ, NR3B3, SHP, KLF4, PKM, and EED (**Figure 2C**). Regarding the downregulated hub miRNAs, the predicted TFs included NR3B3, SHP, CDKN1A, KLF3, USP7, and HSF2 (**Figure 2D**). Detailed information of the TFs predicted for the upregulated or downregulated DE-miRNAs is also listed (**Supplementary Tables 6,7**). Apart from the predicted TFs, we also predicted 5,772 target genes of the upregulated hub miRNAs and 5,600 target genes of the downregulated hub miRNAs. Upregulated hub miRNA-target gene network and downregulated hub miRNA-target gene network are presented in **Figures 2E,F**, respectively.

DE-mRNAs Identification

The expression levels of all samples in the GSE131329 dataset before and after normalization are visualized in **Supplementary Figures 3A,B**, respectively. PCA results prior to and after excluding the outlier (GSM3770543) are also shown (**Supplementary Figures 3C,D**). We then obtained 580 upregulated DE-mRNAs and 790 downregulated DE-mRNAs, which are presented via volcano plot in **Figure 3A**. The detailed information of these respective DE-mRNAs is listed (**Supplementary Tables 8,9**). Subsequently, we intersected the target genes of upregulated hub miRNAs and downregulated

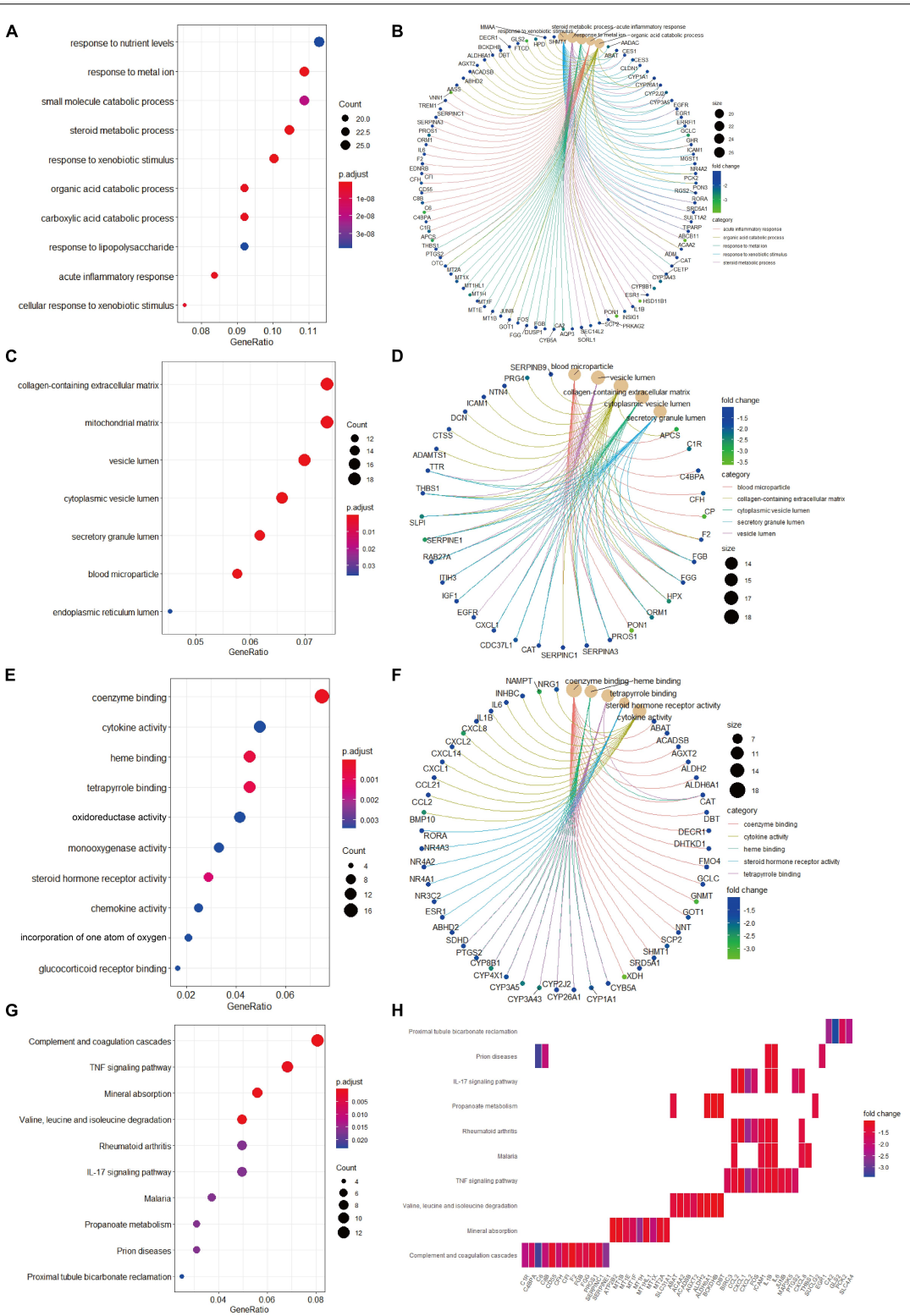

FIGURE 4 | Gene Ontology terms and KEGG pathway enrichment analyses of the downregulated candidate genes. **(A)** The enriched GO-BP terms based on downregulated candidate genes. **(B)** The downregulated candidate genes and their enriched GO-BP terms. **(C)** The enriched GO-CC terms based on downregulated candidate genes. **(D)** The downregulated candidate genes and their enriched GO-CC terms. **(E)** The enriched GO-MF terms based on downregulated candidate genes. **(F)** The downregulated candidate genes and their enriched GO-MF terms. **(G)** KEGG pathway analysis showing the enriched pathways based on downregulated candidate genes. **(H)** Heatmap showing specific downregulated candidate genes and their enriched pathways. BP, biological process; CC, cellular component; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function.

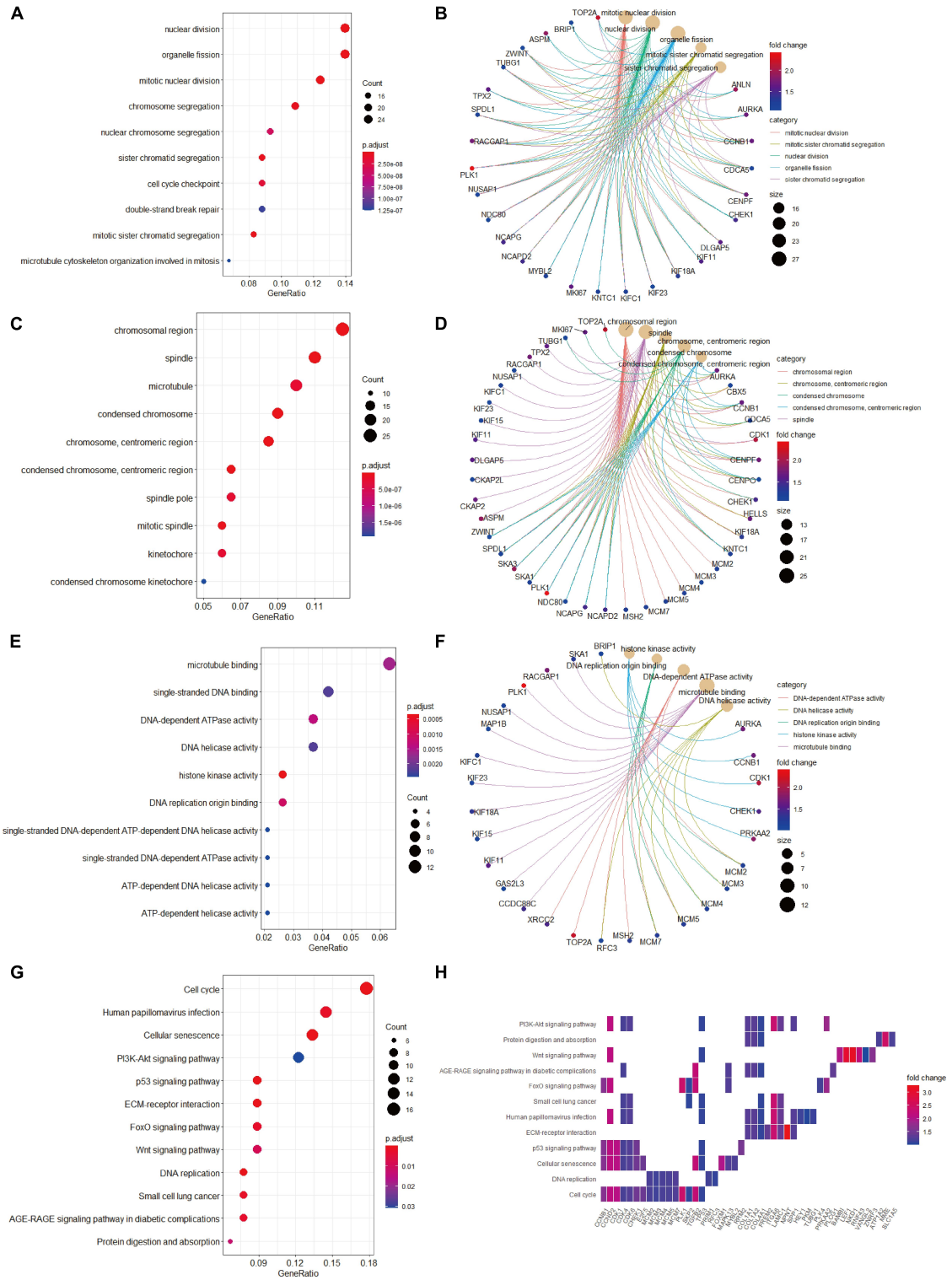

FIGURE 5 | Gene Ontology terms and KEGG pathway enrichment analyses of the upregulated candidate genes. **(A)** The enriched GO-BP terms based on upregulated candidate genes. **(B)** The upregulated candidate genes and their enriched GO-BP terms. **(C)** The enriched GO-CC terms based on upregulated candidate genes. **(D)** The upregulated candidate genes and their enriched GO-CC terms. **(E)** The enriched GO-MF terms based on upregulated candidate genes. **(F)** The upregulated candidate genes and their enriched GO-MF terms. **(G)** KEGG pathway analysis showing the enriched pathways based on upregulated candidate genes. **(H)** Heatmap showing specific upregulated candidate genes and their enriched pathways. BP, biological process; CC, cellular component; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function.

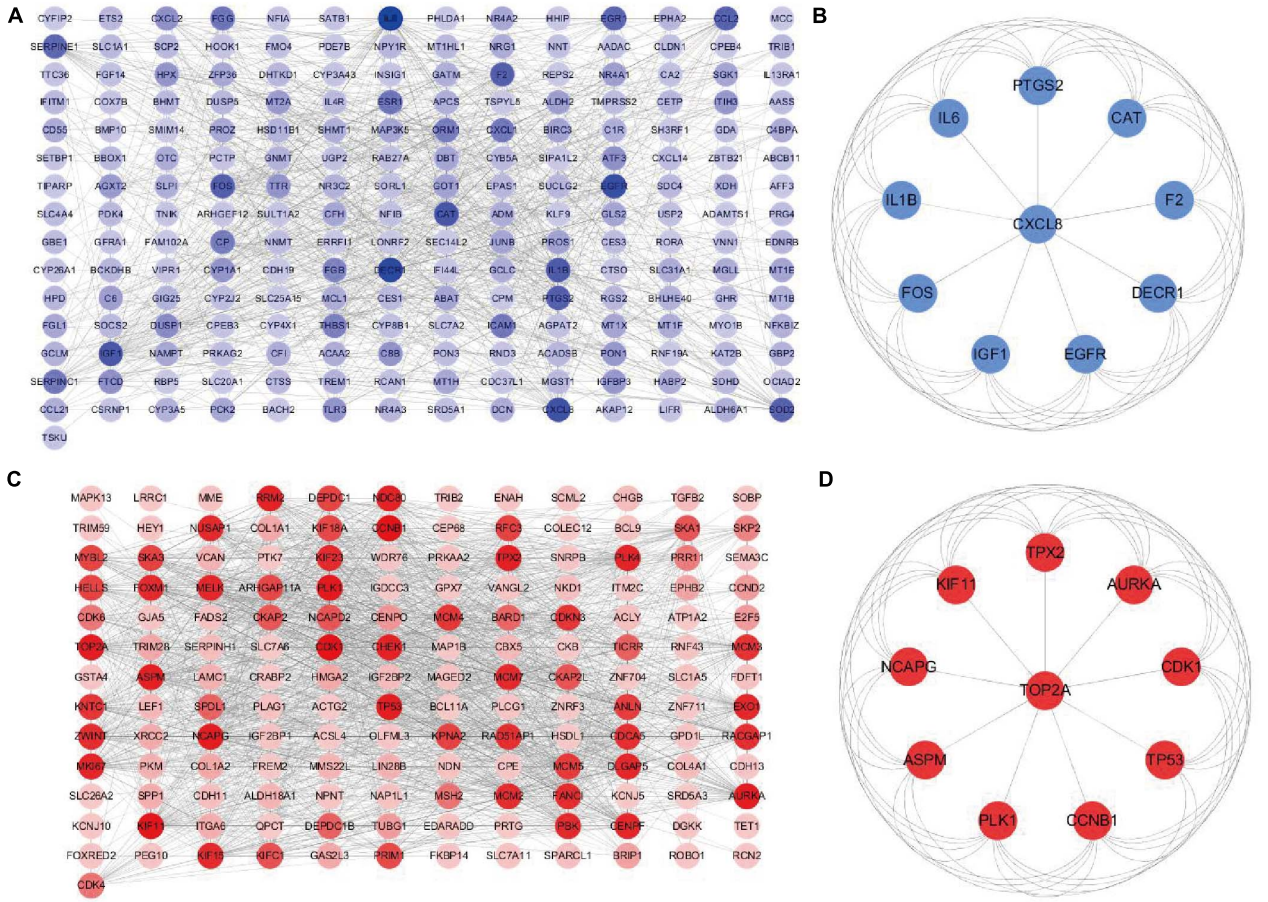

FIGURE 6 | Construction of PPI network based on downregulated or upregulated candidate genes. PPI network of **(A)** downregulated or **(B)** upregulated candidate genes. **(C)** The top 10 hub genes of downregulated candidate genes based on the node degree. **(D)** The top 10 hub genes of upregulated candidate genes based on the node degree. PPI, protein–protein interactions.

DE-mRNAs, resulting in a total of 250 downregulated candidate genes (**Figure 3B**). In addition, the intersection of target genes of downregulated hub miRNAs and upregulated DE-mRNAs resulted in 202 upregulated candidate genes (**Figure 3C**). Detailed information of these downregulated and upregulated candidate genes is also listed (**Supplementary Tables 10, 11**).

Functional Annotation Enrichment Analyses

Biological process analysis indicated that enriched GO terms for downregulated candidate genes included response to nutrient levels, response to metal ion, small molecule catabolic process, and steroid metabolic process (**Figure 4A**). CC analysis showed that the candidate genes were markedly enriched in collagen-containing extracellular matrix, mitochondrial matrix, vesicle lumen, cytoplasmic vesicle lumen, secretory granule lumen, and blood microparticle (**Figure 4C**). In the process of MF analysis, the candidate genes were markedly enriched in coenzyme binding, cytokine activity, heme binding, tetrapyrrole binding, and oxidoreductase activity (**Figure 4E**). The complex relationships between these candidate genes and their related GO terms were visualized using the *cnetplot* R package (**Figures 4B,D,F**). Moreover, KEGG analysis identified complement and coagulation cascades, TNF signaling pathway, mineral absorption, and valine, leucine and isoleucine degradation as markedly enriched pathways (**Figure 4G**). Besides, the enriched pathways and their associated candidate genes were unveiled, which are shown as heatmap in **Figure 4H**.

We then conducted GO terms analysis based on the upregulated candidate genes. BP analysis revealed that nuclear division, organelle fission, mitotic nuclear division, and chromosome segregation served as the top enriched GO terms (**Figure 5A**). CC analysis identified chromosomal region, spindle, microtubule, and condensed chromosome as significantly enriched GO terms (**Figure 5C**). In addition, MF analytic results revealed microtubule binding, single-stranded DNA binding, DNA-dependent ATPase activity, DNA helicase activity, and histone kinase activity as markedly enriched GO terms (**Figure 5E**). The complex relationships between the aforementioned enriched GO terms and their associated candidate genes are also shown (**Figures 5B,D,F**). KEGG

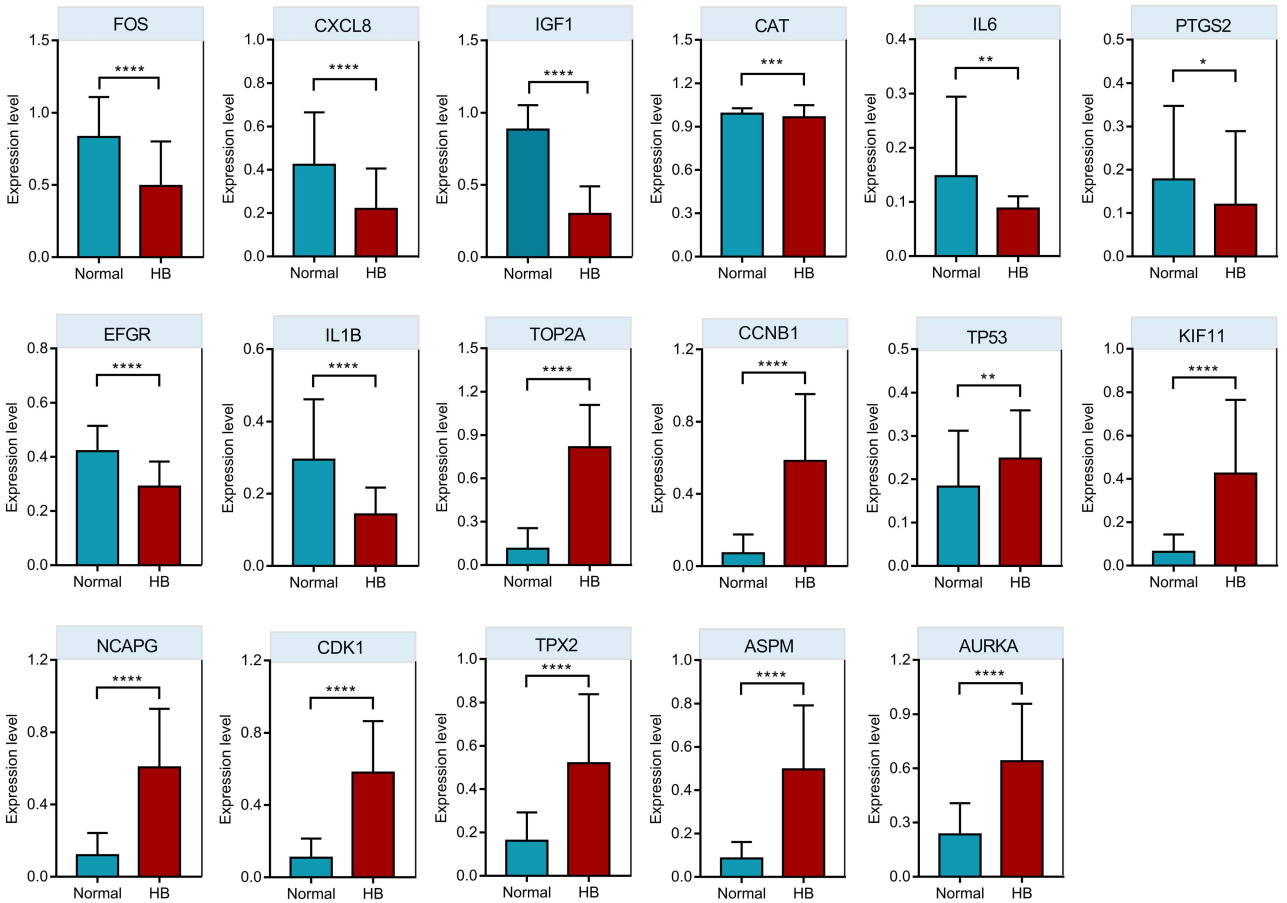

FIGURE 7 | The comparison of hub genes expression levels in HB and normal liver tissue samples from Oncopression database. After UPC-normalization, the expression levels range from 0 to 1 where 0 and 1 indicate no expression and the highest expression, respectively. *P value < 0.05; **P value < 0.01; ***P value < 0.001; ****P value < 0.0001.

analysis identified cell cycle, cellular senescence and PI3K-AKT signaling as significantly enriched pathways (**Figure 5G**), and the associations of these pathways and their related candidate genes were visualized using heatmap (**Figure 5H**).

PPI Network Construction and Hub Genes Screening

The downregulated or upregulated candidate genes were loaded into the STRING database, resulting in the construction of respective PPI network. A total of 211 node pairs were obtained for the downregulated candidate genes (**Figure 6A**), while 157 node pairs were obtained for the upregulated candidate genes (**Figure 6C**). The node pairs were input into Cytoscape software to visualize genes in respective PPI network. The respective top 10 hub genes were detected via Cytoscape plugin CytoHubba (**Figures 6B,D**). Specifically, the top 10 upregulated hub genes were CDK1, CCNB1, KIF11, PLK1, NCAPG, TOP2A, AURKA, TP53, ASPM, and TPX2, while the top 10 downregulated hub genes included IL6, DECR1, EGFR, CXCL8, CAT, IGF1, IL1B, F2, PTGS2, and FOS.

Hub Genes Verification via Oncopression Database

Oncopression database was utilized to validate the expression values of respective top 10 hub genes in the PPI network. As shown in **Figure 7**, eight of the top 10 downregulated hub genes (IL6, EGFR, CXCL8, CAT, IGF1, IL1B, PTGS2, and FOS) had significantly lower expression levels in HB tissue samples compared to normal liver tissue samples, while nine of the top 10 upregulated hub genes (CDK1, CCNB1, KIF11, NCAPG, TOP2A, AURKA, TP53, ASPM, and TPX2) had markedly higher expression levels in HB tissues in comparison to normal liver tissues.

Hub Genes Verification via qRT-PCR

Through literature search, we identified AURKA (Zhang Y. et al., 2018; Tan et al., 2020) and CDK1 (Tian et al., 2021) as previously reported oncogenic genes in HB. In contrast, the role of the other 15 hub genes in HB has not been reported to date or remains controversial. Based on HB and matched normal liver tissue samples in eight children with HB, the mRNA expression levels of these 15 hub genes were

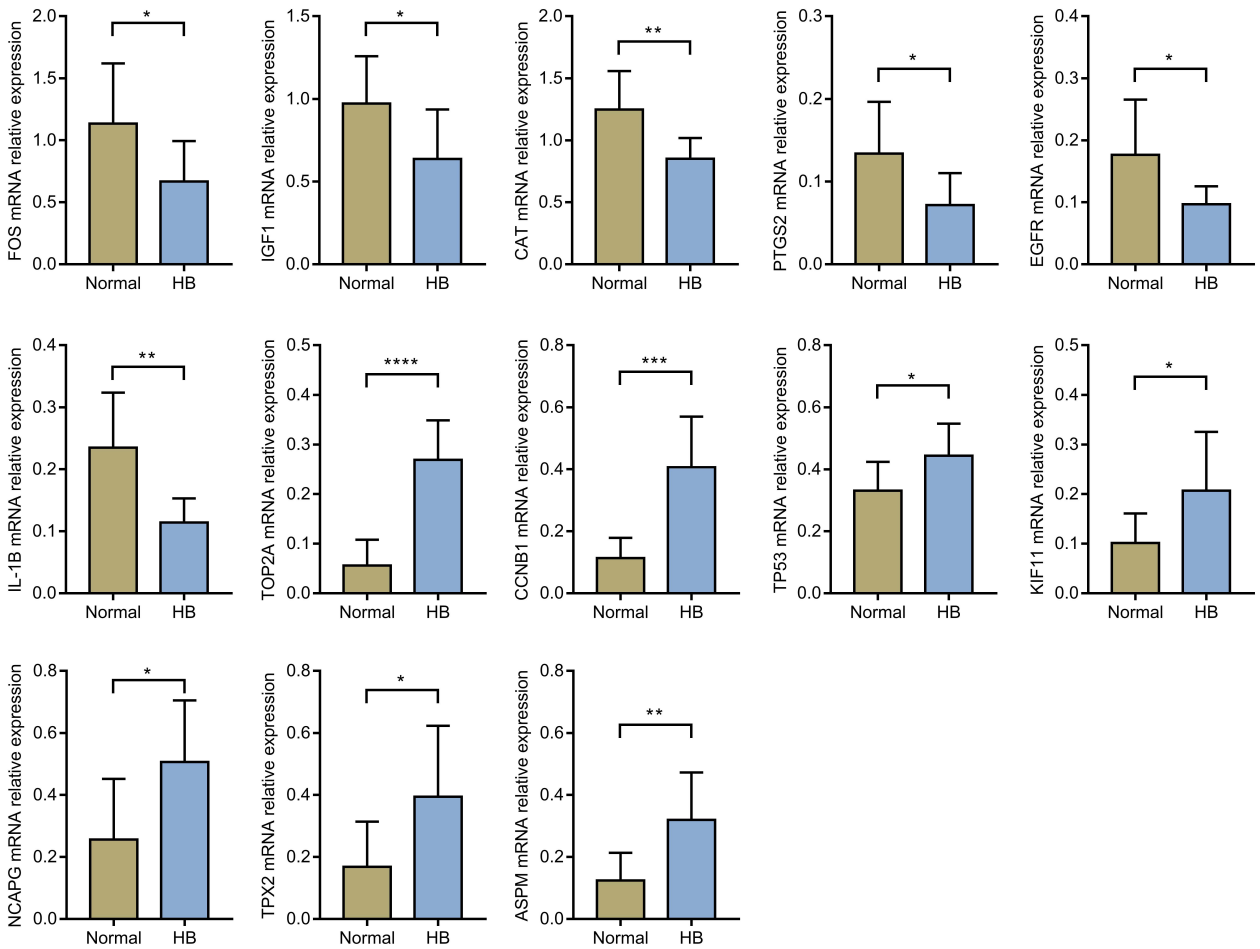

FIGURE 8 | The comparison of hub genes expression levels in HB and normal liver tissues in eight children with HB from our hospital. HB, hepatoblastoma. *P value < 0.05; **P value < 0.01; ***P value < 0.001; ****P value < 0.0001.

validated by qRT-PCR. The expression levels of EGFR, CAT, IGF1, IL1B, PTGS2, and FOS were significantly lower for HB tissues when compared with normal liver tissues. On the other hand, the expression values of CCNB1, KIF11, NCAPG, TOP2A, TP53, ASPM, and TPX2 in HB tissues were significantly higher than those in the normal liver tissues (**Figure 8**). Lastly, according to the predicted miRNA–mRNA pairs and the final verification results, we constructed the potential miRNA–mRNA interaction network involved in HB (**Figure 9**).

DISCUSSION

The extreme rarity of HB has hindered our understanding of its underlying molecular mechanisms, and the majority of potential hub genes, DE-miRNAs and TFs in this study were reported for the first time in HB pathogenesis. Therefore, our work may serve as an important resource for future studies to unveil the underlying mechanisms of these key biomarkers and/or therapeutic targets involved in HB.

The dysregulation of miRNA–mRNA interaction network in liver is associated with various liver diseases, such as liver regeneration (Wang et al., 2019) and hepatocellular carcinoma (Zhang and Du, 2017; Lou et al., 2019). In the context of HB, previous studies have identified multiple miRNAs as promising therapeutic targets (von Frowein et al., 2018; Cui et al., 2019a; Liu et al., 2019). In order to provide an overall picture of miRNA–mRNA interaction network in HB pathogenesis, we performed a comprehensive bioinformatic analysis on the basis of two independent GEO datasets. Among the TFs predicted for DE-miRNAs in this study, CDKN1A was reported to regulate the G1/S transition and affect replication and damage repair of DNA during mitosis (Tokumoto et al., 2003). In addition, AP2 negatively controls the growth of HepG2 HB cells through CDKN1A activation (Zeng et al., 1997). USP7, another predicted TF in this study, was reported to promote proliferation, migration and invasion of HB cell lines through activation of PI3K/AKT signaling (Ye et al., 2021). Apart from CDKN1A and USP7, HNF4A was reported to be essential for Smad2/3 binding regions in HepG2 HB cells, thus affecting transcription regulated by TGF-β (Mizutani et al., 2011). Future studies are needed to validate

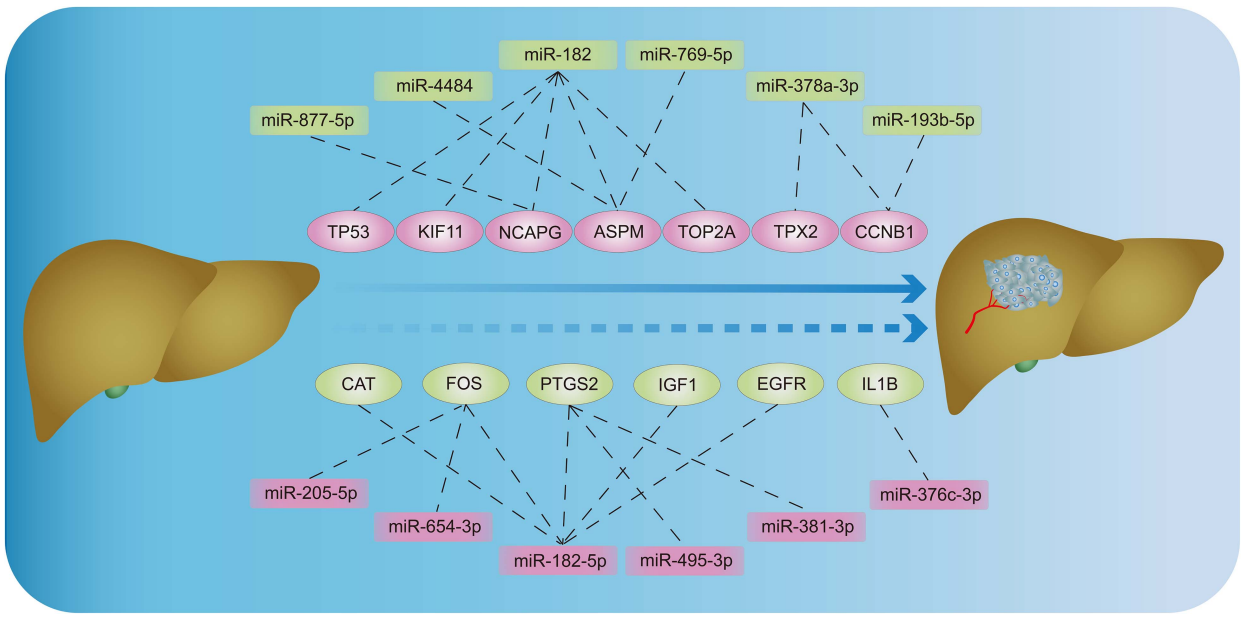

FIGURE 9 | The potential miRNA–mRNA interaction network involved in HB. HB: hepatoblastoma.

the molecular mechanisms of KLF4, PKM, and other TFs in the pathogenesis of HB development.

In the process of functional annotation enrichment analyses, GO-MF analysis on the basis of downregulated candidate genes identified enriched terms relating to oxidative stress injury such as oxidoreductase activity (**Figure 4E**). Similar to our findings, previous study also reported that oxidative stress injury plays an essential role in HB development (Tang et al., 2018). Our KEGG analysis of the upregulated candidate genes revealed that the PI3K/AKT pathway is another crucial pathway involved in HB (**Figure 5G**). In human embryonal tumors, the PI3K/AKT pathway is perhaps the most frequently reported pathway with hyperactivation (Vivanco and Sawyers, 2002; Zhang et al., 2004; Hartmann et al., 2006). In HB cells, it was reported that additive anti-tumor effects can be achieved after combination chemotherapy with PI3K inhibitors (Hartmann et al., 2009).

The cell cycle is composed of the interphase and the mitotic phase. The interphase, including G1, S, and G2 phases, is characterized by the synthesis of DNA and proteins (Norbury and Nurse, 1992). Uncontrolled cell cycle is recognized as a hallmark of tumor and, therefore, constitutes a major therapeutic target for the development of anti-tumor agents. Our KEGG enrichment analysis of the upregulated candidate genes identified cell cycle as the most significantly enriched pathway in HB (**Figure 5G**). For BP within the GO analysis, we found that the upregulated candidate genes played vital roles in multiple cell cycle events, including mitotic nuclear division and chromosome segregation (**Figure 5A**). Cellular defects that affect chromosome separation may increase aneuploidy, which in turn accelerate tumor progression (Pines, 2006). Moreover, other key events that interfere with the cell cycle were also observed in CC and MF within the GO analysis (**Figures 5C,E**).

In addition to the functional annotation enrichment analyses, almost all the upregulated hub genes obtained in this study, including CCNB1, KIF11, NCAPG, TOP2A, ASPM, and TPX2, have been previously reported to be implicated in regulating cell cycle progression. Indeed, the upregulated cell cycle-related proteins can accelerate cellular proliferation in human tumors (Malumbres and Barbacid, 2009). Moreover, the progression through distinct cell cycle phases is monitored by checkpoints that allow or prohibit the progression from one stage to another. Abnormal cell cycle check point hampers the detection and repair of genetic damage, leading to uncontrolled cell division and tumorigenesis. The majority of tumor cells exhibit cell cycle checkpoint defects, among which G1/S phase checkpoint defect is the most typical one (Zhao et al., 2012). CCNB1 is a regulatory protein involved in the G2/M cell cycle transition, and CCNB1 overexpression promotes chromatin bridging by suppressing separase activation (Nam and van-Deursen, 2014). In addition, the proliferation of human HB cell line HepG2 is suppressed by lycorine in a dose-dependent manner through downregulating cyclin A, CCNB1 and CDK1 (Liu et al., 2018). Centrosome linker refers to the protein that concatenates centrosomes during interphase. In the complex of mitotic spindle assembly, the dissolution of the centrosome linker is driven by KIF11 (Hata et al., 2019), a motor protein capable of hydrolyzing ATP. Besides, TPX2 can also regulate mitotic spindle assembly through kinetochore dependent microtubule nucleation and AURKA localization (Moss et al., 2009). NCAPG serves as the regulatory subunit of the condensin complex, which is essential for the conversion between interphase chromatin and mitotic chromosome in the presence of topoisomerases (Kimura et al., 2001). TOP2A, one type of nuclear enzyme, is critical for removing topological barriers left on DNA

during mitosis (Linka et al., 2007). ASPM is implicated in the regulation of mitotic spindle and the orchestration of mitotic processes. Also, the microtubule dynamics at spindle poles are modulated by ASPM with the help of the katanin complex (Jiang et al., 2017).

TP53 is famous for its tumor suppressive role in a variety of human tumors (Aubrey et al., 2018). Interestingly, our results demonstrated that TP53 plays an oncogenic role in HB development. Actually, there are two types of TP53, namely, mutant TP53 (mutp53) and wild type TP53 (wtp53). Missense mutation is the predominant form of mutp53 and expresses full-length mutp53 protein (Olivier et al., 2004). Mutp53 cannot activate the target genes of wtp53 or induce MDM2 expression, leading to the accumulation of mutp53 proteins in HB (Yue et al., 2017). Loss of heterozygosity (LOH) represents the phenomenon that mutp53 may inhibit the function of wtp53 and provide tumor cells with oncogenic functions (Baker et al., 1990). In addition, gain of function (GOF) is defined as the effect of mutp53 on promoting proliferation, metastasis, and anti-apoptosis of tumor cells. A greater number of metastatic tumors was observed for mice expressing mutp53 when compared with $TP53^{-/-}$ mice (Lang et al., 2004; Olive et al., 2004). The expression of mutp53 has been associated with chemoresistance in certain tumors due to GOF and the loss of wtp53 pro-apoptotic function (Yue et al., 2017). Patients with Li-Fraumeni syndrome and mutp53 were reported to have earlier development of tumors compared with those with Li-Fraumeni syndrome and TP53 deletion (Bougeard et al., 2008). Mutp53 can promote oncogenic cellular changes and alter cellular transcriptional profile. Therefore, to the best of our knowledge, the more likely scenario in this study was that most of over-expressed TP53 proteins in HB may belong to mutp53, thereby exerting oncogenic functions.

It is common that one dataset consists of a combination of paired and independent observations, and the terminology for this described scenario is "partially paired data" (Guo and Yuan, 2017). It should be noted that there are partially paired data in both datasets used for the present study. However, we did not take the inherent pairing structure into consideration in the DE analyses, which can lead to suboptimal results (Kuan and Huang, 2013). When analyzing partially paired data, the optimal pooled *t*-test, the test based on the modified maximum likelihood estimator, or the paired *t*-test, is to be recommended under different conditions in order to improve the statistical power (Guo and Yuan, 2017). Therefore, ignoring the matching for partially matched samples is one of the limitations of this study.

In conclusion, our results identify a variety of DE-miRNAs, TFs, and hub genes as potential regulators in the pathogenesis of HB. In addition, the miRNA–mRNA interaction network, PPI modules, and pathways may suggest putative diagnostic biomarkers or therapeutic targets for future HB theranostics.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the medical research ethics committee of Shanghai Children's Hospital, Shanghai Jiao Tong University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

ZL, JL, and LT study design. TC, LT, JC, XZ, JZ, and TG analysis and visualization of data. TC, QS, and LZ manuscript writing. TC, LT, and JL manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by National Natural Science Foundation of China (81871194).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.655703/full#supplementary-material

Supplementary Figure 1 | Data preprocessing of fetal-type tumor and normal liver samples in the GSE153089 dataset. Data of fetal-type tumor and normal liver samples **(A)** prior to and **(B)** after normalization. PCA of fetal-type tumor and normal liver samples **(C)** prior to and **(D)** after excluding the outlier sample. PCA, principal component analysis.

Supplementary Figure 2 | Data preprocessing of embryonal-type tumor and normal liver samples in the GSE153089 dataset. Data of embryonal-type tumor and normal liver samples **(A)** before and **(B)** after normalization. PCA of embryonal-type tumor and normal liver samples **(C)** prior to and **(D)** after excluding the outlier sample. PCA, principal component analysis.

Supplementary Figure 3 | Data preprocessing of HB and normal liver samples in the GSE131329 dataset. Data of HB and normal liver samples **(A)** before and **(B)** after normalization. PCA of HB and normal liver samples **(C)** prior to and **(D)** after excluding the outlier sample. HB: hepatoblastoma; PCA: principal component analysis.

REFERENCES

Aghajanzadeh, T., Tebbi, K., and Talkhabi, M. (2020). Identification of potential key genes and miRNAs involved in Hepatoblastoma pathogenesis and prognosis. *J. Cell Commun. Signal.* 15, 131–142. doi: 10.1007/s12079-020-00584-1

Aronson, D. C., Czauderna, P., Maibach, R., Perilongo, G., and Morland, B. (2014). The treatment of hepatoblastoma: its evolution and the current status as per the

SIOPEL trials. *J. Indian Assoc. Pediatr. Surg.* 19, 201–207. doi: 10.4103/0971-9261.142001

Aubrey, B. J., Kelly, G. L., Janic, A., Herold, M. J., and Strasser, A. (2018). How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ.* 25, 104–113. doi: 10.1038/cdd.2017.169

Baker, S. J., Preisinger, A. C., Jessup, J. M., Paraskeva, C., Markowitz, S., Willson, J. K., et al. (1990). p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res.* 50, 7717–7722.

Bougeard, G., Sesboüé, R., Baert-Desurmont, S., Vasseur, S., Martin, C., Tinat, J., et al. (2008). Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. *J. Med. Genet.* 45, 535–538. doi: 10.1136/jmg.2008.057570

Carvalho, B. S., and Irizarry, R. A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 26, 2363–2367. doi: 10.1093/bioinformatics/btq431

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8(Suppl. 4):S11. doi: 10.1186/1752-0509-8-s4-s11

Cui, X., Liu, X., Han, Q., Zhu, J., Li, J., Ren, Z., et al. (2019a). DPEP1 is a direct target of miR-193a-5p and promotes hepatoblastoma progression by PI3K/Akt/mTOR pathway. *Cell Death Dis.* 10:701. doi: 10.1038/s41419-019-1943-0

Cui, X., Wang, Z., Liu, L., Liu, X., Zhang, D., Li, J., et al. (2019b). The long non-coding RNA ZFAS1 Sponges miR-193a-3p to modulate Hepatoblastoma growth by targeting RALY via HGF/c-Met pathway. *Front. Cell. Dev. Biol.* 7:271. doi: 10.3389/fcell.2019.00271

Fan, Y., Siklenka, K., Arora, S. K., Ribeiro, P., Kimmins, S., and Xia, J. (2016). miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic Acids Res.* 44, W135–W141. doi: 10.1093/nar/gkw288

Guo, B. B., and Yuan, Y. (2017). A comparative review of methods for comparing means using partially paired data. *Stat. Methods Med. Res.* 26, 1323–1340. doi: 10.1177/0962280215577111

Hartmann, W., Digon-Söntgerath, B., Koch, A., Waha, A., Endl, E., Dani, I., et al. (2006). Phosphatidylinositol 3'-kinase/AKT signaling is activated in medulloblastoma cell proliferation and is associated with reduced expression of PTEN. *Clin. Cancer Res.* 12, 3019–3027. doi: 10.1158/1078-0432.Ccr-05-2187

Hartmann, W., Küchler, J., Koch, A., Friedrichs, N., Waha, A., Endl, E., et al. (2009). Activation of phosphatidylinositol-3′-kinase/AKT signaling is essential in hepatoblastoma survival. *Clin. Cancer Res.* 15, 4538–4545. doi: 10.1158/1078-0432.Ccr-08-2878

Hata, S., Pastor, P. A., Panic, M., Liu, P., Atorino, E., Funaya, C., et al. (2019). The balance between KIFC3 and EG5 tetrameric kinesins controls the onset of mitotic spindle assembly. *Nat. Cell Biol.* 21, 1138–1151. doi: 10.1038/s41556-019-0382-6

Hiyama, E., Ueda, Y., Kurihara, S., Kawashima, K., Ikeda, K., Morihara, N., et al. (2019). Gene expression profiling in hepatoblastoma cases of the Japanese study group for pediatric liver tumors-2 (JPLT-2) trial. *Eur. J. Mol. Cancer* 1, 2–8. doi: 10.31487/j.EJMC.2018.01.004

Honda, S., Chatterjee, A., Leichter, A. L., Miyagi, H., Minato, M., Fujiyoshi, S., et al. (2020). A microRNA cluster in the DLK1-DIO3 imprinted region on chromosome 14q32.2 Is dysregulated in metastatic hepatoblastomas. *Front. Oncol.* 10:513601. doi: 10.3389/fonc.2020.513601

Jiang, K., Rezabkova, L., Hua, S., Liu, Q., Capitani, G., Altelaar, A. F. M., et al. (2017). Microtubule minus-end regulation at spindle poles by an ASPM-katanin complex. *Nat. Cell Biol.* 19, 480–492. doi: 10.1038/ncb3511

Kimura, K., Cuvier, O., and Hirano, T. (2001). Chromosome condensation by a human condensin complex in Xenopus egg extracts. *J. Biol. Chem.* 276, 5417–5420. doi: 10.1074/jbc.C000873200

Kuan, P. F., and Huang, B. (2013). A simple and robust method for partially matched samples using the p-values pooling approach. *Stat. Med.* 32, 3247–3259. doi: 10.1002/sim.5758

Kuang, Y., Zheng, X., Zhang, L., Ai, X., Venkataramani, V., Kilic, E., et al. (2020). Adipose-derived mesenchymal stem cells reduce autophagy in stroke mice by extracellular vesicle transfer of miR-25. *J. Extracell. Vesicles* 10:e12024. doi: 10.1002/jev2.12024

Lambert, M., Jambon, S., Depauw, S., and David-Cordonnier, M. H. (2018). Targeting Transcription factors for cancer treatment. *Molecules* 23:1479. doi: 10.3390/molecules23061479

Lang, G. A., Iwakuma, T., Suh, Y. A., Liu, G., Rao, V. A., Parant, J. M., et al. (2004). Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. *Cell* 119, 861–872. doi: 10.1016/j.cell.2004.11.006

Lee, J., and Choi, C. (2017). Oncopression: gene expression compendium for cancer with matched normal tissues. *Bioinformatics* 33, 2068–2070. doi: 10.1093/bioinformatics/btx121

Linabery, A. M., and Ross, J. A. (2008). Trends in childhood cancer incidence in the U.S. (1992-2004). *Cancer* 112, 416–432. doi: 10.1002/cncr.23169

Linka, R. M., Porter, A. C., Volkov, A., Mielke, C., Boege, F., and Christensen, M. O. (2007). C-terminal regions of topoisomerase IIalpha and IIbeta determine isoform-specific functioning of the enzymes in vivo. *Nucleic Acids Res.* 35, 3810–3822. doi: 10.1093/nar/gkm102

Liu, L., Wang, L., Li, X., Tian, P., Xu, H., Li, Z., et al. (2019). Effect of miR-21 on apoptosis in hepatoblastoma cell through activating ASPP2/p38 signaling pathway in vitro and in vivo. *Artif. Cells Nanomed. Biotechnol.* 47, 3729–3736. doi: 10.1080/21691401.2019.1664561

Liu, W., Zhang, Q., Tang, Q., Hu, C., Huang, J., Liu, Y., et al. (2018). Lycorine inhibits cell proliferation and migration by inhibiting ROCK1/cofilin-induced actin dynamics in HepG2 hepatoblastoma cells. *Oncol. Rep.* 40, 2298–2306. doi: 10.3892/or.2018.6609

Liu, Y., Bailey, J. T., Abu-Laban, M., Li, S., Chen, C., Glick, A. B., et al. (2020). Photocontrolled miR-148b nanoparticles cause apoptosis, inflammation and regression of Ras induced epidermal squamous cell carcinomas in mice. *Biomaterials* 256:120212. doi: 10.1016/j.biomaterials.2020.120212

Lou, W., Liu, J., Ding, B., Chen, D., Xu, L., Ding, J., et al. (2019). Identification of potential miRNA-mRNA regulatory network contributing to pathogenesis of HBV-related HCC. *J. Transl. Med.* 17:7. doi: 10.1186/s12967-018-1761-7

Malumbres, M., and Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* 9, 153–166. doi: 10.1038/nrc2602

Mizutani, A., Koinuma, D., Tsutsumi, S., Kamimura, N., Morikawa, M., Suzuki, H. I., et al. (2011). Cell type-specific target selection by combinatorial binding of Smad2/3 proteins and hepatocyte nuclear factor 4alpha in HepG2 cells. *J. Biol. Chem.* 286, 29848–29860. doi: 10.1074/jbc.M110.217745

Moss, D. K., Wilde, A., and Lane, J. D. (2009). Dynamic release of nuclear RanGTP triggers TPX2-dependent microtubule assembly during the apoptotic execution phase. *J. Cell Sci.* 122, 644–655. doi: 10.1242/jcs.037259

Nakra, T., Roy, M., Yadav, R., Agarwala, S., Jassim, M., Khanna, G., et al. (2020). Cytomorphology of hepatoblastoma with histological correlation and role of SALL4 immunocytochemistry in its diagnosis, subtyping, and prognostication. *Cancer Cytopathol.* 128, 190–200. doi: 10.1002/cncy.22231

Nam, H. J., and van-Deursen, J. M. (2014). Cyclin B2 and p53 control proper timing of centrosome separation. *Nat. Cell Biol.* 16, 538–549. doi: 10.1038/ncb2952

Norbury, C., and Nurse, P. (1992). Animal cell cycles and their control. *Annu. Rev. Biochem.* 61, 441–470. doi: 10.1146/annurev.bi.61.070192.002301

Olive, K. P., Tuveson, D. A., Ruhe, Z. C., Yin, B., Willis, N. A., Bronson, R. T., et al. (2004). Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell* 119, 847–860. doi: 10.1016/j.cell.2004.11.004

Olivier, M., Hussain, S. P., Caron de Fromentel, C., Hainaut, P., and Harris, C. C. (2004). TP53 mutation spectra and load: a tool for generating hypotheses on the etiology of cancer. *IARC Sci. Publ.* 157, 247–270.

Pines, J. (2006). Mitosis: a matter of getting rid of the right protein at the right time. *Trends Cell Biol.* 16, 55–63. doi: 10.1016/j.tcb.2005.11.006

Roy, S., Bantel, H., Wandrer, F., Schneider, A. T., Gautheron, J., Vucur, M., et al. (2017). miR-1224 inhibits cell proliferation in acute liver failure by targeting the antiapoptotic gene Nfib. *J. Hepatol.* 67, 966–978. doi: 10.1016/j.jhep.2017.06.007

Schnater, J. M., Köhler, S. E., Lamers, W. H., von Schweinitz, D., and Aronson, D. C. (2003). Where do we stand with hepatoblastoma? A review. *Cancer* 98, 668–678. doi: 10.1002/cncr.11585

Semeraro, M., Branchereau, S., Maibach, R., Zsiros, J., Casanova, M., Brock, P., et al. (2013). Relapses in hepatoblastoma patients: clinical characteristics and outcome–experience of the international childhood liver tumour strategy Group (SIOPEL). *Eur. J. Cancer* 49, 915–922. doi: 10.1016/j.ejca.2012.10.003

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Spector, L. G., and Birch, J. (2012). The epidemiology of hepatoblastoma. *Pediatr. Blood Cancer* 59, 776–779. doi: 10.1002/pbc.24215

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003

Tan, J., Xu, W., Lei, L., Liu, H., Wang, H., Cao, X., et al. (2020). Inhibition of aurora kinase a by alisertib reduces cell proliferation and induces apoptosis and autophagy in huh-6 human hepatoblastoma cells. *Oncol. Targets Ther.* 13, 3953–3963. doi: 10.2147/ott.S228656

Tang, S., Liu, H., Yin, H., Liu, X., Peng, H., Lu, G., et al. (2018). Effect of 2, 2', 4, 4'-tetrabromodiphenyl ether (BDE-47) and its metabolites on cell viability, oxidative stress, and apoptosis of HepG2. *Chemosphere* 193, 978–988. doi: 10.1016/j.chemosphere.2017.11.107

Tian, L., Chen, T., Lu, J., Yan, J., Zhang, Y., Qin, P., et al. (2021). Integrated protein-protein interaction and weighted gene co-expression network analysis uncover three key genes in hepatoblastoma. *Front. Cell. Dev. Biol.* 9:631982. doi: 10.3389/fcell.2021.631982

Tokumoto, M., Tsuruya, K., Fukuda, K., Kanai, H., Kuroki, S., Hirakata, H., et al. (2003). Parathyroid cell growth in patients with advanced secondary hyperparathyroidism: vitamin D receptor and cyclin-dependent kinase inhibitors, p21 and p27. *Nephrol. Dial. Transplant.* 18 Suppl 3, iii9–iii12. doi: 10.1093/ndt/gfg1003

Tong, Z., Cui, Q., Wang, J., and Zhou, Y. (2019). TransmiR v2.0: an updated transcription factor-microRNA regulation database. *Nucleic Acids Res.* 47, D253–D258. doi: 10.1093/nar/gky1023

Vivanco, I., and Sawyers, C. L. (2002). The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat. Rev. Cancer* 2, 489–501. doi: 10.1038/nrc839

von Frowein, J., Hauck, S. M., Kappler, R., Pagel, P., Fleischmann, K. K., Magg, T., et al. (2018). MiR-492 regulates metastatic properties of hepatoblastoma via CD44. *Liver Int.* 38, 1280–1291. doi: 10.1111/liv.13687

Wagner, A., Schwarzmayr, T., Häberle, B., Vokuhl, C., Schmid, I., von Schweinitz, D., et al. (2020). SP8 promotes an aggressive phenotype in hepatoblastoma via FGF8 activation. *Cancers* 12:2294. doi: 10.3390/cancers12082294

Wang, G., Guo, X., Cheng, L., Chu, P., Chen, M., Chen, Y., et al. (2019). An integrated analysis of the circRNA-miRNA-mRNA network reveals novel insights into potential mechanisms of cell proliferation during liver regeneration. *Artif. Cells Nanomed. Biotechnol.* 47, 3873–3884. doi: 10.1080/21691401.2019.1669623

Weng, Y. S., Tseng, H. Y., Chen, Y. A., Shen, P. C., Al Haq, A. T., Chen, L. M., et al. (2019). MCT-1/miR-34a/IL-6/IL-6R signaling axis promotes EMT progression, cancer stemness and M2 macrophage polarization in triple-negative breast cancer. *Mol. Cancer* 18:42. doi: 10.1186/s12943-019-0988-0

Ye, M., He, J., Zhang, J., Liu, B., Liu, X., Xie, L., et al. (2021). USP7 promotes hepatoblastoma progression through activation PI3K/AKT signaling pathway. *Cancer Biomark.* 31, 107–117. doi: 10.3233/CBM-200052

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16, 284–287. doi: 10.1089/omi.2011.0118

Yue, X., Zhao, Y., Xu, Y., Zheng, M., Feng, Z., and Hu, W. (2017). Mutant p53 in cancer: accumulation, gain-of-function, and therapy. *J. Mol. Biol.* 429, 1595–1606. doi: 10.1016/j.jmb.2017.03.030

Zeng, Y. X., Somasundaram, K., and el-Deiry, W. S. (1997). AP2 inhibits cancer cell growth and activates p21WAF1/CIP1 expression. *Nat. Genet.* 15, 78–82. doi: 10.1038/ng0197-78

Zhang, C. J., and Du, H. J. (2017). Screening key miRNAs for human hepatocellular carcinoma based on miRNA-mRNA functional synergistic network. *Neoplasma* 64, 816–823. doi: 10.4149/neo_2017_602

Zhang, J., Hu, S., Schofield, D. E., Sorensen, P. H., and Triche, T. J. (2004). Selective usage of D-type cyclins by Ewing's tumors and rhabdomyosarcomas. *Cancer Res.* 64, 6026–6034. doi: 10.1158/0008-5472.Can-03-2594

Zhang, J., Liu, P., Tao, J., Wang, P., Zhang, Y., Song, X., et al. (2019). TEA domain transcription factor 4 is the major mediator of yes-associated protein oncogenic activity in mouse and human hepatoblastoma. *Am. J. Pathol.* 189, 1077–1090. doi: 10.1016/j.ajpath.2019.01.016

Zhang, L., Jin, Y., Zheng, K., Wang, H., Yang, S., Lv, C., et al. (2018). Whole-genome sequencing identifies a novel variation of WAS Gene coordinating with heterozygous germline mutation of APC to enhance hepatoblastoma oncogenesis. *Front. Genet.* 9:668. doi: 10.3389/fgene.2018.00668

Zhang, Y., Zhao, Y., Wu, J., Liangpunsakul, S., Niu, J., and Wang, L. (2018). MicroRNA-26-5p functions as a new inhibitor of hepatoblastoma by repressing lin-28 homolog B and aurora kinase a expression. *Hepatol. Commun.* 2, 861–871. doi: 10.1002/hep4.1185

Zhao, L., Bode, A. M., Cao, Y., and Dong, Z. (2012). Regulatory mechanisms and clinical perspectives of miRNA in tumor radiosensitivity. *Carcinogenesis* 33, 2220–2227. doi: 10.1093/carcin/bgs235

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership