LONG NON-CODING RNAS AND IMMUNITY

EDITED BY: Adam Williams and Jorge Henao-Mejia
PUBLISHED IN: Frontiers in Immunology

IUIS

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-917-2
DOI 10.3389/978-2-88966-917-2

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

LONG NON-CODING RNAS AND IMMUNITY

Topic Editors:
Adam Williams, Jackson Laboratory, United States
Jorge Henao-Mejia, University of Pennsylvania, United States

We acknowledge the initiation and support of this Research Topic by the International Union of Immunological Societies (IUIS). We hereby state publicly that the IUIS has had no editorial input in articles included in this Research Topic, thus ensuring that all aspects of this Research Topic are evaluated objectively, unbiased by any specific policy or opinion of the IUIS.

Citation: Williams, A., Henao-Mejia, J., eds. (2021). Long Non-Coding RNAs and Immunity. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-917-2

Table of Contents

frontiers in Immunology

EDITORIAL
published: 09 October 2019
doi: 10.3389/fimmu.2019.02378

Editorial: Long Non-coding RNAs and Immunity

Grace J. Kwon[1,2], *Jorge Henao-Mejia*[3,4,5] *and Adam Williams*[1,2]*

[1] Department of Genetics and Genome Sciences, University of Connecticut Health Center, Farmington, CT, United States, [2] The Jackson Laboratory for Genomic Medicine, Farmington, CT, United States, [3] Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, United States, [4] Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States, [5] Division of Transplant Immunology, Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, United States

Keywords: long non-coding RNA, immunity, circular RNA, X-chromosome inactivation, viral infection, rickettsial infection, GWAS

Editorial on the Research Topic

Long Non-coding RNAs and Immunity

Next-generation sequencing has shown that the majority of the human genome does not encode proteins, yet many such "non-coding" regions are actively transcribed into RNA. In particular, long non-coding RNAs (lncRNAs) have received considerable interest for their essential roles in numerous biological processes. Defined as non-protein-coding transcripts longer than 200 nucleotides, lncRNAs are frequently expressed in a tissue-specific or developmental stage-specific manner (1, 2). Multiple studies have shown lncRNAs to be dysregulated in disease, and genome-wide association studies (GWAS) have also identified numerous single nucleotide polymorphisms (SNPs) within lncRNAs (3–5). These features distinguish them as attractive candidates for therapeutic targets or biomarkers. While still in its infancy, the dissection of lncRNA biology has unveiled critical roles in immune cell development and function. However, as the functions of most lncRNAs are yet to be determined, the field must fill a considerable void to understand the breadth of their roles in immunity. This Research Topic focuses on how lncRNAs contribute to immune function in a variety of different contexts, and what key questions drive this rapidly expanding field.

DOSAGE COMPENSATION AND X-LINKED IMMUNE GENES

Females have a more responsive immune system than males (6, 7), potentially driving a higher predisposition to autoimmune diseases such as systemic lupus erythematosus (SLE) (8). The X chromosome carries a significant number of immune-related genes, but how this translates to sex disparities in immune diseases remains to be determined. *Xist* is one of the best characterized lncRNAs, and functions as an effector of dosage compensation in female cells. *Xist* maintains inactivation of one X chromosome by binding along its length and recruiting repressive factors to silence transcription. Immune cells, however, regulate X chromosome inactivation through diverse mechanisms that are not limited to the use of *Xist* RNA, as recently revealed by Syrett et al. For instance, the inactivated X in female murine plasmacytoid dendritic cells (pDCs) lacks both *Xist* RNA and the repressive chromatin modification H3K27me3. Interestingly, pDCs from female mice that spontaneously develop SLE-like disease show increased biallelic expression of the X-linked gene *Tlr7*, as compared to healthy female mice. Further delineating the mechanisms of X chromosome inactivation in immune cells will be necessary before we can develop a complete understanding of how sex and sex-linked genes contribute to autoimmune diseases.

OPEN ACCESS

Edited and reviewed by:
Francesca Granucci,
University of Milano Bicocca, Italy

***Correspondence:**
Adam Williams
adam.williams@jax.org

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *03 September 2019*
Accepted: *23 September 2019*
Published: *09 October 2019*

Citation:
Kwon GJ, Henao-Mejia J and Williams A (2019) Editorial: Long Non-coding RNAs and Immunity. Front. Immunol. 10:2378. doi: 10.3389/fimmu.2019.02378

LNCRNAS IN RICKETTSIAL INFECTION

The *Rickettsia* genus of Gram-negative obligate intracellular bacteria is transmitted to humans via arthropod vectors, including ticks, fleas, and lice, and can infect the vascular endothelium to cause illnesses such as Rocky Mountain spotted fever. There are increasing concerns over the prevalence of *Rickettsia* due to climate change and the spread of vector populations. Understanding rickettsial infection and developing novel therapeutics to mitigate its spread through the vascular endothelium will thus be crucial in the near future. Chowdhury et al. have begun to address the roles of lncRNAs in the host response to *Rickettsia* species using a mouse model of *R. conorii* infection. The authors identified two enhancer lncRNAs affecting expression of nearby genes, *Id2* and *Apol10b*, which have not been previously studied in rickettsial infection. Furthermore, these two lncRNAs exhibit differential expression in infected macrophages compared to endothelial cells, emphasizing the cell type-specificity of lncRNAs. The resulting genome-wide analysis of lncRNA expression during rickettsial infection establishes a foundation to further study the host immune response to *Rickettsia* species and reveals cell type-specific signaling pathways mediated by enhancer lncRNAs.

VIRAL AND HOST LNCRNAS IN ANTIVIRAL IMMUNITY

Some of the most interesting advancements in our understanding of lncRNAs in immunity come from the study of their roles in viral infection. While it would be intuitive to assume host lncRNAs exist to suppress viral infections, it has been shown that host lncRNAs can actually increase viral replication and pathogenesis during certain infections (9). This begets the need to further examine what purpose both host and viral lncRNAs serve, and how it may change in the context of non-infected and infected states. Wang reviews in detail the known lncRNAs involved in multiple steps of the viral life cycle and infection.

LNCING GWAS SNPS TO DISEASE PATHOGENESIS

GWAS have provided a more complete understanding of complex diseases. Of the more than 70,000 variant-trait associations cataloged from GWAS (10), a majority of these map to non-coding regions of the genome, including those containing lncRNAs. Despite this, the impact of only a few SNPs on lncRNA function has been described. A major limitation in the field is that the function of most lncRNAs is unknown. In addition, while SNPs in lncRNAs likely affect their secondary and tertiary structure, we cannot currently predict how this will alter their function. Castellanos-Rubio and Ghosh review the disease-associated lncRNA SNPs that have been described thus far. Further inquiry and development of tools to interpret lncRNA structure and function will be necessary before the impact of lncRNA SNPs (and other genetic variants) on disease can be fully realized.

CIRCULAR RNAS AS A NEW CLASS OF IMMUNE REGULATORS

While lncRNAs typically display features associated with linear messenger RNA, such as 5′ capping, alternative splicing, and polyadenylation, non-coding RNA loci can also be transcribed to form circular RNA (circRNAs). These circRNAs form closed loops without 5′ capping or polyadenylation and can contain multiple exons. Similar to lncRNAs, circRNAs are frequently expressed in a tissue-specific manner, undergo chemical modifications, and function through varied mechanisms. Yang et al. review the mechanisms of circRNA biogenesis, summarize the current understanding of their function in immunity, and present approaches and limitations to studying circRNA function.

CONCLUSION

lncRNAs are now an established class of functional molecules with physiological relevance shown in nearly every human organ system. Their roles in immunity have been identified in immune cell differentiation, maintenance, and effector function. Given that only a fraction of lncRNAs have been functionally characterized a significant wealth of insight awaits discovery. Moreover, as personalized medicine begins to drive the use of genetic tools in the clinic, we can expect the identification of clinically-relevant lncRNA variants and the development of therapeutics that target lncRNA function.

AUTHOR CONTRIBUTIONS

AW and GK wrote and edited the manuscript. JH-M edited the manuscript.

FUNDING

GK was supported by NIH F30-HL143986. AW was supported by NIH R01AI141609, R21AI13522, and R21AI133440.

REFERENCES

1. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* (2012) 22:1775–89. doi: 10.1101/gr.132159.111
2. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* (2011) 25:1915–27. doi: 10.1101/gad.17446611
3. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, et al. A long noncoding RNA associated with susceptibility to celiac disease. *Science.* (2016) 352:91–5. doi: 10.1126/science.aad0467

4. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet*. (2013) 9:e1003201. doi: 10.1371/journal.pgen.1003201
5. Mirza AH, Kaur S, Brorsson CA, Pociot F. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS ONE*. (2014) 9:e105723. doi: 10.1371/journal.pone.0105723
6. Ghosh S, Klein RS. Sex drives dimorphic immune responses to viral infections. *J Immunol*. (2017) 198:1782–90. doi: 10.4049/jimmunol.1601166
7. Spolarics Z. The X-files of inflammation: cellular mosaicism of X-linked polymorphic genes and the female advantage in the host response to injury and infection. *Shock*. (2007) 27:597–604. doi: 10.1097/SHK.0b013e31802e40bd
8. Libert C, Dejager L, Pinheiro I. The X chromosome in immune functions: when a chromosome makes a difference. *Nat Rev Immunol*. (2010) 10:594–604. doi: 10.1038/nri2815
9. Wang Z, Fan P, Zhao Y, Zhang S, Lu J, Xie W, et al. NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription. *Cell Mol Life Sci*. (2017) 74:1117–31. doi: 10.1007/s00018-016-2398-4
10. Buniello A, MacArthuer JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. (2019) 47:D1005–D1012. doi: 10.1093/nar/gky1120

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 Immunology

REVIEW
published: 18 December 2018
doi: 10.3389/fimmu.2018.02977

Circular RNAs and Their Emerging Roles in Immune Regulation

Lan Yang[1], *Jinrong Fu*[1] *and Yufeng Zhou*[1,2]*

[1] *Children's Hospital and Institute of Biomedical Sciences, Fudan University, Shanghai, China,* [2] *Key Laboratory of Neonatal Diseases, Ministry of Health, Shanghai, China*

Circular ribonucleic acid (RNA) molecules (circRNAs) are covalently closed loop RNA molecules with no 5′ end caps or 3′ poly (A) tails, which are generated by back-splicing. Originally, circRNAs were considered to be byproducts of aberrant splicing. However, in recent years, development of high-throughput sequencing has led to gradual recognition of functional circRNAs, and increasing numbers of studies have elucidated their roles in cancer, neurologic diseases, and cardiovascular disorders. Nevertheless, studies of the functions of circRNAs in the immune system are relatively scarce. In this review, we detail relevant research on the biogenesis and classification of circRNAs, describe their functional mechanisms and approaches to their investigation, and summarize recent studies of circRNA function in the immune system.

Keywords: circRNA, biogenesis, function, research approaches, immune regulation

OPEN ACCESS

Edited by:
Adam Williams,
Jackson Laboratory, United States

Reviewed by:
Grace Chen,
Yale University, United States
Yi Zhao,
Sichuan University, China

***Correspondence:**
Yufeng Zhou
yfzhou1@fudan.edu.cn

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *18 September 2018*
Accepted: *04 December 2018*
Published: *18 December 2018*

Citation:
Yang L, Fu J and Zhou Y (2018) Circular RNAs and Their Emerging Roles in Immune Regulation. Front. Immunol. 9:2977. doi: 10.3389/fimmu.2018.02977

INTRODUCTION

Circular RNA (circRNA) is a covalently closed loop molecular form of RNA, and was discovered more than two decades ago (1, 2). Initially, circRNA was considered to be an aberrant byproduct of splicing (3–6). Recently, numerous circRNAs have been identified as a consequence of rapid developments in bioinformatics and high-throughput sequencing. Jeck et al. detected >25,000 circRNAs in fibroblasts using a genome-wide RNase R enrichment strategy (7). Memczak et al. identified 1950 circRNAs in humans, 1903 in mice, and 724 in *Caenorhabditis elegans* using RNA-sequencing data combined with analyses of the human leukocyte database (8). circRNAs are also expressed in fungi, plants, and protists (9–12).

Current research on circRNAs focuses on their role in cancer, neurologic diseases, and cardiovascular disorders. In this review, we cover relevant research on the biogenesis and classification of circRNAs and their functional mechanisms and methodological approaches to their study, along with summarizing recent investigations of the roles of circRNAs in the immune system.

BIOGENESIS OF CIRCRNAS

Rather than canonical splicing, circRNAs are generated through back-splicing (13, 14) (**Figure 1**). Back-splicing can be accompanied by transcription (15) or may occur after transcription has been completed (16). Three main mechanisms have been reported to produce circRNAs: exon skipping, intron pairing-driven circularization, and RNA binding protein (RBP)-driven circularization.

Exon Skipping

Skipping of exons leads to the formation of a lariat structure containing the skipped exon. If introns are spliced before the lariat is unraveled by debranching enzymes, a stable exon-containing circRNA

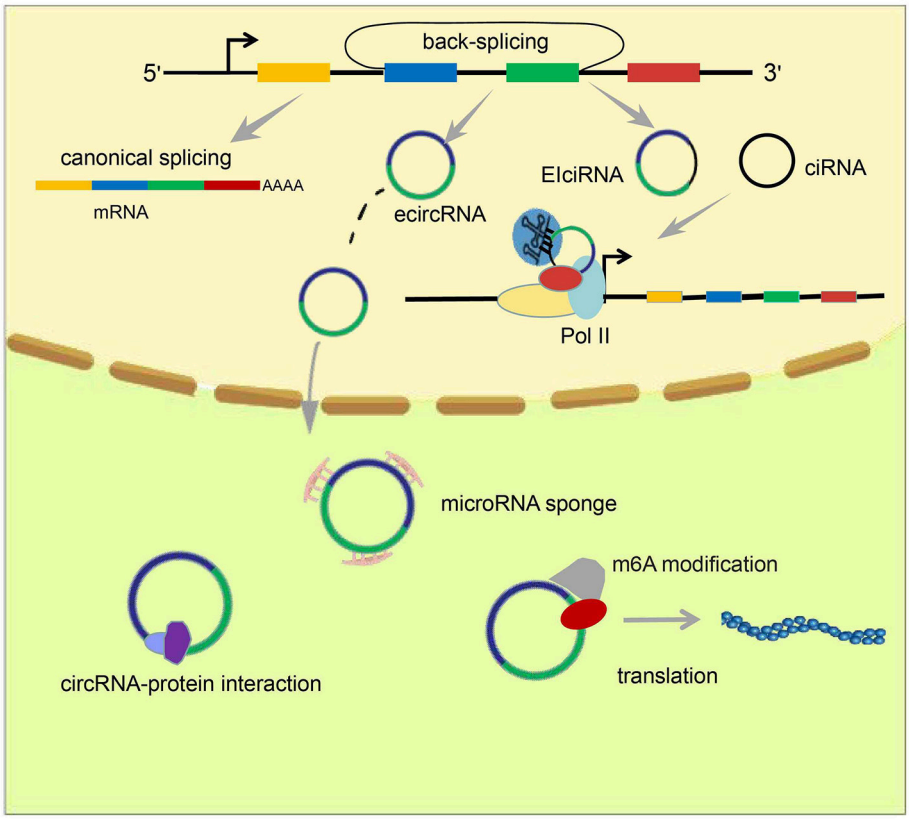

FIGURE 1 | Formation and classification of circRNAs and models of circRNA function**.** circRNAs are formed by back-splicing. There are three major types of circRNA: ecircRNA, ciRNA, and EIciRNA. EcircRNAs function mainly in the cytoplasm through the "microRNA sponge" mechanism, whereas EIciRNA and ciRNA accumulate in the nucleus and facilitate transcription of their parent genes *via cis*-regulatory effects. In addition, some circRNAs have been reported to act by interacting with or encoding proteins.

will be produced (7), along with a linear transcript excluding the skipped exon(s). This mechanism was an early understanding of how circRNAs are formed because the linear transcripts that are produced alongside circRNAs were described some years ago (17, 18).

Intron Pairing-Driven Circularization

Intron pairing-driven circularization is based on reverse complementary matches (RCMs) within flanking introns, and was established independently of the exon-skipping mechanism of circularization. RCMs can induce base-pairing between flanking introns, promoting a hairpin formation, bringing the 5′ and 3′ termini of an exon into spatial proximity, and inducing "head-to-tail" splicing. Intriguingly, the protein adenosine deaminases acting on RNA (ADAR) also participate in this process, together with RCMs. ADAR can unzip double-stranded RNA molecules by converting adenosine residues within them into creatinine molecules, thereby reducing the formation of circRNAs (19). Furthermore, Jeck et al. determined that, unlike sequence regions involved in generation of non-circular transcripts, the 200 bp upstream or downstream of a back-splice site contain canonical complementary ALU repeats, suggesting that intronic pairing may drive the circularization of exonic circRNAs (ecircRNAs). Jeck et al. also found that the length of introns flanking ecircRNAs are greater than those flanking linear RNA exons (7).

RBP-Driven Circularization

In addition to the involvement of introns and exons, some RNA-binding proteins also contribute to the formation of circRNAs. For example, in *Drosophila*, the splicing factor muscleblind (MBL) drives the circularization of the second exon of *MBL* to produce circMBL, by binding specifically to MBL binding sites within introns flanking circMBL sequences (15). More importantly, the interaction between MBL and circMBL contributes to regulation of the levels of MBL protein. When MBL is present in excess, it reduces the production of *MBL* mRNA by promoting circMBL formation. A protein called "quaking" (QKI) was identified by Conn et al. and shown to promote the production of circRNAs during human epithelial–mesenchymal transition by binding to a motif within circRNA flanking introns (20). RBM20 is also an RNA-binding protein that regulates numerous cardiac-specific gene-editing processes. Mutation of RMB20 is involved in dilated cardiomyopathy

through its effects on the generation of circRNAs from *Titin* (21). The RNA-binding protein FUS can mediate circRNA formation by mediating RNA back-splicing in neurons (22). Moreover, HNRNPL, an RNA splicing factor, participates in regulation of circRNA formation in prostate cancer (23). In conclusion, the formation of circRNAs is dependent upon the regulation of cis elements and trans-factors (24).

CIRCRNA CLASSIFICATION

Studies have identified three main types of circRNA: ecircRNA, circular intronic RNA (ciRNA), and exon–intron circRNA (EIciRNA) (**Figure 1**). Published data suggest that ecircRNAs function mainly through the microRNA "sponge" mechanism, proposed first by Memczak et al. who found there were 63 microRNA-7 binding sites on CDR1as, and hence designated CDR1as a "microRNA sponge." EcircRNA can enhance levels of microRNA target genes through adsorption of microRNA molecules. Unlike ecircRNAs, intron-containing circRNAs (ciRNAs or EIciRNAs), in general, reside in the nucleus and regulate gene transcription (25–27). Chen et al. found that the function of ci-ankrd52, which is derived from the second intron of *ANKRD52*, may depend on a consensus motif containing a 7-nucleotide (nt) GU-rich element near its 5′ splice site and an 11-nt C-rich element close to the branchpoint to avoid being debranched. Ci-ankrd52 accumulates mainly in the nucleus and promotes transcription of *ANKRD52* through a *cis*-regulatory effect of RNAPol II (25).

CIRCRNA CHARACTERISTICS

circRNAs are not degraded readily by RNase R, and are more stable than their cognate linear mRNAs, with half-lives exceeding 48 h (7). circRNAs also exhibit a degree of conservation between species. Jeck et al. identified ecircRNAs from 14.4% of actively transcribed genes in human fibroblasts. Furthermore, they identified 69 ecircRNAs from murine testes that mapped to precisely orthologous chromosome locations compared with human circRNAs (7).

Numerous circRNAs exhibit specificity to tissue, developmental period, or cell type. Memczak et al. found that different cells express specific circRNAs through identification of circRNAs in four human cell types: cluster of differentiation (CD)19$^+$ leukocytes, CD34$^+$ leukocytes, neutrophils, and HEK293 cells. Analogously, they noted that several nematode circRNAs were expressed in oocytes but absent in 1- or 2-cell embryos (8).

Although most circRNAs do not have the potential for translation (28), several recent studies have confirmed that circRNAs can be translated into proteins. Legnini et al. screened differentially expressed circRNAs in Duchenne muscular dystrophy by RNA sequencing. They determined that circ-ZNF609 can affect muscle formation through regulation of myoblast proliferation. Surprisingly, they also found that circ-ZNF609 can encode a protein (29), but whether this protein has a role in myoblast proliferation is not known. Yang et al. demonstrated that modification of N^6-methyladenosine (m6A), which is induced by the methyltransferase METTL3/14 and suppressed by the demethylase FTO, promotes protein translation through recruiting an initiation factor, eIF4G2, and a m6A reader, YTHDF3. Furthermore, they found that m6A-containing circRNAs, with the potential to be translated into proteins, are common in the human transcriptome (30). Yang et al. also demonstrated that circ-FBXW7 can be translated into the protein FBXW7-185aa, thereby co-regulating the stability of c-Myc, along with its parent gene-encoded protein, FBXW7, to inhibit the progress of malignant glioma (31). Pamudurti et al. described the discovery of numerous circRNA-translated proteins or peptides based on ribosome "footprinting" experiments in *Drosophila* brains. They demonstrated that ribosomes can bind at the stop codon of circMbl and that a protein encoded by circMbl can be identified by protein spectroscopy (32).

Several reports have detailed many chemical modifications present on DNA and RNA. Surprisingly, recent breakthrough investigations of chemical modification of circRNAs have demonstrated that a m6A modification occurs in circRNA. This was discovered first by Yang et al. who further demonstrated that this modification promotes protein translation through recruitment of the initiation factor eIF4G2, and the m6A reader YTHDF3 (30) (**Figure 1**). Zhou et al. further confirmed the existence of the m6A modification in circRNAs, and also proposed some characteristics of m6A-modified circRNAs (33).

MECHANISMS OF CIRCRNA FUNCTION

MicroRNA Sponge

MicroRNAs produced by Dicer processing enzymes from single-stranded RNA precursors with a hairpin structure of ∼70–90 nt are single-stranded small RNAs of ∼21–23 nt (34, 35). Through initiation of RNA-induced silencing complex (RISC), mRNAs are degraded or their translation is hindered by base pairing with target gene transcripts (36, 37). Memczak et al. was the first to propose the microRNA sponge model. They found that there were 63 microRNA-7 binding sites within the circular transcript CDR1as (ciRS-7). Hansen et al. determined that a testes-specific circRNA called Sry can also serve as a sponge for microRNA-138 (38). Recent studies have shown that various circRNAs can function to adsorb microRNAs, thereby releasing target mRNAs (**Figure 1**). This process is associated with several diseases. In 2017, Han et al. demonstrated that circMTO1 upregulates p21 by competitively binding to microRNA-9, thereby inhibiting the proliferation of hepatocellular carcinoma (39). Zhong et al. clarified that the circRNA MYLK acts as a competing endogenous RNA (ceRNA) to bind directly to microRNA-29a, thereby promoting the development of bladder cancer by activating VEGFA/VEGFR2 signaling pathway and the Ras-Erk pathway (40). The circRNA circEPSTI1 was discovered to be an endogenous competitive RNA and sponge for microRNAs. Through binding to microRNA-4753 and microRNA-6809, circEPSTI1 upregulates expression of B cell CLL/lymphoma 11A (BCL11A), promotes the proliferation of triple-negative breast cancer cells, and inhibits their apoptosis (41).

Interactions Between circRNAs and Proteins

Rather than acting as microRNA sponges in the cytoplasm, some circRNAs act by interacting with corresponding proteins (**Figure 1**). In 2016, circ-foxo3 was found to impede cell-cycle progression by forming a ternary complex with p21 and Cdk2 proteins (42). In addition, another circRNA, circ-Amotl1, which is highly expressed in cancerous cell lines, can increase nuclear retention of the oncogenic protein c-myc to promote its stability, and increase its affinity for binding to several promoters, thereby upregulating c-myc targets such as HIF1α, Cdc25a, ELK-1, and JUN (43). This observation reveals a novel function of circRNAs in tumorigenesis. Conceivably, there are circRNAs other than circ-Amotl1 that act by similar mechanisms. In 2017, Abdelmohsen et al. was the first to report competition between a circRNA and its cognate mRNA for the RNA-binding protein HuR, which has been studied extensively and can regulate protein expression through interaction with a wide range of RNAs. They proposed that high levels of circPABPN1, a circRNA derived from *PABPN1*, repress HuR binding to PABPN1 mRNA through binding to HuR itself, causing a reduction in PABPN1 mRNA translation (44). This mechanism, based on the interaction between a circRNA and its cognate mRNA, provided new insights and spurred further investigation of the roles of circRNAs in the nucleus.

Regulation of Transcription in *cis*

Multiple non-coding RNAs, including HOTAIR and MALAT1, are known to regulate gene transcription in *trans*; that is, they influence the transcription of genes other than their parent genes. However, there have been studies that have shown circRNAs, mainly intron-containing circRNAs (ciRNAs or EIciRNAs), can have *cis*-regulatory effects on gene expression. In 2013, Chen et al. reported that ci-ankrd52, which is generated from the second intron of ANKRD52, accumulates primarily in the nucleus and promotes transcription of ANKRD52 *via* the *cis*-regulatory effects of RNAPol II (25). In 2014, Li et al. proposed a model for the *cis*-regulatory effects of EIciRNAs based on their findings from investigation of two EIciRNAs: circEIF3J and circPAIP2. They proposed that EIciRNAs may interact with proteins such as U1 snRNP *via* RNA–RNA interplay between U1 snRNA and EIciRNA. Then, EIciRNA–U1 snRNP complexes could interact with Pol II at the promoter regions of parental genes, thereby enhancing their transcription (**Figure 1**). This phenomenon produces a positive feedback loop because, once transcription has been initiated, EIciRNA generation will increase, further promoting gene transcription (26).

circRNAs as Biomarkers

In addition to being detected inside cells, circRNA has also been reported to be present in extracellular fluids. Li et al. enriched exosomes in serum samples from patients suffering from colon cancer to examine circRNA expression in exosomes. Compared with healthy controls, hundreds of circRNAs were expressed differentially in serum exosomes from patients with colon cancer (45). In 2016, Guarnerio et al. was the first to report that chromosomal translocations can produce fusion circRNAs in acute promyelocytic leukemia. Furthermore, using an MLL/AF9-AML model, they found that f-circM9 could contribute to progression of acute myeloid leukemia (46). Based on those findings, Tan et al. confirmed that the common fusion gene in non-small-cell lung cancer (NSCLC), *EML4-ALK*, can produce a fused circRNA called F-circEA, and demonstrated that F-circEA promotes the proliferation and migration of cells. F-circEA was also detected in the serum of *EML4-ALK*-positive patients, indicating that this circRNA is highly likely to be useful as a diagnostic marker for *EML4-ALK*-positive NSCLC (47). With further research it is likely that disease-specific circRNAs will be developed as disease biomarkers.

TABLE 1 | Identification algorithms for circRNA.

Tool	Mapper	*De novo*?	Annotation information?
Find_circ	Bowtie2	Yes	No
CIRCexplorer	Bowtie1 and 2	No	Yes
CIRI	Bwa	Yes	No
Mapsplice (51)	Bowtie1	No	Yes

APPROACHES TO STUDIES OF CIRCRNA

Identification Tools of circRNA

Identification of circRNA after sequencing is the first step in circRNA research. Various identification tools have been developed. Whether the identification of circRNAs is accurate and comprehensive is dependent upon the rigor and reliability of the algorithm. The Find_circ algorithm uses bowtie2 to map the original reads to the reference genome, discards all mapped sequences, takes 20 nt of each unmapped read as an anchor, and then determines the location of the anchor in the genome again to identify whether the splicing of circRNA is present (8). The CIRCexplorer algorithm uses the TopHat algorithm to map the RNA-sequencing reads to the human hg19 reference genome, and then maps the unmapped reads with the TopHat-Fusion algorithm. Such reads, unmapped with TopHat but mapped with TopHat-Fusion on the same chromosome in a back-spliced order, are extracted as candidate back-spliced junctionreads (48). The CIRI algorithm proposes paired chiastic clipping (PCC) signals to identify circRNAs. The PCC signal is detected by collecting and comparing the alignment information of all the segments of a read. PCC signals do not rely on existing annotation information, so back-splicing can be identified from zero to predict various types of circRNA, including intronic and intergenic circRNAs (49). In 2015, Hansen et al. found that short circRNAs (especially circRNAs <200 nt in length) are degraded readily by RNase R, and that a circRNA predicted by a single algorithm specifically has a higher false-positive rate. Conclusively, they suggested that identification of circRNAs can be done using a combination of different tools or RNA libraries with linear RNA being removed (50). We have summarized the four commonly used circRNA algorithms in **Table 1.**

Enrichment and Verification of circRNAs

RNase R is a member of the *Escherichia coli* RNR superfamily. It can cleave RNA in the 3′-5′ direction and digest almost all linear RNAs, but it cannot digest circRNAs readily (52). High-throughput sequencing for genome-wide identification of circRNAs requires enrichment of circRNA through treatment of samples with RNase R to remove linear RNAs, followed by enhancement of the concentration of circRNAs to facilitate their identification. Two approaches are used for the identification of circRNAs: reverse transcription-polymerase chain reaction (RT-PCR) and northern blotting. In RT-PCR, after digestion by RNase R, cDNA samples are amplified with divergent primers and convergent primers. Subsequent agarose-gel electrophoresis reveals that amplification with divergent primers generates a band in the RNase R(+) group, whereas the convergent primers produce no band. Divergent and convergent primers generate bands in the RNase R(−) group, indicating that circRNAs are present and resistant to digestion by RNase R (31, 53). Northern blotting is conducted using probes specific for circRNA and mRNA. The results show that linear mRNA cannot be detected in the RNase R(+) group, whereas its corresponding circRNA is visible, indicating that mRNAs are digested, whereas circRNAs are not (54). Notably, some circRNAs will also decrease in abundance after long-term digestion by RNase R, probably because of susceptibility to RNase R.

Knockdown and Knockout of circRNAs

Technology based on RNA interference is used widely to knockdown circRNA expression. To eliminate non-specific knockdown effects on cognate linear RNA, a specific small interfering RNA (siRNA) or short-hairpin RNA (shRNA) targeting circRNA must be directed to the back-splicing site. This strategy limits the design of siRNA sequences, and siRNA or shRNA targeting back-splicing sites will be partially complementary to the cognate linear RNA, raising the possibility of unwanted effects on expression of linear RNA. To solve this problem, Li et al. proposed that the control sequence should be partially replaced (∼10 nt) by a siRNA sequence targeting the back-splicing site (24). Similarly, knockout of circRNAs in animals risks simultaneously influencing expression of the cognate linear RNA and completely knocking out a gene is highly likely to affect expression of neighboring genes (55, 56). Based on the mechanism of circRNA production, Zhang et al. achieved knockout of circGCN1LI in human PA1 cells by removing an intron complementary sequence using CRISPR/Cas9 (16).

circRNA Overexpression

Plasmids used commonly to overexpress circRNAs are universal loop-forming framework vectors and gene-specific flanking sequence vectors. For example, Liang et al. cloned the exon2/3 and flanking sequences of *ZKSCAN1* to construct a vector, and then overexpressed circ-ZKSCAN1. Then, the flanking sequence was modified, and exons2/3 of ZKSCAN1 was replaced with a polyclonal restriction site to construct an empty vector: pcDNA3.1(+) CircRNA Mini Vector (57). Overexpression using a gene-specific flanking vector sequence is consistent with generation of natural circRNAs *in vivo*. Based on the characteristics of circRNA flanking sequences, length-appropriate flanking repeats (Alu elements) and circRNA sequences are cloned to generate a eukaryotic expression vector construct. In 2015, Lu et al. found that tRNA introns can form circular RNAs (tricRNAs) (58) and, in 2016, Schmidt et al. created a new circRNA expression vector based on tRNA splicing and transformation. First, they designed restriction enzyme sites in the two bulge–helix–bulge (BHB) regions of the *tRNA* intron and inserted the circRNA sequence for expression (59). In brief, the vector produced a circRNA by joining BHB motifs *via* an RtcB enzyme, which was not dependent upon a reverse-complementary sequence.

In general, different circRNAs with different loop-forming characteristics exhibit variation in looping efficiency. Moreover, circRNA production by overexpression vectors is accompanied by the production of linear isoforms, which is an additional challenge for strategies employing circRNA overexpression.

CIRCRNAS IN IMMUNITY

circRNAs in Anti-virus Immunity

Few reports have elucidated the role of circRNA in immune responses. Nevertheless, Chen et al. attempted to construct circRNAs to transfect cells based on *in vitro* transcription followed by auto-splicing circulation. Surprisingly, they found that *in vitro* circRNAs induced activation of cellular immune response pathways and inhibited RNA virus infection, which is mediated by retinoic acid-inducible gene-I (RIG-I); however, endogenous circRNAs did not induce this pathway because of their binding to specific RNA-binding proteins (60) (**Figure 2**). These findings suggest that exogenous introduction of circRNAs could be used to activate antiviral immune responses for therapeutic purposes. Nevertheless, some interesting questions must be solved: How do RBPs recognize self and non-self circRNAs, and then induce activation of non-self circRNAs on the RIG-1 pathway?

In addition, Li et al. found that circRNAs are involved in viral infection *via* the immune response factor NF90/NF110 (61). First, they applied a genome-wide siRNA screening strategy targeting all unique human genes with a Dox-inducible circmCherry expression vector to profile proteins involved in circRNA biogenesis. Consequently, they determined that NF90/NF110, encoded by the interleukin enhancer binding factor 3 (*ILF3*) gene, promoted circRNA formation by stabilizing the intron complementary sequence. Upon viral infection, NF90/NF110 was exported from the nucleus to the cytoplasm to inhibit virus replication. Meanwhile, circRNA expression in the nucleus was decreased (**Figure 2**). In this way, circRNAs may be used as "molecular indicators" of NF90/NF110 to prompt antiviral immune responses. However, the detailed mechanisms by which antiviral proteins promote circRNA biogenesis and key elements within circRNAs to interact with antiviral proteins have yet to be explored.

The two studies mentioned above provide evidence to support circRNA involvement in antiviral immunity through interactions with specific antiviral proteins (62). However, one of the studies elucidated the function of exogenous circRNAs in

antiviral immunity, and the other study suggested endogenous circRNAs may be used as molecular indicators of antiviral proteins to prompt antiviral immune responses. In addition to interacting with antiviral proteins, circRNAs may have other roles during antiviral immune responses. Through sequencing of the whole transcriptome, Shi et al. found that expression of 536 circRNAs was dysregulated significantly in herpes simplex virus 1 (HSV-1) -infected cells in contrast to uninfected human fibroblasts. Similarly, they screened differentially expressed genes and microRNAs in HSV-1-infected cells. Furthermore, they undertook analyses of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Their results suggested that these differentially expressed genes were very enriched in the pathways of immune responses, such as the NOD-like receptor signaling pathway and JAK-STAT signaling pathway (63). An integrated analysis of the circRNA–microRNA–gene axis revealed circRNAs can regulate the genes associated with host immune responses, which was mediated by microRNAs. These data suggest that circRNAs can regulate host antiviral immune responses through interactions with the corresponding microRNAs (**Figure 2**). However, these results based on large-scale bioinformation analyses should be validated experimentally.

Interestingly, several studies have demonstrated that some viruses can encode their own microRNAs, which facilitates the entry, replication, and virulence of viruses by targeting host transcripts, including some antiviral signaling molecules (64–67). For example, microRNA-BHRF1-3 encoded by the Epstein–Barr virus diminishes the levels of CXC-chemokine ligand 11, a chemoattractant in immune responses (68). In recent years, ceRNAs, including long non-coding RNA and circRNA (which can combine with microRNAs completely with a protein-coding target) have been studied widely. Ghosal et al. established a database named "HumanViCe" in which circRNAs that can sponge virus microRNAs can be predicted, and these circRNAs are enriched in pathways associated with the entry and replication of viruses and host immune responses (69). In summary, HumanViCe can aid exploration of the roles of circRNAs in viral infection but also circRNAs may act as potential antiviral targets.

circRNAs Acting Against Bacterial Infections

Ng et al. identified a lipopolysaccharide (LPS)-inducible circRNA generated from linear RasGEF1B, named mcircRasGEF1B, the expression of which is dependent upon the LPS-Toll-like receptor-4-nuclear factor-kappa B (LPS-TLR4-NF-κB) pathway. mcircRasGEF1B is cell type-specific, exhibits evolutionary conservation between mice and humans, and localizes preferentially to the cytoplasm (70). Also, its knockdown using the corresponding antisense oligodeoxynucleotide (ASOs) reduces intercellular adhesion molecule (ICAM-1) expression in the LPS/TLR4 signaling pathway by affecting the stability of the mature ICAM-1 mRNA, but not mRNA splicing (70) (**Figure 2**). In the immune system, ICAM-1 recruits leukocytes to sites of tissue inflammation, as well as enhancing adhesion between antigen-presenting cells and T cells (71, 72). In addition, ICAM-1 has been reported to inhibit M2 polarization of macrophages in tumors (73). Therefore, we speculate that mcircRasGEF1B may contribute to suppression of polarization of M2 macrophages during immune responses. These discoveries broaden our understanding and suggest that circRNAs may be important for the "fine tuning" of immune responses, and may help to protect cells from microbial infection.

circRNA as a ceRNA in Tumor Immunity

The relationship between microRNAs and immunity has been well-studied, leading to the hypothesis that circRNA may contribute to immune regulation through interactions with microRNAs. Zhang et al. reported that hsa_circ_0020397 can upregulate expression of PD-L1 (the target gene of microRNA-138) by binding to microRNA-138 in colorectal cancer cells. The consequent increase in PD-L1 levels contributes to tumor escape from immune responses (74, 75) (**Figure 2**). This information provides new insights for "checkpoint therapy" in cancer patients. Zheng et al. demonstrated that circHIPK3 rescues the downregulation of microRNA-124 on expression of interleukin (IL)-6R (76), implying that circHIPK3 may function in tumor immunity response. There are several bioinformatics databases that can be used to predict circRNAs that could bind to microRNAs, including circbase, starBasev2.0, and circinteractome (77).

circRNAs Regulate Immunity *via* Proteins

In addition to microRNAs, circRNAs can interact directly with proteins that function in immune responses. Similarly, we can predict such circRNAs using bioinformatics databases, then verify the predictions using *in vitro* and *in vivo* experiments. Using the circinteractome database, we found that the host gene of hsa_circ_0032139 is *HIF1A*, which plays an important part in inflammation *via* NF-κB and mitogen-activated protein kinase (MAPK) pathways (78). Under hypoxia, degradation of HIF1A protein is prevented, leading to its accumulation, and association with HIF1B to exert transcription regulation on target genes, including pro-inflammatory cytokines, most glycolytic enzymes, and glucose transporters, among others (79–83). Presumably, hsa_circ_0032139 can regulate inflammation through its association with *HIF1A*. Another circRNA, hsa_circ_0038481, has been predicted to associate with *TLR4*, a classical pattern-recognition receptor. This implies that hsa_circ_0038481 may be involved in LPS-stimulated signaling pathways, such as the NF-κB and MAPK.

circRNAs in Immune-Related Diseases

Rheumatoid arthritis (RA) is a chronic, inflammatory synovitis-based systemic disease of unknown etiology (84). Zheng et al. screened the top-ten upregulated, and downregulated circRNAs in RA patients based on analyses of peripheral-blood mononuclear cell chips, and selected the top-five corresponding microRNAs for each circRNA (85). We can speculate that these differentially expressed circRNAs may function in RA by acting as sponges of the corresponding microRNAs, which have been reported to be associated with RA. More importantly,

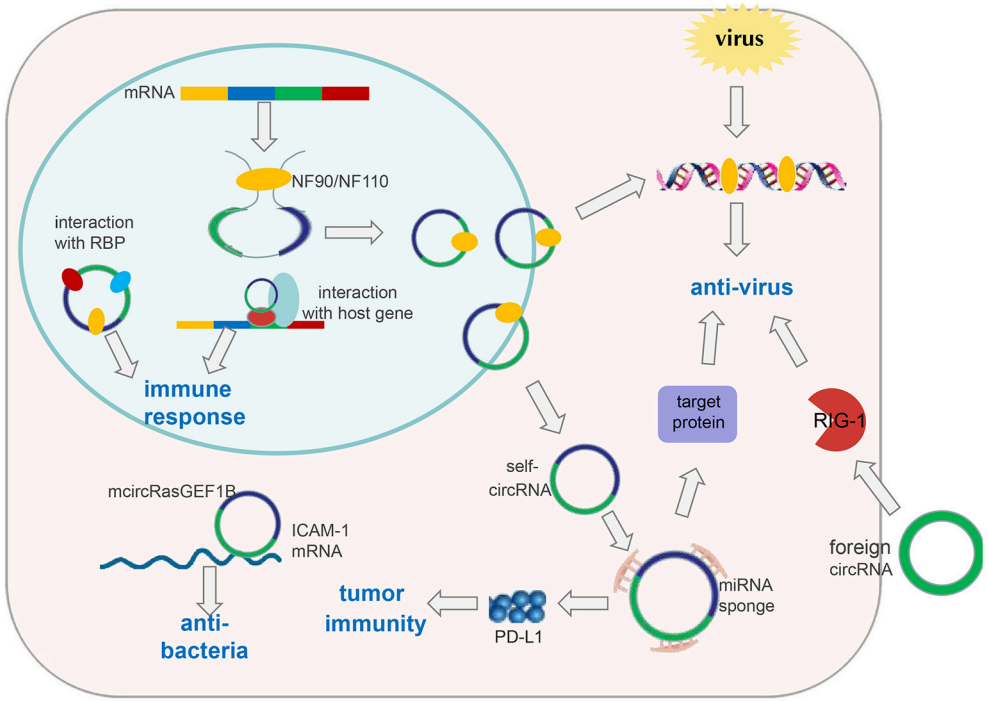

FIGURE 2 | CircRNAs in immune responses. Exogenous circRNAs can activate the RIG-1 cellular immune response pathway to suppress viral replication. The immune factor NF90/NF110 can promote circRNA formation by stabilizing the intron complementary sequence. Under viral infection, NF90/NF110 is exported from the nucleus to the cytoplasm to inhibit virus replication. Meanwhile, circRNA expression in the nucleus is decreased. Thus, endogenous circRNAs can be used as "molecular indicators" of antiviral proteins to prompt antiviral immune responses. circRNAs can also function as "microRNA sponges" to increase expression of target proteins involved in antiviral responses or tumor immunity, such as PD-L1. mcircRasGEF1B can help to protect cells from bacterial infection by enhancing the stability of mature ICAM-1 mRNA. In addition, circRNAs can be involved in immune responses by interacting with proteins or their host genes.

the study by Zheng et al. provides clues that circRNAs may have the same effects as their host genes in RA. For example, hsa_circ_0038644, one of the dysregulated circRNAs in RA, is spliced from *PRKCB*, which is associated with LPS-induced activation of the NF-κB signaling pathway (86). Therefore, we hypothesized that hsa_circ_0038644 can aggravate inflammation in RA patients.

Type-2 diabetes mellitus (T2DM) is characterized by hyperglycemia, insulin resistance, and chronic inflammation (87). Fang et al. found circANKRD36 to be upregulated markedly in the peripheral blood cells of T2DM patients. Furthermore, an association was noted between circANKRD36 expression and inflammatory factors (88). Therefore, circANKRD36 may serve as a potential biomarker and be involved in inflammation in T2DM.

Similarly, Li et al. measured circRNA expression in the T cells of patients with systemic lupus erythematosus (SLE). They found hsa_circ_0045272 to be downregulated significantly. Furthermore, they demonstrated that hsa_circ_0045272 regulated apoptosis and IL-2 production negatively (89). However, the mechanisms underlying the involvement of hsa_circ_0045272 in SLE pathogenesis merits further exploration.

Numerous studies have shown that a high proportion of tumor-infiltrating lymphocytes (TILs) in the tumor microenvironment can improve clinical outcomes. Weng et al. showed that high expression of hsa_circ_0064428 is associated with a low proportion of TILs, poor survival, large tumor volume, and tumor metastasis in patients with hepatocellular carcinoma (90). These observations suggest that hsa_circ_0064428 can act as a potential immune-associated prognosis biomarker for hepatocellular carcinoma.

Overall, the roles of circRNAs in immune diseases have been studied based on large-scale microarray and RNA sequencing analyses, by which differentially expressed circRNAs were screened to further verify their functions experimentally. However, the mechanisms by which circRNAs regulate disease development merit further exploration.

PERSPECTIVES

Although circRNAs have become a hot topic in RNA research in recent years, several important areas merit investigation. For example, Dong et al. showed that circRNA can be reverse-transcribed *in vivo* and fused into the genome to generate a pseudogene (91). Further research is needed to elucidate the molecular mechanisms underlying circRNA *trans*-transcriptional translocation and the effects of pseudogenes derived from circRNAs. In addition, unlike linear mRNA, construction of overexpression vectors for circRNA is

challenging because it requires splicing of the two termini of amplified fragments to generate loop structures. Moreover, data in the mouse circRNA database is incomplete, which poses specific obstacles for *in vivo* experiments. Finally, circRNAs in the cytoplasm, which act as microRNA sponges, have been studied widely, but circRNAs that can function through this mechanism are in the minority. Most circRNAs rarely contain so many miRNA binding sites and have low expression (24, 28), so future research should explore other mechanisms of circRNAs.

Few studies have focused on circRNAs in the immune response. Reminiscent of studies in cancer, circRNA research in immunity also requires measuring circRNA expression, and then elucidating the function of differentially expressed circRNAs and the mechanisms underlying it. Notably, circRNA localization could aid exploration of the mechanisms by which circRNAs regulate immune responses. circRNAs in the cytoplasm are likely to act as microRNA sponges, whereas circRNAs in the nucleus are likely to interact with proteins, or elicit effects by promoting/suppressing the role of their host genes (**Figure 2**). Excitingly, circRNAs are promising markers or drug targets for some immune diseases, which could facilitate rapid diagnosis and treatment.

AUTHOR CONTRIBUTIONS

LY the first author, contributed to collection of references and manuscript preparation. JF and YZ contributed to manuscript modifications.

FUNDING

This work was supported by grants from the National Key R&D Program of China (2016YFC1305102), The National Natural Science Foundation of China (NSFC, 81671561), 1000 Young Talents Plan Program of China, Initial Funding for New PI, Fudan Children's Hospital and Fudan University, and Shanghai Municipal Planning Commission of Science and Research Fund (201740065) (to YZ). Shanghai Pujiang Program 16PJ1401600 (to JF).

REFERENCES

1. Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell* (1993) 73:1019–30. doi: 10.1016/0092-8674(93)90279-Y
2. Chen CY, Sarnow P. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* (1995) 268:415–7. doi: 10.1126/science.7536344
3. Cocquerelle C, Mascrez B, Hétuin D, Bailleul B. Mis-splicing yields circular RNA molecules. *FASEB J*. (1993) 7:155–60. doi: 10.1096/fasebj.7.1.7678559
4. Pasman Z, Been MD. Garcia-Blanco MA. Exon circularization in mammalian nuclear extracts. *RNA* (1996) 2:603–10.
5. Cocquerelle C, Daubersies P, Majérus MA, Kerckaert JP, Bailleul B. Splicing with inverted order of exons occurs proximal to large introns. *EMBO J*. (1992) 11:1095–8. doi: 10.1002/j.1460-2075.1992.tb05148.x
6. Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, et al. Scrambled exons. *Cell* (1991) 64:607–13. doi: 10.1016/0092-8674(91)90244-S
7. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* (2013) 19:141–57. doi: 10.1261/rna.035667.112
8. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* (2013) 495:333–8. doi: 10.1038/nature11928
9. Barrett SP, Wang PL, Salzman J. Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife* (2015) 4:e07540. doi: 10.7554/eLife.07540
10. Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC. Strand-specific RNA sequencing in *Plasmodium falciparum* malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* (2015) 16:454. doi: 10.1186/s12864-015-1603-4
11. Lu T, Cui L, Zhou Y, Zhu C, Fan D, Gong H, et al. Transcriptome-wide investigation of circular RNAs in rice. *RNA* (2015) 21:2076–87. doi: 10.1261/rna.052282.115
12. Wang PL, Bao Y, Yee MC, Barrett SP, Hogan GJ, Olsen MN, et al. Circular RNA is expressed across the eukaryotic tree of life. *PLoS ONE* (2014) 9:e90859. doi: 10.1371/journal.pone.0090859
13. Lasda E, Parker R. Circular RNAs: diversity of form and function. *RNA* (2014) 20:1829−42. doi: 10.1261/rna.047126.114
14. Wilusz JE. A 360 degrees view of circular RNAs: From biogenesis to functions. *Wiley Interdisc Rev RNA* (2018) 9:e1478. doi: 10.1002/wrna.1478
15. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, et al. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* (2014) 56:55–66. doi: 10.1016/j.molcel.2014.08.019
16. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang JL, et al. The biogenesis of nascent circular RNAs. *Cell Rep*. (2016) 15:611–24. doi: 10.1016/j.celrep.2016.03.058
17. Zaphiropoulos PG. Exon skipping and circular RNA formation in transcripts of the human cytochrome P-450 2C18 gene in epidermis and of the rat androgen binding protein gene in testis. *Mol Cell Biol*. (1997) 17:2985–93. doi: 10.1128/MCB.17.6.2985
18. Zaphiropoulos PG. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. *Proc Natl Acad Sci USA*. (1996) 93:6536–41. doi: 10.1073/pnas.93.13.6536
19. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep*. (2015) 10:170–7. doi: 10.1016/j.celrep.2014.12.019
20. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, et al. The RNA binding protein quaking regulates formation of circRNAs. *Cell* (2015) 160:1125–34. doi: 10.1016/j.cell.2015.02.014
21. Khan MA, Reckman YJ, Aufiero S, van den Hoogenhof MM, van der Made I, Beqqali A, et al. RBM20 regulates circular RNA production from the titin gene. *Circ Res*. (2016) 119:996–1003. doi: 10.1161/CIRCRESAHA.116.30956810.1161/CIRCRESAHA.116.309568
22. Errichelli L, Dini Modigliani S, Laneve P, Colantoni A, Legnini I, Capauto D, et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat Commun*. (2017) 8:14741. doi: 10.1038/ncomms14741
23. Fei T, Chen Y, Xiao T, Li W, Cato L, Zhang P, et al. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc Natl Acad Sci USA*. (2017) 114:E5207–15. doi: 10.1073/pnas.1617467114
24. Li X, Yang L, Chen LL. The biogenesis, functions, and challenges of circular RNAs. *Mol Cell* (2018) 71:428–42. doi: 10.1016/j.molcel.2018.06.034
25. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, et al. Circular intronic long noncoding RNAs. *Mol Cell* (2013) 51:792–806. doi: 10.1016/j.molcel.2013.08.017
26. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol*. (2015) 22:256–64. doi: 10.1038/nsmb.2959

27. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, et al. Corrigendum: exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol*. (2017) 24:194. doi: 10.1038/nsmb0217-194a
28. Guo JU, Agarwal V, Guo H, Bartel DP. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol*. (2014) 15:409. doi: 10.1186/s13059-014-0409-z
29. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, et al. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol Cell* (2017) 66:22–37.e9. doi: 10.1016/j.molcel.2017.02.017
30. Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, et al. Extensive translation of circular RNAs driven by N-methyladenosine. *Cell Res*. (2017) 27:626–41. doi: 10.1038/cr.2017.31
31. Yang Y, Gao X, Zhang M, Yan S, Sun C, Xiao F, et al. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J Natl Cancer Inst*. (2018) 110, 304–15. doi: 10.1093/jnci/djx166
32. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, et al. Translation of CircRNAs. *Mol Cell* (2017) 66:9–21.e7. doi: 10.1016/j.molcel.2017.02.021
33. Zhou C, Molinie B, Daneshvar K, Pondick JV, Wang J, Van Wittenberghe N, et al. Genome-wide maps of m6A circRNAs identify widespread and cell-type-specific methylation patterns that are distinct from mRNAs. *Cell Rep*. (2017) 20:2262–76. doi: 10.1016/j.celrep.2017.08.027
34. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* (2004) 116:281–97. doi: 10.1016/S0092-8674(04)00045-5
35. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* (2009) 136:215–33. doi: 10.1016/j.cell.2009.01.002
36. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* (2006) 6:259–69. doi: 10.1038/nrc1840
37. Jansson MD, Lund AH. MicroRNA and cancer. *Mol Oncol*. (2012) 6:590–610. doi: 10.1016/j.molonc.2012.09.006
38. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. Natural RNA circles function as efficient microRNA sponges. *Nature* (2013) 495:384–8. doi: 10.1038/nature11993
39. Han D, Li J, Wang H, Su X, Hou J, Gu Y, et al. Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression. *Hepatology* (2017) 66:1151–64. doi: 10.1002/hep.29270
40. Zhong Z, Huang M, Lv M, He Y, Duan C, Zhang L, et al. Circular RNA MYLK as a competing endogenous RNA promotes bladder cancer progression through modulating VEGFA/VEGFR2 signaling pathway. *Cancer Lett*. (2017) 403:305–17. doi: 10.1016/j.canlet.2017.06.027
41. Chen B, Wei W, Huang X, Xie X, Kong Y, Dai D, et al. circEPSTI1 as a prognostic marker and mediator of triple-negative breast cancer progression. *Theranostics* (2018) 8:4003–15. doi: 10.7150/thno.24106
42. Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res*. (2016) 44:2846–58. doi: 10.1093/nar/gkw027
43. Yang Q, Du WW, Wu N, Yang W, Awan FM, Fang L, et al. A circular RNA promotes tumorigenesis by inducing c-myc nuclear translocation. *Cell Death Differ*. (2017) 24:1609–20. doi: 10.1038/cdd.2017.86
44. Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, et al. Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol*. (2017) 14:361–9. doi: 10.1080/15476286.2017.1279788
45. Li Y, Zheng Q, Bao C, Li S, Guo W, Zhao J, et al. Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res*. (2015) 25:981–4. doi: 10.1038/cr.2015.82
46. Guarnerio J, Bezzi M, Jeong JC, Paffenholz SV, Berry K, Naldini MM, et al. Oncogenic role of fusion-circRNAs derived from cancer-associated chromosomal translocations. *Cell* (2016) 166:1055–6. doi: 10.1016/j.cell.2016.07.035
47. Tan S, Gou Q, Pu W, Guo C, Yang Y, Wu K, et al. Circular RNA F-circEA produced from EML4-ALK fusion gene as a novel liquid biopsy biomarker for non-small cell lung cancer. *Cell Res*. (2018) 28:693–5. doi: 10.1038/s41422-018-0033-7
48. Zhang XO, Wang HB, Zhang Y, Lu X, Chen LL, Yang L. Complementary sequence-mediated exon circularization. *Cell* (2014) 159:134–47. doi: 10.1016/j.cell.2014.09.001
49. Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for *de novo* circular RNA identification. *Genome Biol*. (2015) 16:4. doi: 10.1186/s13059-014-0571-3
50. Hansen TB, Venø MT, Damgaard CK, Kjems J. Comparison of circular RNA prediction tools. *Nucleic Acids Res*. (2016) 44:e58. doi: 10.1093/nar/gkv1458
51. Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL, et al. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res*. (2010) 38:e178. doi: 10.1093/nar/gkq622
52. Cheng ZF, Deutscher MP. Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J Biol Chem*. (2002) 277:21624–9. doi: 10.1074/jbc.M202942200
53. Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A, Mayeda A. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res*. (2006) 34:e63. doi: 10.1093/nar/gkl151
54. Rybak-Wolf A, Stottmeister C, GlaŽar P, Jens M, Pino N, Giusti S, et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol Cell* (2015) 58:870–85. doi: 10.1016/j.molcel.2015.03.027
55. Joung J, Engreitz JM, Konermann S, Abudayyeh OO, Verdine VK, Aguet F, et al. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* (2017) 548:343–6. doi: 10.1038/nature23451
56. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. CRISPRi-based genome-scale identification of functional long non-coding RNA loci in human cells. *Science* (2017) 355:aah7111. doi: 10.1126/science.aah7111
57. Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev*. (2014) 28:2233–47. doi: 10.1101/gad.251926.114
58. Lu Z, Filonov GS, Noto JJ, Schmidt CA, Hatkevich TL, Wen Y, et al. Metazoan tRNA introns generate stable circular RNAs *in vivo*. *RNA* (2015) 21:1554–65. doi: 10.1261/rna.052944.115
59. Schmidt CA, Noto JJ, Filonov GS, Matera AG. A method for expressing and imaging abundant, stable, circular RNAs *in vivo* using tRNA splicing. *Methods Enzymol*. (2016) 572:215–36. doi: 10.1016/bs.mie.2016.02.018
60. Chen YG, Kim MV, Chen X, Batista PJ, Aoyama S, Wilusz JE, et al. Sensing self and foreign circular RNAs by intron identity. *Mol Cell* (2017) 67:228–38 e5. doi: 10.1016/j.molcel.2017.05.022
61. Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, et al. Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. *Mol Cell* (2017) 67:214–27 e7. doi: 10.1016/j.molcel.2017.05.023
62. Cadena, C, Hur S. Antiviral immunity and circular RNA: no end in sight. *Mol Cell* (2017) 67:163–4. doi: 10.1016/j.molcel.2017.07.005
63. Shi J, Hu N, Mo L, Zeng Z, Sun J, Hu Y. Deep RNA Sequencing reveals a repertoire of human fibroblast circular RNAs associated with cellular responses to herpes simplex virus 1 infection. *Cell Physiol Biochem*. (2018) 47:2031–45. doi: 10.1159/000491471
64. Dunn W, Trang P, Zhong Q, Yang E, van Belle C, Liu F. Human cytomegalovirus expresses novel microRNAs during productive viral infection. *Cell Microbiol*. (2005) 7:1684–95. doi: 10.1111/j.1462-5822.2005.00598.x
65. Nair V, Zavolan M. Virus-encoded microRNAs: novel regulators of gene expression. *Trends Microbiol*. (2006) 14:169–75. doi: 10.1016/j.tim.2006.02.007
66. Pfeffer S, Zavolan M, Grässer FA, Chien M, Russo JJ, Ju J, et al. Identification of virus-encoded microRNAs. *Science* (2004) 304:734–6. doi: 10.1126/science.1096781
67. Umbach JL, Kramer MF, Jurak I, Karnowski HW, Coen DM, Cullen BR. MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* (2008) 454:780–3. doi: 10.1038/nature07103
68. Xia T, O'Hara A, Araujo I, Barreto J, Carvalho E, Sapucaia JB, et al. EBV MicroRNAs in primary lymphomas and targeting of CXCL-11 by ebv-mir-BHRF1-3. *Cancer Res*. (2008) 68:1436–42. doi: 10.1158/0008-5472.CAN-07-5126
69. Ghosal S, Das S, Sen R, Chakrabarti J. HumanViCe: host ceRNA network in virus infected cells in human. *Front Genet*. (2014) 5:249. doi: 10.3389/fgene.2014.00249
70. Ng WL, Marinov GK, Liau ES, Lam YL, Lim YY, Ea CK. Inducible RasGEF1B circular RNA is a positive regulator of ICAM-1 in the TLR4/LPS pathway. *RNA Biol*. (2016) 13:861–71. doi: 10.1080/15476286.2016.1207036

71. Staunton DE, Merluzzi VJ, Rothlein R, Barton R, Marlin SD, Springer TA. A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell* (1989) 56:849–53. doi: 10.1016/0092-8674(89)90689-2
72. Hubbard AK, Rothlein R. Intercellular adhesion molecule-1 (ICAM-1) expression and cell signaling cascades. *Free Radic Biol Med.* (2000) 28:1379–86. doi: 10.1016/S0891-5849(00)00223-9
73. Yang M, Liu J, Piao C, Shao J, Du J. ICAM-1 suppresses tumor metastasis by inhibiting macrophage M2 polarization through blockade of efferocytosis. *Cell Death Dis*. (2015) 6:e1780. doi: 10.1038/cddis.2015.144
74. Zhang XL, Xu LL, Wang F. Hsa_circ_0020397 regulates colorectal cancer cell viability, apoptosis and invasion by promoting the expression of the miR-138 targets TERT and PD-L1. *Cell Biol Int*. (2017) 41:1056–64. doi: 10.1002/cbin.10826
75. Xu Z, Li P, Fan L, Wu M. The potential role of circRNA in tumor immunity regulation and immunotherapy. *Front Immunol*. (2018) 9:9. doi: 10.3389/fimmu.2018.00009
76. Zheng Q, Bao C, Guo W, Li S, Chen J, Chen B, et al. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun.* (2016) 7:11215. doi: 10.1038/ncomms11215
77. Liu J, Liu T, Wang X, He A. Circles reshaping the RNA world: from waste to treasure. *Mol Cancer* (2017) 16:58. doi: 10.1186/s12943-017-0630-y
78. Frede S, Stockmann C, Freitag P, Fandrey J. Bacterial lipopolysaccharide induces HIF-1 activation in human monocytes via p44/42 MAPK and NF-kappaB. *Biochem J*. (2006) 396:517–27. doi: 10.1042/BJ20051839
79. Dengler VL, Galbraith M, Espinosa JM. Transcriptional regulation by hypoxia inducible factors. *Crit Rev Biochem Mol Biol.* (2014) 49:1–15. doi: 10.3109/10409238.2013.838205
80. Cramer T, Yamanishi Y, Clausen BE, Förster I, Pawlinski R, Mackman N, et al. HIF-1alpha is essential for myeloid cell-mediated inflammation. *Cell* (2003) 112:645–57. doi: 10.1016/S0092-8674(03)00154-5
81. Nizet V, Johnson RS. Interdependence of hypoxic and innate immune responses. *Nat Rev Immunol*. (2009) 9:609–17. doi: 10.1038/nri2607
82. Papandreou I, Cairns RA, Fontana L, Lim AL, Denko NC. HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. *Cell Metab*. (2006) 3:187–97. doi: 10.1016/j.cmet.2006.01.012
83. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab*. (2008) 7:11–20. doi: 10.1016/j.cmet.2007.10.002
84. Firestein GS. Evolving concepts of rheumatoid arthritis. *Nature* (2003) 423:356–61. doi: 10.1038/nature01661
85. Zheng F, Yu X, Huang J, Dai Y. Circular RNA expression profiles of peripheral blood mononuclear cells in rheumatoid arthritis patients, based on microarray chip technology. *Mol Med Rep*. (2017) 16:8029–36. doi: 10.3892/mmr.2017.7638
86. Sheng YJ, Gao JP, Li J, Han JW, Xu Q, Hu WL, et al. Follow-up study identifies two novel susceptibility loci PRKCB and 8p11.21 for systemic lupus erythematosus. *Rheumatology (Oxford)* (2011) 50:682–8. doi: 10.1093/rheumatology/keq313
87. Seidell JC. Obesity, insulin resistance and diabetes–a worldwide epidemic. *Br J Nutr*. (2000) 83 (Suppl. 1):S5–8. doi: 10.1017/S00071145000 0088X
88. Fang Y, Wang X, Li W, Han J, Jin J, Su F, et al. Screening of circular RNAs and validation of circANKRD36 associated with inflammation in patients with type 2 diabetes mellitus. *Int J Mol Med.* (2018) 42:1865–74. doi: 10.3892/ijmm.2018.3783
89. Li LJ, Zhu ZW, Zhao W, Tao SS, Li BZ, Xu SZ, et al. Circular RNA expression profile and potential function of hsa_circ_0045272 in systemic lupus erythematosus. *Immunology* (2018) 155:137–49. doi: 10.1111/imm. 12940
90. Weng Q, Chen M, Li M, Zheng YF, Shao G, Fan W, et al. Global microarray profiling identified hsa_circ_0064428 as a potential immune-associated prognosis biomarker for hepatocellular carcinoma. *J Med Genet*. (2018) doi: 10.1136/jmedgenet-2018-105440. [Epub ahead of print].
91. Dong R, Zhang XO, Zhang Y, Ma XK, Chen LL, Yang L. CircRNA-derived pseudogenes. *Cell Res*. (2016) 26:747–50. doi: 10.1038/cr.2016.42

Conflict of Interest Statement: 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 Immunology

ORIGINAL RESEARCH
published: 08 January 2019
doi: 10.3389/fimmu.2018.03087

Diversity of Epigenetic Features of the Inactive X-Chromosome in NK Cells, Dendritic Cells, and Macrophages

*Camille M. Syrett, Vishal Sindhava, Isabel Sierra, Aimee H. Dubin, Michael Atchison and Montserrat C. Anguera**

Department of Biomedical Sciences, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, United States

In females, the long non-coding RNA Xist drives X-chromosome Inactivation (XCI) to equalize X-linked gene dosage between sexes. Unlike other somatic cells, dynamic regulation of Xist RNA and heterochromatin marks on the inactive X (Xi) in female lymphocytes results in biallelic expression of some X-linked genes, including *Tlr7*, *Cxcr3*, and *Cd40l,* implicated in sex-biased autoimmune diseases. We now find that while Xist RNA is dispersed across the nucleus in NK cells and dendritic cells (DCs) and partially co-localizes with H3K27me3 in bone marrow-derived macrophages, it is virtually absent in plasmacytoid DCs (p-DCs). Moreover, H3K27me3 foci are present in only 10–20% of cells and we observed biallelic expression of *Tlr7* in p-DCs from wildtype mice and NZB/W F1 mice. Unlike in humans, mouse p-DCs do not exhibit sex differences with interferon alpha production, and interferon signature gene expression in p-DCs is similar between males and females. Despite the absence of Xist RNA from the Xi, female p-DCs maintain dosage compensation of six immunity-related X-linked genes. Thus, immune cells use diverse mechanisms to maintain XCI which could contribute to sex-linked autoimmune diseases.

Keywords: Xist RNA, X-chromosome inactivation, long non-coding RNA, plasmacytoid dendritic cells, macrophages, sex differences, NK cells, interferon alpha

OPEN ACCESS

Edited by:
Adam Williams, Jackson Laboratory, United States

Reviewed by:
Dan Anthony Mitchell, University of Warwick, United Kingdom
Stefan Pinter, University of Connecticut Health Center, United States

****Correspondence:***
Montserrat C. Anguera
anguera@vet.upenn.edu

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *02 August 2018*
Accepted: *13 December 2018*
Published: *08 January 2019*

Citation:
Syrett CM, Sindhava V, Sierra I, Dubin AH, Atchison M and Anguera MC (2019) Diversity of Epigenetic Features of the Inactive X-Chromosome in NK Cells, Dendritic Cells, and Macrophages. Front. Immunol. 9:3087. doi: 10.3389/fimmu.2018.03087

INTRODUCTION

In the immune system, long non-coding RNAs (lncRNAs) are being increasingly recognized as important regulators of gene expression for both innate and adaptive immune responses (1). Indeed, lncRNAs can function as regulators of immune cell differentiation, lymphocyte activation, and inflammatory responses. For example, the lncRNA Morrbid is abundantly expressed in nuclei of neutrophils, eosinophils, and monocytes, and Morrbid deletion reduces the numbers of these short-lived myeloid cells (2). Similar to Morrbid, lnc-DC is also upregulated during differentiation of common myeloid progenitors into dendritic cells (DCs), and regulates DC differentiation through cytoplasmic interactions with the transcription factor STAT3 (3). Activation of DCs and macrophages through specific TLRs results in dramatic upregulation of lincRNA-Cox2, which regulates over 500 genes encoding inflammatory molecules (4).

One of the best characterized lncRNAs is Xist, which is required for silencing the X-chromosome during X-chromosome Inactivation (XCI). Females use XCI for dosage compensation of X-linked genes between the sexes. XCI is initiated during early female mammalian embryonic development

(5) by allele-specific upregulation of Xist from the future inactive X (Xi) (6–8). Xist RNA functions in *cis* to recruit chromatin complexes that deposit heterochromatic modifications (including H3K27me3 and H2a-ubiquitin) across the X, resulting in transcriptional silencing (9–11). During XCI maintenance, these epigenetic modifications are enriched on the Xi and contribute to its transcriptional silencing after cell division, to ensure dosage compensation of X-linked genes. In differentiating embryonic stem cells, *Xist* is continuously expressed from the Xi throughout the cell cycle, and Xist RNA remains tethered to the Xi of its origin throughout mitosis (12).

The majority of somatic cells maintain XCI through continuous expression of *Xist* from the Xi, and enrichment of Xist RNA transcripts and heterochromatin marks on the Xi are cytologically visible. Surprisingly, we have shown that mature naive T and B cells from female mice and humans lack these epigenetic modifications on the Xi. However, Xist RNA and some heterochromatin modifications are present on the Xi in *in vitro* activated lymphocytes (13, 14), suggesting that XCI is dynamically regulated in lymphocytes. Using RNA FISH, Xist RNA localization patterns in lymphocytes can be categorized into four classes: Type I Xist RNA patterns exhibit robust signals, Type II patterns have dispersed signals within the X-chromosome territory, Type III patterns have diffuse signals across the nucleus, and Type IV patterns lack detectible signal (14, 15). This dynamic localization of Xist RNA and heterochromatin marks suggests relaxed transcriptional silencing on the Xi, which is supported by recent observations by our group and others of biallelic expression of the X-linked genes *Tlr7*, *Cxcr3*, and *Cd40l* in mouse and human T and B cells (14, 16).

Based on our findings in lymphocytes, we assessed Xist RNA localization patterns on the Xi in terminally differentiated myeloid and lymphoid-derived cells. We found that NK cells and dendritic cells (DCs) have Xist RNA transcripts dispersed across the nucleus, while bone marrow derived macrophages (BMDMs) have Xist RNA pinpoints clustered at the Xi, and exhibit co-localization of Xist RNA and the heterochromatin mark H3K27me3. Interestingly, resting and activated plasmacytoid DCs (p-DCs) lack Xist RNA localization at the Xi, and most cells also lack H3K27me3. Additionally, we observed biallelic expression of *Tlr7* in wildtype and disease-stage NZB/W F1 p-DCs, yet there were no sex differences with interferon alpha production, unlike in human cells. Together, these data reveal that immune cells use diverse mechanisms to maintain XCI that could contribute to sex-linked autoimmune diseases.

MATERIALS AND METHODS

Mice

Female mice (aged 2–6 months) of various backgrounds (C57BL/6, BALB/c, NZB × NZW F1) were purchased from Jackson Laboratories, and used to isolate bone marrow derived macrophages (BMDM), NK cells, dendritic cells (DCs), and plasmacytoid DCs. All mice were maintained at the Penn Vet animal facility. Animal experiments were approved by the University of Pennsylvania Institutional Animal Care and Use Committee (IACUC). Euthanasia via carbon dioxide was used for animal sacrifice prior to spleen isolation.

Fluorescence Activated Cell Sorting (FACS) Isolation of NK Cells, Lymphoid and Myeloid Dendritic Cells From Spleen

Spleens were harvested on ice in FACS buffer (PBS/3%FCS) and single-cell suspensions were prepared by meshing cells through 40-um strainers, then cells were stained with antibodies for fluorescence activated cell sorting (FACS) analyses. Briefly, cells were stained with fluorochrome-conjugated or biotinylated antibodies to mouse. Staining was performed in PBS/1%BSA containing mouse IgG Fc fragments (Jackson Immunoresearch, Cat # 115-006-020). Dead cells and doublets were excluded and sorting was performed on a FACS Aria II machine using the following markers at a concentration of 1:100 unless otherwise specified: NK cells: TCRb+CD19 (H57-597/6D5, BioLegend), NK1.1 (PK138, BD Pharmingen), NKP46 (29A1.4, eBiosciences). m-DCs: CD11c (N418, BioLegend), CD11b (M1/70, eBiosciences, 1:200). L-DCs: CD8a (53-6.7, eBiosciences). Data were analyzed using FlowJo software.

Isolation and *in vitro* Stimulation of Plasmacytoid Dendritic Cells (p-DCs) and Bone Marrow Derived Macrophages

Plasmacytoid dendritic cells (p-DCs) were isolated from spleen and peripheral lymph nodes by negative selection using a plasmacytoid dendritic cell isolation kit (#130-107-093, Miltenyi Biotec). p-DCs were cultured in RPMI-1640 containing 2 mM L-glutamine, 10% FCS, 1% Pen/Strep and 50 μM β-mercaptoethanol. P-DCs were stimulated with 1 μM CpG (ODN 1826, InvivoGen) and cultured for 3 days.

Bone marrow was isolated from female 6 week old C57BL/6J mice and cultured in complete DMEM (10% FBS, 1% NaPyruvate, 1% HEPES, 30% L929 conditioned medium) and re-fed on day 4. Macrophages were isolated 8 days after differentiation by washing petri culture dishes with Mg^{2+} and Ca^{2+} EDTA-free 4C PBS. Under these culture conditions we estimate that the population of BMDMs is 98% pure using flow cytometry (data not shown). Cells were re-plated with complete DMEM with 10% L929 conditioned media and stimulated with either 1 μM CpG (ODN 1826, InvivoGen) or 1 μg/mL LPS (Sigma) for 3 days.

Xist RNA FISH, Tlr7 RNA FISH, and Immunofluorescence Detection of H3K27me3

Sequential RNA fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF) for immune cells was performed following established protocols for splenocytes (14, 15), where

Abbreviations: Xi, inactive X; XCI, X-chromosome Inactivation; DC, dendritic cell; p-DC, plasmacytoid dendritic cell; IF, immunofluorescence; BMDM, bone marrow derived macrophages; L-DC, lymphoid dendritic cell; M-DC, myeloid dendritic cell.

Xist RNA FISH was performed first followed by IF for the same locations on the slides. For Xist RNA FISH, two Cy3-labeled 20-nucleotide oligo probes were designed to recognize regions within Xist RNA exon 1 (synthesized by IDT). For IF, cells were blocked with 0.2% PBS-Tween, 0.5% BSA. Histone H3K27me3 (Active Motif; Cat. #39155) was diluted 1:100 for IF. Single-molecule RNA FISH for Tlr7 was performed according to Stellaris protocols, using Cy3-labeled oligo probes for exonic regions (Stellaris), and FITC-labeled oligo probes for intronic regions (Stellaris). Images were obtained using a Nikon Eclipse microscope and were categorized by the four types of Xist RNA localization patterns as described previously (14, 15). Statistical significance was calculated using two-tailed *t*-tests and ANOVA.

Analysis of Gene Expression and IFN-α Protein Production in p-DCs

To determine levels of X-lined and IFNα gene expression, p-DCs were isolated from spleens of (NZB × NZW) F1 [NZB/W F1] mice by negative selection using a plasmacytoid dendritic cell isolation kit (#130-107-093, Miltenyi Biotec). p-DCs were cultured as described above and stimulated with 10 μg/mL R848 (Resiquimod, Sigma Aldrich) or 1 μM CpG (InvivoGen) and cultured for 6 h. Cells and supernatants were collected after 6 h of culturing (NS: unstimulated; R848, CPG: activated). RNA was isolated using TRIzol (Invitrogen), and cDNA was synthesized with qScript cDNA SuperMix (Quanta). qRT-PCR was performed using the following primer pairs [For Seq 5′-3′; Rev Seq 5′-3′]: Rpl13a [AGC CTA CCA GAA AGT TTG CTTAC; GCT TCT TCT TCC GAT AGT GCA TC], Xist [GCT GGT TCG TCT ATC TTG TGGG; CAG AGT AGC GAG GAC TTGA AGAG], Cxcr3 [TAC CTT GAG GTT AGT GAA CGT CA; CGC TCT CGT TTT CCC CAT AATC], Cfp [TTC ACC CAG TAT GAG GAG TCC; GCTG ACC ATT GTG GAG ACCT], Irak1 [TCC TCC ACC AAG CAG TCA AG; AAA ACC ACC CTC TCC AAT CCT], Il2rg [CTC AGG CAA CCA ACC TCAC; GCT GGA CAA CAA ATG TCT GGT AG], Msn [GGCT TCC CGT GGA GTG AAA TC; GTC CGG GGC CTT TTT GTC AA], Tlr7 [ATG TGG ACA CGG AAG AGA CAA; GGT AAG GGT AAG ATT GGT GGTG], Ifna2 [TAC TCA GCA GACC TTG AAC CT; CAG TCT TGG CAG CAA GTT GAC], Ccl4 [TTC CTG CTG TTT CTC TTA CACCT; CTG TCT GCC TCT TTT GGT CAG], Irf7 [CTC CTG AGC GCA GCC TTG; GTT CTTAC TGC TGG GGC CAT], Ifit2 [GGA GAG CAA TCT GCG ACAG; GCT GCC TCA TTT AGA CCT CTG].

To determine expression levels, the housekeeping gene Rpl13a was used for normalization ($2^{\wedge}\Delta\Delta CT$). Combined qRT-PCR results are shown from three independent experiments.

Total serum IFNα from *in vitro* p-DCs cell culture supernatants was measured using a VeriKine Mouse IFN alpha ELISA Kit (42120, pbl assay science). Supernatants were collected after 6 h of culture and were undiluted for ELISA. The plate was read at 450 nm immediately after development and was analyzed using protein standards provided in the kit (400, 200, 100, 50, 25, 12.5, 0 pg/mL).

RESULTS

NK Cells Predominantly Lack Xist RNA on the Xi and Xist RNA Is Dispersed Across the Nucleus in Dendritic Cells

To determine if XCI is dynamically regulated in NK cells, lymphoid-DCs (L-DCs), and myeloid DCs (m-DCs) each of these cell types were isolated from female mouse spleens (**Figure 1A**). These cells are derived from common lymphoid progenitors and common myeloid progenitors, which reside in the bone marrow, and are known to have robust Xist RNA "clouds" on the Xi (15). We used fluorescence activated cell sorting (FACS) to isolate each population following the surface marker profiling shown in **Figure 1B**, then immediately cytospun and fixed the cells on glass slides, which preserves nuclear RNA signals. We used Cy3-labeled short oligo probes for Xist to perform RNA FISH, and classified the percentage of cells for each localization pattern (Types I-IV) (14, 15). NK cells predominantly lacked detectible Xist RNA signals (Type IV) and 20–50% of cells exhibited Type III patterns with diffuse Xist RNA pinpoints dispersed across the nucleus (**Figures 1C,D**). M-DCs and L-DCs had about 10% Type II cells, where Xist RNA pinpoints are localized in a nuclear territory encompassing the inactive X (Xi), and 40–90% Type III cells (**Figure 1D**). These results show that NK cells have less Xist RNA localized to the Xi compared to m-DCs and L-DCs, and suggest that NK cells may have more genes that escape XCI than DCs.

Xist RNA and H3K27me3 Modifications Are Localized to the Xi in Bone Marrow Derived Macrophages (BMDM)

Cytokine production and phagocytic activity of macrophages exhibits sex-related differences (17, 18). As the expression of X-linked genes could contribute to these functional differences, we asked whether Xist RNA and H3K27me3 are localized to the Xi in macrophages. We cultured BMDMs for 8 days after isolation, and then activated the cells for 3 days using CpG or LPS. Unstimulated BMDMs had mostly Type II Xist RNA patterns (40–90%) and some Type III (∼5–10%) (**Figures 2A,B**). Stimulation with either CpG or LPS increased the number of Type I Xist RNA patterns for about 5–10% of cells, yet the percentage of Type II and Type III patterns did not significantly change (**Figures 2A,B**). The number of Type I cells decreased by day 3 for both CpG and LPS stimulation, and Xist RNA signal persisted longer and in more cells with CpG stimulation (**Figures 2A,B**). Next, we examined the co-localization of Xist RNA signals with H3K27me3 foci using sequential RNA FISH followed by IF. As shown in **Figures 2C,D**, Xist RNA signals co-localized with a focus of H3K27me3 in 30–50% of BMDMs, and *in vitro* stimulation did not change the level of co-localization (**Figure 2C**). About 20–50% of cells had an Xist RNA signal yet lacked H3K27me3 foci, and very few cells (5–12%) had a H3K27me3 focus and lacked Xist RNA signal (**Figures 2C,D**). These results suggest that Xist RNA localization at the Xi is necessary for H3K27me3 enrichment on this chromosome in BMDMs.

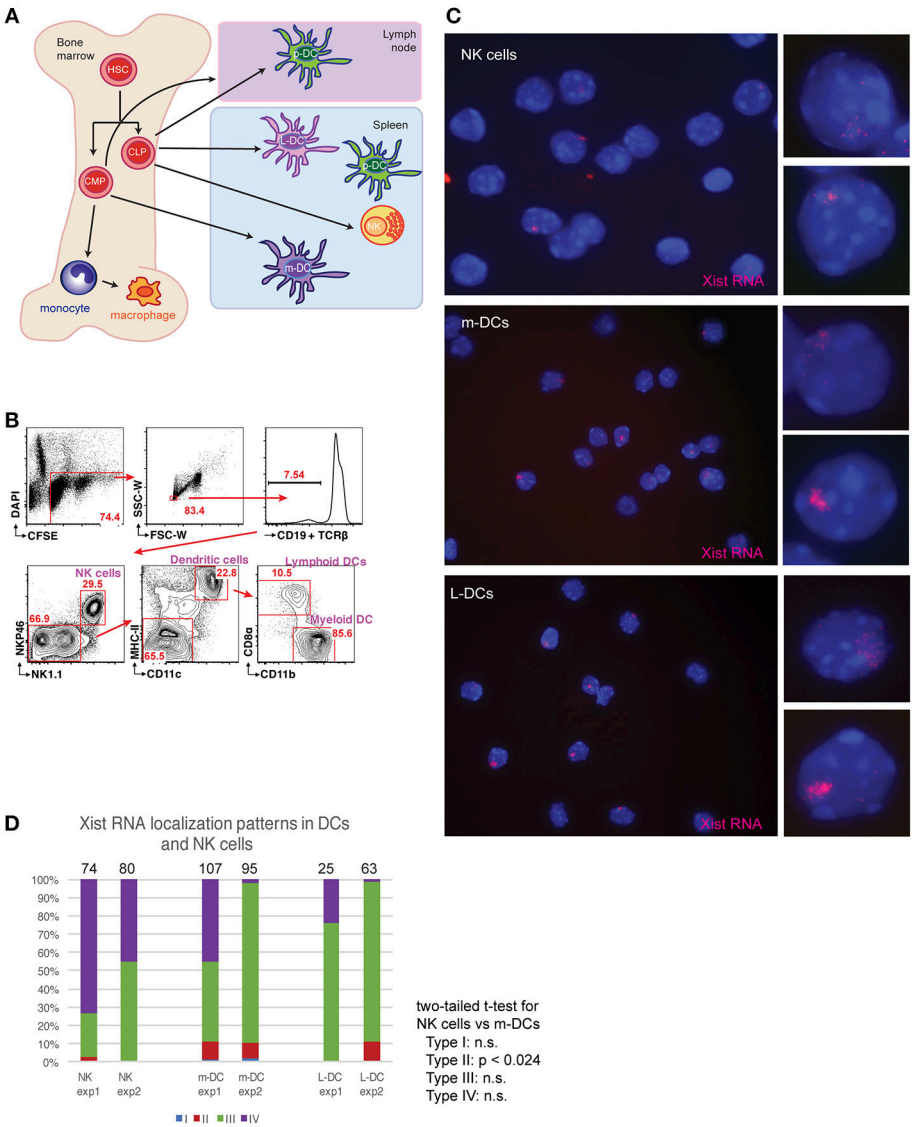

FIGURE 1 | Xist RNA transcripts are mostly absent from the Xi in female NK cells and DCs. **(A)** Schematic showing the origin for the immune cells examined here. Hematopoietic stem cells (HSCs); common lymphoid progenitors (CLPs); common myeloid progenitors (CMPs); plasmacytoid dendritic cells (p-DCs); myeloid-derived DCs (m-DCs); lymphoid-derived DCs (L-DCs). **(B)** Sorting strategy for isolation of m-DCs (MHC-II$^+$, CD11c$^+$, CD11b$^+$, CD8a^{lo}), L-DCs (MHC-II$^+$, CD11c$^+$, CD11b$^+$, CD8a^{hi}), and NK cells (NK1.1$^+$) using FACS. Spleens from two female mice were pooled for each experiment (repeated twice), and flow results from experiment 1 are shown. **(C)** Xist RNA FISH analyses of NK cells, m-DCs, L-DCs, using Cy3 labeled oligo probes. **(D)** Quantification of Xist RNA localization patterns (Types I–IV) for each experiment. The total number of nuclei counted for each cell type is shown above the column. Statistical significance was determined comparing each type of Xist RNA pattern (Types I–IV) for each cell type, using a two-tailed *t*-test. The comparison between NK cells and m-DCs for Type II patterns was the only significant difference ($p < 0.024$). NK cells and L-DCs had no significant differences in Xist RNA localization patterns.

Plasmacytoid DCs Lack Xist RNA and H3K27me3 foci on the Xi and Biallelically Express *Tlr7*

Plasmacytoid DCs (p-DCs) are a distinct lineage of DCs that produce interferon (IFN) in response to viral nucleic acids detected by TLR7 and TLR9 (19). TLR7-mediated stimulation of female plasmacytoid DCs (p-DCs) from human females results in higher levels of IFN regulatory factor 5 (IRF5) and IFNα compared to p-DCs from males (20, 21). TLR7 is an X-linked gene that is prone to escape XCI in female B and T cells (14, 16), and exhibits elevated expression in some female immune cells (22). In female Systemic Lupus Erythematosus (SLE) patients, p-DCs are a major source of aberrant IFN production that contributes to disease progression (23). We asked whether Xist RNA was localized to the Xi in p-DCs pooled from lymph nodes and spleen using RNA FISH. Surprisingly, we did not detect any Xist RNA signal in p-DCs, and 100% of the cells were Type IV (**Figures 3A,B**). Xist RNA signals

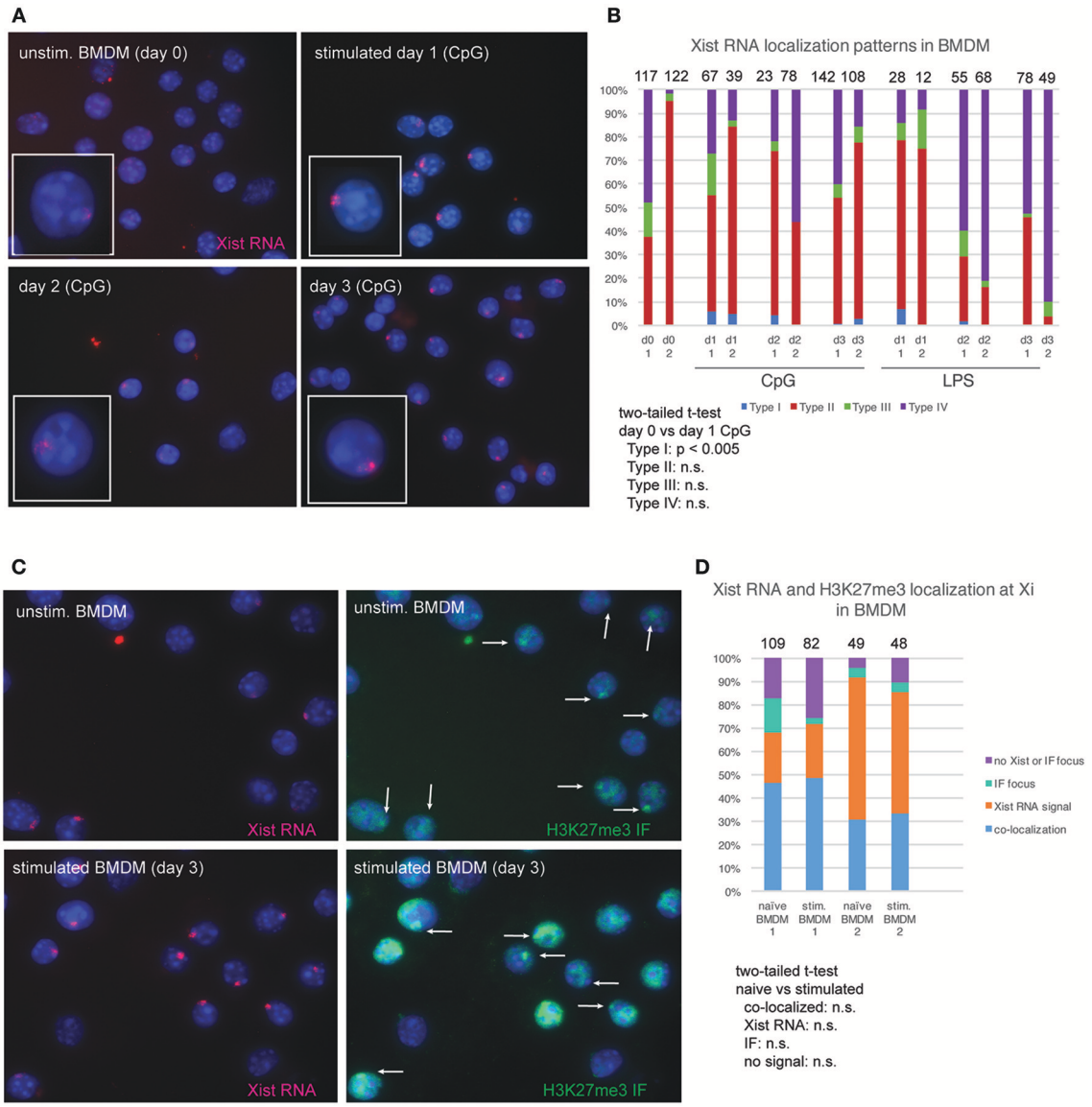

FIGURE 2 | Xist RNA and H3K27me3 foci are localized to the Xi in most female BMDMs. **(A)** Xist RNA FISH for resting BMDMs and *in vitro* stimulated cells (using 1 μM CpG), collected 3 days after stimulation. **(B)** Quantification of Xist RNA localization patterns for BMDMs stimulated with either 1 μM CpG or 1 μg/mL LPS. The total number of nuclei counted for each cell type is shown above the column. Statistical significance for comparisons of resting (day 0) vs. stimulated cells was performed for each type of Xist RNA localization pattern (Types I–IV) using a two-tailed *t*-test, and the only significant difference was for CpG-stimulated Type I cells ($p < 0.005$). **(C)** Sequential Xist RNA and H3K27me3 IF for resting and stimulated BMDMs. White arrows indicate H3K27me3 foci. **(D)** Quantification of co-localization patterns for Xist RNA and H3K27me3 foci. Results from two independent experiments are shown. The total number of nuclei counted for each cell type is shown above the column. Statistical significance for comparisons of resting (day 0) vs. stimulated cells was performed for each type of localization pattern using a two-tailed *t*-test, and the only significant difference was for CpG-stimulated Type I cells ($p < 0.005$).

were also absent from the Xi and the nucleus in LPS or CpG-stimulated p-DCs (**Figures 3A,B**). Next, we investigated whether the repressive chromatin modification H3K27me3, which localizes to the Xi in fibroblasts and some activated lymphocytes, was present in p-DCs. Using sequential RNA FISH followed by immunofluorescence (IF) detection, we found that the majority of p-DCs lacked H3K27me3 foci (**Figure 3C**) and that 10–20% of p-DCs had a detectible focus of H3K27me3 (**Figure 3D**). In sum, p-DCs lack Xist RNA localization to the Xi and enrichment of H3K27me3, suggesting that the chromatin of the Xi may be prone to reactivation of some X-linked genes.

To determine whether the absence of Xist RNA localization to the Xi affects *Tlr7* expression in p-DCs, we performed RNA FISH using oligo probes specific for the exonic and intronic regions of *Tlr7*. Resting p-DCs had low yet detectible signals for Tlr7 RNA: the majority of cells lacked Tlr7 RNA pinpoints, yet we

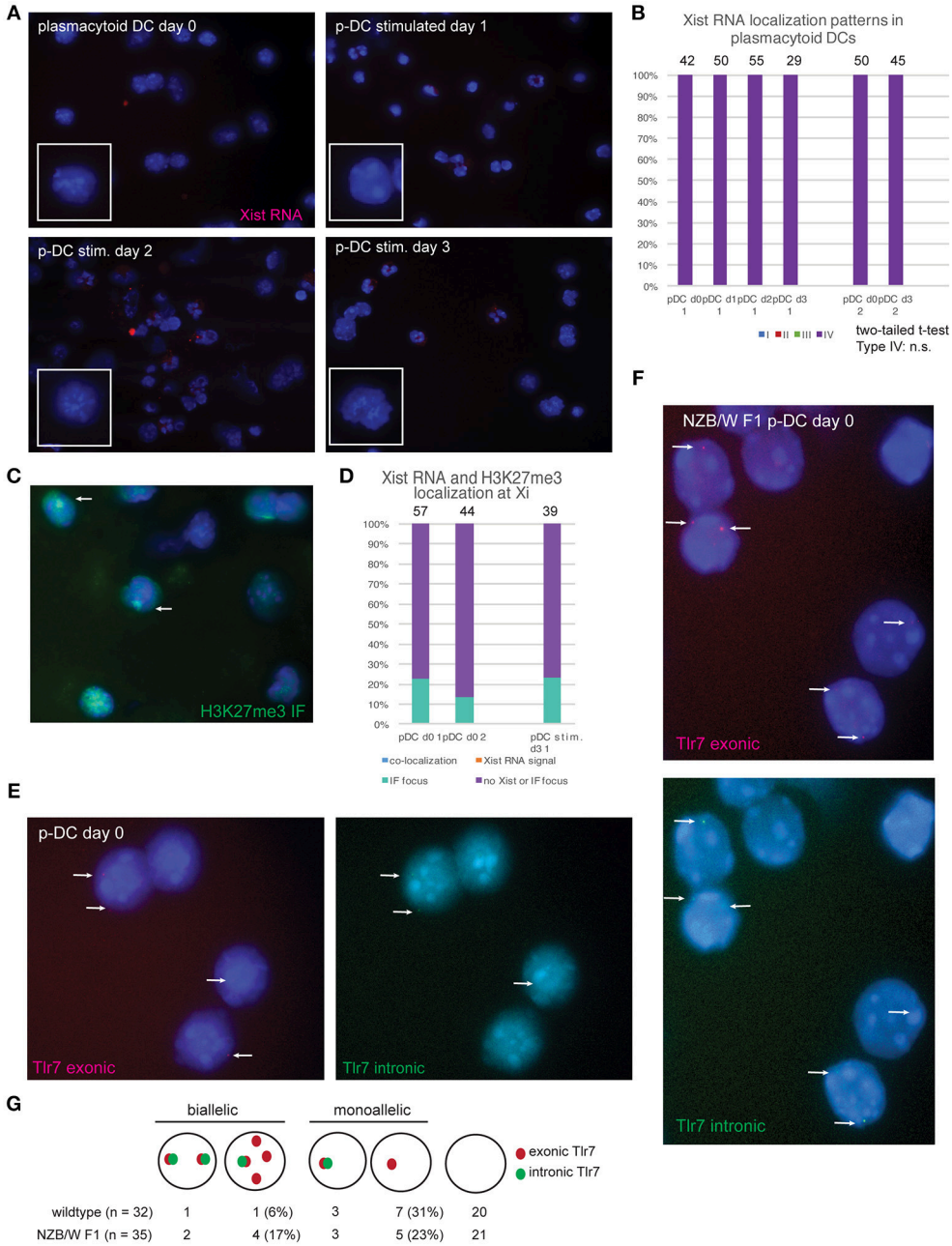

FIGURE 3 | Female plasmacytoid DCs lack Xist RNA at the Xi and exhibit biallelic expression of *Tlr7* in some cells. **(A)** Xist RNA FISH for resting p-DCs and *in vitro* activated p-DCs after days 1–3 of culture. P-DCs were isolated from spleen and lymph nodes from female mice, in two independent experiments and stimulated with CpG. For the second isolation, cells were stimulated for 3 days before collection for RNA FISH. **(B)** Quantification of Xist RNA localization patterns for p-DCs showing that all p-DCs are missing Xist RNA on the Xi. The total number of nuclei counted for each cell type is shown above the column. **(C)** Sequential Xist RNA FISH followed by immunofluorescence (IF) for H3K27me3 enrichment at the Xi. **(D)** Quantification of co-localization patterns for Xist RNA and H3K27me3 foci. Results from two independent experiments are shown. The total number of nuclei counted for each cell type is shown above the column. **(E)** Single-molecule RNA FISH for Tlr7 transcripts using wildtype p-DCs from healthy mice. Oligo probes specific for exonic *Tlr7* were Cy3-labeled (red), and intronic *Tlr7* probes were FITC-labeled (green). White arrows indicate pinpoint signals for nascent *Tlr7* expression from the X-chromosome, with signals from both exonic and intronic probes. **(F)** single-molecule *Tlr7* RNA FISH in p-DCs from NZB/W F1 mice with SLE-like disease. Disease development was assessed by proteinuria and DNA autoantibodies prior to p-DC isolation from spleen and lymph nodes. White arrows indicate pinpoint signals for nascent *Tlr7* expression from the X-chromosome, with signals from both exonic and intronic probes. **(G)** Schematic for counting allele-specific expression, and quantification of monoallelic and biallelic Tlr7 expression in wildtype and NZB/W F1 p-DCs. The percentages of total numbers for biallelic and monoallelic expressing cells are shown in parentheses.

could identify some cells with monoallelic (one pinpoint) and also biallelic (two pinpoints) expression (**Figure 3E**). Next, we isolated p-DCs from (NZB × NZW) F1 female mice (NZB/W F1), which is a high quality model of spontaneous SLE-like disease with a strong female bias. Disease was assessed by proteinuria, serum double strand DNA antibodies, and sudden weight loss (24). NZB/W F1 female mice have increased numbers of p-DCs and produce more IFNα compared to healthy female C57BL/6 mice (25). The p-DCs from diseased female mice had more robust exonic *Tlr7* signals and more biallelic Tlr7 expression compared to healthy mice (**Figures 3F,G**). Thus, the absence of Xist RNA and H3K27me3 enrichment on the Xi in p-DCs correlates with elevated expression of *Tlr7* during SLE-like disease.

Female p-DCs do not Require Xist RNA Localized on the Xi for X-linked Gene Dosage Compensation

Because female p-DCs lacked Xist RNA and H3K27me3 enrichment at the Xi (**Figure 3**), we first asked whether Xist RNA was transcribed in these cells. We isolated splenic p-DCs from males and females, and stimulated cells using two methods (CPG or R848), then isolated total RNA for qRT-PCR. We also included naïve and stimulated B cells for comparison, as Xist RNA levels are similar between naïve B cells that lack Xist RNA localization at the Xi and stimulated cells which have robust Xist RNA clouds (15). *Xist* is expressed at similar levels in unstimulated and stimulated female pDCs, regardless of the method of activation (**Figure 4A**). Thus, *Xist* transcription and localization are uncoupled in pDCs, and this may account for the absence of H3K27me3 enrichment on the Xi in these cells.

We next asked if female pDCs from NZB/W F1 mice produced more of the X-linked *Tlr7* gene than males, and if this resulted in higher levels of IFNα. To determine whether p-DCs exhibited sex-biased gene expression of IFNα and IFNα signature genes, we isolated splenic p-DCs from male and female NZB/W F1 mice at early and late stage disease. We cultured the cells in the presence or absence of the Tlr7 agonist R848 for 6 h, then harvested cells for RNA isolation. We used qRT-PCR to determine the steady state levels for the IFNα signature genes *Ifna2, Ccl4, Irf7*, and *Ifit2*, which are expressed in p-DCs (26, 27). We saw no significant sex differences between the expression of these genes in female and male p-DCs stimulated with R848 (**Figure 4B**).

Higher levels of IFNα production have been reported in p-DCs from human females compared to males (20, 21). Next, we asked whether pDCs from NZB/W F1 mice exhibited sex differences with IFNα protein levels. To determine if female p-DCs produced more IFNα than male p-DCs, we determined the IFNα concentration in supernatants of cultured male and female p-DCs (resting and *in vitro* activated using R848) by ELISA. While the IFNα concentrations were variable, we saw no significant increase in IFNα production in female p-DCs (**Figure 4C**), suggesting that, unlike in humans, murine female and male p-DCs produce similar levels of IFNα.

Next, we asked whether female p-DCs, which lack Xist RNA at the Xi, exhibit greater expression of X-linked genes known to be subject to XCI. We performed qRT-PCR for six X-linked immune genes expressed in p-DCs from male and female NZB/W F1 mice. We did not observe any significant sex differences with the expression of *Cxcr3, Cfp, Irak1, Il2rg, Msn,* and *Tlr7* (**Figure 4D**). Together, these results suggest that female mouse p-DCs are capable of maintaining X-linked gene dosage compensation in the absence of Xist RNA localized at the Xi.

DISCUSSION

Taken together, our findings reveal wide diversity in the localization of the epigenetic modifications Xist RNA and H3K27me3 at the Xi in myeloid and lymphoid lineages. This new insight may have important implications for understanding how X-linked gene expression from the Xi is regulated in diverse immune cell populations. NK cells have faint and dispersed Xist RNA signals across the nucleus (Type III) and nuclei that lack Xist RNA (Type IV), which suggests that some X-linked genes in these cells may be prone to reactivation. Dosage of the X-linked gene *XIAP* affects NK cell function in patients with X-linked lymphoproliferative syndrome presenting with chronic inflammatory bowel disease (28), which underscores the importance of X-linked gene expression in NK cells. We found that resting BMDMs, unlike lymphocytes, have predominantly Type II Xist RNA patterns, and that *in vitro* stimulation with CpG generates few Type I cells. Thus, the epigenetic features of the Xi in female BMDMs more closely resembles that of female fibroblasts, but with less robust Xist RNA clouds. Xist RNA localization on the Xi is correlated with H3K27me3 foci in BMDMs, which is observed in fibroblasts (9), differentiating mouse embryonic stem cells (29), and activated B cells (15).

We also found that splenic m-DCs and L-DCs have more robust and detectible Xist RNA signals compared to NK cells, with most of these DCs classified as Type III with dispersed Xist RNA across the nucleus and some cells being Type II with clustered Xist RNA pinpoints. In contrast, p-DCs are completely distinct from other DCs, lymphocytes, and BMDMs as they lack detectible Xist RNA and are exclusively Type IV. However, we observed that 10–20% of p-DCs have H3K27me3 foci, which suggests that Xist RNA localization at the Xi is not required for H3K27me3 enrichment in these cells. It is possible that the 80–90% of p-DCs that lack Xist RNA/H3K27me3 enrichment are primed for gene-specific reactivation from the Xi. In support, we observe biallelic expression of *Tlr7* in some p-DCs from both healthy and disease-state NZB/W F1 female mice. These results support a model where Xist RNA and heterochromatin marks localized on the Xi promote transcriptional silencing, and gene reactivation may occur from the Xi more readily when these epigenetic modifications are missing.

Despite the absence of Xist RNA transcripts on the Xi, it was surprising that female mouse p-DCs maintained dosage compensation of six X-linked immune genes, including *Tlr7*. We hypothesize that the fidelity of transcriptional silencing of these genes on the Xi is likely maintained by DNA methylation and additional heterochromatin marks (besides H3K27me3) in female p-DCs. It is possible that there are X-linked genes besides

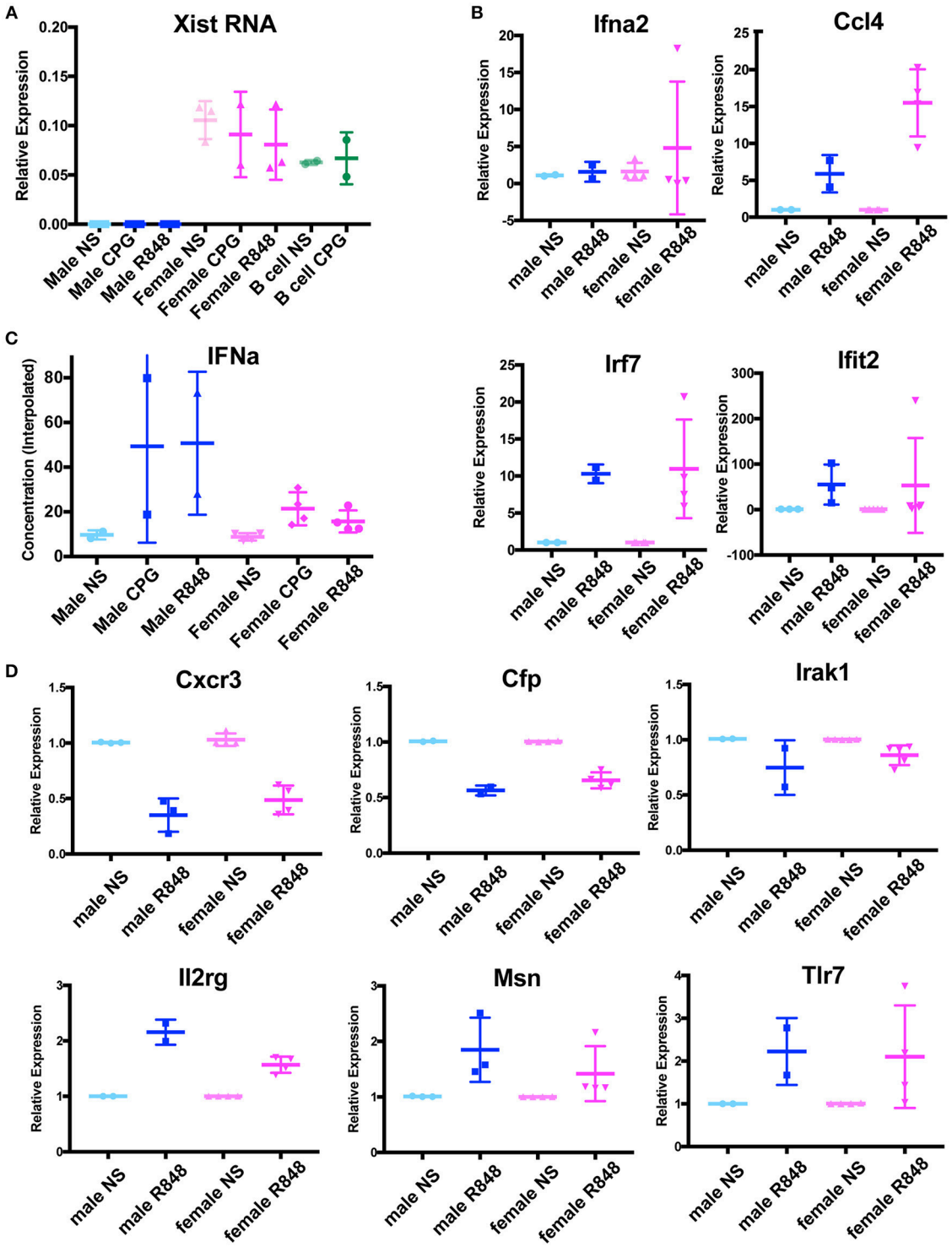

FIGURE 4 | p-DCs do not exhibit a sex difference with IFNα production and X-linked genes are dosage compensated in the absence of Xist RNA localization at the Xi. **(A)** Relative quantity ($2^{\wedge}\Delta Ct$) of Xist RNA in unstimulated male and female pDCs, and cells stimulated with CPG, R848. Female B cells (naïve and CPG stimulated) were included as positive controls. **(B)** Relative quantity ($2^{\wedge}\Delta\Delta Ct$) of four IFNα signature genes. P-DCs were activated with R848 for 6 h or were unstimulated (NS). The housekeeping gene Rpl13a was used for normalization, and male unstimulated samples (NS) were normalized to 1. **(C)** Concentration of IFNα protein produced by cultured male and female p-DCs from NZB/W F1 mice measured by ELISA. **(D)** Relative quantity ($2^{\wedge}\Delta\Delta Ct$) of six X-linked immune genes from male and female splenic p-DCs from NZB/W F1 mice. P-DCs were activated with R848 for 6 h or were unstimulated (NS). The housekeeping gene Rpl13a was used for normalization, and unstimulated samples (NS) were normalized to 1.

the six examined here that specifically escape XCI in female mouse p-DCs. Future experiments that detect allele-specific expression from the Xi will reveal whether Xist RNA localization influences gene reactivation in p-DCs. It has been reported that human female p-DCs have elevated *TLR7* expression and increased IFNα production compared to male cells (21). We were surprised to find that male and female p-DCs from NZB/W F1 mice express similar levels of *Tlr7*, and that IFNα concentrations from *in vitro* cultured cells did not exhibit sex differences. Our findings suggest that the Xi in female pDCs is more transcriptionally silent than the human Xi in p-DCs, which is supported by the observations that the human Xi from various tissues contains more genes that escape XCI (15–25% X-linked genes escape XCI) compared to the mouse Xi (30, 31). We cannot exclude the possibility that human p-DCs may also lack Xist RNA on the Xi, which contributes to female-specific overexpression of *TLR7* in human p-DCs, and that increased expression could come from both the Xa and Xi. *TLR7* has been recently shown to escape XCI in healthy human B cells (16), and it is possible that *TLR7* might be bi-allelically expressed in human p-DCs. Additional studies examining the allelic expression profiles of X-linked genes in human p-DCs are necessary to reveal the origins of female-biased *TLR7* expression. In conclusion, our results demonstrate that female murine immune cells use diverse mechanisms to maintain XCI, which may underlie sex differences with some immune responses and the observed sex-bias in predisposition to autoimmune diseases.

AUTHOR CONTRIBUTIONS

CS and IS isolated the plasmacytoid DCs and BMDMs. CS performed RNA FISH and quantification of Xist RNA localization patterns. CS performed the IF, single molecule RNA FISH for *Tlr7,* and gene expression analyses. VS performed flow cytometry and cell sorting for NK cells and DCs. AD performed the ELISA. MA provided funding for flow cytometry and for VS, MCA, and CS wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by a University of Pennsylvania Research Foundation grant and NIH grants AI124084 and HD085848 (MCA), and T32-GM007229, T32-HD083185, and F31-GM123604 (CS). We would like to thank T. Laufer, M. May, and L. King for discussion and suggestions, and A. DeLaney for isolation of BMDMs.

REFERENCES

1. Chen YG, Satpathy AT, Chang HY. Gene regulation in the immune system by long noncoding RNAs. *Nat Immunol.* (2017) 18:962–72. doi: 10.1038/ni.3771
2. Kotzin JJ, Spencer SP, Mccright SJ, Kumar DBU, Collet MA, Mowel WK, et al. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. *Nature* (2016) 537:239–43. doi: 10.1038/nature19346
3. Wang P, Xue Y, Han Y, Lin L, Wu C, Xu S, et al. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Science* (2014) 344:310–3. doi: 10.1126/science.1251456
4. Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* (2013) 341:789–92. doi: 10.1126/science.1240925
5. Payer B, Lee JT. X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet.* (2008) 42:733–72. doi: 10.1146/annurev.genet.42.110807.091711
6. Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, Mccabe VM, et al. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* (1991) 351:329–31. doi: 10.1038/351329a0
7. Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, et al. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* (1991) 349:38–44. doi: 10.1038/349038a0
8. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. Requirement for Xist in X chromosome inactivation. *Nature* (1996) 379:131–7. doi: 10.1038/379131a0
9. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, et al. Role of histone H3 lysine 27 methylation in X inactivation. *Science* (2003) 300:131–5. doi: 10.1126/science.1084274
10. Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, et al. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* (2003) 4:481–95. doi: 10.1016/S1534-5807(03)00068-6
11. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* (2008) 322:750–6. doi: 10.1126/science.1163045
12. Jonkers I, Monkhorst K, Rentmeester E, Grootegoed JA, Grosveld F, Gribnau J. Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. *Mol Cell Biol.* (2008) 28:5583–94. doi: 10.1128/MCB.02269-07
13. Savarese F, Flahndorfer K, Jaenisch R, Busslinger M, Wutz A. Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Mol Cell Biol.* (2006) 26:7167–77. doi: 10.1128/MCB.00810-06
14. Wang J, Syrett CM, Kramer MC, Basu A, Atchison ML, Anguera MC. Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc Natl Acad Sci USA.* (2016) 113:E2029–38. doi: 10.1073/pnas.1520113113
15. Syrett CM, Sindhava V, Hodawadekar S, Myles A, Liang G, Zhang Y, et al. Loss of Xist RNA from the inactive X during B cell development is restored in a dynamic YY1-dependent two-step process in activated B cells. *PLoS Genet.* (2017) 13:e1007050. doi: 10.1371/journal.pgen.1007050
16. Souyris M, Cenac C, Azar P, Daviaud D, Canivet A, Grunenwald S, et al. TLR7 escapes X chromosome inactivation in immune cells. *Sci Immunol.* (2018) 3:eaap8855. doi: 10.1126/sciimmunol.aap8855
17. Spitzer JA. Gender differences in some host defense mechanisms. *Lupus* (1999) 8:380–3. doi: 10.1177/096120339900800510
18. Marriott I, Bost KL, Huet-Hudson YM. Sexual dimorphism in expression of receptors for bacterial lipopolysaccharides in murine macrophages: a possible mechanism for gender-based differences in endotoxic shock susceptibility. *J Reprod Immunol.* (2006) 71:12–27. doi: 10.1016/j.jri.2006.01.004
19. Gilliet M, Cao W, Liu YJ. Plasmacytoid dendritic cells: sensing nucleic acids in viral infection and autoimmune diseases. *Nat Rev Immunol.* (2008) 8:594–606. doi: 10.1038/nri2358
20. Seillet C, Laffont S, Tremollieres F, Rouquie N, Ribot C, Arnal JF, et al. The TLR-mediated response of plasmacytoid dendritic cells is positively regulated by estradiol *in vivo* through cell-intrinsic estrogen receptor alpha signaling. *Blood* (2012) 119:454–64. doi: 10.1182/blood-2011-08-371831
21. Griesbeck M, Ziegler S, Laffont S, Smith N, Chauveau L, Tomezsko P, et al. Sex differences in plasmacytoid dendritic cell levels of IRF5 drive higher IFN-alpha production in women. *J Immunol.* (2015) 195:5327–36. doi: 10.4049/jimmunol.1501684

22. Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol.* (2016) 16:626–38. doi: 10.1038/nri.2016.90
23. Ronnblom L, Alm GV. An etiopathogenic role for the type I IFN system in SLE. *Trends Immunol.* (2001) 22:427–31. doi: 10.1016/S1471-4906(01)01955-X
24. Morel L, Wakeland EK. Susceptibility to lupus nephritis in the NZB/W model system. *Curr Opin Immunol.* (1998) 10:718–25. doi: 10.1016/S0952-7915(98)80094-0
25. Zhou Z, Ma J, Xiao C, Han X, Qiu R, Wang Y, et al. Phenotypic and functional alterations of pDCs in lupus-prone mice. *Sci Rep.* (2016) 6:20373. doi: 10.1038/srep20373
26. Karrich JJ, Jachimowski LC, Uittenbogaart CH, Blom B. The plasmacytoid dendritic cell as the Swiss army knife of the immune system: molecular regulation of its multifaceted functions. *J Immunol.* (2014) 193:5772–8. doi: 10.4049/jimmunol.1401541
27. Wimmers F, Subedi N, Van Buuringen N, Heister D, Vivie J, Beeren-Reinieren I, et al. Single-cell analysis reveals that stochasticity and paracrine signaling control interferon-alpha production by plasmacytoid dendritic cells. *Nat Commun.* (2018) 9:3317. doi: 10.1038/s41467-018-05784-3
28. Opat S, Hearps AC, Thia K, Yuen A, Rogers B, Chachage M, et al. Adaptive reprogramming of NK cells in X-linked lymphoproliferative syndrome. *Blood* (2018) 131:699–702. doi: 10.1182/blood-2017-08-803668
29. Sarma K, Cifuentes-Rojas C, Ergun A, Del Rosario A, Jeon Y, White F, et al. ATRX directs binding of PRC2 to Xist RNA and polycomb targets. *Cell* (2014) 159:869–83. doi: 10.1016/j.cell.2014.10.019
30. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* (2005) 434:400–4. doi: 10.1038/nature03479
31. Balaton BP, Cotton AM, Brown CJ. Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biol Sex Differ.* (2015) 6:35. doi: 10.1186/s13293-015-0053-7

Conflict of Interest Statement: 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 Immunology

ORIGINAL RESEARCH
published: 09 January 2019
doi: 10.3389/fimmu.2018.03014

Enhancer Associated Long Non-coding RNA Transcription and Gene Regulation in Experimental Models of Rickettsial Infection

Imran H. Chowdhury[1]*, *Hema P. Narra*[1], *Abha Sahni*[1,2], *Kamil Khanipov*[3], *Yuriy Fofanov*[3] *and Sanjeev K. Sahni*[1,2]*

[1] Department of Pathology, University of Texas Medical Branch, Galveston, TX, United States, [2] Institute for Human Infections and Immunity, University of Texas Medical Branch, University Boulevard, Galveston, TX, United States, [3] Department of Pharmacology, University of Texas Medical Branch, University Boulevard, Galveston, TX, United States

OPEN ACCESS

Edited by:
Adam Williams,
Jackson Laboratory, United States

Reviewed by:
Kushagra Bansal,
Harvard Medical School, United States
Michael F. Minnick,
University of Montana, United States

***Correspondence:**
Imran H. Chowdhury
ihchowdh@utmb.edu
Sanjeev K. Sahni
sksahni@utmb.edu

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *24 July 2018*
Accepted: *05 December 2018*
Published: *09 January 2019*

Citation:
Chowdhury IH, Narra HP, Sahni A, Khanipov K, Fofanov Y and Sahni SK (2019) Enhancer Associated Long Non-coding RNA Transcription and Gene Regulation in Experimental Models of Rickettsial Infection. Front. Immunol. 9:3014. doi: 10.3389/fimmu.2018.03014

Recent discovery that much of the mammalian genome does not encode protein-coding genes (PCGs) has brought widespread attention to long noncoding RNAs (lncRNAs) as a novel layer of biological regulation. Enhancer lnc (elnc) RNAs from the enhancer regions of the genome carry the capacity to regulate PCGs in *cis* or in *trans*. Spotted fever rickettsioses represent the consequence of host infection with Gram-negative, obligate intracellular bacteria in the Genus *Rickettsia*. Despite being implicated in the pathways of infection and inflammation, the roles of lncRNAs in host response to *Rickettsia* species have remained a mystery. We have profiled the expression of host lncRNAs during infection of susceptible mice with *R. conorii* as a model closely mimicking the pathogenesis of human spotted fever rickettsioses. RNA sequencing on the lungs of infected hosts yielded reads mapping to 74,964 non-coding RNAs, 206 and 277 of which were determined to be significantly up- and down-regulated, respectively, in comparison to uninfected controls. Following removal of short non-coding RNAs and ambiguous transcripts, remaining transcripts underwent in-depth analysis of mouse lung epigenetic signatures H3K4Me1 and H3K4Me3, active transcript markers (POLR2A, p300, CTCF), and DNaseI hypersensitivity sites to identify two potentially active and highly up-regulated elncRNAs NONMMUT013718 and NONMMUT024103. Using Hi-3C sequencing resource, we further determined that genomic loci of NONMMUT013718 and NONMMUT024103 might interact with and regulate the expression of nearby PCGs, namely Id2 (inhibitor of DNA binding 2) and Apol10b (apolipoprotein 10b), respectively. Heterologous reporter assays confirmed the activity of elncRNAs as the inducers of their predicted PCGs. In the lungs of infected mice, expression of both elncRNAs and their targets was significantly higher than mock-infected controls. Induced expression of NONMMUT013718/Id2 in murine macrophages and NONMMUT024103/Apol10b in endothelial cells was also clearly evident during *R. conorii* infection *in vitro*. Finally, shRNA mediated knock-down of NONMMUT013718 and NONMMUT024103 elncRNAs resulted in reduced expression of endogenous Id2 and Apl10b, demonstrating the

regulatory roles of these elncRNAs on their target PCGs. Our results provide very first experimental evidence suggesting altered expression of pulmonary lncRNAs and elncRNA-mediated regulation of PCGs involved in immunity and during host interactions with pathogenic rickettsiae.

Keywords: *Rickettsia*, long non-coding (lnc) RNA, enhancer long non-coding (elnc) RNA, RNA sequencing, transcription start site, inhibitor of DNA binding 2 protein, apolipoprotein L 10b, host immune responses

INTRODUCTION

Arthropod-borne *Rickettsia* species include obligate intracellular, Gram-negative bacteria known to cause spotted fever and typhus groups of rickettsial dieases in humans (1). The clinical spectrum of spotted fever group (SFG) rickettsioses varies in severity from mild to fatal cases of Rocky Mountain spotted fever (RMSF) caused by *Rickettsia rickettsii*, Mediterranean spotted fever (MSF) due to *R. conorii*, and Queensland tick typhus following infection with *R. australis* (1). A majority of human rickettsial diseases involve transmission from arthropod vectors, for example naturally circulating infected ticks in case of *R. rickettsii* and *R. conorii*. Due mainly to the predilection of pathogenic rickettsiae to target endothelial cells lining the microvasculature in their mammalian hosts and cell-to-cell spread during the course of infection (1, 2), a prominent feature of pathogenesis is the innate immune activation and inflammatory perturbations of microvascular endothelium, leading to complications such as ocular inflammation or retinitis, myocarditis, endocarditis, pulmonary, and cerebral edema due to fluid imbalance associated with the derangements of endothelial barrier, and multi-organ failure in severe cases (1–6). Employing both patient samples and established experimental models of infection, a number of studies have delved into the definition of host responses during rickettsial infections (7, 8), but the mechanisms underlying the activation and regulation of such immune mechanisms remain largely unknown.

Functional annotation of the mammalian genome (FANTOM) and Encyclopedia of DNA Elements (ENCODE) projects have challenged the central dogma of molecular biology by suggesting that non-protein-coding regions carry multiple overlapping codes that profoundly affect gene expression and other cellular processes. Notably, protein-coding sequences occupy <2% of the genome in mammals, whereas a much larger fraction is transcribed into non-coding RNAs (ncRNAs) (9–12). A majority of ncRNA transcripts are functionally active RNAs broadly classified into short non-coding RNAs (miRNA) of less than and long non-coding (lnc) RNAs of more than 200 nucleotides (13, 14). Short ncRNAs are now established as highly versatile molecules capable of interacting with other RNAs, DNA, or a vast repertoire of proteins, highlighting their regulatory potential (15). In recent years, lncRNAs have also been implicated in diverse major biological processes, including immune regulation, cell cycle, apoptosis, post-transcriptional, and translational regulation, epigenetic modification, and nuclear genome organization, highlighting their regulatory activities in the determination of host-pathogen interactions (12). Among these, an important sub-class of lncRNAs derived from the enhancer loci of the genome are designated as enhancer long non-coding (elnc) RNAs or eRNAs (16–19). Active enhancers are traditionally considered as the principal regulatory components of the genome capable of enabling cell or tissue type and cell-cycle specific gene expression in *cis* and *trans*. As such, elncRNA(s) have received considerable attention by virtue of their ability to control protein coding genes (PCGs) by locus control mechanism (20). Potential elncRNAs are generally characterized by higher occupancy of chromatin monomethylation of histone H3 at lysine 4 (H3K4Me1) (signature of enhancer loci) when compared to trimethylation of histone H3 at lysine 4 (H3K4Me3) (signature of promoter loci) and other epigenetic signatures such as RNA pol II, DNaseI hypersensitivity site, and p300 binding sites at or around the transcription start site (TSS). Enhancer elements in the genome play an active role in controlling the transcription of PCGs by stabilizing enhancer-promoter interactions (16, 21).

Although NCBI chromatin and epigenetic gene expression omnibus (GEO) databases have enabled the identification of cell- and tissue-specific active enhancers in both human and mouse (22, 23), active elncRNAs are currently characterized in only a limited numbers of cells and tissues and their functional roles in the host responses and pathogenesis of rickettsial diseases remain poorly understood. In the present study, we have elucidated lncRNA signatures of the host lungs in a murine model of rickettsial disease and identified two elncRNAs that may be involved in the host response to infection.

Abbreviations: SFG, spotted fever group; RMSF, Rocky Mountain spotted fever; MSF, Mediterranean spotted fever; FANTOM, Functional annotation of the mammalian genome; ENCODE, Encyclopedia of DNA Elements; ncRNAs, non-coding RNAs; lnc, long non-coding; elnc, enhancer long non-coding; plnc, promoter long non-coding; PCGs, protein coding genes; TSS, transcription start site; GEO, gene expression omnibus; RIN, RNA integrity number; RKPM, reads per kilobase million; qRT-PCR, Quantitative real-time PCR; LINC, long intergenic non-coding; TES, transcription end site; CTCF, CCCTC binding protein; SVEC, SV40-transformed mouse endothelial cells; shRNA, short hairpin RNA; RNA-seq, RNA sequencing; rRNA, ribosomal RNA; tRNA, transfer RNA; ChIP-Seq, Chromatin Immunoprecipitation sequencing; H3K4Me1, monomethylation of histone H3 at lysine 4; H3K4Me3, Trimethylation of histone H3 at lysine 4; Apol, apolipoprotein; ISGs, interferon-stimulated genes; Id, Inhibitor of DNA binding; bHLH, basic helix-loop-helix.

MATERIALS AND METHODS

Preparation of *R. conorii* Stocks

Monolayers of cultured Vero cells as the host were infected with *R. conorii* (Strain Malish 7) to allow for intracellular growth and replication of rickettsiae. Heavily-infected cells (infection

of ≥80% of cells with ≥50 intracellular rickettsiae) were gently lysed using glass beads for the isolation and purification of rickettsial stocks by differential centrifugation. The rickettsial preparations were stored at −80°C by slow freezing as aliquots of ≤0.5 ml and gently thawed on ice to avoid loss of viability. The infectious titer of stocks thus prepared was determined by rickettsial citrate synthase (gltA)-based quantitative PCR and plaque formation assays using standard protocols and procedures (24, 25).

Mice and Infection

To identify lncRNA transcripts expressed during *R. conorii* infection, we employed an established mouse model of infection (26). C3H/HeN mice were obtained from the Jackson Laboratory and housed in an ABSL3 laboratory suite. Following acclimatization, the animals were infected with a high dose of *R. conorii* (2.25×10^5 pfu/mouse) administered through the tail vein injection. The control group of animals received identical volume of saline intravenously (26). The animals were then monitored at least once daily for overt signs of disease (ruffled fur, hunched posture, and photophobia) and the body weights were recorded. On day 3 post-infection, mice were euthanized and the lungs were removed aseptically. The tissues thus collected were either snap-frozen or stored at −20°C in RNAlater™ solution. All the animal procedures were performed in accordance to the National Institutes of Health Guide for the Care and use of Laboratory Animals, and were maintained by the approval of Institutional Animal Care and Use Committee at the University of Texas Medical Branch (UTMB) (protocol #1109042). The University has a file with the Office of Laboratory Animal Welfare and an approved Assurance Statement (#A3314-01). Use of any cell line in this study was exempt by Institutional Review Board (IRB), and approved by Institutional Biosafety committee (IBC), UTMB.

RNA Extraction and cDNA Library Preparation

Total RNA from lung tissues was extracted using TRIzol reagent according to the manufacturer's instructions (Invitrogen). RNA samples were subjected to treatment with DNaseI (NEB) to remove any contaminating DNA and then enriched with Ribo-Zero rRNA Removal kit (Illumina). Concentration of RNA in sample preparations using the MultiSkan Go UV/Vis instrument for microsample analysis (Thermo Scientific) and the quality of RNA was evaluated on a bioanalyzer (Agilent Technologies). The samples with an RNA integrity number (RIN) of >9 were subjected to RNA-sequencing (12). Briefly, RNA fragments of 50 bases were generated by incubating purified total RNA in a fragmentation buffer (Ambion) and fragmented RNA was then ligated with 5′ and 3′-adaptors using a T4 RNA ligase (NEB). Adaptor-ligated RNAs were reverse transcribed and subjected to PCR amplification with barcoded primers (Illumina) (27, 28). Finally, amplified cDNA libraries were purified using standard gel purification procedures.

RNA Sequencing, Mapping, and Data Analysis

RNA sequencing was performed on an Illumina HiSeq 1500 at the Next Generation Sequencing Core facility at the UTMB. Briefly, 50 base long reads were obtained from the RNA derived from the lungs of *R. conorii*-infected and uninfected control ($n = 3$ for each) mice. The first 14 bases of the reads were trimmed and only reads with high base quality (phred score >15) were used for downstream analysis. All high quality reads were then grouped according to their designation as infected or control. To identify ncRNAs, all reads were first mapped to *Mus musculus* Ref-seq (mm9) genome (to remove reads from mRNAs), and the remaining unmapped reads were then mapped to known mouse ncRNA transcripts in the NONCODE (NONCODE_V4) database with an allowance of up to two mismatches employing CLC Genomic Workbench 9.0.1 (http://www.clcbio.com) RNA-sequencing Analysis tool. The RNA-sequencing data were normalized by calculating "reads per kilobase million" (RKPM) as described earlier (12). Expression of all mRNA and ncRNA transcripts was determined in each infected sample by dividing the normalized reads from *R. conorii*-infected sample with those from the corresponding mock-infected sample. Mann–Whitney *U*-Test was used to compare the differences in relative abundance of identified lncRNA and mRNA transcripts between groups. We next applied Min/Max method to identify the expression of potential lncRNA candidates and their nearby PCGs. Up-regulation of lncRNA transcripts and/or PCGs was determined as the ratio of the lowest normalized reads in the infected group to the highest normalized reads in the control group ($n = 3$). Conversely, down-regulation of lncRNA transcripts and/or PCGs was ascertained by dividing the highest normalized reads in the infected group with the lowest normalized reads of the sample in the uninfected control group ($n = 3$ for each group). The FASTQ files for RNA sequencing data were submitted to GenBank (Accession number GSE121808).

Quantitative Real-Time PCR (qRT-PCR)

Approximately 1 μg of RNA from mock and *R. conorii*-infected lungs was reverse transcribed using a cDNA Synthesis Kit (Applied Biosystems). The cDNA was subjected to qPCR using SYBR Green as the reporter on a StepOnePlus instrument (Applied Biosystems). PCR reactions were performed in triplicate using the primer sequences listed in **Supplementary Table 3**. The datasets were normalized using 18S RNA as the housekeeping gene. The levels of expression and relative quantification were determined via calculations based on the $2^{-\Delta\Delta Ct}$ method (12).

Cataloging of lncRNAs

To catalog lncRNAs, we captured the strand of origin, nature of origin, chromosomal origin, number of exons, and lengths of all the differentially expressed lncRNAs from NONCODE_V4 database. We grouped lncRNAs based on their strand of origin (sense or anti-sense), source of origin (chromosome number 1–20 and mitochondrial DNA), nature of origin [sense-exonic, sense intronic, antisense, antisense-exonic, antisense intronic and LINC (long intergenic non-coding) RNA], exonic

composition (uni-exonic, bi-exonic and multi-exonic), and the length of transcripts (length ∼200–500, 501–2,000, 2,001–5,000, and ≥5,001 bp) as described earlier (12).

TSS Evidence and Filtering of Up-Regulated lncRNA Transcripts

We utilized UCSC genome browser to further categorize up-regulated lncRNA transcripts based on their origin and orientation to the nearby PCGs. We cataloged them into "head to head," "head to tail," "tail to tail," and "tail to head" orientation, and all these classifications were utilized for selection of transcripts for further downstream analysis (29).

The TSSs of lncRNAs and nearby PCGs as reported in the NONCODE database and UCSC genome browser (www.genome.ucsc.edu), respectively, were used to compute the distance of lncRNAs to the nearest PCGs for downstream filtering of transcripts. To identify up-regulated elncRNAs, lncRNA transcripts originating from sense-exonic, sense intronic, antisense to the PCGs, antisense-exonic, antisense intronic and LINC transcripts for which TSS are within a 2 kb window of TSS or transcription end site (TES) of nearby PCGs (29, 30), were excluded from the analysis. The remaining lncRNA transcripts were analyzed for chromatin and epigenetic signatures as described below.

Analysis of ChIP-Seq GEO Data for Chromatin and Epigenetic Signatures

We performed quantitative assessment of chromatin signatures H3K4me1 and H3K4me3 around the TSS of up-regulated lncRNAs to identify elncRNAs. Briefly, ChIP-Seq datasets for H3K4me1 (GSM769013) and H3K4me3 (GSM769012) in the mouse lung were downloaded from the NCBI GEO database (http://www.ncbi.nlm.nih.gov/gds). The genomic sequence of 4 kb around the TSS of filtered lncRNAs was extracted and mapped to the ChIP-seq datasets using Samtools. The reads were normalized by subtracting the reads of lung input (GSM 918739) within same 4 kb window. Based on the profile of chromatin signatures around the TSS of lncRNAs, we cataloged them into either enhancer or promoter (or canonical) lncRNAs. To identify active elncRNAs in the mouse lung, we utilized GEO data for RNA Pol II (GSM918724), p300 co-activator binding site (GSM722862), DNaseI hypersensitivity site (GSM1014194), and CTCF (CCCTC binding protein) binding sites (GSM918722). The mapping of reads was performed with an allowance of upto two base mismatches, and all reads mapping within the 4 kb window around the TSS were used for identification of active elncRNAs as described earlier (31).

Analysis of Hi-3C GEO Data and qRT-PCR

To determine the interaction(s) between an active elncRNA and its nearby PCGs, we applied a combinatorial approach based on the analysis of high throughput sequencing of Chromosome Conformation Capture (Hi-C) and virtual 4C profiles by 3D Epigenome browser (www.3dgenome.org). We uploaded the Hi-C tracks to identify the signals for interactions between the coordinates of elncRNAs and potential PCGs. For further confirmation, we utilized visualization of virtual 4C profiles to identify the location of genomic contact loci of elncRNA in relation to the anchoring point for the promoter of nearby PCGs (32, 33). We next performed qRT-PCR for active elncRNAs and their targets on the RNA from infected mouse lungs to investigate the possibility of correlative changes in their expression. The primer sequences for qRT-PCR are listed in **Supplementary Table 3**.

Cell Culture and Infection

Murine RAW264.7 macrophages, NIH3T3 fibroblasts, and SV40-transformed mouse endothelial cells (SVEC) 4-10 were maintained at 37°C with 5% CO_2 in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (Aleken Biologicals), 10 mM L-glutamine (Thermo Fisher Scientific), 100 Units each of Penicillin and Streptomycin according to previously published procedures (34–36). For experiments with rickettsiae, the Penicillin-Streptomycin mix was removed from the culture medium a minimum of 24 h prior to and during the infection. All experiments were performed with exponentially growing cells at relatively low passage numbers of 5–15. Because endothelial cells and macrophages are the major targets of rickettsial infection, we infected RAW264.7 macrophages and SVECs with *R. conorii* for 3 and 24 h (MOI = 5).

Reporter Constructs and Transient Transfection

Genomic loci of active elncRNAs flanking the boundaries of epigenetic signature were PCR amplified using Phusion High Fedelity DNA polymerase (NEB). The purified PCR fragments were cloned in both sense and antisense orientation upstream of the SV40 promoter in a pGL3 firefly promoter plasmid (Promega). Simultaneously, we picked two other genomic coordinates flanking Chr18: 60429728-60430126 and Chr13: 60430551-60430958 with a minimal ratio of H3K4Me1/H3K4Me3 and negligible peaks of other epigenetic signatures (RNA pol II, P300, and DNaseI hypersensitivity site) and cloned them into the pGL3 promoter plasmid to serve as negative controls in our experiments. All inserts in the promoter plasmid were confirmed by DNA sequencing at the UTMB Molecular Genomics Core facility. The primer sequences and restriction sites of inserts are also listed in **Supplementary Table 3**. Transfection-grade, endotoxin-free plasmids were prepared using an EndoFree® Plasmid Purification kit (Qiagen). We transfected these plasmids along with a pRL-SV40 plasmid as an internal control in mouse NIH3T3 fibroblasts and RAW264.7 macrophages at about 80% confluence. For each assay, 1 μg of blank plasmid (pGL3 promoter plasmid) or elncRNA constructs or negative controls, and 200 ng of pRL-SV40 were co-transfected using Lipofactamine 3000 (Invitrogen). After 24 h of transfection, cell lysates were prepared and dual luciferase assay was performed according to manufacturer's instructions (Promega). Firefly and Renilla luciferase signals were recorded using a GloMax® 20/20 Single-Tube Luminometer (Promega). The signal ratio in each well was calculated by dividing the luciferase signal by Renilla signal.

Knock-Down of elncRNAs and qRT-PCR

Two distinct targets based on the published guidelines (37) were chosen to design short hairpin RNAs (shRNAs) for elncRNA knock-down using an RNAi consortium designing tool (www. broadinstitute.org). To avoid potential confounding effects of non-specific knock-down, shRNA sequences were further verified through a BLAST search in the NCBI. The shRNAs were then cloned into a pLKO.1 lentivirus puro vector (Addgene plasmid #8453; Addgene, Cambridge, MA, USA), followed by sequencing at the UTMB sequencing core to confirm the orientation of the insert. The shRNA target sites and sequences are listed in the **Supplementary File 1** and **Supplementary Table 3**. Endotoxin-free plasmid preparations and transfection of plasmids carrying shRNA hairpin constructs or scrambled sequences (control) were carried out as detailed above. Transfected cells were allowed to recover for 24 h prior to infection with *R. conorii*. Efficiency of knock-down was confirmed by qPCR assay.

Statistical Analysis

D'Agostino & Pearson omnibus normality test was performed to ensure normal distribution of data. Comparisons between the unmatched groups were done by unpaired *t*-test or Mann–Whitney *U*-test, whereas comparisons amongst the matched groups were performed by paired *t*-test or Wilcoxon signed rank test. The correlative analysis was performed using Spearmann correlation test. GraphPad Prism version 5 (GraphPad Software Inc., San Diego, California) was used for all statistical analyses with $P \leq 0.05$ suggesting statistically significant changes.

RESULTS

To determine changes in the lncRNA profile during rickettsial infection, we performed RNA sequencing (RNA-seq) on the lungs of susceptible mice infected with *R. conorii* on day 3 post-infection. The step-by-step schematics for the methods utilized and decision points are presented in **Supplementary Figure 1**. A total of 152.46 ($n = 3$) and 160.3 ($n = 3$) million reads were obtained from the lungs of mock- and *R. conorii*-infected mice, respectively. We first mapped the libraries to the Ref-Seq to remove reads originating from the annotated *Mus musculus* (mm9) coding transcripts and the remaining 106.71 and 108.08 million reads were then mapped to NONCODE_V4 database containing 74,964 ncRNA transcripts (38). The total number of reads from each cDNA library and the reads mapping to mRNAs and ncRNAs are presented in **Supplementary Table 1**. As expected, the reads mapping neither to Ref-seq genes nor ncRNAs predominantly corresponded to the ribosomal RNA (rRNA) and transfer RNA (tRNA) transcripts. All coding and ncRNA transcripts with undetectable expression were excluded from the analysis. We thus identified a total of 1,168 and 6,216 ncRNAs that were either up- or down-regulated, respectively, in the lungs of infected mice (Cut-off fold ≥ 3, $P \leq 0.05$) (**Supplementary Figure 2**). We next applied Min/Max method to identify ncRNA transcripts exhibiting a high degree of regulation in response to infection, restricting the number of up- and down-regulated ncRNAs in our datasets to 206 and 277, respectively, of which further removal of any ncRNA transcripts with a length of <200 bp allowed us to retain 179 up-regulated and 271 down-regulated lncRNA transcripts (**Figure 1A**). To validate our results from RNA-seq, we randomly selected 4 up-regulated ncRNA transcripts (NONMMUT007594, NONMMUT019215, NONMMUT024102, and NONMMUT029515) to independently determine their expression levels by qRT-PCR. Notably, the expression of all of these transcripts in the lungs of *R. conorii*-infected mice was significantly higher, albeit to varying degrees in terms of average fold-induction, than the corresponding controls (**Figure 1B**). Finally, comparison of our data from RNA sequencing and qRT-PCR using Spearmann correlation analysis revealed high level of correlation, confirming excellent agreement between the findings from two independent approaches ($r^2 = 0.98$) (**Figure 1C**).

We next cataloged differentially expressed lncRNAs based on their strand of origin, classification, chromosomal distribution, number of exons, and length. Of up-regulated lncRNAs, the distribution on sense and antisense strands was determined to be 87 (48.6%) and 92 (51.4%), whereas a total of 140 (51.7%) and 131 (48.3%) down-regulated lncRNAs were found to be transcribed from the sense and anti-sense strands, respectively, (**Figure 2A**). Based on the specifics of their origin, we categorized them into different classes, namely sense-exonic, sense intronic, antisense, antisense-exonic, antisense intronic and LINC (**Supplementary Figure 3A**). Of 179 up-regulated lncRNA transcripts, a majority (113 or 63.1%) were sense-exonic and the remaining included 37 (20.7%) LINC, 12 antisense-exonic (6.7%), 12 antisense (6.7%), 3 sense-intronic (1.7%), and another 2 antisense-intronic (1.1%). The down-regulated lncRNA transcripts were represented by 121 LINC (44.6%), 63 sense-exonic (23.2%), 48 antisense (17.7%), 36 sense-intronic (13.3%), and 3 antisense-exonic (1.1%) (**Figure 2B**). Further analysis suggested that most of the differentially expressed lncRNA transcripts are transcribed from chromosome 1, 4, 6, 7, 11, and 13. Majority of the up-regulated lncRNA transcripts are sense-exonic in nature and mainly transcribed from chromosome 4, 6, 7, 11, 16, and 19. However, majority of the down-regulated transcript are LINC in nature and predominantly transcribed from chromosome 1, 2, 6, 7, 15, and 16 (**Figures 2C,D**). Based on exonic composition, about 27 (15.08%) of up-regulated lncRNA transcripts are mono-exonic, 54 (30.17%) are bi-exonic, and remaining 98 (54.75%) transcripts are multi-exonic. On the otherhand, about 31 (11.44%), 59 (21.77%), and 181 (66.79%) down-regulated transcripts are mono-exonic, bi-exonic and multi-exonic, respectively, (**Supplementary Figure 3B**). Next, we performed size based cataloging of differentially regulated lncRNAs. Majority of the regulated lncRNA transcripts range from 501 to 2,000 nucleotides ($n = 106$, 59.22% for up-regulated and $n = 148$, 54.61% for down-regulated), followed by those ranging from 200 to 500 nucleotides including both up-regulated ($n = 37$, 20.67%) and down-regulated ($n = 59$, 21.77%) transcripts. About 28 (15.64%) and 46 (16.97%) up- and down-regulated transcripts are within the range of 2,001–5,000 nucleotides. As expected, only a low number of

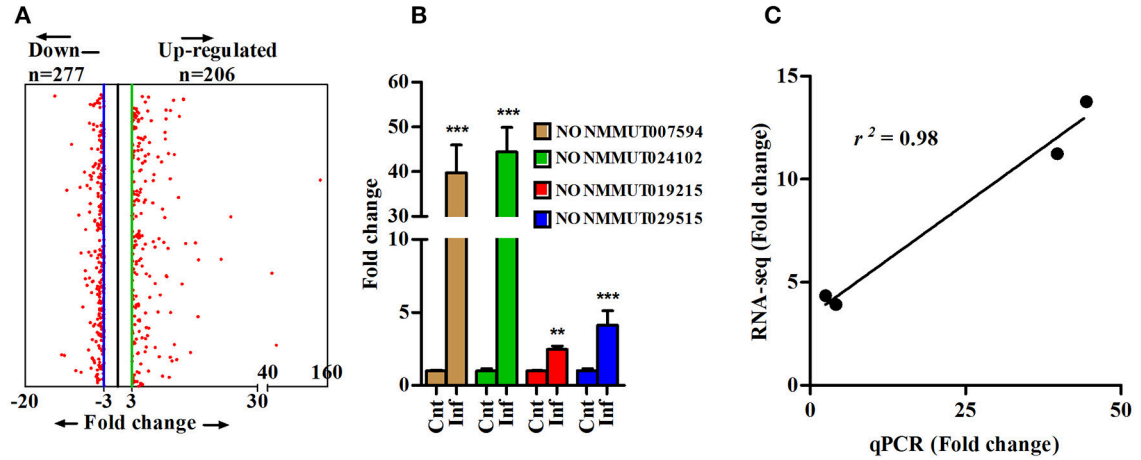

FIGURE 1 | Differentially expressed lncRNA transcripts in lungs of mouse infected with *R. conorii*. **(A)** Differentially expressed lncRNA transcripts ($n = 3$, control and infected mice) based on Min/Max method (3-fold cut-off); **(B)** Validation of expression of four randomly selected up-regulated lncRNA transcripts (from RNA-sequencing analysis) by qPCR in infected mouse lungs ($n = 4$). The error bars represents standard error of mean (SEM) and the level of significance are shown as $^{**}P \leq 0.01$, and $^{***}P \leq 0.001$; and **(C)** Spearmann correlation of fold change of 4 lncRNA transcripts as determined by RNA sequencing and qPCR methods ($r^2 = 0.98$, $P < 0.001$).

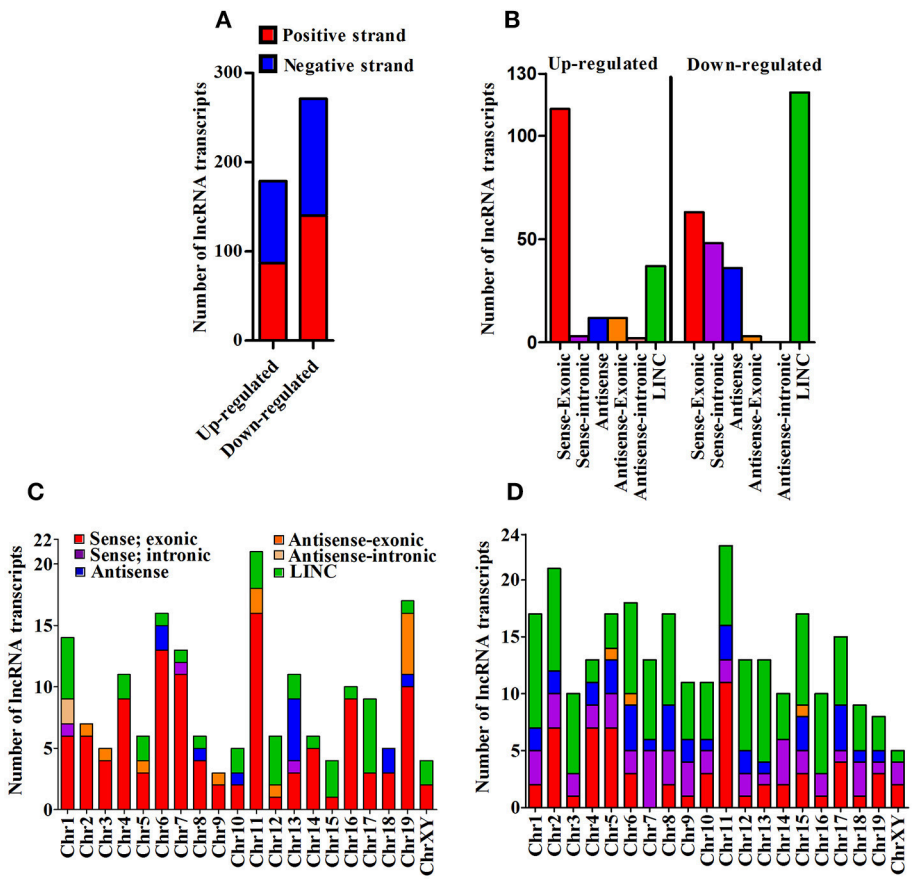

FIGURE 2 | Cataloging of differentially expressed lncRNA transcripts. **(A)** Strand specific origin of differentially expressed lncRNA transcripts; **(B)** Cataloging of lncRNA transcripts based on their origin; **(C)** Chromosomewise distribution of category of up-regulated lncRNA transcripts; and **(D)** Chromosomewise distribution of category of down-regulated lncRNA transcripts.

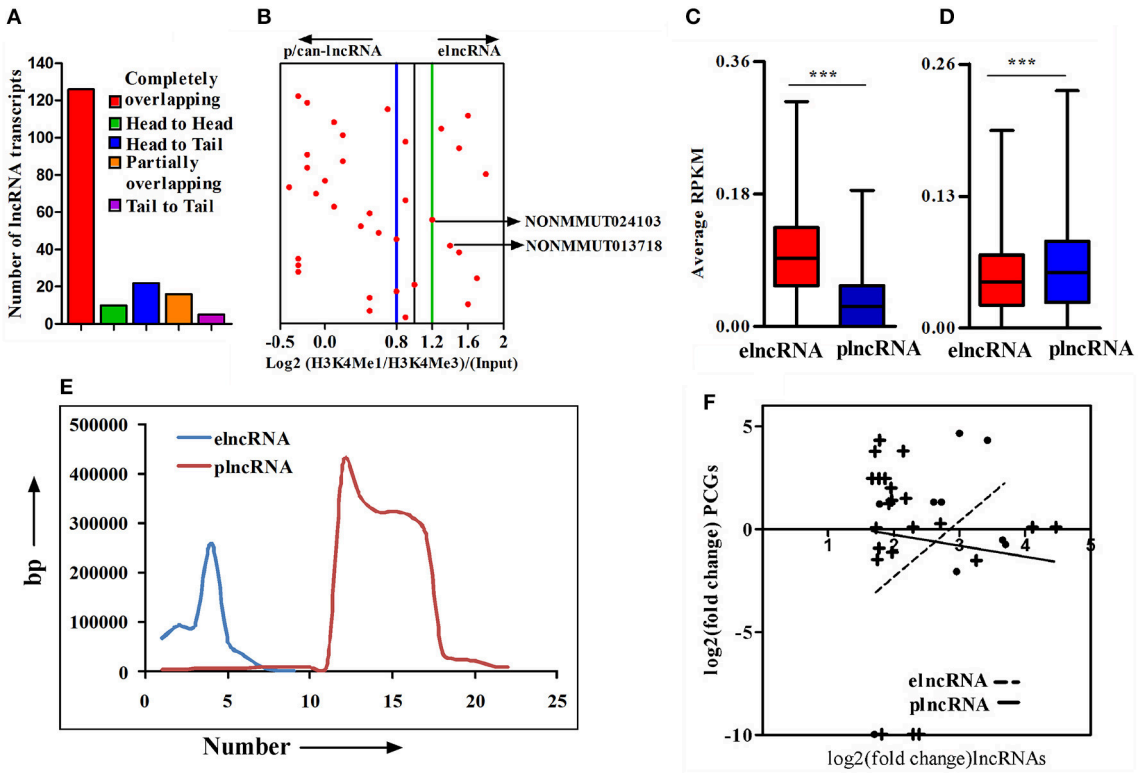

FIGURE 3 | Cataloging of up-regulated lncRNA transcripts based on chromatin signatures. **(A)** cataloging of up-regulated lncRNA transcripts based on their orientation with respect to nearby PCGs; **(B)** Representative plot showing log2 ratio of chromatin signatures H3K4Me1 and H3K4Me3 (after normalization with input) around 4 window of transcription start site (TSS) of lncRNAs; **(C)** Average normalized RPKM (reads per kilobase million) values of H3K4Me1 in elncRNAs and plncRNAs; **(D)** Average normalized RPKM (reads per kilobase million) values of H3K4Me3 in elncRNAs and plncRNAs; **(E)** Distribution of distances of elncRNAs and plncRNAs from their respective nearby genes; and **(F)** Correlation between fold changes of elncRNAs and plncRNAs and their closest protein coding genes (PCGs), respectively. The linear regression curves of the best fit are shown as dotted line for elncRNA and solid line for plncRNA. ***$P \leq 0.001$.

up-regulated (n = 8, 4.47%) and down-regulated (n = 18, 6.64%) transcripts belong to the category of ≥5,001 nucleotides (**Supplementary Figure 3C**).

An important chromatin signature for identification of enhancers is the combination of H3K4Me1 and p300 binding in the absence of H3K4Me3, a mark that has been classically associated with active or poised TSSs. Accordingly, the ratio of H3K4me1/H3K4me3 around TSSs is a useful indicator to segregate lncRNAs into enhancer-associated (elnc) or promoter-associated (plnc) RNAs. We, therefore, recorded the TSSs of up-regulated lncRNAs and their nearest PCGs from the NONCODE database and UCSC genome browser, respectively. The combinatorial evidence of TSS, chromosomal origin, and relative position of lncRNAs to PCGs was then utilized to classify up-regulated lncRNAs into completely overlapping, partially overlapping, head to head, head to tail, and tail to tail category (**Supplementary Figure 4A**). A majority of up-regulated lncRNAs (126 out of 179) were determined to be completely overlapping and a few (n = 16, 8.92%) were from loci partially overlapping with PCGs. Of those remaining, 10 (5.92%), 22 (12.29%), and 5 (2.79%) transcripts were belonging to head to head, head to tail, and tail to tail orientation, respectively, (**Figure 3A**). Based on all the cataloging, we excluded lncRNA transcripts belong to classes sense-exonic, antisense-exonic and sense-intronic lncRNAs, to prevent the confounding influence of reads of chromatin and epigenetic signatures from the overlapping mRNAs. Furthermore, orientation of lncRNAs with respect to nearest PCGs also eliminated any ambiguous LINC lncRNAs, for which the TSSs are located within 2 kb region of both the head and tail ends of nearby PCGs. We, however, retained antisense lncRNAs for which the location of TSSs was 2 kb beyond either end of nearby PCGs. We next employed ChIP-seq data for enriched chromatin state around a 4 kb window of the TSS to determine the Log2(H3K4me1: H3K4me3) ratio for classifying the up-regulated lncRNAs as elncRNA (ratio of ≥1.2) or plnc/can-lncRNA (ratio of ≤0.8). Such quantitative analysis of H3K4Me1/H3K4Me3 ratio allowed for the designation of a total of 9 and 22 transcripts, respectively, as elnc and plncRNAs, whereas 4 transcripts did not clearly qualify for either category (**Figure 3B**). Genomic annotations of elncRNAs, plncRNAs and their respective nearby PCGs are presented in **Supplementary Table 2**.

We further compared average normalized reads for a 4 kb window around the TSSs of elnc and plncRNAs to ensure that

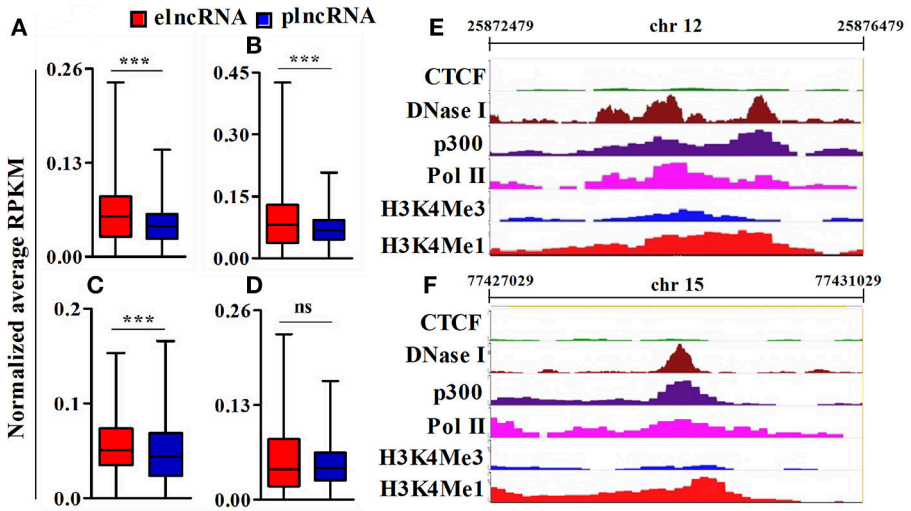

FIGURE 4 | Analysis of epigenomic signatures around transcription start site (TSS) of elncRNAs and plncRNAs. **(A–D)** Average normalized RPKM (reads per kilobase million) values of RNA polII, p300, DNaseI hypersensitivity site, and CCCTC binding factor (CTCF) in elncRNAs and plncRNAs, respectively; **(E,F)** Contrast chromatin (H3K4Me1 and H3K4Me3) and epigenetic (RNA polII, p300, DNaseI hypersensitivity and CTCF binding site) landscapes (in mouse lungs) around TSS of NONMMUT013718 and NONMMUT024103 elncRNAs, respectively. $^{***}P \leq 0.001$ and ns = non-significant.

elncRNAs and plncRNAs are enriched with H3K4Me1 and H3K4Me3, respectively. As expected, normalized average read densities for H3K4Me1 in elncRNA were significantly higher than plncRNA, while those for H3K4Me3 were significantly higher for plncRNA (**Figures 3C,D**). Normalized RPKM values against the positions of 4 kb window are presented in the **Supplementary Figures 4B,C**. Because mechanistic investigations of the regulation of PCGs have implicated elncRNA interactions with looping factors to facilitate chromosomal looping between the enhancer and the promoter(s) of target gene(s), we captured the distances of up-regulated elncRNAs and plncRNAs in relation to the position of proximal PCGs. This analysis suggested that on an average, elncRNAs were located in close proximity to the PCGs, when compared to plncRNAs (**Figure 3E**). Moreover, to test whether transcription of elncRNAs is responsible for the regulation of nearby PCGs, we estimated the correlation between changes in the expression of elncRNAs ($r^2 = 0.3$) and plncRNAs ($r^2 = 0.01$) with their nearby PCGs. As shown, correlation of expression of elncRNAs and their nearby PCGs was determined to be stronger than plncRNAs (**Figure 3F**).

To identify active elncRNAs, we further subjected our dataset for the analysis of other epigenetic signatures, namely RNA pol II, p300, DNaseI hypersensitivity site, and CTCF, within the same 4 kb window around the TSSs of elnc and plncRNAs. The normalized average RPKM values suggested significantly higher read densities of RNA pol II, p300, and DNaseI hypersensitivity sites for elncRNAs in comparison to plncRNAs (**Figures 4A–C**), but not for CTCF (**Figure 4D**). Normalized RPKM values against the positions of 4 kb window are presented in the **Supplementary Figures 5A–D**. Based on epigenetic landscapes, we identified 3 active elncRNA transcripts NONMMUT013718 (**Figure 4E**), NONMMUT024103 (**Figure 4F**), and NONMMUT013717 (a splice variant of NONMMUT013718).

To delineate the possibility of interactions between elncRNAs and the promoters of their nearby PCGs, we sequentially analyzed both Hi-C and virtual 4C profiles by constructing a window around the genomic coordinates of active elncRNAs by 3D Epigenome browser. The triangular heatmap for elncRNA NONMMUT013718 demonstrates the potential for interactions with the PCG ID2 (**Figure 5A**). To further validate this observation and to identify elncRNA regions in contact with the promoter of Id2, we applied virtual 4C profiles supported by 3D genome browser. There are several points of contact for the promoter of Id2 within a 1 Mb window around the anchoring point, including the highest peak representing the potential for interactions with NONMMUT013718 elncRNA (**Supplementary Figure 6A**). On the other hand, NONMMUT024103 elncRNA regulatory region is just 2.4 kb upstream and relatively adjacent to the promoter of nearby Apol10b gene and a signal of interaction between these two loci is evident in Hi-C heat map (**Figure 5B**). Similarly, virtual 4C profile analysis also indicates that the promoter of Apol10b lies in apparent contact with NONMMUT024103 elncRNA (**Supplementary Figure 6B**). Finally, to ascertain whether or not higher expression of elncRNAs NONMMUT013718 and NONMMUT024103 correlates with the expression of their respective target genes ID2 and Apol10b, we performed qRT-PCR on the RNA from lungs of mice infected with *R. conorii*. We observed higher expression of both elncRNAs NONMMUT013718 and NONMMUT024103 as well as their

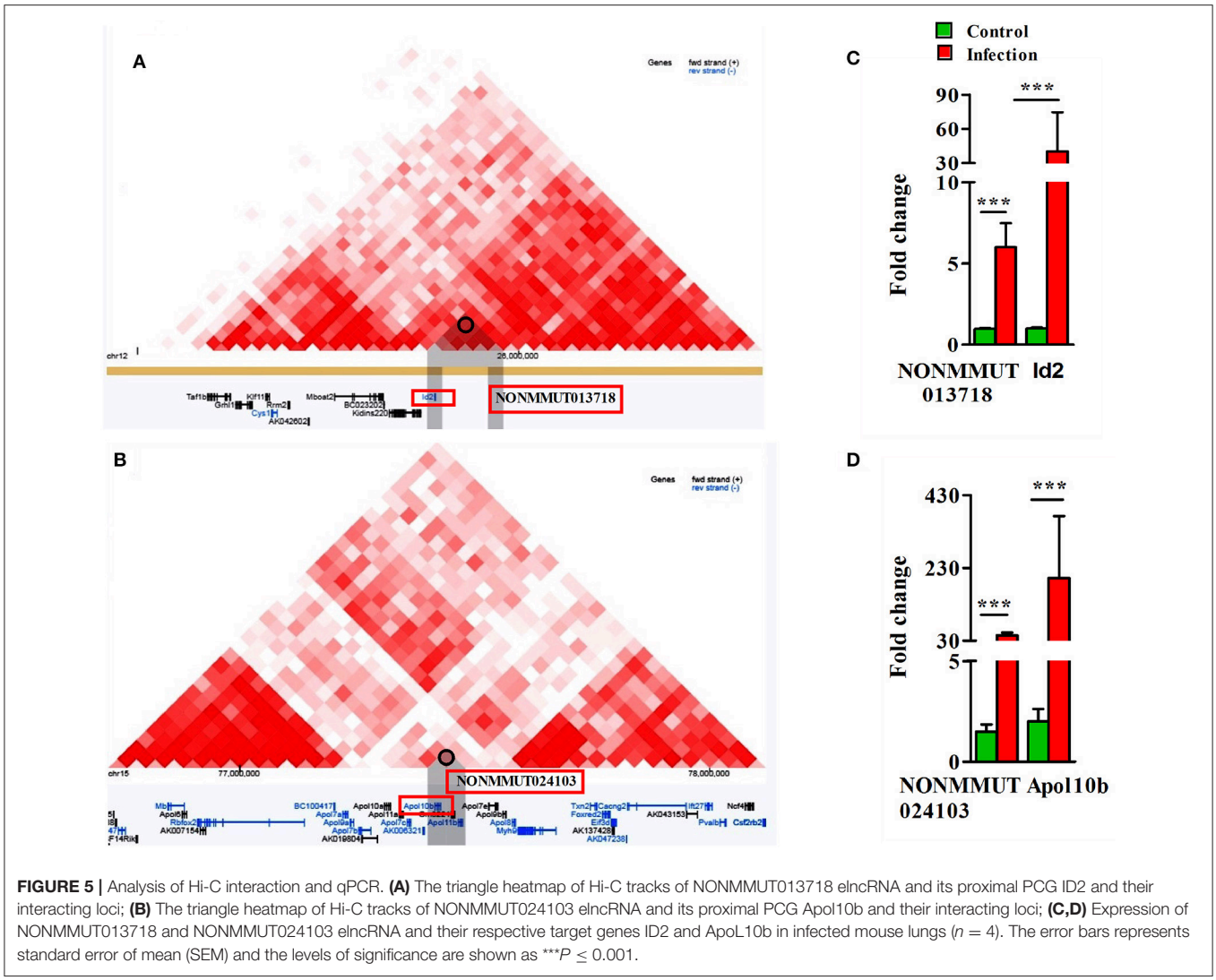

FIGURE 5 | Analysis of Hi-C interaction and qPCR. **(A)** The triangle heatmap of Hi-C tracks of NONMMUT013718 elncRNA and its proximal PCG ID2 and their interacting loci; **(B)** The triangle heatmap of Hi-C tracks of NONMMUT024103 elncRNA and its proximal PCG Apol10b and their interacting loci; **(C,D)** Expression of NONMMUT013718 and NONMMUT024103 elncRNA and their respective target genes ID2 and ApoL10b in infected mouse lungs ($n = 4$). The error bars represents standard error of mean (SEM) and the levels of significance are shown as $^{***}P \leq 0.001$.

target genes ID2 and Apol10b in the mouse lungs in response to infection (**Figures 5C,D**).

To assess the potential regulatory effects of NONMMUT013718 and NONMMUT024103 elncRNAs on nearby PCGs, we performed dual luciferase assay using NIH3T3 and RAW264.7 cells in light of their high transfection efficiency. For NONMMUT013718 elncRNA, we observed a significant increase of luciferase activity, suggesting its ability to drive the downstream PCG in comparison to the blank plasmid (pGL3) as well as negative controls in both NIH3T3 and RAW264.7 cells. Similarly, increased luciferase signal activity was also evident in case of NONMMUT024103 elncRNA, when compared with blank pGL3 and corresponding negative controls (**Figures 6A–D**). To further investigate whether transcriptional activation of nearby PCGs by elncRNAs is orientation independent, we cloned the genomic regions of both elncRNAs in reverse orientation in pGL3 promoter plasmid. Consistent with the findings above, both elncRNAs (NONMMUT013718 and NONMMUT024103) in the reverse orientation significantly enhanced luciferase signal in comparison to the blank plasmid and negative controls (**Figures 6E–H**).

Importantly, pulmonary vascular cells (endothelial cells from different vascular structures, smooth muscle cells, and adventitial fibroblasts) comprise one of the main functional and structural cell types of the lung and resident macrophages are located in close proximity to the epithelial surface of the respiratory system. Since endothelial cells and macrophages are the predominant targets of rickettsial infection, we carried out q-RT-PCR measurements on RNA isolated from murine RAW264.7 macrophages and endothelial cells (SVECs) infected with *R. conorii* for 3 and 24 h. The expression of both elncRNA NONMMUT013718 and its target ID2 in infected macrophages was significantly higher than mock-infected controls at both 3 and 24 h post-infection (**Figure 7A**), while that of NONMMUT024103 elncRNA and its target Apol10b both were below the range of detection at either 3 or 24 h post- infection (data not shown). On the other hand, expression of elncRNA NONMMUT024103 and its target Apol10b was

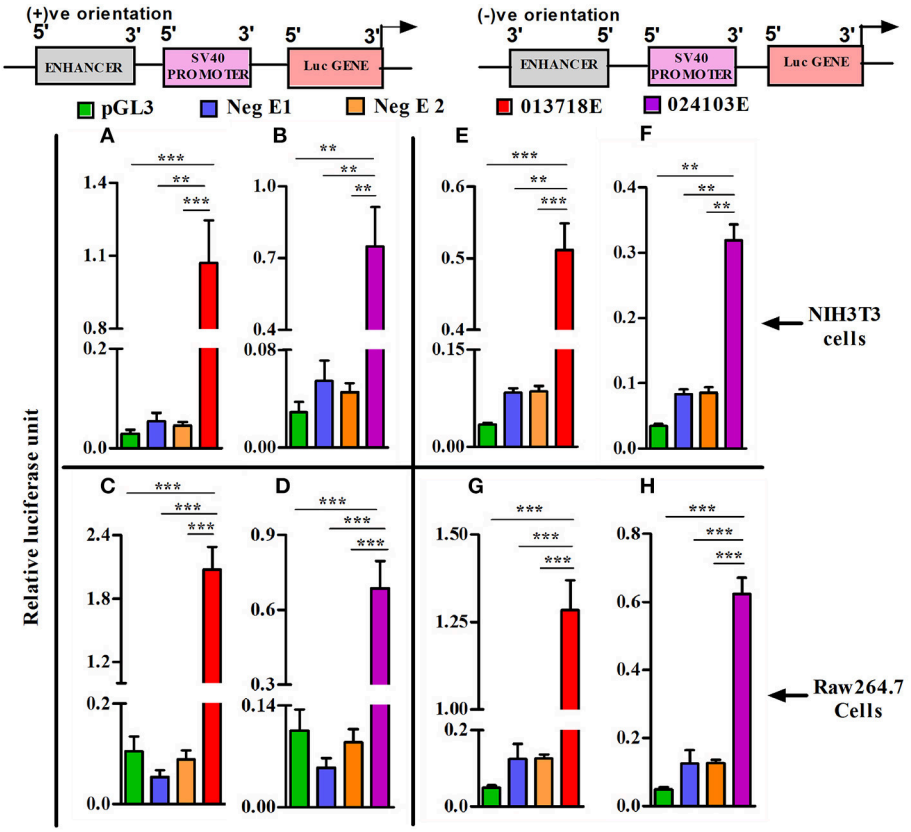

FIGURE 6 | Heterologous reporter assays demonstrating enhancer activity of elncRNA loci. Co-transfection of enhancer reporter construct (sense or antisense orientation) and Renilla reporter plasmid in NIH3T3 cells and Raw 264.7 murine macrophages. **(A,B)** Transfection of NONMMUT013718 and NONMMUT024103 elncRNA constructs (sense orientation) in NIH3T3 cells, and **(C,D)** in Raw 264.7 murine macrophages, respectively; **(E,F)** Transfection of NONMMUT013718 and NONMMUT024103 elncRNA constructs (anti-sense orientation) in NIH3T3 cells, and **(G,H)** in Raw 264.7 murine macrophages, respectively. The data are presented as Mean ± SEM for six independent ($n = 6$) experiments. The negative controls are blank promoter plasmid or inserts with minimal H3K4Me1/H3K4Me3 ratio (Neg E1 and E2). The levels of significance are shown as $^{**}P \leq 0.01$; $^{***}P \leq 0.001$.

significantly higher in SVECs at 3 h post-infection, while only modest increase in NONMMUT024103 was noticeable at 24 h (**Figure 7B**). Expression of NONMMUT013718 elncRNA remained below the limit of detection, whereas its target ID2 remained below the level of control at both 3 and 24 h of infection in SVECs (**Supplementary Figure 7**).

To further confirm the functional role of NONMMUT013718 and NONMMUT024103 in activation of their respective target genes ID2 and Apol10b, we constructed shRNA plasmids against the target site of both NONMMUT013718 and NONMMUT024103 elncRNAs. We chose RAW264.7 macrophages and SVECs for shRNA mediated knock-down based upon higher expression of NONMMUT013718 and NONMMUT024103 elncRNAs, respectively. We verified transfection and knock-down efficiency of NONMMUT013718 and NONMMUT024103 elncRNAs in RNA isolated from RAW264.7 macrophages and SVECs infected with *R. conorii* for 3 and 24 h, respectively. After infection following transfection, expression level of NONMMUT013718 and NONMMUT024103 was found to be significantly lower. Remarkably, knock-down of NONMMUT013718 and NONMMUT024103 down-regulated the expression of ID2 gene in macrophages, and Apol10b in SVECs (**Figure 8**).

DISCUSSION

One of the most remarkable findings of the Human Genome Project is that only about 2% of the DNA accounts for ∼20,000 protein-coding genes. Accordingly, it has become increasingly apparent within the past few years that noncoding genome plays an important role in the regulation of coding genome and substantial progress has been made in assessing the contributions of small, single-stranded noncoding microRNAs as regulatory determinants of host responses following infection, immunization, and autoimmunity (39, 40). lncRNAs have been estimated to constitute about 70–90% of the genomic dark matter, which is generically defined as the transcribed yet untranslated component of the human genome. Although loss-of-function approaches have implicated lncRNAs in the biology of both innate and adaptive immune cells during inflammatory insults, their involvement in host responses and pathogenesis

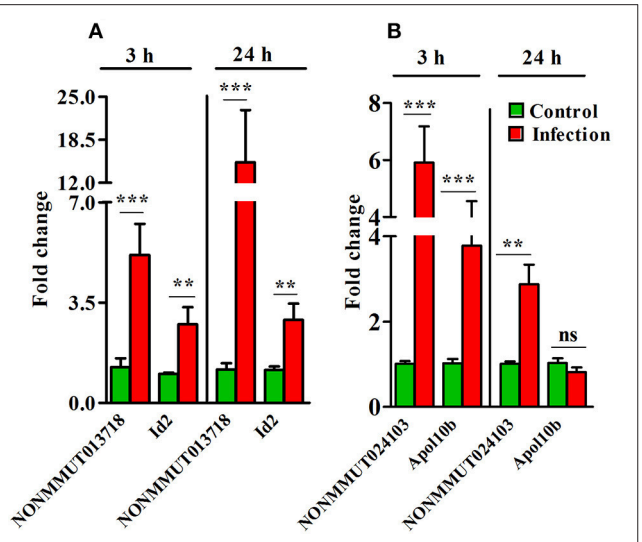

FIGURE 7 | Expression of elncRNAs and their targets in an *in vitro* model of *R. conorii* infection. **(A)** Expression of NONMMUT013718 and its target Id2 in murine Raw264.7 macrophages after infection with *R. conorii* for 3 and 24 h ($n = 6$); **(B)** Expression of NONMMUT024103 and its target Apol10b in murine SVEC endothelial cells after infection with *R. conorii* for 3 and 24 h ($n = 4$). The error bars represent standard error of mean (SEM) and the levels of statistical significance are shown as $^{**}P \leq 0.01$; $^{***}P \leq 0.001$; and ns, non-significant.

of intracellular bacteria remains to be explored to enhance our understanding of their roles in microbial infections. In the present study, we have investigated the global lncRNA profile in the lungs of susceptible murine hosts during infection with *R. conorii*, the causative agent of Mediterranean spotted fever. For this first investigation, we employed a well-established mouse model based on its documented versatility to understand the pathogenesis of human spotted fever group rickettsioses and focused on the lungs as one of the major organ systems targeted by rickettsiae during *in vivo* infection (1, 41). Surprisingly, our RNA sequencing analysis revealed a relatively large number of differentially expressed non-coding transcripts in the lungs during *R. conorii* infection, including 179 up-regulated and 271 down-regulated transcripts of >200 bases length annotated as the lncRNAs in the NONCODE (v4.0) database, an integrated web-based resource dedicated to the analysis of non-coding RNAs (excluding tRNAs and rRNAs). To ensure the consistency of observations, we further validated increased expression of four randomly selected lncRNAs via an independent quantitative PCR-based approach. Thus, although our global analysis was suggestive of potentially important roles for lncRNAs in the determination of coding transcriptome of the host cells or organ systems following infection, an obvious next step emerging from the initial studies was to identify lncRNA candidates with either known functions or potential functional implications. To this end, in-depth analysis of the patterns of histone protein methylation from ChIP sequencing data in conjunction with high throughput chromosome conformation capture sequencing and visualization of Hi-C data in a virtual 4C format revealed that two elncRNAs NONMMUT013718 and NONMMUT024103 may have downstream roles in the regulation of their respective target genes Id2 and Apol10b during rickettsial infection.

Histone proteins in the eukaryotic genome undergo several covalent post-translational modifications, including acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation. Such modifications have a profound effect on gene expression by altering chromatin structure or recruiting modifiers of chromatin activity. Single and trimethylation of histone protein 3 at lysine 4 (H3K4Me1 and H3K4Me3) are well-established features for cataloging of enhancer and promoter elements in the genome (18, 29). Specifically, active enhancers are highly enriched with the presence of H3K4me1 and p300/CBP transcription co-activator binding sites. p300/CBP are two similar acetyltransferases in humans, which play a central role in the pathways responding to intracellular, extracellular, and intercellular signals. These pathways control key cellular functions via altering expression of target genes, through the action of p300/CBP in the nucleus. Open chromatin in the genome is generally ascertained by DNaseI hypersensitivity mapping, however other regulatory elements of the genome such as the promoters, silencers, and insulators also possess DNaseI hypersensitivity sites, rendering exclusive analysis of these sites insufficient for the identification of enhancers (42). Therefore, DNaseI hypersensitivity sites in conjuntion with the p300/CBP protein in the genomic regions indicate the presence of enhancer elements in the genome. p300/CBP recruits RNA polymerase II at the site of enhancers for transcription of elncRNAs (43) and a previous study has documented that a number of uni- and bidirectional elncRNA transcripts have higher occupancy of H3K4Me1, p300/CBP, and RNA pol II (16, 43). On the other hand, CCCTC-binding protein (CTCF) in the genome is considered a hallmark for potential insulator elements that inhibit transcription. Presence of CTCF binding sites in the same domain of enhancer and promoter of the PCG blocks the interaction between these regulatory elements of the genome (31). Based on the H3K4Me1 to H3K4Me3 ratio, we stringently annotated up-regulated lncRNA transcripts into elncRNAs and plncRNAs. Our analysis showed that most of the elncRNA candidates are from intergenic regions and tend to be closely located to the nearby protein coding genes. Moreover, a pattern of positive correlation of expression between elncRNAs and nearby PCGs indicates that these elncRNAs might be associated with nearby PCGs for their expression and function. In addition, examination of the transcription start sites of NONMMUT013718 and NONMMUT024103 within the Gene Expression Omnibus dataset on mouse lung tissues further suggests the likelihood of active elncRNA functions for these transcripts. Enhancers have been proposed to interact with their target promoters by different mechanisms based on their genomic positions. For example, enhancers interact with the PCGs either by transcription of elncRNAs from distal regulatory loci or by formation of chromatin loops with nearby genes (44, 45). Our findings of the ability of NONMMUT013718 and NONMMUT024103 to drive the expression of downstream luciferase reporter genes in an orientation-independent manner and the effects of shRNA-mediated elncRNA knockdown on the expression of target genes during infection suggest the possibility

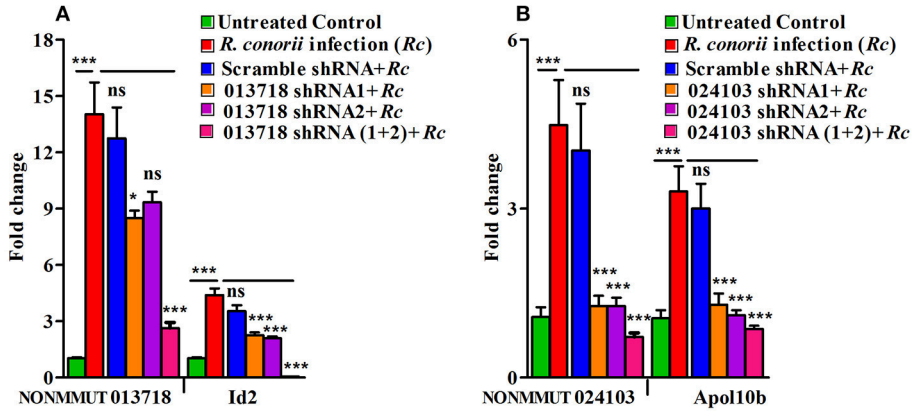

FIGURE 8 | shRNA mediated knock-down of elncRNAs and its target genes during rickettsial infection. Endogenous mRNA expression of Id2 and Apol10b in RAW264.7 macrophages and SVEC endothelial cells after trasfection of NONMMUT013718 shRNA and NONMMUT024103 shRNAs, respectively, following infection with *R. conorii* for 24 and 3 h, respectively. pLKO.1 scramble plasmid was used as a negative control. Knock-down efficiency of **(A)** NONMMUT013718 and its target gene Id2; and **(B)** NONMMUT024103 and its target gene Apol10b were quantified by qPCR. The error bars represented standard error of mean (SEM) and the levels of statistical significance are shown as $^{*}P \leq 0.05$; $^{***}P \leq 0.001$; and ns, non-significant.

of elncRNA interactions with the promoter of PCGs Id2 and Apol10b, respectively. Intriguingly, both of these elncRNAs and their putative targets are found to be highly up-regulated during *R. conorii* infection, which lends further support to the plausible involvement of their potential regulatory roles in the activation of respective proximal PCGs.

In vitro models of infection to delineate interactions between rickettsiae and their target host cells have long been established and routinely used in light of better tractability and direct applicability to decipher the fine details of cellular, molecular, and pathophysiological mechanisms of disease pathogenesis. Although pathogenic rickettisae as intracellular parasites display a predilection to primarily target microvascular endothelial cells lining the small and medium-sized vessels, invasion and infection of macrophages at the site of arthropod feeding during natural transmission to the mammalian hosts and in established needle inoculation-based animal models mimicking human disease is also evident (46). A recent study further documents notable differences in the ability of virulent vis-à-vis avirulent strains of rickettsiae to proliferate in macrophage-like cells *in vitro* as an important determinant of pathogenicity (47). In this context, an intriguing finding of the present study is the up-regulation of elncRNA NONMMUT013718 and Id2 during *R. conorii* infection of RAW264.7 macrophages, whereas only endothelial cells exhibit induced expression of NONMMUT024103 and Apol10b in response to infection. In light of previous evidence indicating that a significant fraction of lncRNAs show lineage-specific expression (48), we interpret these results as the host cell-specific and selective response in regards to the expression of a particular elncRNA.

The proteins belonging to the apolipoprotein (Apol) family are highly conserved across species and are generally thought to be involved in lipid transport and metabolism, due mainly to the association of Apol-1 as a subclass of high-density lipoproteins in human blood. Amongst Apols, the Apol1 in humans is unique in that it can be secreted due likely to its N-terminal signal peptide, accounts for the trypanosome lytic factor activity of human serum, and displays structural and functional similarities with Bcl-2 proteins involved in the regulation of apoptosis and autophagy. In addition, cultured human umbilical vein endothelial cells express CG12_1 (Apol-like) gene as a delayed early marker of inflammation in response to *in vitro* treatment with tumor necrosis factor-α and CG12_1 has been demonstrated to be specifically expressed in endothelial cells lining the normal and atherosclerotic iliac artery and aorta (49). The functions of other members of the Apol family classified on the basis of sequence homology to Apol1, however, are not well-understood. In a recent study, mouse Apol9a and Apol9b have been documented as bonafide interferon-stimulated genes (ISGs) with antiviral activity against Theiler's murine encephalomyelitis virus (50). Our laboratory has previously reported on the ISG response to *R. conorii* infection and its role in the interference with bacterial replication in human microvascular endothelial cells (51), suggesting the possibility of a potential link between higher expression of NONMMUT024103 elncRNA and Apol10b and the type 1 interferon response of host cells.

Inhibitor of DNA binding (Id) proteins, including Id1, Id2, Id3, and Id4, are basic helix-loop-helix (bHLH) transcription regulators. Although other bHLH proteins are known to regulate the transcription of a number of target genes by functioning as homo- or heterodimers, interactions between ubiquitously expressed E protein transcription factors and Id proteins by virtue of their destitute DNA-binding domain inhibit the formation of transcriptionally active complexes (52). Consequently, Id proteins are involved in the control of multiple cellular processes, including differentiation, proliferation, and fate determination (53, 54). Id2 also performs multiple essential functions in the hematopoietic system for the development of dendritic cells, NK cells, intraepithelial T cells, and lymphoid tissue inducer cells (55–57). Although the findings

of this study are the first to demonstrate increased expression of NONMMUT0013718 and Id2 in macrophages but not endothelial cells infected *in vitro*, it remains to be determined whether changes in the lungs during *in vivo Rickettsia* infection are due to increased transcription within target host cells, increased recruitment of inflammatory cells, or possibly a combinatorial effect of both. During *Listeria monocytogenes* infection, Id2 regulates gene expression by $CD8^+$ T cells and determines the magnitude of effector responses, suggesting a mechanism involving Id2 governed and E protein-mediated survival and differentiation of mature T cells (58). Published findings from the mouse model of rickettsiosis employed in this study yield evidence for increased expression of T cell targeting chemokines CXCL9 and CXCL10 in the lungs and infiltration of CD8+ T cells in the perivascular space around *Rickettsia*-infected microvessels (59). In addition, CD8+ T cells have been implicated in protective immunity against rickettsial infections, which is mediated in part by the cytotoxicity toward infected cells. Therefore, further studies to investigate potential regulatory roles of NONMMUT0013718 and Id2 in the determination of host immune responses to rickettsiae with particular attention to T cell mediated immunity are justified and currently ongoing.

CONCLUSION

In conclusion, the present study reports on differential expression of a number of lncRNA transcripts in the lungs as one of the prominent target organs in an established mouse model of rickettsial infection. From this subset of lncRNAs, we have identified two active elncRNAs through systematic application of genomics, epigenomics, and functional analysis to further demonstrate selective, cell-specific regulation of these lncRNAs and their potential target genes in vascular endothelial cells and macrophages as the target host cells *in vitro*. Given the data suggesting contributions of lncRNA-based regulatory networks in the modulation of host gene expression and differentiation as well as homing of T cells, further in-depth mechanistic enquiries of these versatile biological mediators in host-pathogen crosstalk and pathogenesis should unveil new strategies to counteract bacterial infections.

AUTHOR CONTRIBUTIONS

IC designed and performed all *in vitro* experiments. HN and AS performed animal experiments and provided the samples for RNA-seq. IC and KK performed all the bioinformatics analysis. HN and YF assisted in the analysis of RNA sequencing data. IC and SS supervised the study. IC and SS wrote and edited the manuscript and acquired funding to support this work. HN, AS, KK, and YF helped in editing and reviewing the manuscript. SS provided the resources for completion of this study. All authors read and approved the final manuscript.

FUNDING

This work was supported by James W. McLaughlin Postdoctoral Fellowship to IC, Ph.D., at the University of Texas Medical Branch (UTMB), Galveston, TX, USA, exploratory grants R21 AI115231 and R21 AI117483 from the NIAID/NIH, and institutional support funds from the UTMB.

ACKNOWLEDGMENTS

We sincerely thank Dr. Thomas Wood, Ph.D. and Dr. Steve Widen, PhD (UTMB Genomic core facility) for their advice and support with the conduct of sequencing and data analysis. We are grateful to Dr. David H. Walker, MD and Dr. Shinji Makino, Ph.D. for generously providing the cell lines and plasmids used in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2018.03014/full#supplementary-material

Supplementary Table 1 | Total reads from RNA sequencing experiments and the number of reads mapping to coding (mRNA) and non-coding (ncRNA) transcripts.

Supplementary Table 2 | Genomic annotation of up-regulated elncRNAs, plncRNAs and the nearby PCGs.

Supplementary Table 3 | Primer sequences for elncRNA and mRNA expression, and cloning.

Supplementary File 1 | The shRNA target sites are highlighted with colors.

Supplementary Figure 1 | Schematic representation of RNA-sequencing and downstream analysis.

Supplementary Figure 2 | Volcano plot displaying differentially expressed lncRNA transcripts (cut off value 3-fold, *P*-value $-\text{Log}_{10} \leq 0.05 \sim 1.3$). Red and green dots on the x-axis represent positively and negatively regulated transcripts, respectively. Values on the y-axis represent the level of significance.

Supplementary Figure 3 | (A) Strategies applied for cataloging of differentially expressed lncRNA transcripts based on their origin; **(B)** Cataloging of lncRNA transcripts based on their exon numbers; and **(C)** Cataloging of lncRNA transcripts based on their lengths.

Supplementary Figure 4 | (A) Strategies applied to identify orientation of up-regulated lncRNA transcripts with their nearby protein coding genes (PCGs); **(B)** Histogram depicting density of average normalized RPKM (reads per kilobase million) values of H3K4Me1 in elncRNAs and plncRNAs; and **(C)** Histogram depicting density of average normalized RPKM (reads per kilobase million) values of H3K4Me3 in elncRNA and plncRNAs around 4 kb window of transcription start site.

Supplementary Figure 5 | Histogram depicting density of average normalized RPKM (reads per kilobase million) values of **(A)** RNA PolII; **(B)** p300; **(C)** DNaseI hypersensitivity site; and **(D)** CCCTC binding factor (CTCF) around 4 kb window of transcription start site of elncRNAs and plncRNAs, respectively.

Supplementary Figure 6 | Virtual 4C plot of **(A)** NONMMUT013718 and **(B)** NONMMUT024103 elncRNA. The anchoring point (red line) marks the position of **(A)** Id2 and **(B)** Apol10b, respectively. The contact loci are shown with blue line.

Supplementary Figure 7 | Expression of elncRNAs and their targets in an *in vitro* model of *R. conorii* infection. Expression of NONMMUT013718 and its target Id2 in SVECs at 3 and 24 h post-infection.

REFERENCES

1. Sahni SK, Rydkina E. Host-cell interactions with pathogenic *Rickettsia* species. *Future Microbiol.* (2009) 4:323–39. doi: 10.2217/fmb.09.6
2. Sahni SK, Narra HP, Sahni A, Walker DH. Recent molecular insights into rickettsial pathogenesis and immunity. *Future Microbiol.* (2013) 8:1265–88. doi: 10.2217/fmb.13.102
3. Parra-Martinez J, Sancho-Rieger J, Ortiz-Sanchez P, Peset V, Brocalero A, Castillo A, et al. [Encephalitis caused by *Rickettsia conorii* without exanthema]. *Rev Neurol.* (2002) 35:731–4.
4. Masters EJ, Olson GS, Weiner SJ, Paddock CD. Rocky mountain spotted fever: a clinician's dilemma. *Arch Intern Med.* (2003) 163:769–74. doi: 10.1001/archinte.163.7.769
5. Shah SS, McGowan JP. Rickettsial, ehrlichial and *Bartonella* infections of the myocardium and pericardium. *Front Biosci.* (2003) 8:e197–201. doi: 10.2741/995
6. Premaratna R, Chandrasena TG, Dassayake AS, Loftis AD, Dasch GA, de Silva HJ. Acute hearing loss due to scrub typhus: a forgotten complication of a reemerging disease. *Clin Infect Dis.* (2006) 42:e6–8. doi: 10.1086/498747
7. Radulovic S, Price PW, Beier MS, Gaywee J, Macaluso JA, Azad A. Rickettsia-macrophage interactions: host cell responses to *Rickettsia akari* and *Rickettsia typhi*. *Infect Immun.* (2002) 70:2576–82. doi: 10.1128/IAI.70.5.2576-2582.2002
8. Woods ME, Olano JP. Host defenses to *Rickettsia rickettsii* infection contribute to increased microvascular permeability in human cerebral endothelial cells. *J Clin Immunol.* (2008) 28:174–85. doi: 10.1007/s10875-007-9140-9
9. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* (2009) 10:155–9. doi: 10.1038/nrg2521
10. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* (2009) 23:1494–504. doi: 10.1101/gad.1800909
11. Rashid F, Shah A, Shan G. Long non-coding RNAs in the cytoplasm. *Genomics Proteomics Bioinform.* (2016) 14:73–80. doi: 10.1016/j.gpb.2016.03.005
12. Chowdhury IH, Narra HP, Sahni A, Khanipov K, Schroeder CLC, Patel J, et al. Expression profiling of long noncoding RNA splice variants in human microvascular endothelial cells: lipopolysaccharide effects *in vitro*. *Mediators Inflamm.* (2017) 2017:3427461. doi: 10.1155/2017/3427461
13. Cui H, Xie N, Tan Z, Banerjee S, Thannickal VJ, Abraham E, et al. The human long noncoding RNA lnc-IL7R regulates the inflammatory response. *Eur J Immunol.* (2014) 44:2085–95. doi: 10.1002/eji.201344126
14. Krawczyk M, Emerson BM. p50-associated COX-2 extragenic RNA (PACER) activates COX-2 gene expression by occluding repressive NF-kappaB complexes. *Elife* (2014) 3:e01776. doi: 10.7554/eLife.01776
15. Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? *Front Genet.* (2015) 6:2. doi: 10.3389/fgene.2015.00002
16. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell* (2010) 143:46–58. doi: 10.1016/j.cell.2010.09.001
17. Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, et al. eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol Cell* (2013) 49:524–35. doi: 10.1016/j.molcel.2012.11.021
18. Ilott NE, Heward JA, Roux B, Tsitsiou E, Fenwick PS, Lenzi L, et al. Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. *Nat Commun.* (2014) 5:3979. doi: 10.1038/ncomms4979
19. Ounzain S, Pezzuto I, Micheletti R, Burdet F, Sheta R, Nemir M, et al. Functional importance of cardiac enhancer-associated noncoding RNAs in heart development and disease. *J Mol Cell Cardiol.* (2014) 76:55–70. doi: 10.1016/j.yjmcc.2014.08.009
20. Kim YW, Lee S, Yun J, Kim A. Chromatin looping and eRNA transcription precede the transcriptional activation of gene in the beta-globin locus. *Biosci Rep.* (2015) 35:e00179. doi: 10.1042/BSR20140126
21. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* (2013) 494:497–501. doi: 10.1038/nature11884
22. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* (2009) 457:854–8. doi: 10.1038/nature07730
23. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol.* (2015) 16:144–54. doi: 10.1038/nrm3949
24. Rydkina E, Turpin LC, Sahni SK. *Rickettsia rickettsii* infection of human macrovascular and microvascular endothelial cells reveals activation of both common and cell type-specific host response mechanisms. *Infect Immun.* (2010) 78:2599–606. doi: 10.1128/IAI.01335-09
25. Schroeder CL, Narra HP, Rojas M, Sahni A, Patel J, Khanipov K, et al. Bacterial small RNAs in the genus *Rickettsia*. *BMC Genomics* (2015) 16:1075. doi: 10.1186/s12864-015-2293-7
26. Rydkina E, Turpin LC, Sahni A, Sahni SK. Regulation of inducible heme oxygenase and cyclooxygenase isozymes in a mouse model of spotted fever group rickettsiosis. *Microb Pathog.* (2012) 53:28–36. doi: 10.1016/j.micpath.2012.03.010
27. Narra HP, Schroeder CL, Sahni A, Rojas M, Khanipov K, Fofanov Y, et al. Small regulatory RNAs of *Rickettsia conorii*. *Sci Rep.* (2016) 6:36728. doi: 10.1038/srep36728
28. Schroeder CL, Narra HP, Sahni A, Rojas M, Khanipov K, Patel J, et al. Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. *Front Microbiol.* (2016) 7:859. doi: 10.3389/fmicb.2016.00859
29. Marques AC, Hughes J, Graham B, Kowalczyk MS, Higgs DR, Ponting CP. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. *Genome Biol.* (2013) 14:R131. doi: 10.1186/gb-2013-14-11-r131
30. Mao AP, Shen J, Zuo Z. Expression and regulation of long noncoding RNAs in TLR4 signaling in mouse macrophages. *BMC Genomics* (2015) 16:45. doi: 10.1186/s12864-015-1270-5
31. Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* (2012) 488:116–20. doi: 10.1038/nature11243
32. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* (2012) 485:376–80. doi: 10.1038/nature11082
33. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D genome browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol.* (2018) 19:151. doi: 10.1186/s13059-018-1519-9
34. Walker DH, Olano JP, Feng HM. Critical role of cytotoxic T lymphocytes in immune clearance of rickettsial infection. *Infect Immun.* (2001) 69:1841–6. doi: 10.1128/IAI.69.3.1841-1846.2001
35. Koo SJ, Chowdhury IH, Szczesny B, Wan X, Garg NJ. Macrophages promote oxidative metabolism to drive nitric oxide generation in response to *Trypanosoma cruzi*. *Infect Immun.* (2016) 84:3527–41. doi: 10.1128/IAI.00809-16
36. Xu YY, Jin R, Zhou GP, Xu HG. Involvement of GATA1 and Sp3 in the activation of the murine STING gene promoter in NIH3T3 cells. *Sci Rep.* (2017) 7:2090. doi: 10.1038/s41598-017-02242-w
37. Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, et al. Lentivirus-delivered stable gene silencing by *RNA*i in primary cells. *RNA* (2003) 9:493–501. doi: 10.1261/rna.2192803
38. Xie C, Yuan J, Li H, Li M, Zhao G, Bu D, et al. NONCODEv4: exploring the world of long non-coding RNA genes. *Nucleic Acids Res.* (2014) 42(Database issue):D98–103. doi: 10.1093/nar/gkt1222
39. Pauley KM, Cha S, Chan EK. MicroRNA in autoimmunity and autoimmune diseases. *J Autoimmun.* (2009) 32:189–94. doi: 10.1016/j.jaut.2009.02.012
40. Sahni A, Narra HP, Patel J, Sahni SK. MicroRNA signature of human microvascular endothelium infected with *Rickettsia rickettsii*. *Int J Mol Sci.* (2017) 18:1471. doi: 10.3390/ijms18071471
41. Sahni A, Patel J, Narra HP, Schroeder CLC, Walker DH, Sahni SK. Fibroblast growth factor receptor-1 mediates internalization of pathogenic spotted fever *Rickettsiae* into host endothelium. *PLoS ONE* (2017) 12:e0183181. doi: 10.1371/journal.pone.0183181
42. Crawford GE, Holt IE, Whittle J, Webb BD, Tai D, Davis S, et al. Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Res.* (2006) 16:123–31. doi: 10.1101/gr.4074106
43. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature* (2010) 465:182–7. doi: 10.1038/nature09033

44. West AG, Fraser P. Remote control of gene transcription. *Hum Mol Genet.* (2005) 14 Spec No 1:R101–11. doi: 10.1093/hmg/ddi104
45. Sexton T, Bantignies F, Cavalli G. Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. *Semin Cell Dev Biol.* (2009) 20:849–55. doi: 10.1016/j.semcdb.2009.06.004
46. Smalley C, Bechelli J, Rockx-Brouwer D, Saito T, Azar SR, Ismail N, et al. *Rickettsia australis* activates inflammasome in human and murine macrophages. *PLoS ONE* (2016) 11:e0157231. doi: 10.1371/journal.pone.0157231
47. Curto P, Simoes I, Riley SP, Martinez JJ. Differences in intracellular fate of two spotted fever group *Rickettsia* in macrophage-like cells. *Front Cell Infect Microbiol.* (2016) 6:80. doi: 10.3389/fcimb.2016.00080
48. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* (2015) 47:199–208. doi: 10.1038/ng.3192
49. Horrevoets AJ, Fontijn RD, van Zonneveld AJ, de Vries CJ, ten Cate JW, Pannekoek H. Vascular endothelial genes that are responsive to tumor necrosis factor-alpha *in vitro* are expressed in atherosclerotic lesions, including inhibitor of apoptosis protein-1, stannin, and two novel genes. *Blood* (1999) 93:3418–31.
50. Kreit M, Vertommen D, Gillet L, Michiels T. The interferon-inducible mouse apolipoprotein L9 and prohibitins cooperate to restrict theiler's virus replication. *PLoS ONE* (2015) 10:e0133190. doi: 10.1371/journal.pone.0133190
51. Colonne PM, Sahni A, Sahni SK. *Rickettsia conorii* infection stimulates the expression of ISG15 and ISG15 protease UBP43 in human microvascular endothelial cells. *Biochem Biophys Res Commun.* (2011) 416:153–8. doi: 10.1016/j.bbrc.2011.11.015
52. Sikder HA, Devlin MK, Dunlap S, Ryu B, Alani RM. Id proteins in cell growth and tumorigenesis. *Cancer Cell* (2003) 3:525–30. doi: 10.1016/S1535-6108(03)00141-7
53. Yuen HF, Chan YP, Chan KK, Chu YY, Wong ML, Law SY, et al. Id-1 and Id-2 are markers for metastasis and prognosis in oesophageal squamous cell carcinoma. *Br J Cancer* (2007) 97:1409–15. doi: 10.1038/sj.bjc.6604035
54. Roschger C, Cabrele C. The Id-protein family in developmental and cancer-associated pathways. *Cell Commun Signal.* (2017) 15:7. doi: 10.1186/s12964-016-0161-y
55. Hashimoto D, Miller J, Merad M. Dendritic cell and macrophage heterogeneity *in vivo*. *Immunity* (2011) 35:323–35. doi: 10.1016/j.immuni.2011.09.007
56. Best JA, Blair DA, Knell J, Yang E, Mayya V, Doedens A, et al. Transcriptional insights into the CD8(+) T cell response to infection and memory T cell formation. *Nat Immunol.* (2013) 14:404–12. doi: 10.1038/ni.2536
57. Verykokakis M, Krishnamoorthy V, Iavarone A, Lasorella A, Sigvardsson M, Kee BL. Essential functions for ID proteins at multiple checkpoints in invariant NKT cell development. *J Immunol.* (2013) 191:5973–83. doi: 10.4049/jimmunol.1301521
58. Foulds KE, Rotte MJ, Seder RA. IL-10 is required for optimal CD8 T cell memory following Listeria monocytogenes infection. *J Immunol.* (2006) 177:2565–74. doi: 10.4049/jimmunol.177.4.2565
59. Valbuena G, Bradford W, Walker DH. Expression analysis of the T-cell-targeting chemokines CXCL9 and CXCL10 in mice and humans with endothelial infections caused by rickettsiae of the spotted fever group. *Am J Pathol.* (2003) 163:1357–69. doi: 10.1016/S0002-9440(10)63494-3

Conflict of Interest Statement: 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 Immunology

REVIEW
published: 25 January 2019
doi: 10.3389/fimmu.2018.03138

The Opening of Pandora's Box: An Emerging Role of Long Noncoding RNA in Viral Infections

*Pin Wang**

National Key Laboratory of Medical Immunology and Institute of Immunology, Second Military Medical University, Shanghai, China

Emerging evidence has proved that long noncoding RNAs (lncRNAs) participate in various physiological and pathological processes. Recent evidence has demonstrated that lncRNAs are crucial regulators of virus infections and antiviral immune responses. Upon viral infections, significant changes take place at the transcriptional level and the majority of the expression modifications occur in lncRNAs from both the host and viral genomes with dynamic regulatory courses. These lncRNAs exert diverse effects. Some are antiviral either through directly inhibiting viral infections or through stimulating antiviral immune responses, while others are pro-viral through directly promoting virus replication or through influencing cellular status, such as suppressing antiviral mechanisms. Consequently, these dynamic regulations lead to disparate pathophysiological outcomes and clinical manifestations. This review will focus on the roles of lncRNAs in viral infection and antiviral responses, summarize expression patterns of both host- and virally derived lncRNAs, describe their acting stages and modes of action, discuss challenges and novel concepts, and propose solutions and perspectives. Research into lncRNA will help identify novel viral infection-related regulators and design preventative and therapeutic strategies against virus-related diseases and immune disorders.

Keywords: lncRNA, long noncoding RNA, immune response, viral infection, antiviral immunity, RNA-protein interaction

OPEN ACCESS

Edited by:
Jorge Henao-Mejia, University of Pennsylvania, United States

Reviewed by:
Xiaocui He, La Jolla Institute for Allergy and Immunology (LJI), United States
Xiaoyu Hu, Tsinghua University, China

***Correspondence:**
Pin Wang
wangp@immunol.org

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *15 October 2018*
Accepted: *19 December 2018*
Published: *25 January 2019*

Citation:
Wang P (2019) The Opening of Pandora's Box: An Emerging Role of Long Noncoding RNA in Viral Infections. Front. Immunol. 9:3138. doi: 10.3389/fimmu.2018.03138

INTRODUCTION

In the RNA world hypothesis, RNA was proposed to be the original form of life, at least the vital compartment of original life, as its spatial structure possesses two major characteristics that biological functional macromolecules required—diversity and flexibility. However, during subsequent process of life development, RNA transferred its role of information storage to DNA that is more stable, and its catalytic activity to protein which has more sophisticated spatial structure, while RNA itself gradually becomes the intermediate between DNA and protein in life organization. RNA only reserved its activity diversity in some fundamental complexes, such as spliceosome, telomerase, ribosome, signal recognition particle (SRP), some metabolic riboswitches, and ribozymes. This is described as the center dogma of modern molecular genetics and was deeply believed by the academic community until the revealing of large portion of noncoding RNA transcript in the latest annotation of genome sequences and interpretation of transcriptome data. We now know that the role of RNA is much further beyond just the message of genome information, it still preserves its ancient diversity and mystery, leaving us to discover.

Facilitated by fast-developing sequencing techniques and bioinformatics, genomes, and transcription profiling in the early Twenty-first century were conducted in human being and other metazoan species, leading to the unexpected observation that while majority of genome is transcribed only small portion is protein coding sequences (approximately 2% in mammalian genome) (1, 2). More detailed annotation and advanced bioinformatics analysis further helped us to reveal various epigenetic elements and genomic origination of these noncoding genes in different cell types and tissues, through projects such as Encyclopedia of DNA Elements (ENCODE) (3) and The Cancer Genome Atlas (TCGA) Research Network (4). The complexity has led us to shift our understanding of genomic information from linear model to modular model, which combines transcription and function of noncoding RNA with DNA regulatory elements, epigenetic modification and spatial origination.

Amid this progress, as a transcriptional class, lncRNAs were first described in the year of 2002 by Okazaki et al. in the study of large-scale sequencing of full-length cDNA libraries in mouse (5). Actually, lncRNA is an arbitrary category definition mainly referring to RNA transcripts with no obvious peptide coding capacity, usually longer than 200 nucleotides to distinguish from short noncoding RNA, such as microRNA, short interfering RNAs (siRNAs), and Piwi-interacting RNAs (piRNAs). Both long and short noncoding RNA compose regulatory RNAs with diverse unknown functions, in contrast to the housekeeping RNAs with certain functions including ribosomal RNA (rRNA), transfer RNA (tRNA), and small nucleolar RNAs (snoRNAs). However, some recent evidences suggest that some tRNA derived RNA fragments also have regulatory function in diverse aspects, such as intergenerational inheritance (6) and viral infection (7).

The number of lncRNA transcripts being identified keeps increasing these years. Up to now, according to a comprehensive bioinformatics analysis using data from 25 independent studies, 58,648 genes were identified as lncRNAs for human being (8). Interestingly, the number of lncRNAs correlates with the developmental complexity of species, at least in all the annotated eukaryotes, with highest lncRNAs amount in primates followed by mouse and scaling down accordingly to yeast (9). While the transcriptional sequence of lncRNAs are less conserved than that of protein coding genes, their promoter sequences and genomic locations are as conserved as coding genes (10, 11), indicating their expression are under tightly regulatory control and evolutional constrains. This is supported by the observation that lncRNA expression profiles are more tissue specific and cell-type selective than that of coding RNA (8), suggesting lncRNAs prefer to perform subtly function in cell-type specific manners, although majority of them are still less characterized.

Another protagonist of this review is virus. As an anciently derived organization as RNA, virus has some unique connection with RNA molecules. Virus is the only organism on this planet that reserves RNA as the genome and its RNA could be replicated through RNA dependent RNA polymerase and translated into DNA through reverse transcriptase in the life cycle of viral infection, which makes RNA spectrum more diverse in infected cells, including host-derived, virally derived and even some chimeric RNAs. However, host cells have developed mechanisms to distinguish virally derived RNA from its own RNA through pattern recognition receptors (PRRs), such as Toll-like receptors (TLR) and RIG-I-like receptors (RLR). Viral transcribed RNA usually has a 5′ triphosphate uncapped terminal, which could be recognized by host canonical sensor RIG-I (12, 13). Double-stranded RNA (dsRNA) is commonly produced during viral infection as genetic materials or replicating intermediates during virus replication, which will trigger dsRNA sensors in host cells, including canonical sensor MDA5 in the cytoplasm (14) and membrane receptor TLR3 (15) (**Figure 1**). Interaction between virus and its host has never halted since its very beginning. Being the simplest but efficient obligated parasites, virus have evolved a variety of strategies to manipulate hosts to provide material and energy to complete their life cycle, including duplicating their genome, producing their RNAs and proteins, packaging infective particles and finally releasing to infect other host cells. On the other hand, the host has developed an immune system whose activation launches immune responses to eliminate viral infections. Many viral components or their intermediate, such as their nucleic acids, can activate host immune system when being recognized by host cell PRRs and subsequently trigger downstream cascade signaling transduction (**Figure 1**). With the help of adaptor proteins, serine/threonine-protein kinase TBK1 phosphorylates IRF3/7 and TRAF6 activates Nuclear factor-κB (NF-κB) signaling, leading to the transcription of type I interferon (IFN) and inflammatory cytokines, along with many noncoding genes (16, 17) (**Figure 1**). IFN is the most efficient antiviral cytokine. Through JAK/STAT signaling, IFN triggers lots of effecter genes' expression including many noncoding genes, and induces the antiviral status of host cells to defeat invasion. However, uncontrolled antiviral responses and inflammatory status are also detrimental to host cells. So, immune responses must be finely regulated to minimize cytotoxic effect and autoimmunity output, which requires negative feedback regulatory mechanism to control the duration and magnitude of antiviral responses. A large amount of proteins and increasing lncRNAs have been proved to be involved in this regulatory loop. However, during a long time evolution, these regulatory factors are also utilized by some virus to escape host defense in many cases. The competing between host cells and virus has evolved to be a mutual-driving interaction involving more and more regulatory proteins and lncRNAs from both sides. As a research hotspot, interests and studies progress rapidly in recent years. While many reviews have been published on this theme in recent years (9, 18–20), more and more host- or viral- encoded lncRNAs have been characterized and novel action models of lncRNA functions have been revealed. For example, lncRNA directly regulates metabolic activity of the host cells and lncRNA interacts with singling adaptors or sensors to exert functions. This review will focus on these recent advances and cutting-edge technologies in this area to present a comprehensive view of mammalian host- and viral- derived lncRNAs.

VIRALLY DERIVED RNAs

The existence of viral noncoding RNAs has already been known for decades (21, 22). In the year of 1971, it was reported in the plant that viroids, as the smallest infectious pathogens known,

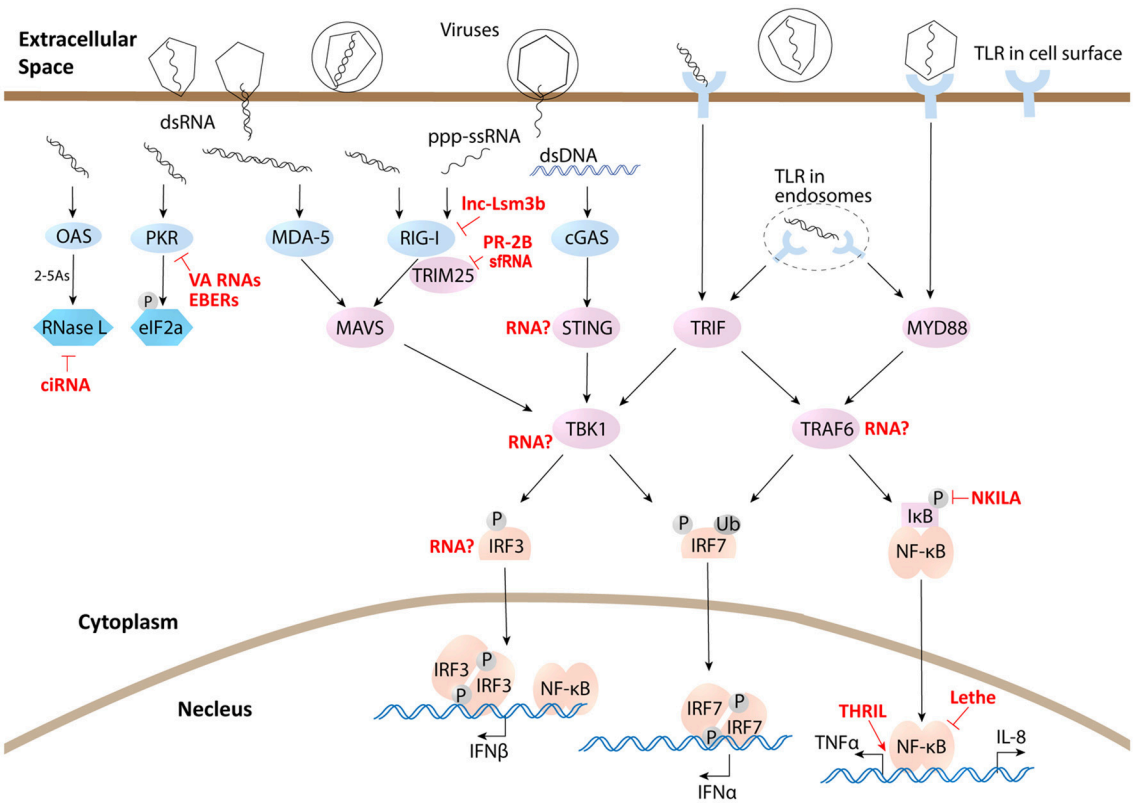

FIGURE 1 | lncRNAs influence immune defense responses through directly interacting with sensors, adaptors, effecters, and transcriptional factors. Light blue represents sensors, cyan represents effecters, light purple represents signaling adaptors, and light brown represents transcriptional factors.

are composed solely of a short strand of circular, single-stranded RNA which are capable of autonomous replication (23, 24). So viroid has been considered to be the living relics of the hypothetical RNA world. The expression of noncoding RNAs in host cells from animal viruses has also been described years ago. Some of them are very abundant after infection, such as PAN (polyadenylated nuclear RNA) from Kaposi's sarcoma-associated herpesvirus (KSHV) (25), EBER1/2 (Epstein-Barr virus encoded small RNA 1 and 2) from Epstein-Barr virus (26, 27) and VA RNA (virus-associated RNA) from adenovirus (28). Aside from duplication by RNA-dependent RNA polymerase, some viral RNAs are transcribed by host polymerase III, such as VA RNAs and EBER1/2, while some are transcribed by polymerase II and polyadenylated, such as PAN RNA (**Table 1**). Some viral lncRNAs in host cells are not generated from canonical working flow, instead they are processed by unique maturation steps and even degradation of host cellular machineries. For example, flavivirus genome RNA is degraded by 5′-3′ exonuclease Xrn1 in host cells as a defense mechanism. However, flavivirus has developed a special secondary or tertiary structure of its RNA to halt Xrn1 processing to the 3′ end. So Xrn1 stalls on this structures and creates a large amount of degradation intermediates in host cells, named subgenomic flavivirus RNA (sfRNA) or Xrn1-resistant RNAs (xrRNAs) (43) (**Table 1**), as a specific feature of flavivirus infection. It has to be underlined that viral RNA amount in host cell is highly correlated with viral infection activity. Some RNAs are only expressed during latency, for example latency-associated transcript (LAT) from herpes simplex virus type 1 (HSV-1), while some are highly expressed during the lytic phase, such as PAN RNA. This trait of expression makes viral RNAs to be potential targets in clinical detection of relative virus infection.

MAMMALIAN lncRNAs

As RNA polymerase I only transcribes ribosomal RNAs in eukaryotes, most lncRNAs discussed here are transcribed by polymerase II (**Table 2**) and undergo similar splicing and modification processing as message RNA (mRNA), such as methylguanosine at the 5′-terminus and a polyadenylated tail at the 3′-terminus. These lncRNAs are often referred to as canonical lncRNA. Genetically, compared with mRNA, lncRNAs harbor fewer of exons pre transcripts and alternately spliced isoforms per gene locus, and the lengths of lncRNA transcripts are more concentrated within the range of 100–1000 nucleotides (66). RNA polymerase III also transcribes some regulatory ncRNAs, such as RNA Alu, 7SK, BC200, B1 and B2 RNAs (67). Compared with canonical lncRNA, these regulatory RNAs are shorter in length, usually no more than 500 nucleotides, and function mainly through interacting with transcription factors and RNA

TABLE 1 | The expressions, functions and mechanisms of viral lncRNAs in viral infection.

Sources	lncRNA	Expression/transcription	Functions/mechanisms	RNA location	References
HCMV	RNA2.7	Highly expressed at early times of infection	Interacting with complex I to prevent GRIM-19 translocalization to stabilizes the mitochondrial membrane potential, resulting in continued ATP production for virus	Cytoplasm	(29, 30)
EBV	oriPts	Expressed during reactivation from EBV latency origin of replication	Modulating paraspeckle-based innate antiviral immune pathway, global viral lytic gene expression, and viral DNA replication during reactivation.	Nucleus	(31)
Group C enterovirus	RNase L ciRNA	Expressed during infection	A competitive inhibitor of the antiviral endoribonuclease RNase L	Cytoplasm	(32)
Flavivirus	sfRNA	Flavivirus genomic RNA degradation intermediates in Xrn1 processing	Oversaturation of Xrn1 degradation and the RNAi machinery	Cytoplasm	(33, 34)
Dengue virus (DENV-2 PR-2B)	PR-2B sfRNA	One of sfRNA	Binding E3 ubiquitin ligase TRIM25 to prevent its deubiquitinylation to unstabilize RIG-I to decrease IFN production and antiviral responses	Cytoplasm	(35)
Dengue virus (DENV-2)	DENV-2 sfRNA	One of sfRNA	Binding to host RNA-binding proteins to antagonize their function in ISG translation, as a molecular sponge of anti-viral effectors.	Cytoplasm	(36)
Adenovirus	VA RNA	Transcribed by polymerase III	Sequestration of several key members of the RNAi pathway and cytoplasmic sensor PKR.	Cytoplasm	(28, 37)
EBV	EBERs	Transcribed by polymerase III	Binding PKR to prevent its dimerization and auto-phosphorylation and signaling to eIF2a, promoting translation of viral proteins	Cytoplasm	(26, 27)
HIV	ASP RNA	Antisense transcript	Recruiting PRC2 to the HIV-1 5' LTR leading to suppressive H3K27 trimethylation and establishment of HIV-1 latency	Nucleus	(38)
KSHV	PAN RNA	Highly expressed during the lytic phase by polymerase II	Guiding PRC2 to the KSHV genome to mediate activation of viral gene expression to produce infectious virus; interacting with H1/H2A, SSBPs, and IRF4 to decrease the expression of IFNγ, IFNα, IL18, and RNase L	Nucleus	(25, 39–42)

polymerase II to regulate transcription (68) or influencing mRNA translation (69).

Genomic location of lncRNA usually closely associates with its molecular function or mode of activity. Based on the relationship with the nearest coding gene in genome, host lncRNAs are classified into four categories. LncRNA genes with a distance further than 5-kb to the nearest coding genes are defined as intergenic lncRNAs, which are also called long intergenic noncoding RNAs (lincRNAs). However, 5-kb distance is an arbitrary threshold by experience, sometime it varies case by case. Intergenic lncRNAs are functionally referring to lncRNAs without overlapping or sharing transcriptional machinery with other genes, which tend to be independent genes at both expression and function levels. So they are easier to be genetically manipulated compared with other lncRNAs (70). Intergenic lncRNAs prefer to function through exerting *in trans* activity far from their transcription site, which is the case for Firre affecting topological organization of multichromosomal regions through interacting with the nuclear-matrix factor heterogeneous nuclear ribonucleoprotein U (hnRNPU) (71). Some intergenic lncRNAs also influence the expression of the nearby genes via promoter competition for a shared set of enhancers (72) or via histone modification regulation (73).

To categorize functionally, lncRNAs with a coding gene in less than 5-kb distance are classified into three biotypes, antisense head to head lncRNA, antisense overlapping lncRNA and sense overlapping lncRNA (**Figure 2**). The antisense head to head lncRNAs or divergent lncRNAs, which means they are transcribed in the antisense direction and positioned head to head to protein-coding genes, account for a significant proportion of host lncRNA (about 20% in mammals and about 70% in lower metazoans) and are predicted to be strongly

TABLE 2 | The expressions, functions, and mechanisms of host lncRNAs in viral infection.

lncRNA	Expression/transcription	Gene locus	Functions/mechanisms	Location	References
NRON	Highly expressed in CD4+ T lymphocytes	Antisense overlapping with coding gene MVB12B	Retaining transcription factor NFAT in the cytoplasm to suppress NFAT-mediated viral gene activation	Cytoplasm	(44)
lncRNA-ACOD1	Induced in many viral infections	An intergenic lncRNA proximal to ACOD1 gene	Binding GOT2 to promote metabolite production to promote viral replication	Cytoplasm	(45)
7SL RNA	Pol III-transcribed SRP RNA	Multicopy gene	Encapsidated into diverse retroviruses and functions as a key cofactor of the antiviral protein A3G	Cytoplasm	(46, 47)
lnc-Lsm3b	Induced by IFN	Sense overlapping with coding gene LSM3	Binding RIG-I to obstruct its conformational shift, prevented downstream signaling, and terminated IFN production	Cytoplasm	(48)
Lethe	Induced by proinflammatory cytokines via NF-κB or glucocorticoid receptor agonist treatment	Intergenic lncRNA	Binding NF-κB activatory subunit RelA to inhibit RelA DNA binding and target gene activation, as a negative feedback of NF-κB.	Cytoplasm	(49)
PACER	Chromatin factor CTCF establishes an open chromatin domain and induces its expression	Antisense head to head lncRNA with gene COX-2	Binding repressive subunit p50 to occlude it from COX-2 promoter, potentially facilitating interaction with active NF-κB dimers p65/p50 to promote COX2 transcription *in cis*	Nucleus	(50)
lincRNA-Cox2	Induced by TLR ligands through MyD88 and NF-κB.	An intergenic lncRNA proximal to Cox2 gene	Binding with hnRNP A/B and A2/B1 to regulates expression of a group of immune response genes	Nucleus	(51)
THRIL/linc1992	Downregulated by TNFα or TLR activation in viral infection	An intergenic lncRNA	Binding hnRNPL to promote transcription of the TNFα gene by binding to its promoter	Nucleus	(52)
NeST/ IfngAS1	Expressed in T cells by NF-κB, STAT4 and T-bet activation	Antisense overlapping with gene IFNG	Promoting IFNγ expression through binding WDR5 and altering histone 3 methylation at the IFNγ locus	Nucleus	(53–56)
LUARIS/ lncRNA#32	Downregulated by IFN	Antisense overlapping with gene HECW1	Binding hnRNPU to activate ATF2 to promote the expression of multiple ISGs	Nucleus	(57)
EGOT	Induced by IFN, HCV, influenza, and SFV	Antisense overlapping with coding gene ITPR1	Inhibiting multiple ISGs' expression as a negative feedback regulatory mechanism of IFN pathway	–	(58)
NRAV	Downregulated during IAV infection	Antisense overlapping with coding gene DYNLL1	Altering epigenetic histone modifications on the promoters of *IFITM3* and *MxA* to attenuate their initial transcription rates	Nucleus	(59)
NRIR/ lncRNA-CMPK2	Upregulated significantly by IFN	An intergenic lncRNA proximal to CMPK2 gene	Repressing expression of many antiviral ISGs probably through interacting with transcription factors or chromatin.	Nucleus	(60)
BISPR	Induced by IFN	Antisense head to head lncRNA with gene BST2	Promoting BST2 expression *in cis* through obstructing PRC2 at the promoter of BST2 to facilitate its transcription and interacting with EZH2 to overlap with enhancer region	Nucleus	(61)
NEAT1	Increased by HSV-1 and HIV	Intergenic lncRNA	Increasing viral gene expression and viral replication for HSV1; negatively regulating viral production for HIV; promoting RIG-I, DDX60 and IL8 expression by removing inhibitory effecter SFPQ to paraspeckles.	Nucleus	(62–65)

related to functions in transcription and development (74). Generally, the expression of antisense lncRNA are likely to be co-expressed with its overlapping coding gene as reported (10), probably because this lncRNA-coding gene pair shares the same chromatin transcriptional loop with a synergistic effect. However, the expressions of sense overlapping lncRNA do not exhibit obvious correlation with coding gene and even in many cases sense lncRNA depresses the expression of proximate coding gene, such as lncRNA Flicr dampening Foxp3 expression in Treg cells (75), probably through competing for shared transcriptional

elements. So it might be summarized that lncRNAs tend to promote its nearby antisense strand transcription while depress the sense nearby gene expression.

There are always some unique lncRNA species that resist common categorization, for example intron-derived lncRNA with snoRNA ends (sno-lncRNAs) (76) and exon-derived circular RNAs (circRNAs) (77), which are also transcribed by RNA polymerase II but do not have 5′ and 3′ terminus of mRNA due to RNA splicing alternation. They are believed to function through altering splicing, affecting parental gene expression, or sponging microRNA binding.

EXPRESSION OF lncRNA DURING VIRAL INFECTIONS

Virally derived RNAs are relative simple in their expression regulations, usually constitutively expressed upon invasion or induced expressed after the latency stage, probably due to our little knowledge on it. However, host-derived lncRNAs have been proved to exhibit more complicated spatial and temporal expression patterns, which this part of review focuses on.

A member of high throughout transcriptomic analyses in both human and mouse have revealed that viral infection induced great changes in host cell transcriptome, which includes large amount of protein coding genes along with lncRNAs (78). Promoter prediction and expression correlation analysis revealed that a great proportion of induced lncRNAs in viral infection are direct targets of IFN signaling or IFN-stimulated genes (ISGs). This is supported by experimental data from expression profiling of mouse macrophages with vesicular stomatitis virus (VSV) infection or recombinant IFNβ stimulation (45) and primary human hepatocytes with direct IFNα treatment (60). Interestingly, host lncRNAs have different expression pattern dynamics—some are induced in minutes while some are expressed in days. Another high throughput study in HuH7 hepatocytes that focused on the late time-point of IFN stimulation (72 h post high dose IFNα2 treatment) revealed another group of IFN-regulated lncRNAs, half being upregulated and half downregulated, while nearly all protein encoding genes changed were upregulated (79, 80). This elegant study clearly indicates that host lncRNAs have more complicated roles in gene expression regulation than what we previously expected.

Aside from IFN-regulated lncRNAs, there are some host lncRNAs that are induced by virus as viral hijacked lncRNA, whose expression does not rely on IFN signaling, as revealed by experiment data from wild type and IFN receptor deficiency macrophages (45). These IFN-independent lncRNAs tend to be manipulated by virus and involved in viral invasion. This is in the case of two intergenic lncRNAs, lncRNA-ACOD1 and VIN (virus inducible lincRNA). LncRNA-ACOD1 expression can be induced by many virus, including VSV, Sendai virus (SeV), HSV-1, and vaccinia virus (VACV), partly dependent on NF-κB signaling, and its induction is attenuated by IRF3 signaling as IRF3 knockout led to a higher expression, indicating lncRNA-ACOD1 is a favorable lncRNA for virus rather than host (45). VIN was identified as viral induced lncRNAs in human lung epithelial cells by several influenza A virus (IAV) strains (H1N1, H3N2, H7N7) and VSV. However, it could not be induced by infection of influenza B virus, treatment of RNA mimics stimulus, or treatment of IFNβ, indicating VIN was selectively utilized by some virus during millions of years' evolution (81). Interestingly, there are some host RNAs identified to be regulated by only one specific virus, such as CSR19, CSR21, CSR26, and CSR34 in hepatitis C virus (HCV) infection (58), indicating their specific role for HCV infection.

It is noteworthy that some host lncRNA's expression also is responsive to viral infection in some specific organs, for example placenta in mammal, which is usually believed to an immune tolerance place. A recent study pointed out that SeV infection of human trophectoderm progenitor cells induced an lncRNA expression, named lncRHOXF1, which was transcribed from the X chromosome. lncRHOXF1 promoted the host response to viral infections (82).

lncRNA WORKING MECHANISMS AT THE MOLECULAR LEVEL

As for the working mode of RNA, aside from riboswitches and ribozymes whose RNA structure alone performs the functional units, most ncRNAs operate as complexes with proteins, such as ribosome, telomerase, snRNP, snoRNP, and RISC complex of microRNAs. LncRNAs also perform in a similar manner, interacting with diverse proteins to perform different functions.

In cell nucleus, lncRNAs usually associate with chromatin modification protein or epigenetic modulator to regulate coding gene expression *in trans* or *in cis,* for example host lncRNA-EPS and lncRNA-COX1 (51, 83). In some cases, the activity of one lncRNA could have more than one target or even the whole chromatin. LncRNA Firre was reported to interact with multiple sites of the genome and influences chromatin topological organization through interacting with the nuclear-matrix factor hnRNPU (71). The well-known lncRNA XIST from X chromosome regulated the expression status of the whole X chromosome through recruitment of the polycomb repressive complexes PRC1/2 (84). Another example is NORAD, a conserved and broadly expressed long noncoding RNA, which preserves the whole genome stability in mitosis by serving as a molecular decoy for PUMILIO proteins (85).

Many lncRNAs utilize base pairing to bind other molecules of nucleic acids, such as microRNAs, as competing endogenous RNAs (ceRNAs) to regulate other RNA transcripts (86, 87). For instance, lncRNA linc-MD1 binds miR-133 and miR-135 in myocytes to liberate the expression of muscle-specific transcription factors MAML1 and MEF2C (88). LncRNAs also function through associating signaling transductors or enzymes. In the cytoplasm of dendritic cells, host lncRNA lnc-DC binds transcription factor STAT3 to protect its phosphorylation on tyrosine-705 through preventing protein phosphatase SHP1 binding (89). Another example is from Song's lab, showing that NF-κB-upregulated lncRNA NKILA binds to NF-κB/IκB complex and directly masks phosphorylation motifs of IκB, thereby inhibiting IKK-induced IκB phosphorylation and thus NF-κB activation (90).

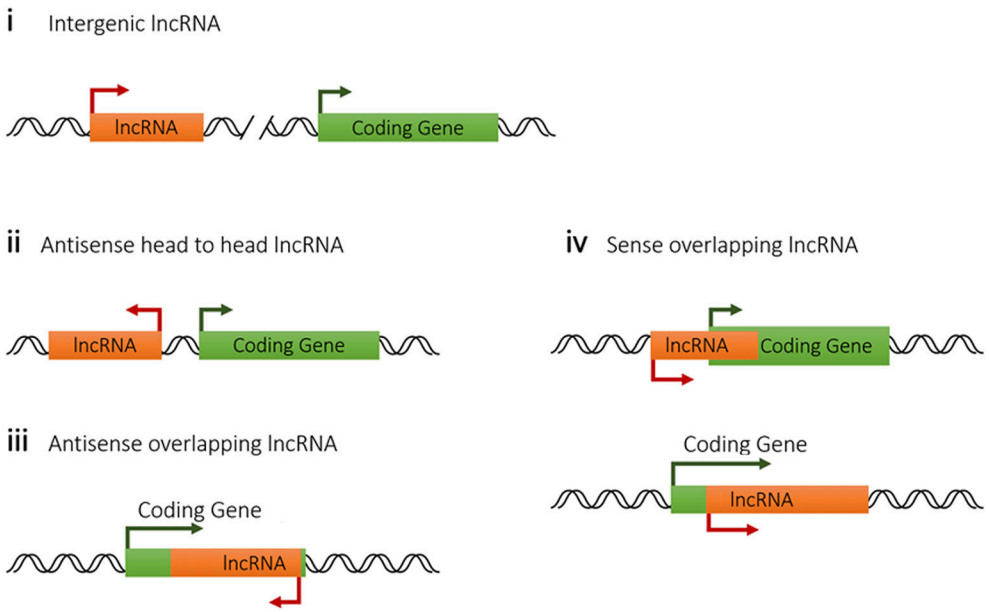

FIGURE 2 | Classification of lncRNAs from host genome according to their positions to the nearest coding genes.

In summary, lncRNAs associate with various molecules through base-pairing or spatial structure interaction to exploit different actions as illustrated in **Figure 3**. However, it is still unknown whether there are other undiscovered modes of action for lncRNA.

lncRNAs REGULATING VIRAL LIFE CYCLE

The course of viral infection cycle in host cells is the trigger of host immune defense and the causes of pathological damages. Taking enveloped virus for example, the life cycle of virus can be summarized as following: First, virus enters the host cell membrane through receptor-mediated endocytosis, followed by viral genome release into the cytoplasm. After passing through a refined procedure of genome integration and latency, or without latency, viral genome utilizes cellular machinery and viral enzymes to synthesize protein component, replicate their genome, and then assemble new progeny virions. Finally infective virions are released and infect other host cells. While large amount of molecules and medicines have been proved to target the process of viral life cycle, increasing number of lncRNAs are revealed to regulate different steps in this process.

Viral Gene Expression

During HSV infection, host paraspeckle lncRNA NEAT1, together with paraspeckle protein P54nrb and PSPC1, associates with HSV-1 genomic DNA and recruits STAT3 to paraspeckle. They facilitate the interaction between STAT3 and viral gene promoters to increase viral gene expression and viral replication, as reported in both human and mouse cell lines (62). However, during human immunodeficiency virus (HIV) infection, NEAT1 and paraspeckle bodies negatively regulate viral production in human cell line through increasing nucleus-to-cytoplasm export of instability element (INS)-containing HIV-1 mRNAs to promote HIV-1 transcript splicing (63). These results indicate that NEAT1 and paraspeckle probably perform different roles in different viral infections. As an important nuclear body for gene expression regulation, paraspeckles are targeted by some viral RNAs. Epstein-Barr virus (EBV) derived lncRNAs oriPtL and oriPtR are bi-directionally transcribed RNAs from EBV latency origin of replication in the nucleus. They bind paraspeckle protein NONO and RNA-editing enzyme ADAR to modulate global viral lytic gene expression and viral DNA replication through an evolutionarily conserved and thermodynamically stable hairpin at their family of repeat (FR) regions (31).

Viral RNA Stability

To affect viral RNA stability or directly degrade viral RNA during infection, host cells have developed many defense mechanisms, including endoribonuclease RNaseL, exoribonuclease Xrn1, and RNAi pathway. However, some viruses have evolved their own strategies to counteract this degradation, such as poliovirus and other group C enterovirus. RNaseL, which degrades single stranded RNAs, is activated by the second messenger 2′-5′-linked oligoadenylate (2–5As) produced by its synthetase OAS, another important sensor for dsRNA in the cytoplasm. Poliovirus and other group C enterovirus have a conserved RNA structure within the open reading frame which functions as a competitive inhibitor of the antiviral endoribonuclease RNaseL (**Figure 1**). Hence, this viral RNA was named the RNaseL competitive inhibitor RNA (RNase L ciRNA) (32). Another example is flavivirus sfRNA. As described in the sections above, flavivirus produces large amount of sfRNAs which can halted exoribonuclease Xrn1 digesting, and sfRNAs can also repress

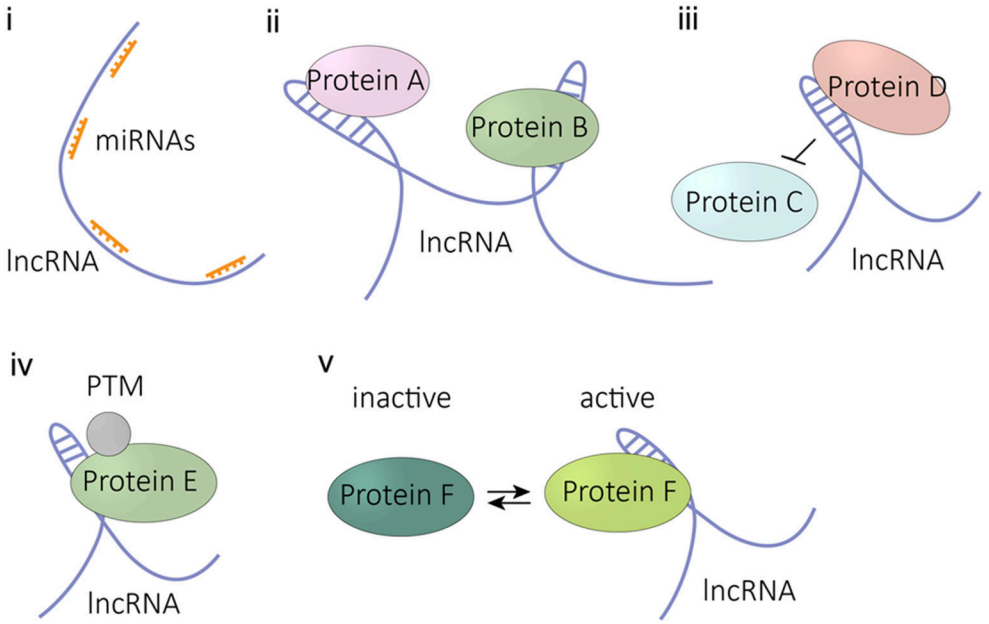

FIGURE 3 | Molecular modes of lncRNA activities. (i) competitive binding microRNA as a sponge RNA, (ii) binding protein as a competitive inhibitor, (iii) binding one protein to prevent its interaction with another protein, (iv) influencing the PTM modification of the binding protein, (v) affecting the activity of the binding protein. PTM, post translational modification.

the RNAi machinery by saturation of endoribonuclease Dicer (33, 34). It is the same for adenovirus VA RNAs whose structure and high abundancy sequestrate several key members of the RNAi pathway, such as Dicer and Exportin 5 (37).

Transition and Rhythm of Viral Life Cycle

Stage transition and rhythm of viral life cycle is closely associated with the course of pathological processing and immune response status. So maintaining viral latency or switching to lytic reactivation is intensively regulated by both virus and host, which involves many lncRNAs from both sides. It has been reported that viral RNA ASP, an HIV antisense transcript, recruits polycomb repression complex 2 (PRC2) to the HIV-1 5′ LTR, resulting in the accumulation of suppressive H3K27 trimethylation to facilitate the establishment of HIV latency (38). Another case is PAN RNA from KSHV. During initiation of the lytic phase, KSHV expresses a highly abundant long noncoding transcript, viral PAN RNA, which guides specific demethylases and PRC2 to the KSHV genome to mediate activation of viral gene expression, leading to the production of infectious virus and lytic infection (39–41). On the other hand, host cells are also trying to influence the life cycle of virus. It has been reported that host lncRNA NRON, highly expressed in resting $CD4^+$ T lymphocytes, maintains HIV-1 latency by retaining transcription factor NFAT in the cytoplasm to suppress NFAT-mediated viral gene activation (44).

Metabolic Regulation

As viral replication requires large amounts of material and energy, it has developed some strategies to hijack metabolic network of host cells to direct metabolite flow to their benefit. However, for a long time we knew little about the underlying molecular mechanism of how virus performs. Recently, a host intergenic lncRNA, lncRNA-ACOD1, was reported to be induced by many viruses as mentioned above and further results from our lab revealed that it promoted viral infection through manipulating host cell metabolism. In the cytoplasm, it directly binds to glutamic-oxaloacetic transaminase 2 (GOT2) near the substrate niche. This RNA-protein interaction promotes the catalytic activity of GOT2 to facilitate metabolite production, such as L-aspartate, α-ketoglutarate and subsequent lipid production, which feeds virus and accelerates viral replication (45). This work marked lncRNA-AOCD1 as a metabolic regulator that is hijacked by virus in infection. The large number of unstudied virally induced transcripts makes it highly likely that future studies will reveal a much greater share for this class of lncRNAs in regulation of viral infection.

Recently more and more studies have revealed that metabolic regulation is vital for host immune regulation. Metabolite is not just the source of energy and nutrition, but also a regulator of host cells immunity, which is the case for itaconate pathway in innate immunity (91) and tetrahydrobiopterin (BH4) in adaptive immunity (92). Although it still at the beginning, metabolic lncRNA probable also has other functions in immune regulation.

It is noteworthy that while lncRNA-ACOD was revealed to be induced in many organs and tissues, including liver, spleen, lung, and lymph nodes, it has a constitutive high-expression level in heart, which is a high energy-consuming organ. Considering lncRNA-ACOD1 as a metabolic regulator, its high expression in heart indicates that it may also participate in the regulation of cardiovascular function.

If host cells cannot take back the control of metabolism, as a protective mechanism for host cells, initiating mitochondria-induced cell death of viral infected cell at early stage of infection can shut down the energy source for virus and restrict viral infection spreading. However, some viruses have evolved to maintain host cell alive at least until they completely fulfill the infective cycle and infect other cells. John H. Sinclair et al. found a 2.7-kilobase viral RNA transcript (RNA2.7 or Beta2.7) from human cytomegalovirus (HCMV) protected host cells from apoptosis, and RNA2.7 accounts for >20% of total viral transcription at the early stage of infection (12 to 24 h) (29). RNA2.7 interacted with complex I and prevented translocation of its essential subunit GRIM-19 to stabilize the mitochondrial membrane potential, resulting in continued adenosine triphosphate (ATP) production to support virus replication (30).

These two studies above represent delicate strategies in modulating the metabolic viability of the infected cell by noncoding RNAs from both virus and host. Researchers have begun to apply these RNAs or the functional fragment in the diagnoses and therapies of clinical relevant disorders. For example, the 800 nucleotide subdomain of RNA2.7, which plays an anti-apoptotic role and maintains a high level of ATP production in neurons, has been exploited in the development of a novel therapeutic for Parkinson's disease (93).

Packaging and Releasing

Packaging and releasing is the last step of life cycle for virus, which is also an excellent chance for host to block infection. It has been reported that cytidine deaminase APOBEC3G (A3G) has a broad antiviral activity against diverse retroviruses and retrotransposons, through inducing C-to-U mutations in the minus-strand viral DNA after its encapsidation into virions (94, 95). However, this process relies on polymerase III-transcribed host 7SL RNA, which is the RNA component of SRP, also known as 4.5S RNA. Studies with HIV-1 infection revealed that A3G selectively interacts with 7SL RNA, then interacted with the RNA-binding domain of HIV-1 Gag protein and was preferentially packaged into virus particles. So 7SL RNA encapsulated into retroviruses functions as a key cofactor of the antiviral protein A3G (46, 47), which proposes a new working model for Pol III-transcribed host noncoding RNAs to participate antiviral immune responses.

lncRNAs REGULATING ANTIVIRAL INNATE IMMUNE RESPONSES

Effective induction of IFN and cytokines along with the downstream effectors expression are crucial for host antiviral response and is known to be orchestrated by multiple mechanisms (96, 97). Nowadays increasing evidences point to the presence of lncRNA-mediated regulatory mechanisms on this pathway. Just like protein regulators, some lncRNAs involved in this mechanism promote this response as positive regulator to strengthen antiviral defense, while others function to attenuate immune responses as a negative feedback regulator to avoid excessive immune pathological effect or are utilized by virus to escape antiviral defense.

Antiviral Innate Immune Signaling and Cytokine Productions

The first step of host antiviral responses is sensing viral invasion. Interestingly, some viral sensors in host are also regulated by lncRNAs, which is the case of canonical sensor RIG-I. A recent research revealed an IFN-induced host lncRNA lnc-Lsm3b competed with viral RNAs in binding RIG-I monomer. The binding of lnc-Lsm3b to RIG-I protein obstructs the conformational shift of RIG-I protein which is essential for its activation. So lnc-Lsm3b prevents downstream signaling and thereby terminates IFN production. As host lnc-Lsm3b is an immune induced gene at the late stage of innate response, so in viral infection it functions as a negative feedback regulator of RIG-I pathway (48) (**Figure 1**). Another example comes from viral RNAs, a particular sfRNA from dengue virus DENV-2 clade (PR-2B). It has been reported to bind E3 ubiquitin ligase TRIM25. This binding prevents the deubiquitinylation of TRIM25 by ubiquitin-specific peptidase 15 (USP15). Ubiquitinated TRIM25 is unable to polyubiquitinate RIG-I and stabilize RIG-I, resulting in a significant decrease in the IFN production and an impaired antiviral responses to facilitate viral infection (35) (**Figure 1**).

Another viral sensor, double-stranded RNA-dependent protein kinase (PKR) is a cytoplasmic sensor of viral RNA, whose activation induces translation inhibition to suppress viral protein synthesis through phosphorylation of eIF2a. However, some viral lncRNAs, such as VA RNAs and EBERs, bind to PKR, but do not induce PKR activation. Instead, they prevent PKR dimerization and auto-phosphorylation. Therefore, signaling through PKR to eIF2a is blocked and translation of viral proteins is properly initiated (27, 28) (**Figure 1**).

The downstream immune signaling is also regulated directly by lncRNAs through RNA-protein interactions. For example, host lncRNA lnc-DC regulates STAT3 signaling as that is described above. Another example is NF-κB, which is believed to be a key director of inflammatory cytokine expression and late stage IFN production in antiviral immune responses. Lethe, an intergenic lncRNA that is also considered as a pseudogene, is selectively induced by inflammatory cytokines and glucocorticoid receptor agonist. Functionally, it interacts with NF-κB active subunit RelA to inhibit RelA DNA binding and target gene activation, as a negative feedback to NF-κB (49). So Lethe could have profound effect in immune responses. Another example for NF-κB is p50-associated COX2 extragenic RNA (PACER), which is an antisense head to head lncRNA with coding gene cyclooxygenase 2 (COX2). PACER associates with p50, a repressive subunit of NF-κB, and occludes it from the COX-2 promoter to facilitate active NF-κB dimer p65/p50 to COX2 promoter to promote its expression (50). Interestingly, there is an intergenic lncRNA nearby Cox2 (Ptgs2) gene in mouse, named lincRNA-Cox2. However, lincRNA-Cox2 does not alter Cox2 (Ptgs2) expression in mouse. As an immune-induced gene, it regulates expression of a group of immune response genes,

including chemokines, chemokine receptors, and ISGs, through binding hnRNP A/B and A2/B1 (51).

Some lncRNA can directly regulate the transcription of cytokines. Inflammatory cytokine tumor necrosis factor α (TNFα) is a potent activator of host immune responses to viral infections. It has been reported its transcription is regulated by an intergenic lncRNA from host genome, named THRIL which is short for TNFα and hnRNPL related immunoregulatory lincRNA. THRIL functions through binding hnRNPL to form a functional lncRNA/hnRNPL complex that binds to TNFα promoter. However, it is downregulated by TNFα or TLR stimulation as a negative feedback mechanism during viral infection (52). As an essential host lncRNA for the formation of nuclear body paraspeckle, NEAT1 regulates many immune-related genes expression, including antiviral cytokine IL8 (64) and host sensor RIG-I and DDX60 (65). NEAT1 transits the splicing factor proline and glutamine-rich protein (SFPQ) to paraspeckle to remove its transcriptional inhibitory effects, promoting the expression of immune responses genes.

Virally encoded RNAs also participate in regulating host IFN and cytokine production as an immune modulator for the sake of virus. KSHV derived PAN RNA was proved in to interact with histone H1/H2A, single-stranded binding proteins (SSBPs) and interferon regulatory factor 4 (IRF4) in infected cells to decrease the expression of IFNγ, IFNα, interleukin-18, and RNase L to facilitate viral infection in primate cell lines (42).

Taken together, despite only small fraction of lncRNAs being studied, existing data points a critical role for this class of lncRNAs in regulation of immune signaling and cytokine expression. Increasing evidence of lncRNA directly interacting with signaling molecules has also been found not only in immune signaling but also in other biomedical flied including p53 signaling (98) and EGFR signaling (99). However, we still do not know how many other sensors, adaptors and transcription factors could interact with lncRNA to be modulated. We just put up some of the remaining questions in antiviral signaling as illustrated in **Figure 1**: Is TBK1-IRF3 activity regulated by some lncRNA? Does lncRNA interact with signaling adaptors, such as STING or TRAF6, to modulate their protein modifications and functions? Resolving these questions will help us deepen our understanding of relationship between lncRNAs and signaling cascades. Furthermore, as many protein factors have been identified as negative or positive regulators of immune signaling, defining the interaction of the regulatory lncRNAs with these proteins in the context of immune responses will yield an uncharted regulatory network of immune cascades.

Expression of Interferon-Stimulated Genes (ISGs)

ISGs comprise antiviral effectors and immune regulators and a number of lncRNAs exert their functions through regulating ISGs to modulate the antiviral effect. There are many such cases for IFN-induced lncRNAs. LncRNA NRIR (negative regulator of IFN response) was identified in primary human hepatocytes as an intergenic lncRNA (60), originated from a locus downstream of the protein-coding ISG CMPK2, and therefore it was also named lncRNA-CMPK2. NRIR could be induced by IFNα, IFNβ, and IFNγ through JAK-STAT pathway to generate a spliced polyadenylated nuclear transcript. Functionally, knocking down NRIR in human hepatocyte cell line significantly increased many antiviral ISGs expression, including CMPK2, viperin, IFIT1, IFIT3, ISG15, MxA, CXCL10, and IFITM1, and subsequently decreased HCV replication, suggesting NRIR acts as a repressor of ISGs expression (60). Although its molecular mechanism yet to be further determined, the case of NRIR indicates that lncRNAs play an important role in the feedback loop of IFN-induced gene expression regulation. In addition, NRIR is reported to have a remarkably high level in livers of patients with HCV infection than that in healthy donors, indicating this negative regulator could be utilized by HCV *in vivo* (60).

Another example is host lncRNA EGOT (Eosinophil Granule Ontogeny Transcript), which is a polyadenylated nuclear conserved lncRNA (100). EGOT was first described to be expressed in eosinophils and is thought to function in mature eosinophils (101). Later studies from GTEx Consortium revealed that the highest levels of EGOT were found in nonhematopoietic tissues such as breast, pancreas, pituitary, vagina and kidney cortex (102). Interestingly, EGOT genomic locus in human being was marked by monomethylation of Lys4 of histone H3 (H3K4), but not trimethylation of H3K4, indicating that EGOT could be an enhancer RNA (103). Its function remained obscure until a recent study using human liver cell line HuH7 cells revealed that the level of EGOT was dramatically induced by viral infection, such as HCV, influenza, and Semliki Forest virus (SFV), and high doses of IFNα stimulation (58), and furthermore, knockdown experiment in HCV infected cells revealed that EGOT negatively regulated antiviral responses through inhibiting a subset of ISGs' expression as a negative feedback regulatory mechanism of IFN pathway. However, this mechanism is often utilized by viruses, such as HCV, influenza, and Semliki Forest virus (SFV) (58).

While described above are two lncRNAs upregulated in viral infection, host lncRNA NRAV (Negative Regulator of Antiviral Response) was markedly reduced in infection with influenza virus and a number of other viruses in several cell lines. NRAV was firstly described in a study on genes expression changes in response to influenza virus H1N1 infection in human alveolar epithelium cell line A549 (59). Importantly, overexpression experiments with cDNA microarray analysis revealed NRAV depressed the expression of many antiviral effecters, including IFIT2, IFIT3, IFITM3, MxA, and OASL (59). Studies in human cell lines or transgenic mice showed that enforced expression of NRAV markedly promoted viral replication while knockdown of NRAV suppresses viral replication in IAV infection model. Furthermore, molecular studies revealed that the spatial structure of NRAV associates with the promoters of *IFITM3* and *MxA* to alter their epigenetic histone modifications to suppress their initial transcription rates (59). It is noteworthy that NRAV is an antisense overlapping lncRNA and locates in the first intron of *DYNLL1* gene encoding human dynein light chain, however, they are transcribed as independent operating unit, which is a unique working model *in trans* for antisense overlapping lncRNA.

Host lncRNA LUARIS (lncRNA up-regulator of antiviral response IFN signaling) is also down-regulated, like NRAV in viral infection, and as it was named, it functions to promote ISGs expression. LUARIS was identified in a screen for IRF3-dependent genes in HuS immortalized human hepatocytes as

an IFN-reduced lncRNA originally named lncRNA#32 (57). It was reported that during HBV or HCV infection in primary hepatocytes, LUARIS associates with hnRNPU and functions through activating transcription factor 2 (ATF2) to promote the expression of multiple ISGs. And silencing of LUARIS dramatically reduced the level of ISGs' expression and increased cellular sensitivity to encephalomyocarditis virus (EMCV) infection (57). These data indicate LUARIS has evolved to control the magnitude of the IFN response through multiple regulatory pathways to prevent possible toxicity of overstimulation.

While many lncRNAs regulate the expressions of multiple ISGs, some lncRNAs can only modulate a single ISG target, for example BISPR (BST2 IFN-Stimulated Positive Regulator). Host lncRNA BISPR belongs to antisense head to head lncRNAs, which is transcribed from a bidirectional promoter shared with BST2/Tetherin. BST2 is an IFN-induced restriction factor that blocks the budding of enveloped viruses by tethering them to the cell surface. Independent studies from two groups revealed that BISPR/BST2 gene-pair is induced by IFN stimulation in many cells, such as human hepatocyte cell line HuH7 (79, 80) and monocyte cell line THP1 (61). Interestingly, the increase of BISPR expression precedes that of BST2 after IFN stimulation, indicating that BISPR induces or facilitates the initiation of BST2 transcription. Future studies revealed that BISPR knockdown reduced BST2 expression and ectopic expression of BISPR RNA enhanced BST2 expression, indicating that BISPR RNA mediates this function, rather than the transcription (61). Mechanistically, BISPR obstructs the repressive activity of PRC2 at the promoter of BST2 to facilitate the transcription. BISPR also interacts with methyltransferase component EZH2 and an enhancer region to promote the formation of enhancer-promoter complex. Since many lncRNAs belong to antisense head to head lncRNAs and many immunity-related genes have bidirectional promoters, the mechanism of BISPR and BST2 study shed light on these antisense lncRNA family's functional and molecular exploration.

Aside from transcription, the translation of ISG mRNA is also regulated by noncoding RNAs. It was reported that conserved host RNA-binding proteins G3BP1, G3BP2, and CAPRIN1, which were required for ISG translation, were targeted by a non-coding RNA from dengue virus. Human pathogen dengue virus is a positive-strand RNA flavivirus and it produces abundant non-coding sfRNA, which directly binds to G3BP1, G3BP2, and CAPRI, as a molecular sponge, to antagonize ISGs translation (36). This mechanism impairs establishment of the antiviral defense of host cells, allowing virus to replicate and escape from the IFN response. Interestingly, Dengue sfRNA response has not been observed in other flaviviruses. The unique association of this sfRNA molecule to dengue viral pathogenesis provides a potential molecular target for clinical diagnosis and future therapeutic options for dengue virus infection.

lncRNAs IN ANTIVIRAL ADAPTIVE IMMUNITY

lncRNA transcriptome profiling in different T cell lineages has been performed and characterized in both humans and mice (104, 105). Thousands of lncRNAs have been identified to closely associate with T lymphocyte differentiation and a number of them are identified as novel T helper (Th) cell lineage-specific lncRNAs. Most of these lncRNAs are intragenic or adjacent to lineage-specific protein-coding genes in the genome. And many were bound and regulated by the key transcription factors T-bet, GATA-3, STAT4, and STAT6 as revealed by RNA-Seq data and ChIP-Seq data (106). Some of these lncRNAs have been proved to have functions for T cells, including an enhancer-like lncRNA called IfngAS1 (also known as Tmevpg1 or NeST) promoting Th1 cytokine IFNγ expression (53), host lncRNA Th2-LCR lncRNA controlling Th2 cytokines IL-4, IL-5, and IL-13 expression (107), and lincR-Ccr2-5′AS regulating the migration of Th2 cells (106). However, few of them have been characterized for molecule mechanism. To our knowledge, it is the only case for IfngAS1 located adjacent to the *IFNG* in both mice and humans (54, 55). IfngAS1 promotes IFNγ expression by binding to WD repeat-containing protein 5 (WDR5), a component of histone H3 lysine 4 (H3K4) methyltransferase complex, and alter histone 3 methylation at the IFNγ locus (53, 56).

Despite these discovery in T lymphocytes have been achieved, many questions are yet to be answered. For example, it is not clear whether T cell activation signaling is regulated by lncRNA as an intrinsic regulation; whether TCR complex involves the interaction with lncRNA during T cell activation and whether B lymphocyte differentiation and function being regulated by lncRNA. The lack of mechanistic insight in this field is due, in part to technical obstacles. For example, cell-specific gene-manipulating for lncRNA genes *in vivo* is much harder than coding genes, because base-pair insertions usually do not lead to functional mutation for noncoding genes. The development of gene editing techniques, such as CRISPR system will provide more convenient approaches for research to manipulate lncRNA genes in specific lymphocytes in the future. Other obstacles include the low abundance of samples, as the number of one specific T cell subtype is very low. With the development of trace amount detection technique, such as SHERLOCK (108) and DETECTR (109), and super-resolution structured illumination microscopy $(SIM)^2$ and Cryo-scanning electron microscopy (Cryo-SEM), RNA molecule will be better detected for small amount or even in single cell.

CHALLENGES AND PERSPECTIVE

lncRNAs from host and virus exert their functions through various mechanisms, such as associating with transcription factors, chromatin modifiers, signaling adapters, enzymes and microRNAs, to influence gene expression, host metabolism, post translational modification, and protein activities. Although great achievements have been obtained in the field of lncRNA, there are still a significant number of concerns to be solved. While the linear sequence of RNA is relative easy to analyze, RNA spatial structure is still difficult to be examined or predicted as the RNA structure is flexible and usually interacts with other molecules *in vivo*. Nevertheless, great effort has been made to interpret the physiological structures of RNAs

using elaborate biochemical methodologies to distinguish single-strand, double-strand, exposed, and buried regions. Some approaches are used to resolve one specific lncRNA structure at a single-nucleotide resolution (110) and some are designed to reveal the whole higher-order transcriptome structure in living cells (111–113). Nevertheless, RNA modification adds another layer of complexity in RNA structural and functional research. One example is that a study from Tao Pan's laboratory demonstrated that RNA local structures are altered by one site of N^6-methyladenosine (m^6A), which is the most abundant modification in eukaryotic RNA (114, 115). This local change in structure increases the binding by heterogeneous nuclear ribonucleoprotein C (hnRNPC) (116). So the RNA functional studies have evolved to the combination of linear sequence, spatial structure and RNA modification, which set a higher level of requirement for experimental exploration and bioinformatics analysis.

Another challenge in RNA mechanistic study is the identification of lncRNA binding molecules. Aside from nucleic acid, protein is so far the only recognized molecule that RNA interacts with. Whether there are other compounds, such as small chemical molecules or metal ions, interacting with lncRNA *in vivo* is largely unknown. Nevertheless, a number of approaches have been developed to get comprehensive profile of RNA-binding proteins (RBPs) *in vivo*. Ultraviolet (UV) crosslinking has been used to covalently stabilize native protein-RNA interactions in living cells. The crosslinked proteins are isolated by oligo (dT) purification for mass spectrometry identification. This approach, named RNA-protein interactome capture, identified over a thousand RBPs within different cells and species, such as human HeLa and HEK293 cells (117, 118), mouse embryonic stem cells (119), Saccharomyces cerevisiae (120), and Caenorhabditis elegans (121). Many proteins identified were not previously recognized to bind RNA, namely unorthodox RBPs, include many metabolic enzymes, regulators of alternative splicing, the E3 ubiquitin ligase, and the FAST kinase domain-containing protein 2 (FASTKD2) (122, 123). Furthermore, to determine how RBPs bind to RNA in living cells, Matthias W. Hentze et al. have improved the resolution of this approach from protein level to RNA binding peptide level, by adding a protease digestion step followed by a second round of oligo (dT) capture and mass spectrometry. They have discovered numerous RNA-binding domains (RBDs) in human HeLa cells and revealed that catalytic centers and protein-protein interacting domains are preferred RNA binding sites (124). Interestingly, nearly half of the RNA-binding sites were mapped to intrinsically disordered protein regions, indicating flexible protein domains are the favorable part for RNA-protein interactions. Recently, a new approach, RNA interactome with click chemistry (RICK), has been developed to capture the nascent RNA-protein interactome. Using this method, newly transcribed RNAs were integrated with 5-ethynyluridine and after UV crosslinking, the RNA-protein complex was labeled with biotin via click chemical reaction and subjected to purification and identification (125). This method allows identification of proteins bound to a wide range of RNA species, including the nonpolyadenylated RNAs that were neglected in the past.

Many techniques are designed to examine the binding proteins of one specific RNA molecule (126). Ci Chu et al. developed the method of chromatin isolation by RNA purification (ChIRP) using antisense DNA oligonucleotides to capture and purify specific lncRNA-chromatin complexes, initially to address lncRNA-binding sites on the genome (127). They further developed the methodology that enabled the identification of lncRNA-binding proteins and RNA-RNA interactions (128, 129). Two other groups have also developed similar approaches independently, RNA antisense purification (RAP) (130) and capture hybridization analysis of RNA targets (Chart) (131), with some differences in cross-linking and chromatin shearing. As ChIRP-like methods use chemical cross-linking, they do not differentiate direct interactions from indirect interactions. Another method named cross-linking ligation and sequencing of hybrids (CLASH) uses UV crosslinking that captures only direct RNA-protein interaction, which is suitable for investigation of RNA-protein interaction in nucleus (132).

Some lncRNAs are not sensitive to siRNA mediated RNA degradation, especially the ones located in the nucleus, as siRNA machinery is mainly located in cytoplasm. So knocking down and knocking out strategies often interfere lncRNA studies. Despite the technical hurdles and difficulties, much effort has been made to inactivate lncRNA genes in mouse models and these studies have made great discoveries. An elegant review by Lingjie Li et al. summarized genetic targeting strategies used to study lncRNA loci *in vivo* (133). However, these gene deletion approaches are difficult to scale up for genome-wide functional screening of lncRNAs, as many lncRNA are induced or suppressed in viral or other immune relative models. Happily, integrated genome-wide CRISPR interference (CRISPRi) and CRISPR activation (GRISPRa), a novel systematic functional screening system of lncRNAs has been developed (134). This technique represses or activates transcription via recruitment of a nuclease-dead Cas9 (dCas9) enzyme to the transcriptional start site (TSS) of genes by a customizable single guide RNA (sgRNA). A study designed a sgRNA library targeting 16,401 lncRNA loci, with 10 sgRNAs against each lncRNA TSS. In seven diverse human cell lines, 499 lncRNA loci were identified to be necessary for robust cellular growth. Surprisingly, the majority of these lncRNA genes showed growth modifying function exclusively in one cell type, and not a single lncRNA modified cell growth across all the cell line tested (135). This is a promising approach for lncRNA functional screening in other models, for example it could be used for the identification of viral specific lncRNAs or cell type-specific antiviral lncRNAs.

As illustrated by the literature, the most significant character that lncRNAs differ from other RNA molecules is that lncRNAs exhibit high cellular- and organ- specific expression patterns. Therefore, lncRNAs modulate protein activities and signaling pathways only in specific cell types. Understanding the mechanisms of specific expression regulation at multiple levels, including but not limited to specific transcriptional factors, epigenetic modification, and local chromatin spatial organization, will help us better understand spatial and temporal regulation of lncRNAs and choose more specific interfering target in certain pathological circumstances, such as viral infection.

However, as the strong correlation between lncRNA expression patterns and its functions, analysis of lncRNA expression profiles in different cells under different physiological and pathological models is a good way to predict the relevant functions of one specific lncRNA or a panel of lncRNAs. Furthermore, combining more omics data at the same time, such as transcriptome, epigenomics, proteinomics, phosphoproteomics and metabolomics, and conducting integrated correlation analysis with coding genes, epigenetic modification, protein modification and metabolite, will provide more detailed function information of lncRNAs and might draw the whole regulatory network draft for us.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

The author thanks Dr. Xuetao Cao for critically reading this manuscript and other staff in our department for their assistance. This work is supported by grants from the National Natural Science Foundation of China (31470871, 81671566, 31722019) and Shanghai Rising-Star Program (16QA1404700).

REFERENCES

1. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* (2014) 157:77–94. doi: 10.1016/j.cell.2014.03.008
2. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science* (2005) 309:1559–63. doi: 10.1126/science.1112014
3. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature* (2012) 489:57–74. doi: 10.1038/nature11247
4. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet.* (2013) 45:1113–20. doi: 10.1038/ng.2764
5. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* (2002) 420:563–73. doi: 10.1038/nature01266
6. Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, et al. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* (2016) 351:397–400. doi: 10.1126/science.aad7977
7. Wang Q, Lee I, Ren J, Ajay SS, Lee YS, Bao X. Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection. *Mol Ther J Am Soc Gene Ther.* (2013) 21:368–79. doi: 10.1038/mt.2012.237
8. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* (2015) 47:199–208. doi: 10.1038/ng.3192
9. Fortes P, Morris KV. Long noncoding RNAs in viral infections. *Virus Res.* (2016) 212:1–11. doi: 10.1016/j.virusres.2015.10.002
10. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* (2011) 147:1537–50. doi: 10.1016/j.cell.2011.11.055
11. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* (2012) 22:1775–89. doi: 10.1101/gr.132159.111
12. Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, et al. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol.* (2004) 5:730–7. doi: 10.1038/ni1087
13. Pichlmair A, Schulz O, Tan CP, Naslund TI, Liljestrom P, Weber F, et al. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. *Science* (2006) 314:997–1001. doi: 10.1126/science.1132998
14. Kato H, Takeuchi O, Mikamo-Satoh E, Hirai R, Kawai T, Matsushita K, et al. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med.* (2008) 205:1601–10. doi: 10.1084/jem.20080091
15. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* (2001) 413:732–8. doi: 10.1038/35099560
16. Liu S, Cai X, Wu J, Cong Q, Chen X, Li T, et al. Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science* (2015) 347:aaa2630. doi: 10.1126/science.aaa2630
17. Shembade N, Ma A, Harhaj EW. Inhibition of NF-kappaB signaling by A20 through disruption of ubiquitin enzyme complexes. *Science* (2010) 327:1135–9. doi: 10.1126/science.1182364
18. Meng XY, Luo Y, Anwar MN, Sun Y, Gao Y, Zhang H, et al. Long non-coding RNAs: emerging and versatile regulators in host-virus interactions. *Front Immunol.* (2017) 8:1663. doi: 10.3389/fimmu.2017.01663
19. Ouyang J, Hu J, Chen JL. lncRNAs regulate the innate immune response to viral infection. *Wiley Interdiscip Rev RNA* (2016) 7:129–43. doi: 10.1002/wrna.1321
20. Liu W, Ding C. Roles of LncRNAs in viral infections. *Front Cell Infect Microbiol.* (2017) 7:205. doi: 10.3389/fcimb.2017.00205
21. Mathews MB. Genes for VA-RNA in adenovirus 2. *Cell* (1975) 6:223–9. doi: 10.1016/0092-8674(75)90013-6
22. Marx JL. "Viroids": a new kind of pathogen? *Science* (1972) 178:734. doi: 10.1126/science.178.4062.734
23. Shimura H, Masuta C. Plant subviral RNAs as a long noncoding RNA (lncRNA): analogy with animal lncRNAs in host-virus interactions. *Virus Res.* (2016) 212:25–9. doi: 10.1016/j.virusres.2015.06.016
24. Diener TO. Viroids: "living fossils" of primordial RNAs? *Biol Direct* (2016) 11:15. doi: 10.1186/s13062-016-0116-7
25. Conrad NK. New insights into the expression and functions of the Kaposi's sarcoma-associated herpesvirus long noncoding PAN RNA. *Virus Res.* (2016) 212:53–63. doi: 10.1016/j.virusres.2015.06.012
26. Mrazek J, Kreutmayer SB, Grasser FA, Polacek N, Huttenhofer A. Subtractive hybridization identifies novel differentially expressed ncRNA species in EBV-infected human B cells. *Nucleic Acids Res.* (2007) 35:e73. doi: 10.1093/nar/gkm244
27. Iwakiri D. Multifunctional non-coding Epstein-Barr virus encoded RNAs (EBERs) contribute to viral pathogenesis. *Virus Res.* (2016) 212:30–8. doi: 10.1016/j.virusres.2015.08.007
28. Vachon VK, Conn GL. Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Res.* (2016) 212:39–52. doi: 10.1016/j.virusres.2015.06.018
29. Gatherer D, Seirafian S, Cunningham C, Holton M, Dargan DJ, Baluchova K, et al. High-resolution human cytomegalovirus transcriptome. *Proc Natl Acad Sci USA.* (2011) 108:19755–60. doi: 10.1073/pnas.1115861108
30. Reeves MB, Davies AA, McSharry BP, Wilkinson GW, Sinclair JH. Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science* (2007) 316:1345–8. doi: 10.1126/science.1142984
31. Cao S, Moss W, O'Grady T, Concha M, Strong MJ, Wang X, et al. New noncoding lytic transcripts derived from the epstein-barr virus latency origin of replication, oriP, are hyperedited, bind the paraspeckle protein, NONO/p54nrb, and support viral lytic transcription. *J Virol.* (2015) 89:7120–32. doi: 10.1128/JVI.00608-15
32. Townsend HL, Jha BK, Silverman RH, Barton DJ. A putative loop E motif and an H-H kissing loop interaction are conserved and functional features in a group C enterovirus RNA that inhibits ribonuclease L. *RNA Biol.* (2008) 5:263–72. doi: 10.4161/rna.7165
33. Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, et al. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells. *J Virol.* (2012) 86:13486–500. doi: 10.1128/JVI.01104-12

34. Schnettler E, Tykalova H, Watson M, Sharma M, Sterken MG, Obbard DJ, et al. Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses. *Nucleic Acids Res.* (2014) 42:9436–46. doi: 10.1093/nar/gku657
35. Manokaran G, Finol E, Wang C, Gunaratne J, Bahl J, Ong EZ, et al. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* (2015) 350:217–21. doi: 10.1126/science.aab3369
36. Bidet K, Dadlani D, Garcia-Blanco MA. G3BP1, G3BP2 and CAPRIN1 are required for translation of interferon stimulated mRNAs and are targeted by a dengue virus non-coding RNA. *PLoS Pathog.* (2014) 10:e1004242. doi: 10.1371/journal.ppat.1004242
37. Andersson MG, Haasnoot PC, Xu N, Berenjian S, Berkhout B, Akusjarvi G. Suppression of RNA interference by adenovirus virus-associated RNA. *J Virol.* (2005) 79:9556–65. doi: 10.1128/JVI.79.15.9556-9565.2005
38. Zapata JC, Campilongo F, Barclay RA, DeMarino C, Iglesias-Ussel MD, Kashanchi F, et al. The human immunodeficiency virus 1 ASP RNA promotes viral latency by recruiting the polycomb repressor complex 2 and promoting nucleosome assembly. *Virology* (2017) 506:34–44. doi: 10.1016/j.virol.2017.03.002
39. Rossetto CC, Tarrant-Elorza M, Verma S, Purushothaman P, Pari GS. Regulation of viral and cellular gene expression by Kaposi's sarcoma-associated herpesvirus polyadenylated nuclear RNA. *J Virol.* (2013) 87:5540–53. doi: 10.1128/JVI.03111-12
40. Rossetto CC, Pari G. KSHV PAN RNA associates with demethylases UTX and JMJD3 to activate lytic replication through a physical interaction with the virus genome. *PLoS Pathog.* (2012) 8:e1002680. doi: 10.1371/journal.ppat.1002680
41. Borah S, Darricarrere N, Darnell A, Myoung J, Steitz JA. A viral nuclear noncoding RNA binds re-localized poly(A) binding protein and is required for late KSHV gene expression. *PLoS Pathog.* (2011) 7:e1002300. doi: 10.1371/journal.ppat.1002300
42. Rossetto CC, Pari GS. Kaposi's sarcoma-associated herpesvirus noncoding polyadenylated nuclear RNA interacts with virus- and host cell-encoded proteins and suppresses expression of genes involved in immune modulation. *J Virol.* (2011) 85:13290–7. doi: 10.1128/JVI.05886-11
43. Charley PA, Wilusz J. Standing your ground to exoribonucleases: function of flavivirus long non-coding rnas. *Virus Res.* (2016) 212:70–7. doi: 10.1016/j.virusres.2015.09.009
44. Li J, Chen C, Ma X, Geng G, Liu B, Zhang Y, et al. Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nat Commun.* (2016) 7:11730. doi: 10.1038/ncomms11730
45. Wang P, Xu J, Wang Y, Cao X. An interferon-independent lncRNA promotes viral replication by modulating cellular metabolism. *Science* (2017) 358:1051–5. doi: 10.1126/science.aao0409
46. Wang T, Tian C, Zhang W, Luo K, Sarkis PT, Yu L, et al. 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J Virol.* (2007) 81:13112–24. doi: 10.1128/JVI.00892-07
47. Zhang W, Du J, Yu K, Wang T, Yong X, Yu XF. Association of potent human antiviral cytidine deaminases with 7SL RNA and viral RNP in HIV-1 virions. *J Virol.* (2010) 84:12903–13. doi: 10.1128/JVI.01632-10
48. Jiang M, Zhang S, Yang Z, Lin H, Zhu J, Liu L, et al. Self-recognition of an inducible host lncRNA by RIG-I feedback restricts innate immune response. *Cell* (2018) 173:906–19.e13. doi: 10.1016/j.cell.2018.03.064
49. Rapicavoli NA, Qu K, Zhang J, Mikhail M, Laberge RM, Chang HY. A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *Elife* (2013) 2:e00762. doi: 10.7554/eLife.00762
50. Krawczyk M, Emerson BM. p50-associated COX-2 extragenic RNA (PACER) activates COX-2 gene expression by occluding repressive NF-kappaB complexes. *Elife* (2014) 3:e01776. doi: 10.7554/eLife.01776
51. Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* (2013) 341:789–92. doi: 10.1126/science.1240925
52. Li Z, Chao TC, Chang KY, Lin N, Patil VS, Shimizu C, et al. The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. *Proc Natl Acad Sci USA.* (2014) 111:1002–7. doi: 10.1073/pnas.1313768111
53. Collier SP, Collins PL, Williams CL, Boothby MR, Aune TM. Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *J Immunol.* (2012) 189:2084–8. doi: 10.4049/jimmunol.1200774
54. Vigneau S, Rohrlich PS, Brahic M, Bureau JF. Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *J Virol.* (2003) 77:5632–8. doi: 10.1128/JVI.77.10.5632-5638.2003
55. Collier SP, Henderson MA, Tossberg JT, Aune TM. Regulation of the Th1 genomic locus from Ifng through Tmevpg1 by T-bet. *J Immunol.* (2014) 193:3959–65. doi: 10.4049/jimmunol.1401099
56. Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. *Cell* (2013) 152:743–54. doi: 10.1016/j.cell.2013.01.015
57. Nishitsuji H, Ujino S, Yoshio S, Sugiyama M, Mizokami M, Kanto T, et al. Long noncoding RNA #32 contributes to antiviral responses by controlling interferon-stimulated gene expression. *Proc Natl Acad Sci USA.* (2016) 113:10388–93. doi: 10.1073/pnas.1525022113
58. Carnero E, Barriocanal M, Prior C, Pablo Unfried J, Segura V, Guruceaga E, et al. Long noncoding RNA EGOT negatively affects the antiviral response and favors HCV replication. *EMBO Rep.* (2016) 17:1013–28. doi: 10.15252/embr.201541763
59. Ouyang J, Zhu X, Chen Y, Wei H, Chen Q, Chi X, et al. NRAV, a long noncoding rna, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell Host Microbe* (2014) 16:616–26. doi: 10.1016/j.chom.2014.10.001
60. Kambara H, Niazi F, Kostadinova L, Moonka DK, Siegel CT, Post AB, et al. Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Res.* (2014) 42:10668–80. doi: 10.1093/nar/gku713
61. Kambara H, Gunawardane L, Zebrowski E, Kostadinova L, Jobava R, Krokowski D, et al. Regulation of interferon-stimulated gene BST2 by a lncRNA transcribed from a shared bidirectional promoter. *Front Immunol.* (2014) 5:676. doi: 10.3389/fimmu.2014.00676
62. Wang Z, Fan P, Zhao Y, Zhang S, Lu J, Xie W, et al. NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription. *Cell Mol Life Sci.* (2017) 74:1117–31. doi: 10.1007/s00018-016-2398-4
63. Zhang Q, Chen CY, Yedavalli VS, Jeang KT. NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. *mBio* (2013) 4:e00596–12. doi: 10.1128/mBio.00596-12
64. Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol Cell* (2014) 53:393–406. doi: 10.1016/j.molcel.2014.01.009
65. Ma H, Han P, Ye W, Chen H, Zheng X, Cheng L, et al. The long noncoding RNA NEAT1 exerts antihantaviral effects by acting as positive feedback for RIG-I signaling. *J Virol.* (2017) 91:e02250-16. doi: 10.1128/JVI.02250-16
66. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* (2008) 18:1433–45. doi: 10.1101/gr.078378.108
67. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. *Trends Genet.* (2007) 23:614–22. doi: 10.1016/j.tig.2007.09.001
68. Diribarne G, Bensaude O. 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biol.* (2009) 6:122–8. doi: 10.4161/rna.6.2.8115
69. Muddashetty R, Khanam T, Kondrashov A, Bundman M, Iacoangeli A, Kremerskothen J, et al. Poly(A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles. *J Mol Biol.* (2002) 321:433–45. doi: 10.1016/S0022-2836(02)00655-1
70. Mattick JS, Rinn JL. Discovery and annotation of long noncoding RNAs. *Nat Struct Mol Biol.* (2015) 22:5–7. doi: 10.1038/nsmb.2942
71. Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol.* (2014) 21:198–206. doi: 10.1038/nsmb.2764

72. Cho SW, Xu J, Sun R, Mumbach MR, Carter AC, Chen YG, et al. Promoter of lncRNA gene PVT1 is a tumor-suppressor DNA boundary element. *Cell* (2018) 173:1398–412. doi: 10.1016/j.cell.2018.03.068
73. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* (2011) 472:120–4. doi: 10.1038/nature09819
74. Luo S, Lu JY, Liu L, Yin Y, Chen C, Han X, et al. Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. *Cell Stem Cell* (2016) 18:637–52. doi: 10.1016/j.stem.2016.01.024
75. Zemmour D, Pratama A, Loughhead SM, Mathis D, Benoist C. Flicr, a long noncoding RNA, modulates Foxp3 expression and autoimmunity. *Proc Natl Acad Sci USA*. (2017) 114:E3472–80. doi: 10.1073/pnas.1700946114
76. Yin QF, Yang L, Zhang Y, Xiang JF, Wu YW, Carmichael GG, et al. Long noncoding RNAs with snoRNA ends. *Mol Cell* (2012) 48:219–30. doi: 10.1016/j.molcel.2012.07.033
77. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol.* (2014) 32:453–61. doi: 10.1038/nbt.2890
78. Josset L, Tchitchek N, Gralinski LE, Ferris MT, Eisfeld AJ, Green RR, et al. Annotation of long non-coding RNAs expressed in Collaborative Cross founder mice in response to respiratory virus infection reveals a new class of interferon-stimulated transcripts. *RNA Biol.* (2014) 11:875–90. doi: 10.4161/rna.29442
79. Carnero E, Barriocanal M, Segura V, Guruceaga E, Prior C, Borner K, et al. Type I Interferon regulates the expression of long non-coding RNAs. *Front Immunol.* (2014) 5:548. doi: 10.3389/fimmu.2014.00548
80. Barriocanal M, Carnero E, Segura V, Fortes P. Long Non-Coding RNA BST2/BISPR is Induced by IFN and regulates the expression of the antiviral factor tetherin. *Front Immunol.* (2014) 5:655. doi: 10.3389/fimmu.2014.00655
81. Winterling C, Koch M, Koeppel M, Garcia-Alcalde F, Karlas A, Meyer TF. Evidence for a crucial role of a host non-coding RNA in influenza A virus replication. *RNA Biol.* (2014) 11:66–75. doi: 10.4161/rna.27504
82. Penkala I, Wang J, Syrett CM, Goetzl L, Lopez CB, Anguera MC. lncRHOXF1, a Long Noncoding RNA from the X Chromosome that suppresses viral response genes during development of the early human placenta. *Mol Cell Biol.* (2016) 36:1764–75. doi: 10.1128/MCB.01098-15
83. Atianand MK, Hu W, Satpathy AT, Shen Y, Ricci EP, Alvarez-Dominguez JR, et al. A long noncoding RNA lincRNA-EPS acts as a transcriptional brake to restrain inflammation. *Cell* (2016) 165:1672–85. doi: 10.1016/j.cell.2016.05.075
84. Almeida M, Pintacuda G, Masui O, Koseki Y, Gdula M, Cerase A, et al. PCGF3/5-PRC1 initiates polycomb recruitment in X chromosome inactivation. *Science* (2017) 356:1081–4. doi: 10.1126/science.aal2512
85. Lee S, Kopp F, Chang TC, Sataluri A, Chen B, Sivakumar S, et al. Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* (2015) 164:69–80. doi: 10.1016/j.cell.2015.12.017
86. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature* (2014) 505:344–52. doi: 10.1038/nature12986
87. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet.* (2016) 17:272–83. doi: 10.1038/nrg.2016.20
88. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* (2011) 147:358–69. doi: 10.1016/j.cell.2011.09.028
89. Wang P, Xue Y, Han Y, Lin L, Wu C, Xu S, et al. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Science* (2014) 344:310–3. doi: 10.1126/science.1251456
90. Liu B, Sun L, Liu Q, Gong C, Yao Y, Lv X, et al. A Cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* (2015) 27:370–81. doi: 10.1016/j.ccell.2015.02.004
91. Dominguez-Andres J, Novakovic B, Li Y, Scicluna BP, Gresnigt MS, Arts RJW, et al. The itaconate pathway is a central regulatory node linking innate immune tolerance and trained immunity. *Cell Metab*. (2018) 29:1–10. doi: 10.1016/j.cmet.2018.09.003
92. Cronin SJF, Seehus C, Weidinger A, Talbot S, Reissig S, Seifert M, et al. The metabolite BH4 controls T cell proliferation in autoimmunity and cancer. *Nature* (2018) 563:564–8. doi: 10.1038/s41586-018-0701-2
93. Poole E, Kuan WL, Barker R, Sinclair J. The human cytomegalovirus non-coding Beta2.7 RNA as a novel therapeutic for Parkinson's disease–Translational research with no translation. *Virus Res.* (2016) 212:64–9. doi: 10.1016/j.virusres.2015.05.007
94. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* (2003) 424:99–103. doi: 10.1038/nature01709
95. Yang B, Chen K, Zhang C, Huang S, Zhang H. Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J Biol Chem.* (2007) 282:11667–75. doi: 10.1074/jbc.M606864200
96. Chen W, Han C, Xie B, Hu X, Yu Q, Shi L, et al. Induction of Siglec-G by RNA viruses inhibits the innate immune response by promoting RIG-I degradation. *Cell* (2013) 152:467–78. doi: 10.1016/j.cell.2013.01.011
97. Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol.* (2014) 32:513–45. doi: 10.1146/annurev-immunol-032713-120231
98. Schmitt AM, Garcia JT, Hung T, Flynn RA, Shen Y, Qu K, et al. An inducible long noncoding RNA amplifies DNA damage signaling. *Nat Genet.* (2016) 48:1370–6. doi: 10.1038/ng.3673
99. Lin A, Li C, Xing Z, Hu Q, Liang K, Han L, et al. The LINK-A lncRNA activates normoxic HIF1alpha signalling in triple-negative breast cancer. *Nat Cell Biol.* (2016) 18:213–24. doi: 10.1038/ncb3295
100. Rose D, Stadler PF. Molecular evolution of the non-coding eosinophil granule ontogeny transcript. *Front Genet.* (2011) 2:69. doi: 10.3389/fgene.2011.00069
101. Wagner LA, Christensen CJ, Dunn DM, Spangrude GJ, Georgelas A, Kelley L, et al. EGO, a novel, noncoding RNA gene, regulates eosinophil granule protein transcript expression. *Blood* (2007) 109:5191–8. doi: 10.1182/blood-2006-06-027987
102. Consortium GT. The genotype-tissue expression (GTEx) project. *Nat Genet.* (2013) 45:580–5. doi: 10.1038/ng.2653
103. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* (2007) 39:311–8. doi: 10.1038/ng1966
104. Pang KC, Dinger ME, Mercer TR, Malquori L, Grimmond SM, Chen W, et al. Genome-wide identification of long noncoding RNAs in CD8+ T cells. *J Immunol.* (2009) 182:7738–48. doi: 10.4049/jimmunol.0900603
105. Spurlock CF 3rd, Shaginurova G, Tossberg JT, Hester JD, Chapman N, Guo Y, et al. Profiles of long noncoding RNAs in human naive and memory T cells. *J Immunol.* (2017) 199:547–58. doi: 10.4049/jimmunol.1700232
106. Hu G, Tang Q, Sharma S, Yu F, Escobar TM, Muljo SA, et al. Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat Immunol.* (2013) 14:1190–8. doi: 10.1038/ni.2712
107. Spurlock CF 3rd, Tossberg JT, Guo Y, Collier SP, Crooke PS 3rd, Aune TM. Expression and functions of long noncoding RNAs during human T helper cell differentiation. *Nat Commun.* (2015) 6:6932. doi: 10.1038/ncomms7932
108. Zhang C, Yang X, Qi Q, Gao Y, Wei Q, Han S. lncRNA-HEIH in serum and exosomes as a potential biomarker in the HCV-related hepatocellular carcinoma. *Cancer Biomarkers A Dis Markers* (2018) 21:651–9. doi: 10.3233/CBM-170727
109. Fu N, Zhao SX, Kong LB, Du JH, Ren WG, Han F, et al. LncRNA-ATB/microRNA-200a/beta-catenin regulatory axis involved in the progression of HCV-related hepatic fibrosis. *Gene* (2017) 618:1–7. doi: 10.1016/j.gene.2017.03.008
110. Spitale RC, Crisalli P, Flynn RA, Torre EA, Kool ET, Chang HY. RNA SHAPE analysis in living cells. *Nat Chem Biol.* (2013) 9:18–20. doi: 10.1038/nchembio.1131
111. Spitale RC, Flynn RA, Zhang QC, Crisalli P, Lee B, Jung JW, et al. Structural imprints *in vivo* decode RNA regulatory mechanisms. *Nature* (2015) 519:486–90. doi: 10.1038/nature14263
112. Wan Y, Qu K, Zhang QC, Flynn RA, Manor O, Ouyang Z, et al. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* (2014) 505:706–9. doi: 10.1038/nature12946

113. Lu Z, Zhang QC, Lee B, Flynn RA, Smith MA, Robinson JT, et al. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* (2016) 165:1267–79. doi: 10.1016/j.cell.2016.04.028
114. Zheng G, Dahl JA, Niu Y, Fu Y, Klungland A, Yang YG, et al. Sprouts of RNA epigenetics: the discovery of mammalian RNA demethylases. *RNA Biol.* (2013) 10:915–8. doi: 10.4161/rna.24711
115. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* (2012) 485:201–6. doi: 10.1038/nature11112
116. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* (2015) 518:560–4. doi: 10.1038/nature14234
117. Baltz AG, Munschauer M, Schwanhausser B, Vasile A, Murakawa Y, Schueler M, et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* (2012) 46:674–90. doi: 10.1016/j.molcel.2012.05.021
118. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* (2012) 149:1393–406. doi: 10.1016/j.cell.2012.04.031
119. Kwon SC, Yi H, Eichelbaum K, Fohr S, Fischer B, You KT, et al. The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol.* (2013) 20:1122–30. doi: 10.1038/nsmb.2638
120. Mitchell SF, Jain S, She M, Parker R. Global analysis of yeast mRNPs. *Nat Struct Mol Biol.* (2013) 20:127–33. doi: 10.1038/nsmb.2468
121. Matia-Gonzalez AM, Laing EE, Gerber AP. Conserved mRNA-binding proteomes in eukaryotic organisms. *Nat Struct Mol Biol.* (2015) 22:1027–33. doi: 10.1038/nsmb.3128
122. Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Alleaume AM, et al. The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun.* (2015) 6:10127. doi: 10.1038/ncomms10127
123. Papasaikas P, Tejedor JR, Vigevani L, Valcarcel J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* (2015) 57:7–22. doi: 10.1016/j.molcel.2014.10.030
124. Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, et al. Comprehensive identification of RNA-binding domains in human cells. *Mol Cell* (2016) 63:696–710. doi: 10.1016/j.molcel.2016.06.029
125. Bao X, Guo X, Yin M, Tariq M, Lai Y, Kanwal S, et al. Capturing the interactome of newly transcribed RNA. *Nat Methods* (2018) 15:213–20. doi: 10.1038/nmeth.4595
126. Chu C, Spitale RC, Chang HY. Technologies to probe functions and mechanisms of long noncoding RNAs. *Nat Struct Mol Biol.* (2015) 22:29–35. doi: 10.1038/nsmb.2921
127. Chu C, Quinn J, Chang HY. Chromatin isolation by RNA purification (ChIRP). *J Vis Exp.* (2012) 25:pii 3912 doi: 10.3791/3912
128. Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, et al. Systematic discovery of xist RNA binding proteins. *Cell* (2015) 161:404–16. doi: 10.1016/j.cell.2015.03.025
129. Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* (2013) 493:231–5. doi: 10.1038/nature11661
130. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* (2013) 341:1237973. doi: 10.1126/science.1237973
131. Simon MD, Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, et al. The genomic binding sites of a noncoding RNA. *Proc Natl Acad Sci USA.* (2011) 108:20497–502. doi: 10.1073/pnas.1113536108
132. Helwak A, Tollervey D. Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). *Nat Protocols* (2014) 9:711–28. doi: 10.1038/nprot.2014.043
133. Li L, Chang HY. Physiological roles of long noncoding RNAs: insight from knockout mice. *Trends Cell Biol.* (2014) 24:594–602. doi: 10.1016/j.tcb.2014.06.003
134. Bester AC, Lee JD, Chavez A, Lee YR, Nachmani D, Vora S, et al. An integrated genome-wide CRISPRa approach to functionalize lncRNAs in drug resistance. *Cell* (2018) 173:649–64.e20. doi: 10.1016/j.cell.2018.03.052
135. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* (2017) 355:35–40. doi: 10.1126/science.aah7111

Conflict of Interest Statement: The author declares 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 Immunology

REVIEW
published: 08 March 2019
doi: 10.3389/fimmu.2019.00420

Disease-Associated SNPs in Inflammation-Related lncRNAs

Ainara Castellanos-Rubio[1,2] *and Sankar Ghosh*[3]*

[1] *Immunogenetics Research Laboratory, Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, UPV/EHU, Leioa, Spain,* [2] *Functional Studies in Celiac Disease Group, BioCruces Health Research Institute, Barakaldo, Spain,* [3] *Department of Microbiology and Immunology, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, United States*

Immune-mediated diseases, such as celiac disease, type 1 diabetes or multiple sclerosis, are a clinically heterogeneous group of diseases that share many key genetic triggers. Although the pathogenic mechanisms responsible for the development of immune mediated disorders is not totally understood, high-throughput genomic studies, such as GWAS and Immunochip, performed in the past few years have provided intriguing hints about underlying mechanisms and pathways that lead to disease. More than a hundred gene variants associated with disease susceptibility have been identified through such studies, but the progress toward understanding the underlying mechanisms has been slow. The majority of the identified risk variants are located in non-coding regions of the genome making it difficult to assign a molecular function to the SNPs. However, recent studies have revealed that many of the non-coding regions bearing disease-associated SNPs generate long non-coding RNAs (lncRNAs). LncRNAs have been implicated in several inflammatory diseases, and many of them have been shown to function as regulators of gene expression. Many of the disease associated SNPs located in lncRNAs modify their secondary structure, or influence expression levels, thereby affecting their regulatory function, hence contributing to the development of disease.

Keywords: lncRNA, linc RNA, inflammation, inflammatory disease, GWAS, SNP

OPEN ACCESS

Edited by:
Jorge Henao-Mejia, University of Pennsylvania, United States

Reviewed by:
Trinath Jamma, Birla Institute of Technology and Science, India
Renata Meirelles Pereira, Federal University of Rio de Janeiro, Brazil

***Correspondence:**
Sankar Ghosh
sg2715@columbia.edu

Specialty section:
This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: *27 September 2018*
Accepted: *18 February 2019*
Published: *08 March 2019*

Citation:
Castellanos-Rubio A and Ghosh S (2019) Disease-Associated SNPs in Inflammation-Related lncRNAs. Front. Immunol. 10:420. doi: 10.3389/fimmu.2019.00420

INTRODUCTION

Immune mediated disorders, such as celiac disease (CeD), inflammatory bowel diseases (IBD), atherosclerosis, rheumatoid arthritis (RA), type 1 diabetes (T1D) or multiple sclerosis (MS) among others, are a group of clinically heterogeneous diseases caused by dysfunction of the immune system. These disorders, share underlying pathogenic mechanisms that are not totally understood, although the general belief is that they develop due to an imbalance in the interaction between genetic and environmental factors (1, 2).

Immune mediated disorders share dysregulation of many key regulatory pathways and techniques, such as genome wide association studies (GWAS) coupled with next generation sequencing (NGS), have significantly increased our knowledge of genetic factors underlying such diseases (3). In the past decade or so, hundreds of risk alleles, both common and disease specific, have been identified by GWAS. Moreover, using the Immunochip platform in which 200,000 polymorphisms in 186 immune disease related regions were analyzed, additional immune disease associated variants were identified that revealed common susceptibility loci for several of these diseases (4–7). While these studies have helped identify immune disease conferring gene variants, the progress toward the understanding of the underlying mechanisms has remained limited. This

difficulty is particularly exacerbated by the fact that around 90% of the SNPs associated with these diseases are located in non-coding regions, making it difficult to link them to specific biological pathways (8–10).

Advances in the sequencing and annotation of the human genome have revealed that many non-coding regions of the genome encode long non-coding RNAs (lncRNAs). The importance of lncRNA molecules in different biological processes is beginning to be appreciated although there is much that remains to be understood. LncRNAs are RNA molecules longer than 200 bp in length with no protein-coding potential. LncRNA expression is generally cell-lineage specific and they have diverse and still not very well-characterized mechanisms of action. The emerging view is that lncRNAs are fundamental regulators of transcription as they have been shown to control every level of the gene expression program. LncRNAs have been shown to control processes like protein synthesis, RNA maturation, and transport to regulate genes post-transcriptionally and they are also involved in transcriptional gene silencing through regulating the chromatin structure (11–13).

Many lncRNAs are enriched for disease-associated SNPs, suggesting that these SNPs might alter the function of lncRNAs e.g., by altering their secondary structures (14). Moreover, expression profile analyses of lncRNAs located in autoimmune disease-associated regions showed that lncRNAs are enriched in these loci, suggesting that lncRNAs may be crucial for interpreting GWAS findings (15). Disease associated SNPs can modify the lncRNA sequence or alter their gene expression levels, affecting their regulatory capacity, and alterations in the structure and function of lncRNAs have been associated with several immune-mediated diseases. However, the precise mechanism by which lncRNA variants contribute to the pathogenesis of disease remains unknown in the majority of cases (16).

As previously done with protein coding genes, intergenic SNPs have been analyzed in the context of lncRNA expression quantitative trait loci (eQTLs), namely, genetic variants that would explain variation in the lncRNA expression levels (17). More than 100 cis-eQTLs have been found in different tissues that appear to regulate the expression of nearby lncRNAs. In general, these eQTLs are lncRNA specific and do not regulate the expression of neighboring protein coding genes, but since lncRNAs can regulate the expression of protein-coding genes, both, located close by or farther away in the genome, it is possible that these SNPs also influence the function of protein-coding genes. Moreover, a considerable number of the lncRNA cis-eQTLs belong to disease-associated SNPs, making lncRNAs a potential link between non-coding SNPs and the expression of protein-coding genes (18–21) (**Figure 1A**).

Disease associated SNPs have been also suggested to be involved in "splicing models," in which the presence of different alleles would influence the splicing of the lncRNA by regulating exon skipping (22) (**Figure 1B**). In this way, different isoforms of the lncRNA would present different affinity for their binding partners affecting the regulation of downstream events. It has been observed, that when human endothelial cells are stimulated with lipopolysaccharide (LPS) several lncRNAs show splice variant-specific expression at different stimulation time points

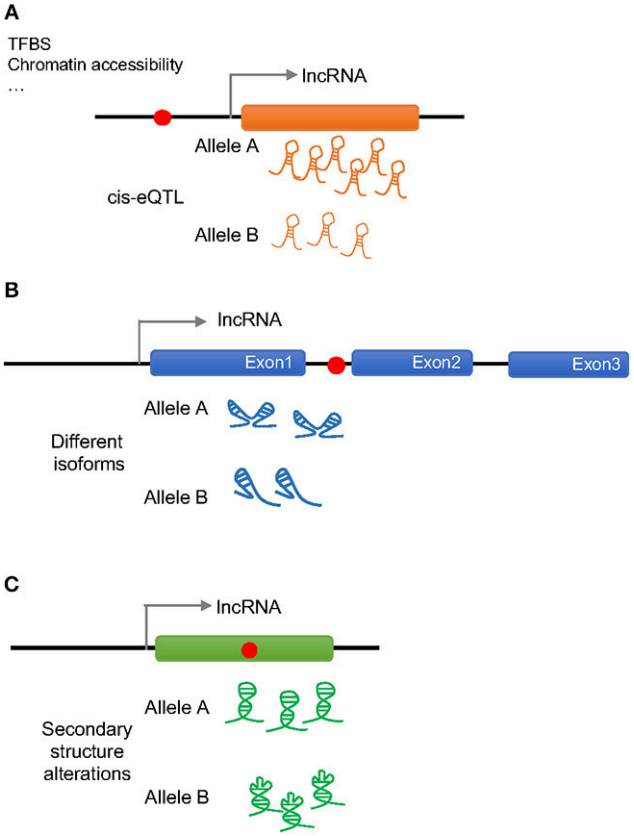

FIGURE 1 | Possible effects of a disease associated SNP on lncRNA regulation and function. Red dots represent disease associated SNPs. **(A)** A SNP in the promoter region of a lncRNA can change the amount of transcribed lncRNA (cis-eQTL) by altering the binding of transcription factor binding sites (TFBS) or the chromatin accessibility altering downstream effects. **(B)** An intronic SNP in a lncRNA can influence the splicing of the lncRNA generating different isoforms that will act differently. **(C)** A SNP in the exonic sequence of a lncRNA can change its secondary structure altering the binding to the molecular partners.

(23), underlining the importance of SNP regulated splicing in lncRNA function.

As it is generally believed that lncRNA molecules adopt specific secondary and tertiary structures to execute their functions, it is likely that disease associated SNPs have an impact on lncRNA structure (**Figure 1C**). The analysis of secondary structures has largely been performed using computational tools, and several tools can predict changes in lncRNA structures caused by the presence of different alleles of a certain SNP (24–26). For example, GWAS SNPs associated with IBD and T1D have been shown to disrupt the structure of an lncRNA implicated in the pathogenesis of both diseases, which associates with the *BACH2* gene (27). However, this field is still in its infancy and the principles that underlie the impact of SNPs on lncRNA structure and function remains to be fully established.

In this article, we have reviewed the link between four intergenic GWAS variants that are located within lncRNA sequences, which have been associated with inflammatory diseases, and we discuss the studies that have been carried

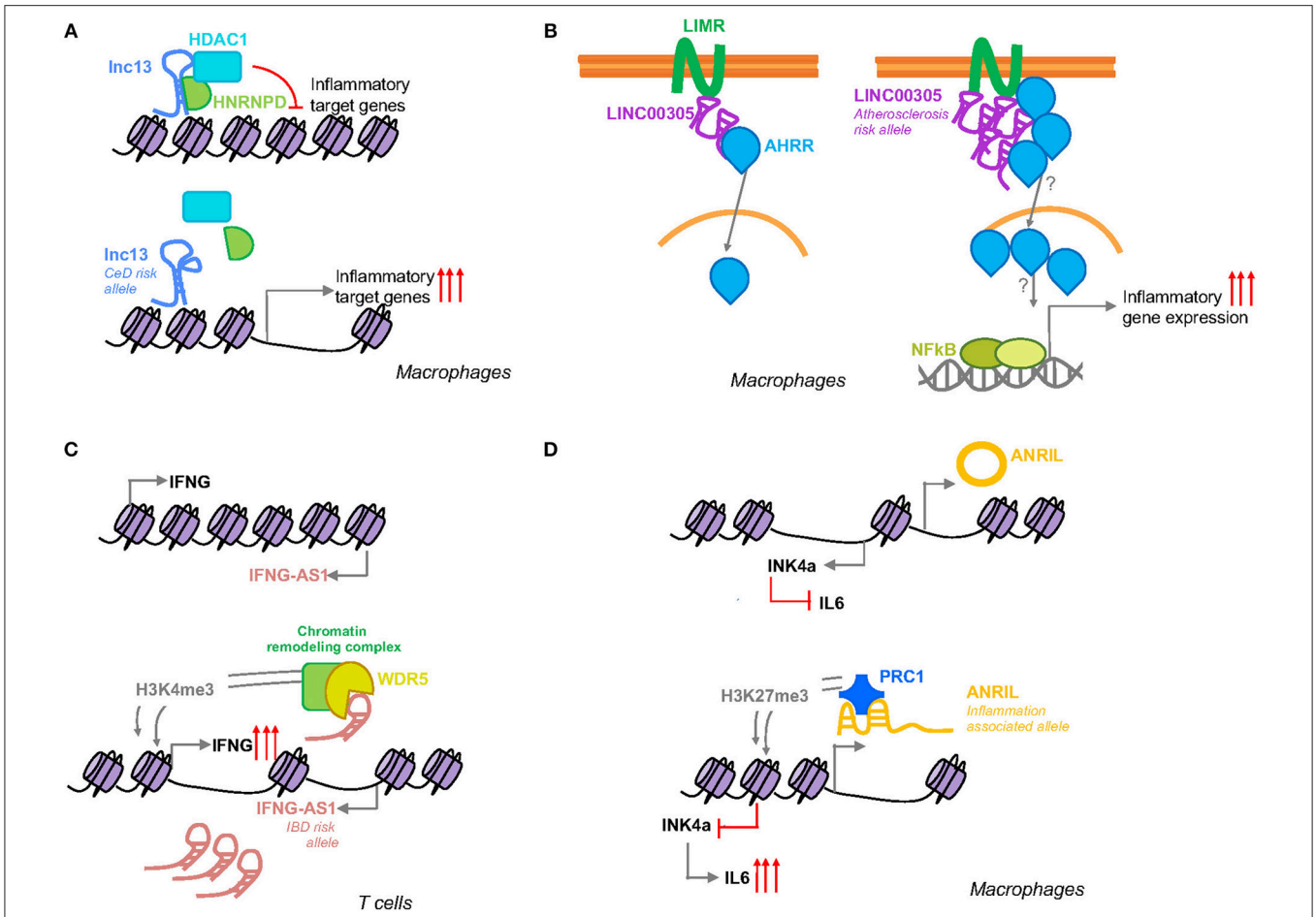

FIGURE 2 | Schematic representation of the function of inflammation associated SNP harboring lncRNAs. **(A)** *Lnc13* harbors a CeD associated SNP that changes the secondary structure of the lncRNA modifying its binding with the proteins hnRNPD and HDAC1 and regulating the expression of disease related inflammatory genes. **(B)** *LINC00305* interacts with the transmembrane protein LIMR facilitating the binding of this protein to AHRR that in turn induces the translocation of the latter to the nucleus inducing NFκB and subsequent inflammatory gene expression. Atherosclerosis patients present higher levels of this lincRNA which could be influenced by a SNP located in the promoter region of *LINC00305*. **(C)** *IFNG-AS1* is closely located to *IFNG* gene. Activation of its transcription leads to induction of *IFNG* by WDR5 mediated H3K4me3 methylation. IBD patients present higher levels of *IFNG-AS1* that could be related to a disease associated SNP located in the enhancer region of the lncRNA. **(D)** Suggested SNP related splicing model for *ANRIL* mediated inflammation regulation. The inflammation associated allele will affect *ANRIL* splicing generating a linear *ANRIL* that will interact with a member of the PRC1 complex mediating an epigenetic transcriptional repression of the *INK4a* gene via H3K27me3.

out to characterize their contribution to the development of disease pathogenesis. As of now, these four inflammatory-disease associated SNPs are the best mechanistically characterized in the context of lncRNA function.

LNC13 AND CELIAC DISEASE SUSCEPTIBILITY

Celiac disease is a complex, chronic, immune-mediated disease that affects ∼1% of the population and develops in genetically susceptible individuals in response to ingested gluten proteins from wheat, barley, and rye (28). The strongest genetic association, around 40% of the genetic risk (29), maps to the human leukocyte antigen (HLA) region in chromosome 6p21, and virtually all CeD patients carry HLA-DQ2 or HLA-DQ8 heterodimers (30, 31).

Two GWA studies, together with the Immunochip project, have identified a total of 39 non-HLA loci associated with the genetic risk of CeD (32–34). Only 3 of the CeD associated SNPs are linked to protein-altering variants located in exonic regions, although some potentially causative coding genes have been proposed, mainly related to the immune response, due to the existence of signals near their 5′ or 3′ regulatory regions. Although some lncRNAs have been related to celiac disease pathogenesis due to the location of an associated SNP within their transcriptional region, and differential expression found in samples from CeD patients (35, 36), the exact mechanism by which they contribute to disease development is not understood.

The only functionally characterized lncRNA harboring a CeD associated intergenic SNP so far has been found linked to the NF-κB pathway (37), which is known to be constitutively active in the CeD mucosa (38, 39). This lncRNA, named *lnc13*, harbors the

SNP rs917997 and it is located in the associated region 2q12, with the sense sequence overlapping the coding gene *IL18RAP* that had been proposed, but never firmly confirmed, as the functional candidate gene in the region (40–42).

This lncRNA is expressed in different human cells and tissues, including mononuclear cells in the lamina propria, where it was observed to be localized in the nucleus. *Lnc13* quantification in small intestinal biopsy samples from celiac patients and controls showed markedly lower levels of this lncRNA in CeD samples, contrary to the expression of the coding mRNA, *IL18RAP* (42). In fact, it is known that *IL18RAP* expression is induced in response to inflammation via NF-κB in certain immune cells (43). The characterization of the regulation, function and mechanisms of action of *lnc13* revealed that under basal conditions *lnc13* represses the expression of certain CeD related genes (*STAT1*, *MYD88*, *IL1RA*, and *TRAF2*) via its interaction with hnRNPD (Heterogeneous Nuclear Ribonucleoprotein D), a nuclear AU1 rich RNA binding protein, and HDAC1 (Histone Deacetylase 1), a histone deacetylase which negatively regulates transcription, proteins on the chromatin (**Figure 2A**). In response to inflammatory stimuli, *lnc13* is degraded by Decapping enzyme 2 (DCP2), releasing the protein complex from chromatin and allowing the expression of the proinflammatory genes (37).

Although the GWAS disease association has been generally attributed to the SNP rs917997 (33), located 1.5 kb away from the coding gene, linkage analysis of the *lnc13* region revealed that there are a total of six SNPs in total linkage disequilibrium within the lncRNA sequence. The nucleotide changes in *lnc13* cause a disruption of the secondary structure of this lncRNA decreasing its affinity to bind hnRNPD and chromatin, resulting in higher expression of the proinflammatory genes (37). Thus, patients with the risk haplotype have a higher basal expression of CeD related inflammatory genes, thereby increasing their predisposition to develop the disease (**Figure 2A**).

The SNPs in *lnc13* have also been associated with other inflammatory diseases, such as T1D, Crohn's disease or rheumatoid arthritis (44–46). Interestingly, while the risk signal in CeD corresponds to the T allele, the C allele is the risk allele in T1D, suggesting that the function of the lncRNA may be cell specific, but equally affected by the SNPs in different inflammatory diseases.

In summary, it is known that *lnc13* and the CeD associated SNP rs917997 contribute to the pathology of celiac disease by regulating expression of certain immune related genes that play a role in the development of inflammation in the intestinal epithelia. However, other cell-type and allele specific functions cannot be excluded, due to the association of this SNP with other inflammatory conditions.

LINC00305 AND INFLAMMATORY RESPONSE IN ATHEROSCLEROSIS

Atherosclerosis is a complex, chronic disease of the arterial wall triggered by multiple factors including amongst others, inflammation and lipid metabolism (47). Monocyte-mediated inflammation plays a critical role in atherosclerosis due to the secretion of proinflammatory cytokines and amplification of local inflammation (48).

Although a locus in the 9p21.3 region is the strongest genetic factor for atherosclerosis described so far, GWAS have led to the identification of a substantial number of additional genetic loci associated with atherosclerosis and atherosclerosis-related complications (49). Analysis of atherosclerosis GWAS SNPs, revealed that the SNP rs2850711 is located within an intronic sequence of a long intergenic non-coding RNA named *LINC00305*. This lincRNA was found to be overexpressed in atherosclerotic plaques and in peripheral blood mononuclear cells (PBMCs) from patients, supporting its role in the disease. Analysis of *LINC00305* in the cell types that composed the plaques revealed that monocytes are the primary cell type expressing this lncRNA and that its expression was induced in response to stimulation with lipopolysaccharide (LPS). Further functional studies, demonstrated that *LINC00305* interacts with the transmembrane protein LIMR (Lipocalin-1 Interacting Membrane Receptor) enhancing its interaction with AHRR (Aryl-Hydrocarbon Receptor Repressor) which at the same time promotes NF-κB activation and subsequent inflammatory gene expression (**Figure 2B**) (50).

The development of atherosclerotic plaques is induced by the change in phenotype of the vascular smooth muscle cells in response to the cytokines secreted by inflammatory cells (51). To investigate the functional significance of *LINC00305* in the pathogenesis of atherosclerosis, the lincRNA was stably overexpressed in human monocytes and these were co-cultured with human aortic smooth muscle cells. It was observed that those muscle cells that were cultured in the presence of the monocytes overexpressing *LINC00305* showed less expression of their basal markers, thus suggesting that they were switching to the pathogenic phenotype (50). Independent studies, have shown that overexpression of this lincRNA induces apoptosis in hypoxia induced endothelial cells (52). Further analysis of the role of this lincRNA in the regulation of apoptosis, revealed that it acts as an endogenous sponge for miR-136 which had been previously related to apoptosis in the context of atherosclerosis (52, 53).

Although it is quite clear that *LINC00305* plays a functional role in development of atherosclerosis by inducing production of inflammatory cytokines in monocytes, and by regulating apoptosis via miR-136, the role of the associated SNP in the function of the lincRNA remains to be elucidated. The GWAS SNP rs2850711 is transmitted in a linkage disequilibrium (LD) block of a total of 16 SNPs, all of which are located within introns, and hence probably do not influence the secondary structure of the lincRNA. Although there are no *in vitro* molecular studies evaluating this possibility, it is noteworthy that one of the associated SNPs lies within an experimentally confirmed USF2 (Upstream Transcription Factor 2) binding region (54). As USF2 is a protein that has been associated with cholesterol metabolism and atherosclerosis development (55), further mechanistic studies assessing the contribution of the SNP alleles in the function and regulation of the lincRNA are necessary to understand how the SNPs in non-coding regions

identified by GWAS influence the inflammatory environment in atherosclerosis.

IFNG-AS1 (NEST OR TMEVPG1) AND ULCERATIVE COLITIS

Inflammatory bowel diseases (IBD) are chronic common inflammatory gastrointestinal disorders clinically comprised of Crohn's disease (CD) and ulcerative colitis (UC) (56). These diseases are believed to develop due to inappropriate inflammatory responses to intestinal microbes and foreign antigens in genetically susceptible individuals (57, 58).

Meta-analyses of multiple GWAS have implicated 163 genetic loci in IBD susceptibility. Although functional analysis of the associated SNPs have revealed multiple pathophysiological mechanisms, the function for many of the genes in close association with these loci are yet to be determined (59, 60). It has been observed that several lncRNAs are differentially expressed in inflammatory bowel disease, and that the expression profiling of lncRNAs can be useful to stratify IBD patients from healthy controls (27, 61, 62).

When comparing the genomic location of differentially expressed lncRNAs with those of IBD susceptibility loci, *IFNG-AS1* (also called *NeST* or *Tmevpg1*) was found to fulfill both criteria (61). The IBD associated SNP rs7134599 is located in the region 12q15 in close proximity to the inflammatory cytokine *IFNG*. This SNP is in total LD with 10 other SNPs within the lncRNA gene. Additionally, *IFNG-AS1* is significantly overexpressed in intestinal samples of ulcerative colitis patients and its expression appears to correlate with the elevated levels of *IFNG*, *IL1*, *IL6*, and *TNF-α* observed in patients (61). Increased expression level of this lncRNA has also been related to other inflammatory diseases, such as Hashimoto's thyroiditis or Sjögren syndrome (63, 64) although the mechanisms by which it contributes to development of these diseases remain unclear.

IFNG-AS1 gene was first related with the immune response in the context of susceptibility to persistent Theiler's virus infection. It was observed that *IFNG-AS1* is expressed in immune cells of mouse and human origin and it was speculated that this lncRNA may regulate the expression of *IFNG* (65). Further studies, demonstrated that *IFNG-AS1* contributes to *IFNG* expression regulation as part of the Th1 differentiation program and that T-bet guides epigenetic remodeling of the lncRNA enhancers, leading to recruitment of stimulus-inducible transcription factors, such as NF-κB (66, 67). More recently, it was observed that different mouse strains with different genotype composition of *IFNG-AS1* were sufficient to confer disparate immune-related phenotypes. Specifically, certain alleles, derived from SJL/J mouse strain, were responsible for the failure to clear Theiler's virus, but at the same time they conferred resistance to lethal infection with *Salmonella enterica Typhimurium* and induced synthesis of *Ifng* in $CD8^+$ T cells. Functional analysis performed in this same study, showed that *IFNG-AS1* is a nuclear lncRNA that can act in trans. *IFNG-AS1* binds WDR5, a component of active chromatin remodeling complexes increasing H3K4me3 methylation which in turn programs an active chromatin state that induces *Ifng* gene transcription (68) (**Figure 2C**).

Although *IFNG-AS1* is differentially expressed in IBD samples and harbors a disease associated SNP (61) the exact impact of the different alleles in the development of the disease has not been assessed so far. *In silico* interaction evaluation of human *IFNG-AS1* lncRNA and WDR5 protein using CatRapid (69), an algorithm that estimates the binding propensity of protein-RNA pairs, states that these two molecules are also able to interact, suggesting that it may act in a similar way as described in mice. Analysis of the location of the SNPs that are in LD with the associated SNP rs7134599 reveals that all of them are located in intronic regions, suggesting that they will most likely not affect the secondary structure of the RNA molecule. However, analysis of the region using HaploReg v4.1 (70) shows that 4 of the SNPs are located within enhancer histone marks and all of them are predicted to disturb a protein binding motif that could change the regulation of the lncRNA expression, thereby influencing the levels of *IFNG*.

Thus, *INFG-AS1* is clearly involved in the immune response and inflammatory processes involved in disease, and although *in silico* data point to a disturbance of lncRNA expression regulation mediated by the IBD associated SNPs uncovered by GWAS, the true relevance of these SNPs have still to be experimentally confirmed.

ANRIL AND INFLAMMATION

The antisense non-coding RNA in the INK4 locus or *ANRIL* was first described in melanoma patients (71) and since its discovery it has been shown to be involved in several types of cancers (72). This lncRNA is located in the 9p21 region, that has been associated by GWAS not only to cancer but also to other diseases that are related with inflammation, such as coronary artery disease (73) or type 2 diabetes (T2D) (74). *ANRIL* is expressed as either linear or circular forms, that have been observed to have opposing effects in disease development (75), making the deciphering of the functionality of this lncRNA and the involvement of its related SNPs highly complicated.

ANRIL has been described to interact with *CBX7* (Chromobox 7), one of the members of the polycomb repressive complex 1 (PRC1). *CBX7* binds both, *ANRIL* and H3K27me3 to mediate an epigenetic transcriptional repression of the *INK4a* (Inhibitor of CDK4) gene, which is located adjacent to the *ANRIL* gene (**Figure 2D**) (76). INK4a is a cell cycle inhibitor that is lost in a wide spectrum of cancers (77), but it has been also been reported to act as an anti-inflammatory molecule that is able to suppress the production of IL-6 in macrophages (78). *ANRIL* itself has been also shown to regulate the inflammatory response by its interaction with the YY1 (Yin Yang 1) protein (79), a transcription repressor involved in cancer development and immune processes (80).

The influence on gene expression of the variants within *ANRIL* region have been analyzed in a variety of tissues and cells, but the results have been inconsistent (75). Several SNPs have been described to be involved in alternative splicing

events, and modifying *ANRIL* structure has been suggested to lead to changes in its function and consequent regulation of downstream inflammatory genes (**Figure 2D**) (22). Although the exact mechanism by which the SNPs within *ANRIL* confer susceptibility to disease has not been firmly established, it seems clear that the disease associated SNPs are related with *ANRIL* expression, suggesting that modulation of its expression mediates disease susceptibility.

One example of SNP-dependent *ANRIL* related inflammation is the correlation between the lead periodontitis associated SNP, rs1333048, and the levels of the C-reactive protein. Periodontitis is a complex, chronic inflammatory disease associated with increased concentration of high-sensitive C-reactive protein (hsCRP), a marker for systemic inflammation. It was found that AA-genotype of *ANRIL* rs1333048 is associated with significantly elevated hsCRP plasma levels in patients with periodontitis (81). However, the functional relationship between the SNP, *ANRIL* and the hsCRP molecule has not been clearly established.

Another disease in which *ANRIL* has been functionally implicated is Type 2 diabetes (T2D). Although the major causes of T2D are insulin resistance and beta cell dysfunction, recent evidence implicates the immune system in the pathogenesis of this disease that can be considered as an autoinflammatory disease (82). T2D associated SNPs in the *ANRIL* locus were evaluated, and it was observed that the risk genotype was correlated with increased levels of *ANRIL* expression. Moreover, although the associated SNPs did not seem to influence insulin secretion, it was observed that they affect human beta cell proliferation index, with homozygous risk alleles showing approximately half of the proliferation capacity observed in the presence of the protective alleles (83). Although this study suggests that ANRIL lncRNA may play a role in human islets and uncovers a link between T2D associated SNPs and beta cell proliferation, once again, the functional relationship between the SNPs and the biological process is still not understood.

Additionally, *ANRIL* is significantly downregulated in the inflamed intestinal mucosa of Crohn's and inflammatory bowel disease patients (84). At the same time its reduced levels in rats have been related to prevention of coronary atherosclerosis due to lower expression of inflammatory factors (85) which are upregulated in patients with coronary artery disease (86).

It therefore seems clear that disease associated SNPs in *ANRIL* lncRNA influence its function in the context of inflammatory diseases. However, the involvement of *ANRIL* in inflammation and the influence of the GWAS SNPs in the function of the different isoforms of *ANRIL* needs further investigation.

CONCLUDING REMARKS

Although our knowledge about the genetic variants contributing to immune mediated diseases has increased considerably in the last decade, the intergenic location of the great majority of the associated SNPs has made it difficult to decipher their functional roles in disease development. As disease associated SNPs are enriched within lncRNAs, and as many of these RNA molecules have been implicated in the regulation of inflammatory processes, a new field of study focused on the influence of disease-associated SNPs in the function of inflammation-related lncRNAs has been opened. Interestingly, such lncRNAs have been linked to major immune-mediated diseases as celiac disease, type 1 diabetes or rheumatoid arthritis. The experimental approaches utilized so far have been mainly focused on the expression analysis of the SNP harboring lncRNA in diseased tissues, but functional studies evaluating the contribution of each allele to lncRNA function, and thus to disease development, is mostly missing. In general, the function of the lncRNA itself, and the mechanisms by which they contribute to inflammatory disease development, are mostly uncharacterized. Analyzing the position and the linkage disequilibrium block of the associated SNP within the lncRNA sequence can help predict the functional impact of the allelic variant. Associated SNPs may not only affect the expression of the lncRNA itself, but also their splicing, their secondary structure or their ability to regulate expression of downstream genes. Thus, approaches that evaluate the functional differences of the lncRNA alleles are necessary in order to understand how the disease-associated SNPs affect the function of such inflammation related lncRNAs.

As our knowledge about the molecular mechanisms by which the inflammation related lncRNAs exert their biological functions increases, so will our understanding of how the disease associated SNPs influence lncRNA function thereby opening up the possibility for targeting such lncRNAs for diagnostic and therapeutic purposes.

AUTHOR CONTRIBUTIONS

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

FUNDING

AC-R is funded by an Ikerbasque Research Fellowship, SG is funded by grant R01 DK102180 (NIH).

REFERENCES

1. Cotsapas C, Hafler DA. Immune-mediated disease genetics: the shared basis of pathogenesis. *Trends Immunol.* (2013) 34:22–6. doi: 10.1016/J.IT.2012.09.001
2. Wu X, Chen H, Xu H. The genomic landscape of human immune-mediated diseases. *J Hum Genet.* (2015) 60:675–81. doi: 10.1038/jhg.2015.99
3. Diaz-Gallo L-M, Martin J. Common genes in autoimmune diseases: a link between immune-mediated diseases. *Expert Rev Clin Immunol.* (2012) 8:107–9. doi: 10.1586/eci.11.90
4. Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet.* (2013) 14:661–73. doi: 10.1038/nrg3502
5. Ricaño-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet.* (2013) 14:325–53. doi: 10.1146/annurev-genom-091212-153450
6. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science.* (2008) 322:881–8. doi: 10.1126/science.1156409

7. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. (2007) 447:661–78. doi: 10.1038/nature05911
8. Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. (2015) 518:337–43. doi: 10.1038/nature13835
9. Tak YG, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin*. (2015) 8:57. doi: 10.1186/s13072-015-0050-4
10. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*. (2009) 106:9362–7. doi: 10.1073/pnas.0903103106
11. Chen YG, Satpathy AT, Chang HY. Gene regulation in the immune system by long noncoding RNAs. *Nat Immunol*. (2017) 18:962–72. doi: 10.1038/ni.3771
12. Mowel WK, Kotzin JJ, McCright SJ, Neal VD, Henao-Mejia J. Control of immune cell homeostasis and function by lncRNAs. *Trends Immunol*. (2018) 39:55–69. doi: 10.1016/j.it.2017.08.009
13. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*. (2011) 43:904–14. doi: 10.1016/j.molcel.2011.08.018
14. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet*. (2013) 9:e1003569. doi: 10.1371/journal.pgen.1003569
15. Hrdlickova B, Kumar V, Kanduri K, Zhernakova DV, Tripathi S, Karjalainen J, et al. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. *Genome Med*. (2014) 6:88. doi: 10.1186/s13073-014-0088-0
16. Li X, Wu Z, Fu X, Han W. lncRNAs: Insights into their function and mechanics in underlying disorders. *Mutat Res Mutat Res*. (2014) 762:1–21. doi: 10.1016/j.mrrev.2014.04.002
17. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci*. (2013) 368:20120362. doi: 10.1098/rstb.2012.0362
18. Popadin K, Gutierrez-Arcelus M, Dermitzakis ET, Antonarakis SE. Genetic and epigenetic regulation of human lincRNA gene expression. *Am J Hum Genet*. (2013) 93:1015–26. doi: 10.1016/j.ajhg.2013.10.022
19. McDowell I, Pai A, Guo C, Vockley CM, Brown CD, Reddy TE, et al. Many long intergenic non-coding RNAs distally regulate mRNA gene expression levels. *bioRxiv*. (2016) 044719. doi: 10.1101/044719
20. Tan JY, Smith AAT, Ferreira da Silva M, Matthey-Doret C, Rueedi R, Sönmez R, et al. cis-Acting complex-trait-associated lincrna expression correlates with modulation of chromosomal architecture. *Cell Rep*. (2017) 18:2280–8. doi: 10.1016/J.CELREP.2017.02.009
21. Kumar V, Westra H-J, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet*. (2013) 9:e1003201. doi: 10.1371/journal.pgen.1003201
22. Aguilo F, Di Cecilia S, Walsh MJ. Long non-coding RNA ANRIL and polycomb in human cancers and cardiovascular disease. *Curr Top Microbiol Immunol*. (2016) 394:29–39. doi: 10.1007/82_2015_455
23. Chowdhury IH, Narra HP, Sahni A, Khanipov K, Schroeder CLC, Patel J, Fofanov et al. Expression profiling of long noncoding RNA splice variants in human microvascular endothelial cells: lipopolysaccharide effects *in vitro*. *Mediators Inflamm*. (2017) 2017:1–18. doi: 10.1155/2017/3427461
24. Gong J, Liu W, Zhang J, Miao X, Guo A-Y. lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res*. (2015) 43:D181–6. doi: 10.1093/nar/gku1000
25. Ren C, An G, Zhao C, Ouyang Z, Bo X, Shu W. Lnc2Catlas: an atlas of long noncoding RNAs associated with risk of cancers. *Sci Rep*. (2018) 8:1909. doi: 10.1038/s41598-018-20232-4
26. Halvorsen M, Martin JS, Broadaway S, Laederach A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet*. (2010) 6:e1001074. doi: 10.1371/journal.pgen.1001074
27. Mirza AH, Kaur S, Brorsson CA, Pociot F. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS ONE*. (2014) 9:e0105723. doi: 10.1371/journal.pone.0105723
28. Husby S, Koletzko S, Korponay-szabó IR, Mearin Ml, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr*. (2012) 54:136–60. doi: 10.1097/mpg.0b013e31821a23d0
29. Gutierrez-Achury J, Zhernakova A, Pulit SL, Trynka G, Hunt KA, Romanos J, et al. Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat Genet*. (2015) 47:577–8. doi: 10.1038/ng.3268
30. Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. *Hum Immunol*. (1992) 35:188–92. doi: 10.1016/0198-8859(92)90104-U
31. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med*. (1989) 169:345–50. doi: 10.1084/JEM.169.1.345
32. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat Genet*. (2007) 39:827–9. doi: 10.1038/ng2058
33. Hunt KA, Zhernakova A, Turner G, Heap GAR, Franke L, Bruinenberg M, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet*. (2008) 40:395–402. doi: 10.1038/ng.102
34. Dubois PCA, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet*. (2010) 42:295–302. doi: 10.1038/ng.543
35. Plaza-Izurieta L, Fernandez-Jimenez N, Irastorza I, Jauregi-Miguel A, Romero-Garmendia I, Vitoria JC, et al. Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur J Hum Genet*. (2015) 23:1100–5. doi: 10.1038/ejhg.2014.244
36. Santin I, Jauregi-miguel A, Velayos T, Castellanos-rubio A, Garcia-etxebarria K, Romero-garmendia I, et al. Celiac disease-associated lncRNA named HCG14 regulates NOD1 expression in intestinal cells. *J Pediatr Gastroenterol Nutr*. (2018) 67:225–31. doi: 10.1097/mpg.0000000000001970
37. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PHR, et al. A long noncoding RNA associated with susceptibility to celiac disease. *Science*. (2016) 352:91–5. doi: 10.1126/science.aad0467
38. Maiuri MC, De Stefano D, Mele G, Fecarotta S, Greco L, Troncone R, et al. Nuclear factor κB is activated in small intestinal mucosa of celiac patients. *J Mol Med*. (2003) 81:373–9. doi: 10.1007/s00109-003-0440-0
39. Fernandez-Jimenez N, Castellanos-Rubio A, Plaza-Izurieta L, Irastorza I, Elcoroaristizabal X, Jauregi-Miguel A, et al. Coregulation and modulation of NF B-related genes in celiac disease: uncovered aspects of gut mucosal inflammation. *Hum Mol Genet*. (2014) 23:1298–310. doi: 10.1093/hmg/ddt520
40. Heap GA, Trynka G, Jansen RC, Bruinenberg M, Swertz MA, Dinesen LC, et al. Complex nature of SNP genotype effects on gene expression in primary human leucocytes. *BMC Med Genomics*. (2009) 2:1. doi: 10.1186/1755-8794-2-1
41. Myhr CB, Hulme MA, Wasserfall CH, Hong PJ, Lakshmi PS, Schatz DA, et al. The autoimmune disease-associated SNP rs917997 of IL18RAP controls IFNγ production by PBMC. *J Autoimmun*. (2013) 44:8–12. doi: 10.1016/J.JAUT.2013.06.001
42. Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernández-Jimenez N, Gutierrez G, CEGEC, Bilbao JR. Revisiting genome wide association studies (GWAS) in coeliac disease: replication study in Spanish population and

expression analysis of candidate genes. *J Med Genet*. (2011) 48:493–6. doi: 10.1136/jmg.2011.089714
43. Sareneva T, Julkunen I, Matikainen S. IFN-alpha and IL-12 induce IL-18 receptor gene expression in human NK and T cells. *J Immunol*. (2000) 165:1933–8. Available online at: http://www.ncbi.nlm.nih.gov/pubmed/10925275 (accessed January 17, 2019).
44. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, Yang JHM, et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med*. (2008) 359:2767–77. doi: 10.1056/NEJMoa0807917
45. Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ, et al. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. *Am J Hum Genet*. (2008) 82:1202–10. doi: 10.1016/J.AJHG.2008.03.016
46. Coenen MJH, Trynka G, Heskamp S, Franke B, van Diemen CC, Smolonska J, et al. Common and different genetic background for rheumatoid arthritis and coeliac disease. *Hum Mol Genet*. (2009) 18:4195–203. doi: 10.1093/hmg/ddp365
47. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature*. (2011) 473:317–25. doi: 10.1038/nature10146
48. Ghattas A, Griffiths HR, Devitt A, Lip GYH, Shantsila E. Monocytes in coronary artery disease and atherosclerosis: where are we now? *J Am Coll Cardiol*. (2013) 62:1541–51. doi: 10.1016/J.JACC.2013.07.043
49. Holdt LM, Teupser D. From genotype to phenotype in human atherosclerosis–recent findings. *Curr Opin Lipidol*. (2013) 24:410–8. doi: 10.1097/MOL.0b013e3283654e7c
50. Zhang D-D, Wang W-T, Xiong J, Xie X-M, Cui S-S, Zhao Z-G, et al. Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF-κB pathway in human monocytes. *Sci Rep*. (2017) 7:46204. doi: 10.1038/srep46204
51. Owens GK, Kumar MS, Wamhoff BR. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. *Physiol Rev*. (2004) 84:767–801. doi: 10.1152/physrev.00041.2003
52. Zhang B-Y, Jin Z, Zhao Z. Long intergenic noncoding RNA 00305 sponges miR-136 to regulate the hypoxia induced apoptosis of vascular endothelial cells. *Biomed Pharmacother*. (2017) 94:238–43. doi: 10.1016/J.BIOPHA.2017.07.099
53. Zhang C, Kang K, Li X, Xie B. MicroRNA-136 promotes vascular muscle cell proliferation through the ERK1/2 pathway by targeting PPP2R2A in atherosclerosis. *Curr Vasc Pharmacol*. (2015) 13:405–12. Available onlinev at: http://www.ncbi.nlm.nih.gov/pubmed/25409743 (accessed May 18, 2018).
54. Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. (2012) 489:57–74. doi: 10.1038/nature11247
55. Pan X, Jiang X-C, Hussain MM. Impaired cholesterol metabolism and enhanced atherosclerosis in clock mutant mice. *Circulation*. (2013) 128:1758–69. doi: 10.1161/CIRCULATIONAHA.113.002885
56. Fusunyan RD, Sanderson IR. Inflammatory bowel disease. In: Peter J, editor. *Encyclopedia of Immunology*. London: Elsevier (1998). p. 1375–81. doi: 10.1006/rwei.1999.0351
57. Eckburg PB, Relman DA. The role of microbes in Crohn's disease. *Clin Infect Dis*. (2007) 44:256–62. doi: 10.1086/510385
58. Weinstock JV. Helminths and mucosal immune modulation. *Ann N Y Acad Sci*. (2006) 1072:356–64. doi: 10.1196/annals.1326.033
59. Ye BD, McGovern DPB. Genetic variation in IBD: progress, clues to pathogenesis and possible clinical utility. *Expert Rev Clin Immunol*. (2016) 12:1091–107. doi: 10.1080/1744666X.2016.1184972
60. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. (2012) 491:119–24. doi: 10.1038/nature11582
61. Padua D, Mahurkar-Joshi S, Law IKM, Polytarchou C, Vu JP, Pisegna JR, et al. A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation. *Am J Physiol Gastrointest Liver Physiol*. (2016) 311:G446–57. doi: 10.1152/ajpgi.00212.2016
62. Wu F, Huang Y, Dong F, Kwon JH. Ulcerative colitis-associated long noncoding RNA, BC012900, regulates intestinal epithelial cell apoptosis. *Inflamm Bowel Dis*. (2016) 22:782–95. doi: 10.1097/MIB.0000000000000691
63. Peng H, Liu Y, Tian J, Ma J, Tang X, Rui K, et al. The long noncoding RNA IFNG-AS1 promotes T helper type 1 cells response in patients with Hashimoto's thyroiditis. *Sci Rep*. (2015) 5:17702. doi: 10.1038/srep17702
64. Wang J, Peng H, Tian J, Ma J, Tang X, Rui K, et al. Upregulation of long noncoding RNA TMEVPG1 enhances T helper type 1 cell response in patients with Sjögren syndrome. *Immunol Res*. (2016) 64:489–96. doi: 10.1007/s12026-015-8715-4
65. Vigneau S, Rohrlich P-S, Brahic M, Bureau J-F. Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *J Virol*. (2003) 77:5632–8. doi: 10.1128/JVI.77.10.5632-5638.2003
66. Collier SP, Collins PL, Williams CL, Boothby MR, Aune TM. Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *J Immunol*. (2012) 189:2084–8. doi: 10.4049/jimmunol.1200774
67. Collier SP, Henderson MA, Tossberg JT, Aune TM. Regulation of the Th1 genomic locus from Ifng through Tmevpg1 by T-bet. *J Immunol*. (2014) 193:3959–65. doi: 10.4049/jimmunol.1401099
68. Gomez JA, Wapinski OL, Yang YW, Bureau J-FO, Gopinath S, Monack DM, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-γ locus. *Cell*. (2013) 152:743–54. doi: 10.1016/j.cell.2013.01.015
69. Agostini F, Zanzoni A, Klus P, Marchese D, Cirillo D, Tartaglia GG. catRAPID omics: a web server for large-scale prediction of protein-RNA interactions. *Bioinformatics*. (2013) 29:2928–30. doi: 10.1093/bioinformatics/btt495
70. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. (2012) 40:D930–4. doi: 10.1093/nar/gkr917
71. Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res*. (2007) 67:3963–9. doi: 10.1158/0008-5472.CAN-06-2004
72. Li Z, Yu X, Shen J. ANRIL: a pivotal tumor suppressor long non-coding RNA in human cancers. *Tumor Biol*. (2016) 37:5657–61. doi: 10.1007/s13277-016-4808-5
73. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. (2007) 316:1488–91. doi: 10.1126/science.1142447
74. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. (2007) 316:1336–41. doi: 10.1126/science.1142364
75. Aarabi G, Zeller T, Heydecke G, Munz M, Schäfer A, Seedorf U. Roles of the Chr.9p21.3 ANRIL locus in regulating inflammation and implications for anti-inflammatory drug target identification. *Front Cardiovasc Med*. (2018) 5:47. doi: 10.3389/fcvm.2018.00047
76. Yap KL, Li S, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell*. (2010) 38:662–74. doi: 10.1016/j.molcel.2010.03.021
77. Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. *Cell*. (2006) 127:265–75. doi: 10.1016/j.cell.2006.10.003
78. Murakami Y, Mizoguchi F, Saito T, Miyasaka N, Kohsaka H. p16INK4a exerts an anti-inflammatory effect through accelerated IRAK1 degradation in macrophages. *J Immunol*. (2012) 189:5066–72. doi: 10.4049/jimmunol.1103156
79. Zhou X, Han X, Wittfeldt A, Sun J, Liu C, Wang X, et al. Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-κB pathway. *RNA Biol*. (2016) 13:98–108. doi: 10.1080/15476286.2015.1122164

80. Lin J, He Y, Chen J, Zeng Z, Yang B, Ou Q. A critical role of transcription factor YY1 in rheumatoid arthritis by regulation of interleukin-6. *J Autoimmun.* (2017) 77:67–75. doi: 10.1016/J.JAUT.2016.10.008
81. Teeuw WJ, Laine ML, Bizzarro S, Loos BG. A lead ANRIL polymorphism is associated with elevated CRP levels in periodontitis: a pilot case-control study. *PLoS ONE.* (2015) 10:e0137335. doi: 10.1371/journal.pone.0137335
82. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. *Nat Rev Immunol.* (2011) 11:98–107. doi: 10.1038/nri2925
83. Kong Y, Sharma RB, Ly S, Stamateris RE, Jesdale WM, Alonso LC. *CDKN2A/B* T2D genome-wide association study risk SNPs impact locus gene expression and proliferation in human islets. *Diabetes.* (2018) 67:872–84. doi: 10.2337/db17-1055
84. Mirza AH, Berthelsen CH, Seemann SE, Pan X, Frederiksen KS, Vilien M, et al. Transcriptomic landscape of lncRNAs in inflammatory bowel disease. *Genome Med.* (2015) 7:39. doi: 10.1186/s13073-015-0162-2
85. Song C-L, Wang J-P, Xue X, Liu N, Zhang X-H, Zhao Z, et al. Effect of circular ANRIL on the inflammatory response of vascular endothelial cells in a rat model of coronary atherosclerosis. *Cell Physiol Biochem.* (2017) 42:1202–12. doi: 10.1159/000478918
86. Haemmig S, Simion V, Feinberg MW. Long non-coding RNAs in vascular inflammation. *Front Cardiovasc Med.* (2018) 5:22. doi: 10.3389/fcvm.2018.00022

Conflict of Interest Statement: 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.

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