A New View of Genome Organization Through RNA Directed Interactions

Gabriel Khelifi [1,2,3] and Samer M. I. Hussein [1,2,3]*

[1] Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, QC, Canada, [2] Université Laval Cancer Research Center, Université Laval, Québec, QC, Canada, [3] Oncology Division, Centre Hospitalier Universitaire (CHU) de Québec-Université Laval Research Center, Québec, QC, Canada

INTRODUCTION

Over the past few decades, we have come to appreciate the complexity of processes regulating chromatin architecture. Ranging from chromatin accessibility (Klemm et al., 2019) to long-range genome organization (Cremer and Cremer, 2001), a wide array of mechanisms is used by the cell to control gene expression. We have recently begun to understand the primordial role of non-coding RNAs (ncRNAs) in such processes and can now declare that several types of RNAs are essential to the regulation of gene expression (Cech and Steitz, 2014). In the nucleus, long non-coding RNAs (lncRNAs) (Bonasio and Shiekhattar, 2014), together with enhancer RNAs (eRNAs) (Li et al., 2016a), stable intronic sequence RNAs (sisRNAs) (Chan and Pek, 2019) and various other classes of transcripts (reviewed in (Li and Fu, 2019)) come together to ensure tight regulation of the chromatin. LncRNAs represent transcripts of more than 200 nucleotides that do not contain any apparent open reading frame (Marchese et al., 2017). While some lncRNAs are localized and active in the cytosol (Noh et al., 2018), many are nuclear and implicated in transcriptional regulation (Vance and Ponting, 2014). These nuclear lncRNAs can modulate the expression of genes through interactions with DNA or chromatin-associated proteins (Bonasio and Shiekhattar, 2014; **Figure 1A**). eRNAs are abundantly transcribed RNAs generated from enhancer regions (Li et al., 2016a). They modulate enhancer activity potentially through interactions with the mediator complex, transcription factors or chromosomal looping factors. Lastly, sisRNAs represent RNAs containing intronic sequences, and increasing evidence shows that several of them act on chromatin regulation (Chan and Pek, 2019). All of these various classes of chromatin-associated RNAs are essential to the regulation of gene expression (Li and Fu, 2019). Some transcripts appear to be "*cis*-acting," influencing the expression of genes within their own chromosome, while others control transcriptional processes on other chromosomes in *trans*. Some *cis*-acting RNAs have been shown to function through the formation of R-loops with the complementary sequence from their transcribed loci and affect local gene expression, as is the case with *GATA3-AS1* and *VIM-AS1* (Boque-Sastre et al., 2015; Gibbons et al., 2018). Both *cis*- and *trans*-acting RNAs can affect gene expression through direct RNA-DNA contacts (i.e., RNA-DNA triplexes), as employed by *MEG3* and *KHPS1* (Mondal et al., 2015; Blank-Giwojna et al., 2019), or indirectly through protein intermediates such as the interaction of *FIRRE* lncRNA with the SAF-A protein (Hacisuleyman et al., 2014; **Figure 1A**). These interactions often require specific RNA "domains," such as motifs recognized by chromatin bound proteins, or RNA-DNA triplex-forming motifs such as polypurine tracts (Li et al., 2016b; Li and Fu, 2019). RNAs exhibiting interactions in both *cis* and *trans* have also

A RNA interactions with chromatin

B Type I RNA-chromatin interactions

C Type II RNA-chromatin interactions

D

Technique	Full Name	Highlights of the technique	Reference
Detection of RNA-DNA Interactions			
MARGI	MApping of RNA–Genome Interactome	- In solution proximity ligation - High number of input cells needed - 100bp reads	Sridhar et al., 2017
GRID-seq	Global RNA Interactions with DNA	- *In-situ* proximity ligation - Low number of input cells needed - 18 to 23bp reads	Li et al., 2017
CHAR-seq	Chromatin-Associated RNA sequencing	- In solution proximity ligation - High number of input cells needed - ~65bp reads	Bell et al., 2018
iMARGI	in-situ MApping of RNA–Genome Interactome	- *In-situ* proximity ligation - Low number of input cells needed - 100bp reads	Wu et al., 2019
RADICL-seq	RNA And DNA Interacting Complexes Ligated	- *In-situ* proximity ligation - Low number of input cells needed - 27bp reads	Bonetti et al., 2020
SPRITE	Split-Pool Recognition of Interactions by Tag Extension	- Protocol for mapping long range DNA interactions - Limited ability to detect RNA-DNA interactions - Proximity-ligation-independent	Quinodoz et al., 2018
Detection of RNA-RNA Interactions			
PARIS	Psoralen Analysis of RNA Interactions and Structures	- AMT-based UV crosslinking - 2D gel clean-up	Lu et al., 2016
LIGR-seq	LIGation of interacting RNA	- AMT-based UV crosslinking - RNase R digestion	Sharma et al., 2016
SPLASH	Sequencing of Psoralen crosslinked, Ligated And Selected Hybrids	- Biotinylated-psoralen-based UV crosslinking - Streptavidn beads clean-up	Aw et al., 2016
MARIO	Mapping RNA-RNA Interactions in-vivO	- UV crosslinking of RNAs to proteins - Biotinylation of total proteins - Immobilization of protein-RNA complexes on Streptavidin beads	Nguyen et al., 2016
COMRADES	CrOss-linking of Matched RNAs And DEep Sequencing	- Psoralen-triethylene glycol azide-based UV crosslinking - Pulldown of target RNA - Biotin-click reaction and enrichment of cross-linked RNAs	Ziv et al., 2018

FIGURE 1 | (A) Potential means of chromatin-RNA association. Interactions can be mediated by protein intermediates or rely on either Hoogsteen base-pairing to form RNA-DNA triplexes, or Watson-Crick base-pairing to form RNA-DNA duplexes such as R-loops. **(B)** Representation of potential type I RNA-chromatin

(Continued)

FIGURE 1 | interactions, where RNAs recruit associated proteins to the chromatin, enabling the modulation of genomic accessibility. For example, recruited proteins can be part of chromatin modifying complexes or act as transcription factors. **(C)** Representation of potential type II RNA-chromatin interactions, where RNAs interact with genetically distant loci and modulate genome architecture, potentially bringing them physically closer. Interactions can be direct associations of the RNA with the DNA, or be mediated either by chromatin-bound proteins, or RNA duplexing. RNA bound protein complex intermediates can also generate indirect interactions. **(D)** Table of referenced techniques for identification of genome and transcriptome-wide DNA-RNA or RNA-RNA interactions, respectively. For more details, we refer the reader to the corresponding publications within the table. AMT, 4′-aminomethyl-4,5′,8-trimethylpsoralen.

been described, for example: the lncRNAs *FIRRE* (Hacisuleyman et al., 2014; Lewandowski et al., 2019) and *ANRIL* (Kong et al., 2018). This hints at the great complexity of processes regulated by chromatin-associated RNAs.

We postulate that RNAs acting on chromatin can be functionally separated in two groups, following how they affect the transcriptional landscape of the cell. First, RNAs can act locally to where they bind (i.e., short-range) on the structure of the chromatin itself, for example, by modifying its accessibility through the recruitment of structural protein factors or protein complexes that establish chromatin marks (Type I) (**Figure 1B**). Various Type I RNAs have been characterized. For example, the lncRNAs *HOTAIR* and *FENDERR* and the sisRNA generated from intronic sequences of *SMYD3*, do so in part by recruiting the PRC2 repressive complex to their genomic binding sites (Guil et al., 2012; Grote et al., 2013; Mozdarani et al., 2020). Second, RNAs can control the organization of the genome, which we define as Type II interactions. They do so by promoting long-range chromatin interactions and the bridging of distant genomic loci (**Figure 1C**). Type II RNAs include examples such as the lncRNAs *LUNAR1* or *Kancr*, both involved in chromosomal looping and activation of genes near the loop anchor points (Trimarchi et al., 2014; Li et al., 2020). Additionally, some lncRNAs can potentially exhibit both types of interactions. For example, the lncRNA *XIST* is transcribed from the X-chromosome and is directly implicated in its inactivation (XCI) (Loda and Heard, 2019). During this process, *XIST* transcripts recruit protein factors necessary for XCI, highlighting a Type I mechanism (Chu et al., 2015). Upon XCI, progressive compaction of the X-chromosome occurs and allows *XIST* to spread on the whole X-chromosome, which results in global heterochromatinization (Engreitz et al., 2013). Moreover, the multiple domains which enable *XIST* to attach to X-chromosome-bound proteins and the *XIST*-dependent X-chromosome conformation, hint at a potential Type II mechanism for this lncRNA in XCI.

While a growing number of RNAs acting on gene expression are being characterized, we still lack a bigger picture on how prevalent these interactions are, and on the importance of the proposed mechanisms. The first major obstacle in delineating which RNAs act on chromatin regulation resides in the immense number of existing transcripts. For example, estimates place the number of potentially functional lncRNAs in the tens of thousands (Marchese et al., 2017), and characterization of functional eRNAs or sisRNAs is still too early to grant accurate estimates (Li et al., 2016a; Chan and Pek, 2019). A tremendous amount of work lies ahead to fully understand the above-mentioned processes. Understanding where and how

RNAs interact with chromatin, and the resulting effect on gene regulation, therefore remains an upcoming challenge in the characterization of the global transcriptional landscape.

GENOME-WIDE CATALOGING OF RNA-CHROMATIN INTERACTIONS

The past 20 years have been marked by the development of several techniques aiming to map long-range chromatin interactions and to decipher genome architecture. Efficient techniques now include direct ligation of proximal DNA fragments (3C) (Kempfer and Pombo, 2020), ligation of barcodes to interacting DNA fragments (SPRITE) (Quinodoz et al., 2018), or physical isolation of thin nuclear sections and analysis of the DNA contained within (GAM) (Beagrie et al., 2017). The 3C-based methodologies have inspired new protocols to resolve RNA-DNA interactions, such as MARGI (Sridhar et al., 2017), GRID-seq (Li et al., 2017), CHAR-seq (Bell et al., 2018), iMARGI (Wu et al., 2019), and RADICL-seq (Bonetti et al., 2020). Moreover, SPRITE was also adapted to reveal RNA-DNA interactions (Quinodoz et al., 2018; **Figure 1D**). Therefore, a vast number of datasets showing RNA-chromatin interactions now exists for diverse cellular contexts. MARGI has already enabled the identification of Type I mechanisms implicating RNA-chromatin interactions in the establishment of various chromatin marks (Sridhar et al., 2017). Additionally, GRID-seq showed that an enrichment of RNA-chromatin interactions is implicated in the role of super-enhancers (Li et al., 2017), representing potential Types I and II RNA-chromatin interactions. GRID-seq was also used to detect the diverse genomic binding sites of *XIST*, both locally to its transcription site, but also throughout the X chromosome. It also highlighted the sites of initiation of XCI via *XIST*. These findings establish the major importance of such high-throughput techniques in characterizing different mechanisms responsible for the tight control of chromatin by RNAs.

LIMITATIONS OF CURRENT METHODOLOGIES FOR PROBING RNA-CHROMATIN INTERACTIONS

However, limitations are present in the developed techniques, and many challenges still lie ahead to fully understand how chromatin is regulated by RNAs. First, the developed techniques generally exhibit a high prevalence for RNA reads that map to introns, revealing a widespread capture of nascent RNAs from loci undergoing transcription (actively transcribed

RNAs) (Li et al., 2017; Bonetti et al., 2020). During their transcription process, RNAs can be captured via fixation to their genomic loci, therefore appearing as mapped interactions. These highly abundant interactions represent a major contaminant of such experiments, potentially overshadowing actual functional interactions. Their presence should be accounted for to prevent them from affecting the sequencing depth achieved for functional interactions and to allow for the detection of lower-abundance RNA-chromatin interactions. To reduce the effect of nascent RNA-bias, RADICL-seq uses a controlled RNase H digestion step (Bonetti et al., 2020). Although this step does indeed reduce this bias, a significant portion of nascent RNA still remains, and it would be necessary to explore additional ways to remove these unwanted RNAs.

Another main limitation in the current protocols and available datasets resides in the relatively small size of the sequenced tags corresponding to the DNA and its interacting RNA (Li et al., 2017; Sridhar et al., 2017; Bell et al., 2018; Quinodoz et al., 2018; Wu et al., 2019; Bonetti et al., 2020; **Figure 1D**). Indeed, their short size results in poor mapping of the obtained DNA-RNA pairs to the genome and transcriptome. This problem is further amplified when working with transcripts containing repeated sequences. For example, lncRNAs generally possess a large number of transposable elements (TEs) and other repeated sequences (Johnson and Guigó, 2014), and 83% of lncRNAs contain one or more TEs, compared to only 6% of mRNAs (Kelley and Rinn, 2012). Interestingly, some TEs are implicated in localization of RNA transcripts to the nucleus (Lubelsky and Ulitsky, 2018), which further highlights a role in genome regulation. These TEs have also emerged as important domains in lncRNA function, with several cases demonstrating that an embedded TE acts as the RNA's functional motif (Johnson and Guigó, 2014). These TEs can enable interactions with complementary sequences in DNA or other RNAs, as seems to be the case with the lncRNAs *ANRIL* (Holdt et al., 2013) and *LEADeR* (Profumo et al., 2019). Therefore, a loss in reads corresponding to repeat elements due to an inability to map them represents a major potential hurdle in the current protocols. This is evident in the relatively low percentage of uniquely mapped reads for protocols such as GRID-seq or RADICL-seq (14 and 45%, respectively) (Li et al., 2017; Bonetti et al., 2020). The higher read mapping observed with RADICL-seq compared to GRID-seq is due to an additional 7 base-pairs in the final read length (Bonetti et al., 2020). These protocols use restriction enzymes to generate sequence lengths of around 20 base-pairs for GRID-seq and 27 for RADICL-seq for both RNA and DNA tags (**Figure 1D**). The increase in read length with RADICL-seq allows mapping of some reads corresponding to TE-containing RNAs. This enabled the authors to determine that transcripts containing TEs are indeed differentially engaged in interactions with chromatin, once again hinting at the importance of TEs in chromatin regulation. Other methodologies, namely MARGI (Sridhar et al., 2017) and iMARGI (Wu et al., 2019), circumvent this limitation through a protocol that preserves the full length of the respective RNA and DNA tags. This enables the generation of libraries containing longer fragments and results in higher mapping of reads.

Thus far, these techniques have relied on "short-read" sequencing technologies, which do not fully overcome the challenge of mapping repeated and complex sequences within RNAs (Dijk et al., 2018). To remediate this problem, one option would be to incorporate "long-read" sequencing technologies to these protocols. Indeed, throughout the last 10 years, the rise in the availability of such sequencing techniques has meant that more and more laboratories can get access to this technology. While it is still in its early stages compared to "short-read" sequencing approaches, the sequencing of long DNA or RNA fragments, ranging from a few hundred nucleotides to tens of kilobases, results in a more accurate alignment of repeat sequences (Dijk et al., 2018). A higher percentage of DNA and RNA pairs should therefore be uniquely mapped, even in the presence of interactions dependent on the complementarity of repeat elements. Additionally, long-read sequencing could provide more detailed information on the RNAs interacting with chromatin. For one, the specific isoforms of RNAs which interact with chromatin could be detected. Also, detection of nascent RNAs would be more precise, as current analysis only take intronic reads into account when counting for nascent RNAs, whereas long-read sequencing will reveal the whole transcript. Therefore, long-read sequencing represents a very promising new tool for further iterations of such protocols.

RNA-RNA INTERACTIONS TO HELP UNDERSTAND LNCRNA-CHROMATIN INTERACTIONS

While it is clear now that chromatin-associated RNAs affect the structure and regulation of chromatin, the role of RNA-RNA interactions (RRIs) is not as well-explored in this context. Due to RNA's inherent ability to base pair and form complex higher order structures, it can simultaneously interact with DNA and multiple RNA and protein molecules (Lu and Chang, 2018). Inter-RNA interactions could enable Type II interactions, through base-pairing between two chromatin-associated RNAs (**Figure 1C**). Meanwhile, intra-RNA interactions, through the RNA's secondary structure, may help in identifying sites available for binding to chromatin or RNA-binding proteins. With these features in mind, we expect that the associations revealed by the techniques probing RNA-chromatin interactions will serve as starting points for integration of RRI networks involved in chromatin regulation. Various protocols recently aimed to investigate RRIs on a transcriptome-wide scale, such as PARIS (Lu et al., 2016), LIGR-seq (Sharma et al., 2016), SPLASH (Aw et al., 2016), and MARIO (Nguyen et al., 2016). Additionally, another technique, COMRADES (Ziv et al., 2018), initially intended to probe RRIs for a single RNA, could also be used for genome-wide RRIs (**Figure 1D**). These protocols crosslink RNA duplexes, to reveal both RRIs and, to some extent, the secondary structure of every RNA. However, the limitations highlighted for RNA-DNA probing techniques, such as read length and unique mapping of repeated sequences, still apply to RRI-probing protocols. These techniques have nevertheless proven to be instrumental in revealing several cellular processes dependent on

RRIs. For example, PARIS highlighted structural folding patterns in the *XIST* A-repeats, which is necessary for binding of SPEN (Lu et al., 2016), a transcriptional repressor involved in X-inactivation (Chu et al., 2015). Analysis of RRIs in combination with RNA localization of *XIST* on the chromatin by GRID-seq (Li et al., 2017) consequently highlights the potential Type I and II mechanisms of *XIST* in XCI. This example demonstrates how combining RRI data to an integrative map of the genome organization extracted from 3C or other related techniques, coupled with existing RNA-chromatin interaction information, will provide a better understanding of the complex mechanisms behind chromatin regulation. Long-range, indirect chromatin interactions mediated by several duplexed RNAs, or by protein complexes exhibiting RNA-binding functions (**Figure 1C**) will only then become more apparent. Overall, these types of studies provide a more complete view on the complexity of genome organization and chromatin structure.

CONCLUDING REMARKS

Taken together, the methods described here represent useful tools for elucidating the role of RNA-DNA and RNA-RNA interactions in gene expression regulation. While various improvements are still needed, the existing datasets represent a comprehensive look of how a genome could be organized through RNA interactions. Several RNA-DNA interactions have now been cataloged either through these genome-wide techniques or through specific RNA directed techniques. Of these, lncRNAs seem to represent an important fraction of the factors that regulate gene expression, chromatin accessibility and genome organization. In addition, these RNAs are not limited only to chromatin binding but may act as conduits to bring in other types of interactors, such as other RNAs, RNA-binding proteins, and transcriptional complexes. All these elements combined together help forge the transcriptional landscape necessary to maintain and transition between defined cell states.

AUTHOR CONTRIBUTIONS

GK and SH conceived and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Aw, J. G. A., Shen, Y., Wilm, A., Sun, M., Lim, X. N., Boon, K.-L., et al. (2016). *In vivo* mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* 62, 603–617. doi: 10.1016/j.molcel.2016.04.028

Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519–524. doi: 10.1038/nature21411

Bell, J. C., Jukam, D., Teran, N. A., Risca, V. I., Smith, O. K., Johnson, W. L., et al. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. *Elife* 7:e27024. doi: 10.7554/eLife.27024.037

Blank-Giwojna, A., Postepska-Igielska, A., and Grummt, I. (2019). lncRNA KHPS1 activates a poised enhancer by triplex-dependent recruitment of epigenomic regulators. *Cell Rep.* 26, 2904.e4–2915.e4. doi: 10.1016/j.celrep.2019.02.059

Bonasio, R., and Shiekhattar, R. (2014). Regulation of Transcription by Long Noncoding RNAs. *Annu. Rev. Genet.* 48, 1–23. doi: 10.1146/annurev-genet-120213-092323

Bonetti, A., Agostini, F., Suzuki, A. M., Hashimoto, K., Pascarella, G., Gimenez, J., et al. (2020). RADICL-seq identifies general and cell type–specific principles of genome-wide RNA-chromatin interactions. *Nat. Commun.* 11:1018. doi: 10.1038/s41467-020-14337-6

Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S., et al. (2015). Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5785–5790. doi: 10.1073/pnas.1421197112

Cech, T. R., and Steitz, J. A. (2014). The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* 157, 77–94. doi: 10.1016/j.cell.2014.03.008

Chan, S. N., and Pek, J. W. (2019). Stable intronic sequence RNAs (sisRNAs): an expanding universe. *Trends Biochem. Sci.* 44, 258–272. doi: 10.1016/j.tibs.2018.09.016

Chu, C., Zhang, Q. C., da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., et al. (2015). Systematic discovery of xist RNA binding proteins. *Cell* 161, 404–416. doi: 10.1016/j.cell.2015.03.025

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301. doi: 10.1038/35066075

Dijk, E. L., van, Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. *Trends Genet* 34, 666–681. doi: 10.1016/j.tig.2018.05.008

Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., et al. (2013). The xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341:1237973. doi: 10.1126/science.1237973

Gibbons, H. R., Shaginurova, G., Kim, L. C., Chapman, N., Spurlock, C. F., and Aune, T. M. (2018). Divergent incRNA GATA3-AS1 regulates GATA3 transcription in T-Helper 2 cells. *Front. Immunol.* 9:2512. doi: 10.3389/fimmu.2018.02512

Grote, P., Wittler, L., Hendrix, D., Koch, F., Währisch, S., Beisaw, A., et al. (2013). The tissue-specific lncRNA fendr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* 24, 206–214. doi: 10.1016/j.devcel.2012.12.012

Guil, S., Soler, M., Portela, A., Carrère, J., Fonalleras, E., Gómez, A., et al. (2012). Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat. Struct. Mol. Biol.* 19, 664–670. doi: 10.1038/nsmb.2315

Hacisuleyman, E., Goff, L. A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., et al. (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206. doi: 10.1038/nsmb.2764

Holdt, L. M., Hoffmann, S., Sass, K., Langenberger, D., Scholz, M., Krohn, K., et al. (2013). Alu Elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet.* 9:e1003588. doi: 10.1371/journal.pgen.1003588

Johnson, R., and Guigó, R. (2014). The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA* 20, 959–976. doi: 10.1261/rna.044560.114

Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 13:R107. doi: 10.1186/gb-2012-13-11-r107

Kempfer, R., and Pombo, A. (2020). Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* 21, 207–226. doi: 10.1038/s41576-019-0195-2

Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-8

Kong, Y., Hsieh, C.-H., and Alonso, L. C. (2018). ANRIL: a lncRNA at the CDKN2A/B locus with roles in cancer and metabolic disease. *Front. Endocrinol.* 9:405. doi: 10.3389/fendo.2018.00405

Lewandowski, J. P., Lee, J. C., Hwang, T., Sunwoo, H., Goldstein, J. M., Groff, A. F., et al. (2019). The firre locus produces a trans-acting RNA molecule that functions in hematopoiesis. *Nat. Commun.* 10:5137. doi: 10.1038/s41467-019-12970-4

Li, W., Notani, D., and Rosenfeld, M. G. (2016a). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* 17, 207–223. doi: 10.1038/nrg.2016.4

Li, W., Shen, W., Zhang, B., Tian, K., Li, Y., Mu, L., et al. (2020). Long non-coding RNA LncKdm2b regulates cortical neuronal differentiation by cis-activating Kdm2b. *Protein Cell* 11, 161–186. doi: 10.1007/s13238-019-0650-z

Li, X., and Fu, X.-D. (2019). Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.* 20, 503–519. doi: 10.1038/s41576-019-0135-1

Li, X., Zhou, B., Chen, L., Gou, L.-T., Li, H., and Fu, X.-D. (2017). GRID-seq reveals the global RNA–chromatin interactome. *Nat. Biotechnol.* 35:nbt3968. doi: 10.1038/nbt.3968

Li, Y., Syed, J., and Sugiyama, H. (2016b). RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem. Biol.* 23, 1325–1333. doi: 10.1016/j.chembiol.2016.09.011

Loda, A., and Heard, E. (2019). Xist RNA in action: past, present, and future. *PLoS Genet.* 15:e1008333. doi: 10.1371/journal.pgen.1008333

Lu, Z., and Chang, H. Y. (2018). The RNA base-pairing problem and base-pairing solutions. *Cold Perspect Biol.* 10:a034926. doi: 10.1101/cshperspect.a034926

Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., et al. (2016). RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* 165, 1267–1279. doi: 10.1016/j.cell.2016.04.028

Lubelsky, Y., and Ulitsky, I. (2018). Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 555, 107–111. doi: 10.1038/nature25757

Marchese, F. P., Raimondi, I., and Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* 18:206. doi: 10.1186/s13059-017-1348-2

Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., et al. (2015). MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* 6:7743. doi: 10.1038/ncomms8743

Mozdarani, H., Ezzatizadeh, V., and Parvaneh, R. R. (2020). The emerging role of the long non-coding RNA HOTAIR in breast cancer development and treatment. *J. Transl. Med.* 18:152. doi: 10.1186/s12967-020-02320-0

Nguyen, T. C., Cao, X., Yu, P., Xiao, S., Lu, J., Biase, F. H., et al. (2016). Mapping RNA–RNA interactome and RNA structure *in vivo* by MARIO. *Nat. Commun.* 7:12023. doi: 10.1038/ncomms12023

Noh, J. H., Kim, K. M., McClusky, W. G., Abdelmohsen, K., and Gorospe, M. (2018). Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip. Rev. RNA* 9:e1471. doi: 10.1002/wrna.1471

Profumo, V., Forte, B., Percio, S., Rotundo, F., Doldi, V., Ferrari, E., et al. (2019). LEADeR role of miR-205 host gene as long noncoding RNA in prostate basal cell differentiation. *Nat. Commun.* 10:307. doi: 10.1038/s41467-018-08153-2

Quinodoz, S. A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J. M., Detmar, E., et al. (2018). Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 174, 744.e24–757.e24. doi: 10.1016/j.cell.2018.05.024

Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B. J. (2016). Global mapping of human RNA-RNA interactions. *Mol. Cell* 62, 618–626. doi: 10.1016/j.molcel.2016.04.030

Sridhar, B., Rivas-Astroza, M., Nguyen, T. C., Chen, W., Yan, Z., Cao, X., et al. (2017). Systematic mapping of RNA-chromatin interactions *In vivo*. *Curr. Biol.* 27, 602–609. doi: 10.1016/j.cub.2017.01.011

Trimarchi, T., Bilal, E., Ntziachristos, P., Fabbri, G., Dalla-Favera, R., Tsirigos, A., et al. (2014). Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell* 158, 593–606. doi: 10.1016/j.cell.2014.05.049

Vance, K. W., and Ponting, C. P. (2014). Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet* 30, 348–355. doi: 10.1016/j.tig.2014.06.001

Wu, W., Yan, Z., Nguyen, T. C., Chen, Z. B., Chien, S., and Zhong, S. (2019). Mapping RNA–chromatin interactions by sequencing with iMARGI. *Nat. Protoc.* 14, 3243–3272. doi: 10.1038/s41596-019-0229-4

Ziv, O., Gabryelska, M. M., Lun, A. T. L., Gebert, L. F. R., Sheu-Gruttadauria, J., Meredith, L. W., et al. (2018). COMRADES determines *in vivo* RNA structures and interactions. *Nat. Methods* 15, 785–788. doi: 10.1038/s41592-018-0121-0

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.

The handling editor declared a shared affiliation with the authors though no other collaboration at time of review.