Commentary

Stem Cell INVESTIGATION

Divergent lncRNAs take the lead on pluripotent cell differentiation

Alessandro Rosa, Monica Ballarino

Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy

Correspondence to: Alessandro Rosa. Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy. Email: alessandro.rosa@uniroma1.it; Monica Ballarino. Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy. Email: monica.ballarino@uniroma1.it.

Received: 27 July 2016; Accepted: 07 September 2016; Published: 19 September 2016.

doi: 10.21037/sci.2016.09.02

View this article at: http://dx.doi.org/10.21037/sci.2016.09.02

A better understanding of the overall transcriptional landscape of human cells has been possible due to the ENCODE and FANTOM projects (1-3). It emerged that mammalian genomes are more intricate than previously suspected and produce a lattice of transcripts, among which only 2% encodes for proteins (4,5). Although alternative splicing and the presence of different Transcription Start Sites can concur to this complexity, a huge part of it can be explained by the existence of short (<30 nt) and long (>200 nt) non-coding RNAs (ncRNAs) which exert their roles without being translated into proteins (6,7). These include the long non-coding RNAs (lncRNAs), which represent a large and diverse class of RNA polymerase II transcripts longer than 200 nucleotides and act as fine-tuners of gene expression by a range of mechanisms (8). LncRNAs have histone-modification profiles, splicing signals, and exon/intron lengths akin to protein-coding genes. However, despite these similarities, lncRNAs are low expressed, preponderant in the nucleus and highly tissue-specific suggesting potential roles in specifying cell identity (9-11) and, when deregulated, in disease (12-14). LncRNAs mode of action can be hinted by their nuclear or cytoplasmic localization due to their unique ability to base-pair, which result in protein coordination and RNA interaction (15). When they are nuclear, lncRNAs can reshape the chromatin by acting in cis (at nearby regions) or in trans (on distant loci). This regulatory potential, in combination with the tissue specificity of lncRNAs, suggests that they can be an active component of a broad epigenetic regulatory network. Indeed, lncRNAs can scaffold distinct histone modification complexes to coordinate discrete functions on specific genomic loci (16-19). Several lines of evidence also suggest that lncRNA-genome interactions can contribute in the organization of three-dimensional nuclear structures (20,21). This is the case of the enhancer RNAs (eRNA), a recent class of lncRNAs which downregulation lead to a reduction in mRNA levels from specific neighbouring genes (22) and a loss of specific enhancer-promoter contacts (23).

Based on their closeness to protein-coding genes, lncRNAs can be classified into (I) sense or antisense, where both overlap another transcript but differ in the selected strand; (II) bidirectional, when its transcript expression starts in close proximity of the neighboring gene; (III) intronic, when their transcription begins from an intron of another transcript; (IV) intergenic, when lncRNA is created between two protein coding genes of an independent transcription unit (24). Although this classification is systematic and limited to the involvement of varied molecular mechanisms in lncRNA expression, it does not provide further information about their modes of action nor for their cellular functions. Evidence suggests that lncRNAs are positively correlated with the expression of the antisense coding genes (9).

Luo and colleagues focused on a particular aspect of lncRNA regulation, which is their ability to act as epigenetic controllers (25). The authors determined the genomic distribution of human and mouse lncRNA genes relative to protein coding loci and focus on antisense head-to-head, also named divergent, lncRNAs (<5 kb to a coding gene). This class of transcripts comprises almost 20% of total lncRNAs. In human, gene ontology analysis revealed that the neighboring protein coding genes associated to divergent lncRNAs were strongly enriched in nuclear functions, as transcription factor activity and sequence-specific DNA binding, as well as embryo development. In line with a possible role in development,

the chromatin environment surrounding the divergent lncRNA loci is particularly enriched for chromatin marks that are characteristic of developmental regulators, as the H3K27me3 and the H3K4me3 modifications.

Previous studies have described mammalian lncRNA/mRNA divergent transcription (26,27); however, Luo and colleagues provided novel insight into lncRNA roles at a functional level (25). Indeed, the knock-down of a subset of these lncRNAs in mouse embryonic stem cells (mESCs) led to a downregulation of nearby genes involved in cell fate specification. The authors provide a more detailed analysis on the regulation of the Even-Skipped Homeobox 1 (*EVX1*) by its divergent partner, *Evx1as*. Evx1as downregulation by RNAi and antisense oligonucleotides (ASOs), or its genomic deletion by CRISPR/Cas9, led to an attenuation of *EVX1* transcription. Conversely, the tethering of Evx1as to *EVX1* promoter mediated by a CRISPR system produced a significant increase of the levels of *EVX1* transcription, indicating a cis regulatory function of the lncRNA in control of *EVX1* gene expression. A deeper investigation of the Evx1as protein interactors revealed that Evx1as facilitates Mediator recruitment and the binding of H3K4me3 and H3K27ac at the promoter, thus shaping an active chromatin state. Single cell analysis during mESC differentiation upon LIF withdrawal revealed that EVX1as is present in most of the cells that activate *EVX1* expression. *EVX1* is involved in the determination of the character of primitive streak derivatives during gastrulation, in a regulatory network with Goosecoid (*GSC*) and *Brachyury* (28). *EVX1* and *Brachyury* stimulate each other's expression, while *EVX1* and *GSC* are mutually repressive. As a result, *EXV1* promotes mesoderm specification at the expenses of endoderm. In their paper, Luo *et al.* showed that both *EVX1* and *Evx1as* were required for proper mesoderm differentiation. Notably, compared to EVX1, depletion of Evx1as seemed to produce a stronger negative effect on the expression of mesoderm genes, leaving an open possibility of a broader function for this divergent lncRNA.

In conclusion, in addition to cis regulatory DNA elements as promoters or enhancers, divergent lncRNAs may provide another layer of transcription regulation. Fine-tuning of gene expression is extremely important during early phases of embryonic development, when key factors must be activated, maintained and eventually repressed in a tightly regulated manner, in space and time. Increasing evidence suggests that the contribution of the non-coding portion of the transcriptome to this fine-tuning is considerable. Certainly, lncRNAs represent flexible, mobile, and transient molecules thus providing a convenient mean to precisely regulate nearby gene expression in a site-specific way. We believe that more examples of lncRNAs with a key function in development and differentiation, such as those described in the paper by Luo and colleagues, will come in the next future.

Acknowledgements

The authors are grateful to C. Xia for critical reading of the paper. This work was partially supported by a grant from Sapienza University to A Rosa.

Footnote

Provenance: This is a Guest Commentary commissioned by Editor-in-Chief Zhizhuang Joe Zhao (Pathology Graduate Program, University of Oklahoma Health Sciences Center, Oklahoma City, USA).
Conflicts of Interest: The authors have no conflicts of interest to declare.

Comment on: Luo S, Lu JY, Liu L, *et al.* Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. Cell Stem Cell 2016;18:637-52.

References

1. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22:1760-74.
2. Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. Science 2005;309:1559-63.
3. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.
4. Kapranov P, Cawley SE, Drenkow J, et al. Large-scale transcriptional activity in chromosomes 21 and 22. Science 2002;296:916-9.
5. Bertone P, Stolc V, Royce TE, et al. Global identification of human transcribed sequences with genome tiling arrays. Science 2004;306:2242-6.
6. Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011;25:1915-27.
7. Guttman M, Amit I, Garber M, et al. Chromatin signature

reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223-7.
8. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011;43:904-14.
9. Derrien T, Johnson R, Bussotti G, 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.
10. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014;15:7-21.
11. Rosa A, Ballarino M. Long Noncoding RNA Regulation of Pluripotency. Stem Cells Int 2016;2016:1797692.
12. Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015;21:1253-61.
13. Ballarino M, Morlando M, Fatica A, et al. Non-coding RNAs in muscle differentiation and musculoskeletal disease. J Clin Invest 2016;126:2021-30.
14. Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell 2013;152:1298-307.
15. Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet 2006;15 Spec No 1:R17-29.
16. Wutz A, Rasmussen TP, Jaenisch R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet 2002;30:167-74.
17. Zappulla DC, Cech TR. Yeast telomerase RNA: a flexible scaffold for protein subunits. Proc Natl Acad Sci U S A 2004;101:10024-9.
18. Tsai MC, Manor O, Wan Y, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science 2010;329:689-93.
19. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature 2012;482:339-46.
20. Quinn JJ, Zhang QC, Georgiev P, et al. Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes Dev 2016;30:191-207.
21. Quinodoz S, Guttman M. Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. Trends Cell Biol 2014;24:651-63.
22. Ørom UA, Derrien T, Beringer M, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010;143:46-58.
23. Lai F, Orom UA, Cesaroni M, et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 2013;494:497-501.
24. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet 2016;17:47-62.
25. Luo S, Lu JY, Liu L, et al. Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. Cell Stem Cell 2016;18:637-52.
26. Hung T, Wang Y, Lin MF, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 2011;43:621-9.
27. Ponjavic J, Oliver PL, Lunter G, et al. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. PLoS Genet 2009;5:e1000617.
28. Kalisz M, Winzi M, Bisgaard HC, et al. EVEN-SKIPPED HOMEOBOX 1 controls human ES cell differentiation by directly repressing GOOSECOID expression. Dev Biol 2012;362:94-103.

doi: 10.21037/sci.2016.09.02
Cite this article as: Rosa A, Ballarino M. Divergent lncRNAs take the lead on pluripotent cell differentiation. Stem Cell Investig 2016;3:47.

