Editorial Commentary

LncRNA expression and regulatory networks across pediatric cancers

Unai Illarregi[1]^, Elixabet Lopez-Lopez[2,3]^

[1]Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Leioa, Spain; [2]Department of Biochemistry and Molecular Biology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Leioa, Spain; [3]Pediatric Oncology Group, Biobizkaia Health Research Institute, Barakaldo, Spain

Correspondence to: Elixabet Lopez-Lopez, PhD. Department of Biochemistry and Molecular Biology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Barrio Sarriena s/n, 48940 Leioa, Spain; Pediatric Oncology Group, Biobizkaia Health Research Institute, Barakaldo, Spain. Email: elixabet.lopez@ehu.eus.

Comment on: Modi A, Lopez G, Conkrite KL, *et al.* Integrative Genomic Analyses Identify LncRNA Regulatory Networks across Pediatric Leukemias and Solid Tumors. Cancer Res 2023;83:3462-77.

Keywords: Long non-coding RNAs (lncRNAs); childhood; cancer; regulatory networks

Submitted Oct 26, 2023. Accepted for publication Jan 11, 2024. Published online Feb 20, 2024.
doi: 10.21037/tp-23-531
View this article at: https://dx.doi.org/10.21037/tp-23-531

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nt with low or no protein-coding potential. Despite this low potential to encode proteins, they fulfill a great variety of regulatory functions; however, they remain poorly characterized. In fact, when they were discovered, lncRNAs were dismissed as transcriptional noise because of their generally poor sequence conservation, low expression, and low visibility in genetic screens (1,2).

More recently, helped by extensive technology development, lncRNAs have been shown to be very versatile molecules that can regulate many biological processes through their ability to bind proteins, DNA and RNA, and there has been an increasing interest in their study. However, mechanistic information has been obtained for only a minority of lncRNAs (1).

In oncology research, lncRNAs have been shown to act as either tumor suppressors or oncogenes. In fact, pan-cancer studies of lncRNA expression in adult cancers have identified lncRNAs of interest in those malignancies (3-6). Nevertheless, a comprehensive characterization of lncRNA expression across pediatric cancers is needed, in order to determine if lncRNAs of interest in adults may also have a role in pediatric cancer, which has different characteristics.

In this context, Modi *et al.* recently published a thorough article entitled "Integrative Genomic Analyses Identify LncRNA Regulatory Networks across Pediatric Leukemias and Solid Tumors" in *Cancer Research* (7). In the article, the authors performed a study of lncRNAs in six different pediatric cancer types, identifying both lncRNAs expressed across different pediatric cancers, and lncRNAs with histotype-specific expression. Moreover, they further explored upstream regulatory processes and downstream targets of those lncRNAs.

The study analyzed 1,044 leukemia and solid tumor samples previously sequenced as part of the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project, including acute myeloid leukemia, B-lymphoblastic leukemia, T-lymphoblastic leukemia, Willms tumors, extracranial rhabdoid tumors, and neuroblastoma (NBL), representing some of the most common pediatric malignances.

Interestingly, 2,657 lncRNAs were identified as robustly expressed across pediatric cancers. This is a very interesting result that sheds light into lncRNAs of interest in pediatric oncology. However, the number of lncRNAs identified is quite low, especially if we compare it to the number of

^ ORCID: Unai Illarregi, 0000-0001-7418-5616; Elixabet Lopez-Lopez, 0000-0002-5309-3031.

Table 1 Number of lncRNAs included in different human reference genome annotations and database versions

Annotation/database	Version	Release date	Number of lncRNA genes
GENCODE	19	2013.12	7,114[†]
	29	2018.10	7,635[†]
	44	2023.07	18,866[‡]
T2T-CHM13	–	2022.01	18,389[‡]
NONCODE	6	2020.11	96,411
LNCipedia	1.2	2012.10	11,390
	5.2	2018.08	56,946 (49,372[§])
lncRNAKB	–	2020.01	72,166

We can see that identification of lncRNA genes has increased a lot in recent years, and that specialized databases include much more lncRNAs than general annotations. [†], genes with the gene biotype lincRNA were considered; [‡], genes with the biotype lncRNA were considered; [§], high confidence lncRNAs. lncRNAs, long non-coding RNAs; lincRNA, long intergenic non-coding RNA.

lncRNAs described in different annotations (*Table 1*) and the number of protein coding genes identified in the study following the same criteria (n=15,588). One explanation for this low number of lncRNAs identified could be the stringent criteria stablished for minimum expression. As mentioned above, in general, lncRNAs present lower expression levels than protein coding genes, and many lncRNAs of interest could have presented expression levels too low to pass the criteria stablished. On the other hand, it needs to be mentioned that only sequencing information from poly(A)-enriched libraries was used, and lncRNAs without a poly(A) tail could not be identified. In this context, Guo *et al.* showed that after annotating genes with ENSEMBL GRCh37.35 version in two different cell lines, poly(A) library samples detected 4,122 and 6,169 potential new transcripts, whereas in total RNA samples from the same cell lines 53,282 and 58,111 potential new transcripts were identified (8). Therefore, it would be interesting to perform similar studies with total RNA sequencing data, which would include all types of lncRNAs, and with a higher sequencing depth in order to also obtain information for lncRNAs with lower expression levels.

On the other hand, a challenge for this kind of studies, as acknowledged by the authors, is the difficulty in obtaining pediatric healthy tissue controls. In the case of the study by Modi *et al.*, the absence of control samples makes it difficult to identify lncRNAs acting as tumor suppressor genes in the cancer types analyzed, as they would be downregulated or silenced in the tumor samples. Taking into account that including control samples for each leukemia and solid tumor type would provide more information about the changes in expression of lncRNAs in cancer, it would be of interest to include new control samples in the cases in which they are available or use publicly available data, in order to obtain a more complete picture of the diseases.

Another relevant outcome of the manuscript is the study of histotype-specificity regarding lncRNA expression in pediatric tumors. In this regard, principal component analysis of lncRNA expression showed differences between leukemias and solid tumors, as well as a better ability of lncRNAs to separate histotypes than the combination of lncRNAs and protein coding genes. Furthermore, tissue specificity index revealed greater tissue specific expression for lncRNAs than for protein coding genes. NBL presented the highest number of tissue-elevated lncRNAs. Using the top five highly expressed tissue elevated lncRNAs per cancer, it was possible to cluster samples of the same histotype. This would be in agreement with the higher tissue specificity of lncRNAs previously observed in normal tissues (9). As pointed out by Modi *et al.*, this higher tissue specificity observed for lncRNAs would make them more ideal candidates as biomarkers than protein coding genes.

Remarkably, taking into account this potential as biomarkers of the lncRNAs and the limited information available regarding lncRNA characterization and function, the authors performed analyses of upstream mechanisms of regulation and downstream targets of lncRNAs.

On the one hand, Modi *et al.* looked into upstream regulation through the analysis of somatic copy number alterations and other structural variants (SVs). This way, lncRNAs with significant differential expression due to somatic copy number alterations and expressed lncRNAs

disrupted by SVs were identified, several of them associated with well-known SVs in the cancer types in which they were identified. This could mean that lncRNAs are relevant mediators of those driver alterations. Of note, 30 lncRNAs with pan-cancer expression and SV breakpoints were also identified, which could point to some general mechanisms across pediatric cancer.

On the other hand, in order to identify the downstream mechanism of action of the identified lncRNAs, they focused on the identification of differentially expressed lncRNAs that attenuated, enhanced or inverted transcriptional factor effect on their target genes. Interestingly, the majority of lncRNA modulators identified acted in trans, and the same lncRNA could have different effect on different transcription factor-target pairs. Moreover, the majority of lncRNA modulators appeared to be active in only one cancer. As a proof of the utility of this approach, the biological pathways regulated through this mechanism were consistent with the type of tumor in which they took place. Therefore, this type of analysis could help stablish new networks of interest in the pathogenesis of pediatric malignancies.

As a proof of concept, a deeper analysis was performed in NBL. In this type of cancer, lncRNA expression varied in association with different cell lineages of origin. Using ChIP-seq and chromatin capture data in NBL cell lines, 547 lncRNA genes were identified as regulated by the adrenergic core transcriptional circuitry.

An integrative multi-omic analysis of all the genetic and epigenetic data obtained, allowed the prioritization of candidate functional lncRNAs. This way, the top ranked lncRNA was *MEG3*, with a known role in NBL (10,11), which demonstrates the ability of the approach to capture relevant lncRNAs. The next prioritized lncRNA was *TBX2-AS1*, which was selected for experimental study. Silencing of *TBX2-AS1* in NBL cell lines led to reduction in cell growth and changes in cell morphology suggestive of neuronal differentiation, demonstrating an important role of this lncRNA that could explain a driver effect in the origin of the disease. In addition, an analysis of the genes deregulated by the silencing of *TBX2-AS1* revealed that there was enrichment in inflammation-associated pathways among genes significantly upregulated, while the E2F target genes was the only significantly enriched gene set across the downregulated genes. In fact, the presence of transcriptional factor motifs or ChIP-seq tracks from ENCODE was confirmed, supporting the regulation of *TBX2-AS1* downstream target genes by E2F transcription factors. Furthermore, it is important to highlight that the NBL cell lines used in the study were originated from a male (NLF) and a female (SKNSH), avoiding the very frequent sex bias in genomic studies (12).

This initial study in NBL supports the need for similar characterization approaches in other pediatric tumor types. Such a comprehensive study of lncRNAs could be helpful for a better understanding of pediatric cancer, and for the identification of new biomarkers and therapeutic targets. Moreover, additional functionalities of lncRNAs could be tested. For instance, studying the role of lncRNAs as miRNA sponges, as performed in other tumor types (13-15) would be an important next step in understanding how lncRNAs impact post-transcriptional regulation in pediatric cancers. In this context, the TARGET project database contains miRNA-seq data for all the studied cancer types except for NBL. Therefore, integration of RNA-seq and miRNA-seq data from available samples could help understand other lncRNA regulation mechanisms and networks in pediatric tumors.

In conclusion, the study performed by Modi *et al.* is of great relevance because it shows evidence for the importance of lncRNAs in pediatric cancers, characterizing the most expressed lncRNAs in six childhood cancer types and predicting lncRNA gene regulatory networks. These data provide insights into novel lncRNA functions and regulatory mechanisms in pediatric cancers, and pave the way for future mechanistic studies.

Acknowledgments

Funding: This work was supported by the University of the Basque Country (UPV/EHU) through a predoctoral fellowship awarded to author U.I.

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Translational Pediatrics*. The article has undergone external peer review.

Peer Review File: Available at https://tp.amegroups.com/article/view/10.21037/tp-23-531/prf

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://tp.amegroups.com/article/view/10.21037/tp-23-531/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Mattick JS, Amaral PP, Carninci P, et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol 2023;24:430-47.
2. Hazan J, Bester AC. CRISPR-Based Approaches for the High-Throughput Characterization of Long Non-Coding RNAs. Noncoding RNA 2021;7:79.
3. Isaev K, Jiang L, Wu S, et al. Pan-cancer analysis of non-coding transcripts reveals the prognostic onco-lncRNA HOXA10-AS in gliomas. Cell Rep 2021;37:109873.
4. Samadi P, Shahnazari M, Shekari A, et al. A pan-cancer analysis indicates long noncoding RNA HAND2-AS1 as a potential prognostic, immunomodulatory and therapeutic biomarker in various cancers including colorectal adenocarcinoma. Cancer Cell Int 2023;23:307.
5. Li Y, Jiang T, Zhou W, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat Commun 2020;11:1000.
6. Yang Y, Xiong Z, Li W, et al. FHIP1A-DT is a potential novel diagnostic, prognostic, and therapeutic biomarker of colorectal cancer: A pan-cancer analysis. Biochem Biophys Res Commun 2023;679:191-204.
7. Modi A, Lopez G, Conkrite KL, et al. Integrative Genomic Analyses Identify LncRNA Regulatory Networks across Pediatric Leukemias and Solid Tumors. Cancer Res 2023;83:3462-77.
8. Guo Y, Zhao S, Sheng Q, et al. RNAseq by Total RNA Library Identifies Additional RNAs Compared to Poly(A) RNA Library. Biomed Res Int 2015;2015:862130.
9. Seifuddin F, Singh K, Suresh A, et al. lncRNAKB, a knowledgebase of tissue-specific functional annotation and trait association of long noncoding RNA. Sci Data 2020;7:326.
10. Ye M, Lu H, Tang W, et al. Downregulation of MEG3 promotes neuroblastoma development through FOXO1-mediated autophagy and mTOR-mediated epithelial-mesenchymal transition. Int J Biol Sci 2020;16:3050-61.
11. Ye M, Gao R, Chen S, et al. Downregulation of MEG3 and upregulation of EZH2 cooperatively promote neuroblastoma progression. J Cell Mol Med 2022;26:2377-91.
12. Özdemir BC, Richters A, Espinosa da Silva C, et al. Cancer researchers' perceptions of the importance of the sex of cell lines, animals, and human samples for cancer biology research. iScience 2023;26:106212.
13. Han J, Chen X, Wang J, et al. Glycolysis-related lncRNA TMEM105 upregulates LDHA to facilitate breast cancer liver metastasis via sponging miR-1208. Cell Death Dis 2023;14:80.
14. Le Hars M, Castro-Vega LJ, Rajabi F, et al. Pro-tumorigenic role of lnc-ZNF30-3 as a sponge counteracting miR-145-5p in prostate cancer. Biol Direct 2023;18:38.
15. Yao F, Huang X, Xie Z, et al. LINC02418 upregulates EPHA2 by competitively sponging miR-372-3p to promote 5-Fu/DDP chemoresistance in colorectal cancer. Carcinogenesis 2022;43:895-907.

Cite this article as: Illarregi U, Lopez-Lopez E. LncRNA expression and regulatory networks across pediatric cancers. Transl Pediatr 2024;13(2):383-386. doi: 10.21037/tp-23-531