Genome scale analysis of alternative splicing events in lncRNA knockdowns

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS AND COMPUTING
IUPUI

Felipe Wendt Porto[1], Swapna Vidhur Daulatabad[1], Sarath Chandra Janga[1, 2, 3]

[1] Department of BioHealth Informatics, Indiana University School of Informatics and Computing, IUPUI. [2] Centre for Computational Biology and Bioinformatics, Health Information and Translational Sciences, Indianapolis. [3] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis.

ABSTRACT

Recent developments in our understanding of the interactions between lncRNA and cellular components have improved treatment approaches for various human diseases including cancer, vascular diseases, and brain diseases (1, 2, 3). Although investigation of specific lncRNAs revealed their role in the metabolism of cellular RNA, our understanding of their contribution to post-transcriptional regulation is relatively limited. In this study, we explore the role of lncRNAs in modulating alternative splicing and their impact on downstream protein-RNA interaction networks. Analysis of alternative splicing events across 39 lncRNA wildtype and knockout RNA-sequencing datasets from three human cell lines: Hela (Cervical Cancer), K562 (Myeloid Leukemia), and U87 (Glioblastoma), resulted in high confidence (fdr < 0.01) identification of 4432 skipped exon events and 2474 retained intron events, implicating 759 genes to be impacted at post-transcriptional level due to the loss of lncRNAs. We observed that a majority of the alternatively spliced genes in a lncRNA knockout were specific to the cell type, in agreement with the finding that genes affected by alternative splicing also displayed enriched functions in a cell type specific manner (4, 5). To understand the mechanism behind this cell-type specific alternative splicing patterns, we analyzed RNA binding protein (RBP)-RNA interaction profiles across the spliced regions. Despite limited RBP binding data across cell lines, alternatively spliced events detected in lncRNA perturbation experiments were associated with RBPs binding in proximal intron-exon junctions, in a cell type specific manner. Based on the RBP binding profiles in Hela and K562 cells, we hypothesize that several lncRNAs are likely to exhibit a sponge effect in disease contexts, resulting in the functional disruption of RBPs, and their downstream functions. We propose that such lncRNA sponges can extensively rewire the post-transcriptional gene regulatory networks by altering the protein-RNA interaction landscape in a cell-type specific manner.

OBJECTIVES

- To understand lncRNA's role in splicing mechanisms from data generated by a novel lncRNA knockdown technique.
- Identify genes affected by lncRNA knockdown inducing alternatively spliced events.
- To examine lncRNA's role in alternative splicing across cell lines.
- Explore RBP binding patterns of alternatively spliced events induced by the lncRNA knockdown.

MATERIALS & METHODS

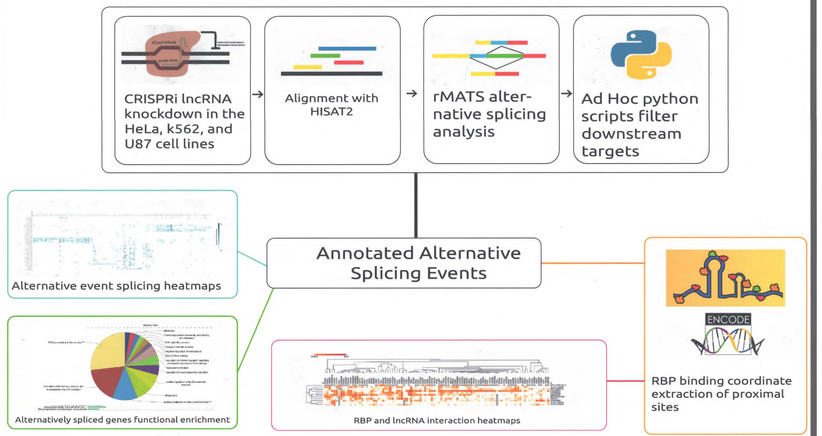

Figure 1. Overall process of the experiment. Scripts were made to automate: downloading of data, alignment of the reads to their controls, and annotate the alternative splicing events. From the annotated events, both alternative splicing event information and proximal RBP binding locations were obtained. Alternative splicing events were then analyzed by generating heatmaps and enriching the functions of the spliced events. RBP binding was examined by overlapping proximal bindings with binding locations annotated by ENCODE.

- Extracted 96 sequences from GEO experiment GSE85011.
- Extracted the RBP binding locations using CLIP-seq data from ENCODE.
- Sequence alignment to reference genome using hisat2.
- Experimental metadata annotation in context of the samples.
- Identified various alternative spliced events across lncRNA knockdown and wildtype using rMATS
- Administered FDR (< .01) filters on the events to extract the most significant alternative splicing events.
- Generated heatmaps of alternatively spliced events induced by the lncRNA knockdowns using the online tool Morpheus.
- Performed functional enrichment analysis using CLUEGO, to annotated clusters of alternatively spliced genes
- Identified the RBPs binding in proximal sites of the recorded alternative spliced events. Proximal sites being: Exon start and end: upstream and downstream
- RBP binding enrichment across lncRNA knockdown samples alternative splicing events were computed using the Fishers exact test.

	Bound	Unbound
RBP's bound motifs	X	Y
All motifs generated from lncRNA KD	A	B

RESULTS

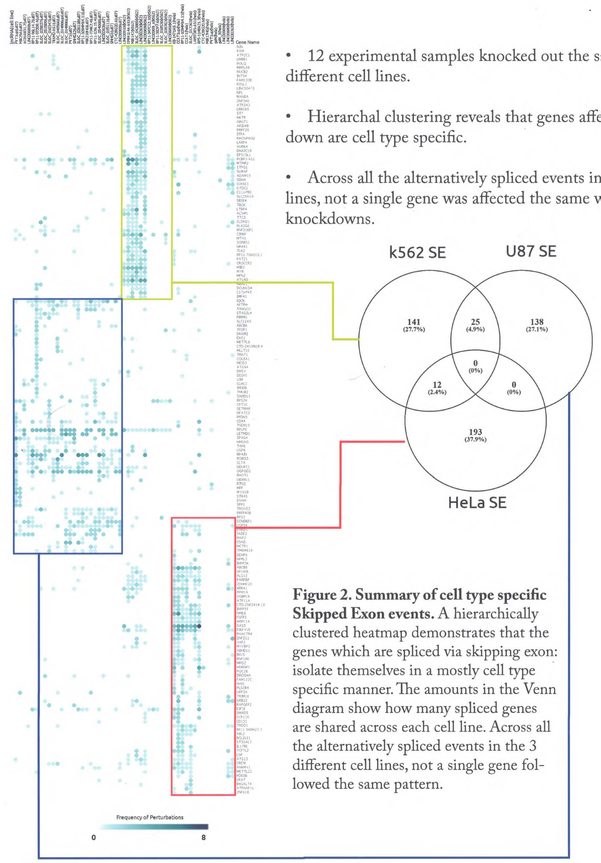

Figure 2. Summary of cell type specific Skipped Exon events. A hierarchically clustered heatmap demonstrates that the genes which are spliced via skipping exon: isolate themselves in a mostly cell type specific manner. The amounts in the Venn diagram show how many spliced genes are shared across each cell line. Across all the alternatively spliced events in the 3 different cell lines, not a single gene followed the same pattern.

- 12 experimental samples knocked out the same lncRNAs across different cell lines.
- Hierarchal clustering reveals that genes affected by each knockdown are cell type specific.
- Across all the alternatively spliced events in the 3 different cell lines, not a single gene was affected the same way by the lncRNA knockdowns.

- Genes were alternatively spliced in a cell type specific manner, so the functions enriched from one cell line were not found in the other cell lines.
- Across the 3 cell lines the most enriched functions were:

  Hela: metabolic processes, vesicle formation, and vesicle movement.
  K562: microfiber construction, cell growth, and apoptosis.
  U87: DNA endonuclease repair, and ribosomal complex formation.

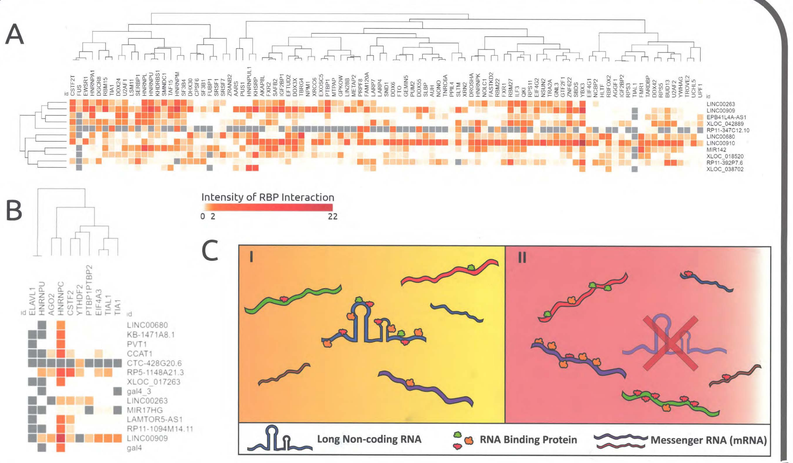

Figure 3. Proximal RNA-Binding Protein analysis. The grey values represent NA values within the heatmaps and the intensities are log scaled FDRs of RBP protein interaction with genes from a lncRNA knockdown. Figures 3A and 3B showcase the lncRNAs knocked down in respective cell lines A) k562 and B) hela and the resulting proximal RBP interaction intensity. C) An illustration of the proposed RBP sponge model, that occurs when a lncRNA is not present. C(I) demonstrates when the lncRNA is present, and C(II) demonstrates the proliferated RBP interactions if the lncRNA is not present. D) Shows the extent of RBPs interacting to lncRNA LINC00909, and it showcases the binding sponge model.

- There were 14 lncRNAs in the HeLa cell line and 11 in K562 cell line which bound proximally to RBP.
- In the HeLa cell line. Knockdowns of LINC00909 and RP5-1148A21.3 were the only lncRNAs which induced interactions with RBPs
- In the K562 cell line LINC00910, LINC00680, RP11-392P7.6, and LINC00909 were the lncRNAs that showed the most significant cases of interactions across many RBPs.
- KHSRP, CSTF2T, YBX3, ZNF622, SAFB2, SRSF1, and QKI were the RBPs that seemed to interact with the most lncRNAs overall in the K562 cell line.

CONCLUSIONS

- Alternative splicing induced by lncRNA knockdown has been shown to be cell type specific and continues to display a similar trend within the context of human cancer cell lines: HeLa, K562, and U87.
- Proximal RBP binding appears to be cell type specific, and can be confirmed as more binding data is available.
- We hypothesize that lncRNAs behave as a binding sponge for RBPs and are curious to see if other alternative splicing experiments represent this behavior as well.
- lncRNAs can be used as a medium of understanding alternative splicing in cancer. Combined with functional enrichment and RBP binding, this could help annotate lncRNA functions.

REFERENCES

1. Budak G., Srivastava R., Janga S.C. (2017) "Seten: a tool for systematic identification and comparison of processes, phenotypes, and diseases associated with RNA-binding proteins from condition-specific CLIP-seq profiles."
2. Elia, L. and M. Quintavalle (2017). "Epigenetics and Vascular Diseases: Influence of Non-coding RNAs and Their Clinical Implications."
3. Uszczynska-Ratajczak, B., et al. (2018). "Towards a complete map of the human long non-coding RNA transcriptome."
4. Gonzalez, I., et al. (2015). "A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature."
5. Liu, S. J., et al. (2017). "CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human

ACKNOWLEDGEMENTS

The authors acknowledge the team at Janga Lab for the guidance, comradery, and feedback which they have provided. This study was supported by the National Institute of General Medical Sciences (NIGMS) under Award Number R01GM123314.