Characterizing the Bidirectional Nature and Transcriptional Activity of Alternative Promoters Across Different Locations in a Gene

Apurva Mishra [1,2†], Helen Feng [2], Dr. Ana Fiszbein [2]
[1] Davis Senior High School, 315 W 14th St. #1914, Davis, CA 95616
[2] Fiszbein Lab, Boston University, Boston, MA 02215
†Corresponding author: apurvagyan@gmail.com

OVERVIEW

INTRODUCTION

In past research, promoters have been found to initiate RNA transcription bidirectionally, meaning that nascent RNA reads exist both in the downstream sense and upstream antisense directions. In addition to this, downstream and hybrid exons (exons that can be used as first and internal exons) have been found to have an increase in usage in comparison to upstream and first exons only (exons that can only be used as first exons), respectively. However, this behavior has been largely unexplored in alternative promoters, and is therefore of great interest for the advancement of the field, and the purpose of our study.

RESEARCH OBJECTIVES

- We hypothesize that in accordance with previous research, downstream and hybrid exons should display greater expression of nascent RNA than upstream and first exons only (FEO).
- To test this, we aim to collect data on the relative transcriptional activity within genes and characterize the patterns of transcription in alternative promoters for different exon classifications.
- In addition, we gain to understanding on the frequency and scale of bidirectional transcription in alternative promoters.

BACKGROUND

BIDIRECTIONAL TRANSCRIPTION

Existing research has displayed that promoters can initiate transcription bidirectionally - in the positive and negative strand simultaneously. This is referred to as downstream 'sense' or upstream 'antisense' transcription respectively.

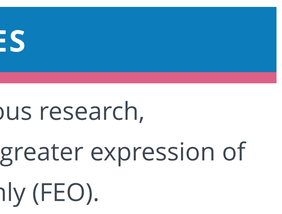

Figure 1: Bidirectional Promoter

TRANSCRIPTIONAL START SITE (TSS)

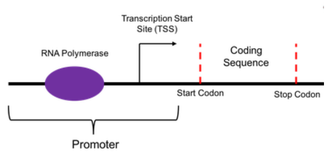

Figure 2: Transcriptional Start Site

The transcriptional start site (TSS) is the location within the core promoter where transcription begins. It provides a reference as to where exons are being transcribed from, and defines downstream versus upstream expression [2].

EXON CHARACTERIZATION

- **FEO vs. Hybrid:** An exon that is only used as a first exon during transcription is a first exon only (FEO). An exon that can be used as both a first and internal exon is a hybrid exon. Previous research has shown these exons to exhibit varying transcription based on relative location from the TSS.
- **Genomic Order:** Genomic order describes the nth exon from upstream to downstream on the strand being viewed. Since a hybrid exon can always be an internal exon, it therefore must be contained in genomic order 2 or above.

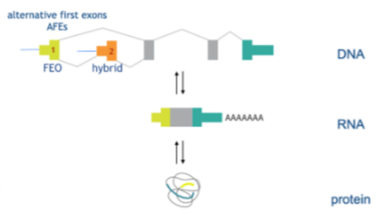

Figure 3: FEO vs. Hybrid Exons

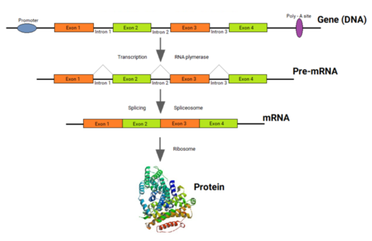

Figure 4: Exon Genomic Order

METHODS

GENE COLLECTION

In order to compare nascent RNA reads for different exon types and genomic orders, we first produced a gene list for which to aggregate read count data. The requirements for gene and exon collection varied for exon type and genomic order respectively:

- Only genes with at least 3 genomic orders (3 exons) were observed
- For each gene, the first 3 exons were selected for analysis, which all were either FEO or hybrid exons

DATA ANALYSIS

1. Up to 400 coordinate locations both downstream and upstream from the TSS were assigned for both the positive and negative strands in the gene list for nucleotide-level comparison
2. Pro-Seq data from nascent RNA read samples of HeLa cells in humans were used to collect read counts for respective genes [(1)]
3. Read counts across all genes for each classification (exon type or genomic order) were then aggregated to determine the total nascent reads per each nucleotide relative to the TSS
4. Read counts for relative exon locations were then visualized to: **(1)** Compare FEO vs. hybrid exons; **(2)** Compare genomic orders; **(3)** Compare downstream vs. upstream transcription

RESULTS

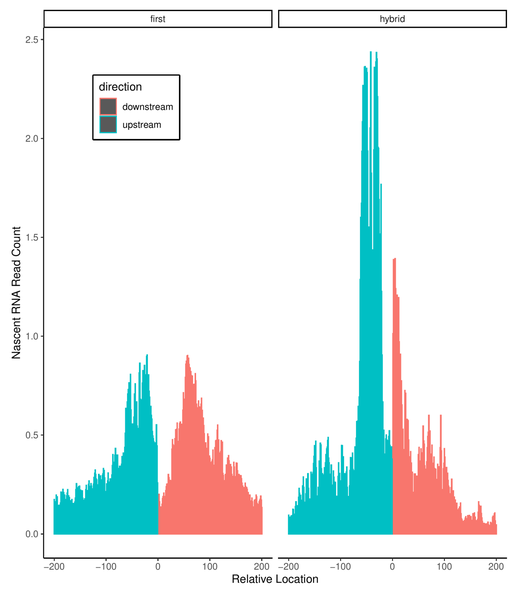

Figure 5: Nascent RNA Reads for Exon Type

- Expression in both the sense and antisense directions is greater for hybrid exons than first exons only
- More comparative upstream expression than downstream expression for hybrid exons

- Results present up to 35x more nascent RNA reads in antisense direction than sense direction for hybrid exons
- Both FEO and hybrid exons have imbalanced bidirectionality; more upstream transcription

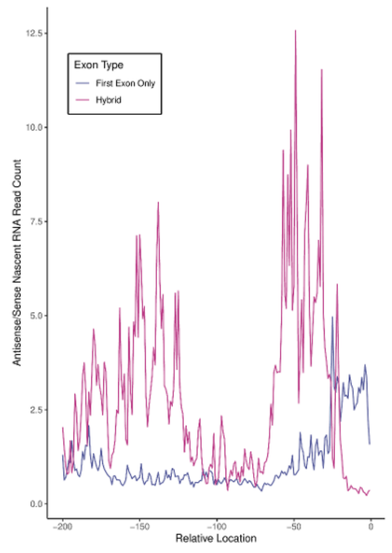

Figure 6: Antisense RNA Reads Proportional to Sense Direction for Exon Type

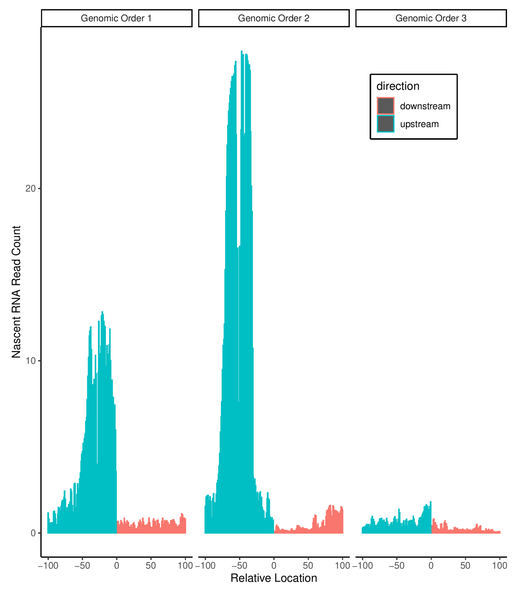

Figure 7: Nascent RNA Reads for Genomic Orders

- Results display far greater expression in the antisense direction for both genomic orders 1 and 2
- Increase in upstream expression from genomic order 1 to 2, and decrease from genomic order 2 to 3

- No visible spike in upstream expression in genomic order 3, up to 200x more antisense reads for genomic order 2
- Greatest activity across all genomic orders near 25th upstream nucleotide from TSS

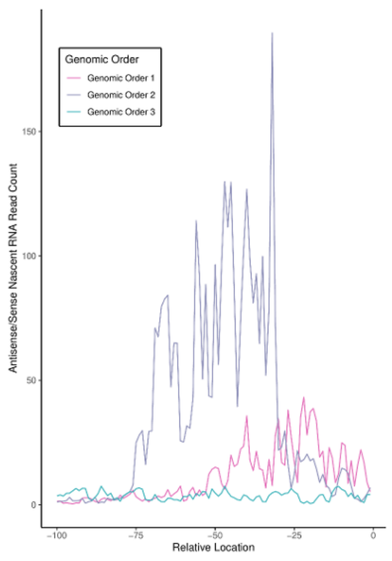

Figure 8: Antisense RNA Reads Proportional to Sense Direction for Genomic Orders

DISCUSSION AND CONCLUSION

In summary, we find that for all genomic order exons in a gene, nascent RNA transcription was more common in the upstream antisense direction than the downstream sense direction. Upstream gene expression increased between the first and second genomic order exons along a gene, but soon dropped in the third genomic order exon. Finally, transcription in both the sense and antisense direction occurred at higher rates in hybrid exons as compared to first exons only (FEO), concurring with our hypothesis and previous research in the field. This study entails further investigation into why transcriptional disparities exist among alternative promoters. The observations provided allow for a greater understanding of how alternative promoters exhibit bidirectional transcription, and take a step in that direction.

ACKNOWLEDGEMENTS

Thank you to my research professor Dr. Ana Fiszbein and direct mentor Helen Feng, and the entirety of the Fiszbein Lab in the Biology Department of Boston University, for their guidance and support in this project. Thank you also to the Research in Science and Engineering (RISE) Program for presenting this opportunity and enabling this research.

REFERENCES

1. Blumberg, A., Zhao, Y., Huang, YF.*et al.*Characterizing RNA stability genome-wide through combined analysis of PRO-seq and RNA-seq data.*BMC Biol***19,**30 (2021). https://doi.org/10.1186/s12915-021-00949-x
2. Jin, V.X., Singer, G.A., Agosto-Pérez, F.J.*et al.*Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs.*BMC Bioinformatics***7,**114 (2006). https://doi.org/10.1186/1471-2105-7-114