Specific Aims:

Rationale: Splicing is a process where pre-mRNAs are processed to remove introns and stitch together exons. This process offers the potential to introduce tremendous diversity into the proteome by changing which sequences from the pre-mRNA are included as exons. Accordingly, more than 90% of human genes are alternatively spliced in different conditions, such as at different developmental stages or across cell types.[26] Errors in this process have been implicated in a wide variety of human diseases including Familial Dysautonomia, early onset Parkinson disease, and cancer.[21] Further recent efforts have produced drugs that modulate splicing outcomes to produce improvements in Spinal muscular Atrophy, Huntington's Disease, Duchenne muscular dystrophy, and cancer.[18] Work to understand the mechanisms of splicing stands to benefit not only our understandings of development and differentiation but also our ability to target these processes with therapeutics. Much progress has been made using massively parallel splicing assays (MPSAs) to probe the mechanisms of splicing.[1–3, 5, 6, 8, 13, 15, 17, 19, 20, 22, 23, 28] These efforts have largely focused on understanding how the sequence characteristics of the pre-RNA dictate the inclusion or exclusion of a potential exon and have made great strides. However, while most alternative splicing events are believed to be local affecting an individual intron or exon, recent work has begun to show that correlated splicing decisions are far more common than was previously understood.[10, 25, 29] A particularly interesting case of correlated splicing is mutually exclusive exons (MXEs). MXEs are clusters of exons that are spliced such that every isoform includes exactly one exon from the cluster. While in other species clusters of more that 2 MXEs are well known, examples in humans have only recently been detected and have never been extensively characterized.[10, 12] A recent analysis of human RNA-seq data identified 629 MXE clusters, of of which less than 25% had an identifiable mechanism for enforcing mutual exclusivity.[10] **I propose to develop a massively parallel splicing assay for characterizing mechanisms of mutual exclusivity in splicing and apply it to investigating new mechanisms of mutual exclusivity in human splicing. I will focus these efforts on Pyruvate Kinase M (PKM) and CD55.**

Aim 1: Characterize the mechanism for mutual exclusivity in Pyruvate Kinase M (PKM) exons 9 and 10. PKM is a central metabolic gene that controls the rate limiting step in glycolysis with two primary isoforms defined by the inclusion of either exon 9 or exon 10. The conversion between these two isoforms is a primary driver of the Warburg effect and is important in a wide variety of cancers.[7, 16] However, while much is known about the regulation of PKM splicing, the mechanism for enforcing exclusivity between exons 9 and 10 is still unknown.

Sub-aim 1.1: Develop a screen to detect deviations from mutual exclusivity in splicing. We will utilize PKM minigene constructs designed by the Krainer lab for studying PKM splicing [27] and adapt MPSA techniques honed by the Kinney lab [11, 28] to develop a massively parallel assay for detecting mutual exclusivity or an MX-MPSA. We will use next-generation sequencing technologies that are capable of sequencing the entire minigene to measure the effects of mutations in the context of varying the inclusion ratios of the relevant MXEs using splice modifying drugs. This will allow us to detect deviations from mutual exclusivity. We call this an MX-MPSA.

Sub-aim 1.2: Identifying sequence elements necessary for mutual exclusivity in PKM splicing. To identify the sequence elements responsible for maintaining the trade off between exons 9 and 10, we propose to perform a mutant screen looking for mutations that break the anti-correlation between exon 9 and 10 across a range of ratios of exon inclusion. Previous work from our group has shown that the elements responsible for PKM mutual exclusivity are located internal to exons 9 and 10, so we will utilize saturation mutagenesis of these exons as well as broader random mutagenesis strategies to map elements that affect the degree of anti-correlation between the exons. Identified regions will be mapped at higher resolution with deep mutagenesis.

Sub-aim 1.3: Mechanistically characterize sequence elements enforcing mutual exclusivity. We will follow up these experiments with mechanistic studies of the sequence elements that enforce mutual exclusivity. In many arthropods and the human ATE1 gene [9, 12, 14] MXEs are enforced by RNA secondary structural elements. We will use mutant cycle analysis on predicted important secondary structural elements to show dependence on structure. We will also use RNA based pull downs as well as motif mutants coupled with RNAi knockdown and over-expression of RNA binding proteins (RBPs) to demonstrate the importance of RBP binding.

Aim 2: Identify mechanisms responsible for large human MXE clusters in CD55. Large scale analysis of RNA-seq datasets has revealed human genes with MXE behavior in clusters of more than two exons for the first time.[10] However, no mechanism for enabling MXEs in clusters of more than 2 exons has been demonstrated for a human gene. In this aim, we will investigate mechanisms that enable large MXE clusters in humans focusing primarily on the 5 MXE cluster identified in CD55. CD55 is important for shielding cells from complement based attack and is implicated in both CHAPEL syndrome and cancer.[4, 24]

Sub-aim 2.1: Generate a minigene that recapitulates MXE behavior in the CD55 large cluster setting. In order to elucidate the mechanisms of MXE splicing in CD55, we will generate a minigene containing the MXE cluster from CD55 and validate that we observe the expected splicing behavior. We will also trial antisense oligonucleotides (ASOs) and risdiplam sensitive 5'-splice sites for inducing different exon inclusion ratios.

Sub-aim 2.2: Identifying sequence elements necessary for MXE behavior in CD55. Bioinformatic analysis of the CD55 MXE cluster has suggested possible secondary structure control elements.[10] We will perform targeted mutagenesis of these elements as well as random mutagenesis over the whole region and screen for regions that disrupt splicing. Identified regions will be further mapped using deeper mutagenesis and ASO targeting experiments.

Sub-aim 2.3: Mechanistically characterize sequence elements related to MXE behavior in CD55. We will follow up these experiments by using mutant cycle analysis, and modulation of predicted binding proteins to gain insight into the mechanisms of mutually exclusive splicing in CD55.

References

[1] Scott I Adamson, Lijun Zhan, and Brenton R Graveley. "Vex-seq: high-throughput identification of the impact of genetic variation on pre-mRNA splicing efficiency". en. In: *Genome Biol.* 19.1 (June 2018), p. 71. ISSN: 1465-6906. DOI: 10.1186/s13059-018-1437-x. URL: http://dx.doi.org/10.1186/s13059-018-1437-x.

[2] Pablo Baeza-Centurion et al. "Combinatorial Genetics Reveals a Scaling Law for the Effects of Mutations on Splicing". en. In: *Cell* 176.3 (Jan. 2019), 549–563.e23. ISSN: 0092-8674, 1097-4172. DOI: 10.1016/j.cell.2018.12.010. URL: http://dx.doi.org/10.1016/j.cell.2018.12.010.

[3] Pablo Baeza-Centurion et al. "Mutations primarily alter the inclusion of alternatively spliced exons". en. In: *Elife* 9 (Oct. 2020). ISSN: 2050-084X. DOI: 10.7554/eLife.59959. URL: http://dx.doi.org/10.7554/eLife.59959.

[4] Rashmi Bharti et al. "CD55 in cancer: Complementing functions in a non-canonical manner". en. In: *Cancer Lett.* 551 (Oct. 2022), p. 215935. ISSN: 0304-3835, 1872-7980. DOI: 10.1016/j.canlet.2022.215935. URL: http://dx.doi.org/10.1016/j.canlet.2022.215935.

[5] Simon Braun et al. "Decoding a cancer-relevant splicing decision in the RON proto-oncogene using high-throughput mutagenesis". en. In: *Nat. Commun.* 9.1 (Aug. 2018), pp. 1–18. ISSN: 2041-1723, 2041-1723. DOI: 10.1038/s41467-018-05748-7. URL: https://www.nature.com/articles/s41467-018-05748-7.

[6] Rocky Cheung et al. "A Multiplexed Assay for Exon Recognition Reveals that an Unappreciated Fraction of Rare Genetic Variants Cause Large-Effect Splicing Disruptions". en. In: *Mol. Cell* 73.1 (Jan. 2019), 183–194.e8. ISSN: 1097-2765, 1097-4164. DOI: 10.1016/j.molcel.2018.10.037. URL: http://dx.doi.org/10.1016/j.molcel.2018.10.037.

[7] Heather R Christofk et al. "The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth". en. In: *Nature* 452.7184 (Mar. 2008), pp. 230–233. ISSN: 0028-0836. DOI: 10.1038/nature06734. URL: https://www.nature.com/articles/nature06734.

[8] Mariela Cortés-López et al. "High-throughput mutagenesis identifies mutations and RNA-binding proteins controlling CD19 splicing and CART-19 therapy resistance". en. In: *Nat. Commun.* 13.1 (Sept. 2022), pp. 1–17. ISSN: 2041-1723, 2041-1723. DOI: 10.1038/s41467-022-31818-y. URL: https://www.nature.com/articles/s41467-022-31818-y.

[9] Brenton R Graveley. "Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures". en. In: *Cell* 123.1 (Oct. 2005), pp. 65–73. ISSN: 0092-8674. DOI: 10.1016/j.cell.2005.07.028. URL: http://dx.doi.org/10.1016/j.cell.2005.07.028.

[10] Klas Hatje et al. "The landscape of human mutually exclusive splicing". en. In: *Mol. Syst. Biol.* 13.12 (Dec. 2017), p. 959. ISSN: 1744-4292. DOI: 10.15252/msb.20177728. URL: http://dx.doi.org/10.15252/msb.20177728.

[11] Yuma Ishigami et al. "Specificity, synergy, and mechanisms of splice-modifying drugs". en. Dec. 2022. DOI: 10.1101/2022.12.30.522303. URL: https://www.biorxiv.org/content/10.1101/2022.12.30.522303v1.

[12] Yongfeng Jin et al. "Mutually exclusive alternative splicing of pre-mRNAs". en. In: *Wiley Interdiscip. Rev. RNA* 9.3 (May 2018), e1468. ISSN: 1757-7004, 1757-7012. DOI: 10.1002/wrna.1468. URL: https://onlinelibrary.wiley.com/doi/10.1002/wrna.1468.

[13] Philippe Julien et al. "The complete local genotype–phenotype landscape for the alternative splicing of a human exon". en. In: *Nat. Commun.* 7.1 (May 2016), pp. 1–8. ISSN: 2041-1723, 2041-1723. DOI: 10.1038/ncomms11558. URL: https://www.nature.com/articles/ncomms11558.

[14] Marina Kalinina et al. "Multiple competing RNA structures dynamically control alternative splicing in the human ATE1 gene". en. In: *Nucleic Acids Res.* 49.1 (Jan. 2021), pp. 479–490. ISSN: 0305-1048, 1362-4962. DOI: 10.1093/nar/gkaa1208. URL: http://dx.doi.org/10.1093/nar/gkaa1208.

[15] Shengdong Ke et al. "Saturation mutagenesis reveals manifold determinants of exon definition". en. In: *Genome Res.* 28.1 (Jan. 2018), pp. 11–24. ISSN: 1088-9051, 1549-5469. DOI: 10.1101/gr.219683.116. URL: http://dx.doi.org/10.1101/gr.219683.116.

[16] Wai Kit Ma et al. "ASO-based PKM Splice-switching Therapy Inhibits Hepatocellular Carcinoma Cell Growth". en. Mar. 2022. DOI: 10.1158/0008-5472.CAN-20-0948. URL: https://repository.cshl.edu/id/eprint/40463/.

[17] Martin Mikl et al. "Dissecting splicing decisions and cell-to-cell variability with designed sequence libraries". en. In: *Nat. Commun.* 10.1 (Oct. 2019), pp. 1–14. ISSN: 2041-1723, 2041-1723. DOI: 10.1038/s41467-019-12642-3. URL: https://www.nature.com/articles/s41467-019-12642-3.

[18] Christopher R Neil et al. "Reprogramming RNA processing: an emerging therapeutic landscape". en. In: *Trends Pharmacol. Sci.* 43.5 (May 2022), pp. 437–454. ISSN: 0165-6147. DOI: 10.1016/j.tips.2022.02.011. URL: https://www.cell.com/trends/pharmacological-sciences/fulltext/S0165-6147(22)00046-3.

[19] Alexander B Rosenberg et al. "Learning the sequence determinants of alternative splicing from millions of random sequences". en. In: *Cell* 163.3 (Oct. 2015), pp. 698–711. ISSN: 0092-8674, 1097-4172. DOI: 10.1016/j.cell.2015.09.054. URL: http://dx.doi.org/10.1016/j.cell.2015.09.054.

[20] Dvir Schirman et al. "A broad analysis of splicing regulation in yeast using a large library of synthetic introns". en. In: *PLoS Genet.* 17.9 (Sept. 2021), e1009805. ISSN: 1553-7390, 1553-7404. DOI: 10.1371/journal.pgen.1009805. URL: http://dx.doi.org/10.1371/journal.pgen.1009805.

[21] Marina M Scotti and Maurice S Swanson. "RNA mis-splicing in disease". en. In: *Nat. Rev. Genet.* 17.1 (Nov. 2015), pp. 19–32. ISSN: 1471-0056. DOI: 10.1038/nrg.2015.3. URL: https://www.nature.com/articles/nrg.2015.3.

[22] Rachel Soemedi et al. "Pathogenic variants that alter protein code often disrupt splicing". en. In: *Nat. Genet.* 49.6 (June 2017), pp. 848–855. ISSN: 1061-4036, 1546-1718. DOI: 10.1038/ng.3837. URL: http://dx.doi.org/10.1038/ng.3837.

[23] Přemysl Souček et al. "High-throughput analysis revealed mutations' diverging effects on SMN1 exon 7 splicing". en. In: *RNA Biol.* 16.10 (Oct. 2019), pp. 1364–1376. ISSN: 1547-6286, 1555-8584. DOI: 10.1080/15476286.2019.1630796. URL: http://dx.doi.org/10.1080/15476286.2019.1630796.

[24] Lorraine Stallard, Iram Siddiqui, and Aleixo Muise. "Beyond IBD: the genetics of other early-onset diarrhoeal disorders". en. In: *Hum. Genet.* (Feb. 2023). ISSN: 0340-6717, 1432-1203. DOI: 10.1007/s00439-023-02524-6. URL: http://dx.doi.org/10.1007/s00439-023-02524-6.

[25] Hagen Tilgner et al. "Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events". en. In: *Nat. Biotechnol.* 33.7 (May 2015), pp. 736–742. ISSN: 1087-0156. DOI: 10.1038/nbt.3242. URL: https://www.nature.com/articles/nbt.3242.

[26] Eric T Wang et al. "Alternative isoform regulation in human tissue transcriptomes". en. In: *Nature* 456.7221 (Nov. 2008), pp. 470–476. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature07509. URL: http://dx.doi.org/10.1038/nature07509.

[27] Z Wang et al. "Exon-centric regulation of pyruvate kinase M alternative splicing via mutually exclusive exons". In: *Journal of molecular* (2012). URL: https://academic.oup.com/jmcb/article-abstract/4/2/79/908162.

[28] Mandy S Wong, Justin B Kinney, and Adrian R Krainer. "Quantitative Activity Profile and Context Dependence of All Human 5' Splice Sites". en. In: *Mol. Cell* 71.6 (Sept. 2018), 1012–1026.e3. ISSN: 1097-2765, 1097-4164. DOI: 10.1016/j.molcel.2018.07.033. URL: http://dx.doi.org/10.1016/j.molcel.2018.07.033.

[29] Chenchen Zhu et al. "Single-molecule, full-length transcript isoform sequencing reveals disease-associated RNA isoforms in cardiomyocytes". en. In: *Nat. Commun.* 12.1 (July 2021), p. 4203. ISSN: 2041-1723. DOI: 10.1038/s41467-021-24484-z. URL: http://dx.doi.org/10.1038/s41467-021-24484-z.