Investigations of *In Silico* Splice Prediction Tools in the Context of *SLC26A4* Related Hearing Loss

Joseph J. Chin, Kevin T. Booth, Richard J.H. Smith, Hela Azaiez

Molecular Otolaryngology and Renal Research Laboratories, Department of Otolaryngology, University of Iowa

Introduction

- *SLC26A4* is the third most common cause of hearing loss (HL)
- Variants in *SLC26A4* are associated with autosomal recessive nonsyndromic hearing loss with enlarged vestibular aqueduct (DFNB4) and Pendred syndrome (HL with thyroid goiter)
- In 14-31% of cases with HL and Mondini malformation/enlarged vestibular aqueduct, only one pathogenic variant is identified, suggesting the presence of an unidentified second pathogenic variant. We hypothesized that in some cases, the second variant is a splice-altering synonymous change.

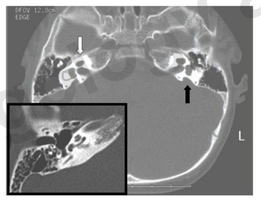

Figure 1: Radiology of *SLC26A4*-related hearing loss. Mondini malformation (white arrow) or an enlarged vestibular aqueduct (black arrow).

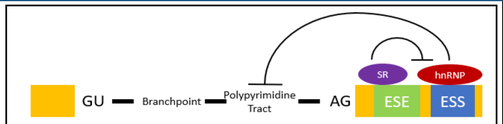

Figure 2: Overview of Exonic Splice Enhancers (ESE) and Exonic Splice Silencers (ESS). ESEs and ESSs are short regions within exons where various regulatory proteins can bind to and thereby regulate splicing. Examples of regulatory proteins include SR (splice enhancing) and hnRNP (splice repressing).

Variant Prioritization

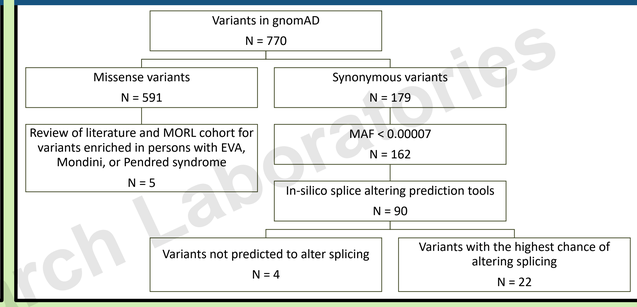

Figure 3: Overview of prioritization strategy. Variants were prioritized by MAF and by *in silico* splice prediction tools (Human Splicing Finder and SpliceAI) to obtain a list of 22 synonymous variants with a high likelihood of impacting splicing. 4 variants that were not predicted to alter splicing were randomly selected.

Minigene Splicing Analysis

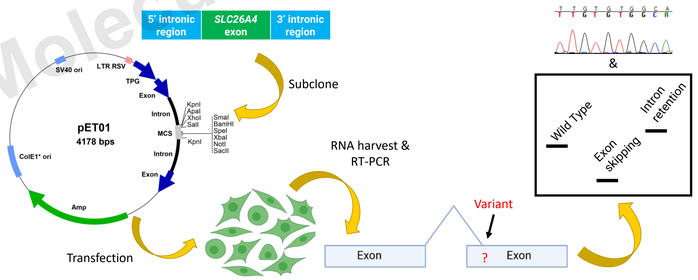

Figure 4: Overview of minigene splicing assay. *SLC26A4* exon of interest was subcloned into a pET01 vector. Vectors were then transfected into HEK293 cells. The impact of the variants was assessed using gel electrophoresis and Sanger sequencing.

Overview of *In Silico* Predictions

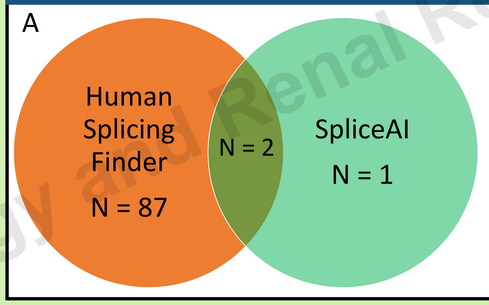

Figure 5: Overview of *in silico* splice predictions. (A) Venn diagram illustrating the number of variants flagged as splice altering by Human Splice Finder (HSF) or SpliceAI. (B) Overview of the distribution of changes in ESE/ESS ratios in Human Splicing Finder predictions.

Variants Tested

Summary of predicted effect	Variant	Exon	*In silico* predictions					Change in splicing
			Human Splicing Finder prediction			SpliceAI effect	Predicted splice alteration	
			ESE/ESS absolute value	Cryptic effect	Net change			
C.S.	c.237A>G	3	n/a	Donor	15%	0	Yes	No
	c.657G>A	6	n/a	Donor	59%	0	Yes	No
	c.1896G>A	17	n/a	Acceptor	62%	0	Yes	No
	c.1935A>G	17	n/a	Donor	15%	0	Yes	No
	c.2022A>G	17	n/a	Donor	49%	0.01	Yes	No
ESE / ESS	c.574C>T	5	10	n/a	n/a	0.01	Yes	No
	c.840C>T	**7**	**6**	**n/a**	**n/a**	**0**	**Yes**	**Yes**
	c.1068C>T	9	8	n/a	n/a	0	Yes	No
	c.1608C>T	14	5	n/a	n/a	0	Yes	No
	c.1614C>T	14	5	n/a	n/a	0	Yes	No
	c.2007C>T	17	10	n/a	n/a	0	Yes	No
ESE / ESS & C.S.	c.273A>G	3	6	Donor	13%	0.45	Yes	No
	c.486C>G	5	3	Donor	18%	0	Yes	No
	c.855T>A	7	2	Donor	28%	0	Yes	No
	c.1050G>A	9	4	Acceptor	71%	0.07	Yes	No
	c.1206G>A	**10**	**2**	**Acceptor**	**14%**	**0.95**	**Yes**	**Yes**
				Donor	**1263%**			
	c.1713A>G	16	7	Donor	55%	0.03	Yes	No
	c.2029C>A	17	5	Donor	71%	0.09	Yes	No
	c.2199G>A	19	4	Donor	21%	0	Yes	No
	c.2331A>G	21	2	Donor	16%	0	Yes	No
ESE / ESS & C.S. & SpliceAI	c.909A>G	7	5	Acceptor	52%	0.69	Yes	No
				Donor	13%			
SpliceAI	c.471C>T	4	n/a	n/a	n/a	0.59	Yes	No
No impact	c.225C>G	3	n/a	n/a	n/a	0	No	No
	c.678T>C	6	n/a	n/a	n/a	0	No	No
	c.1113T>C	9	n/a	n/a	n/a	0	No	No
	c.2163G>C	19	n/a	n/a	n/a	0	No	No

Table 1: Summary of variants tested, their splice predictions, and minigene splicing result. All coordinates are reported on the NM_000441.2 transcript. Variants with a SpliceAI score of >=0.5 were predicted to alter splicing.

Impact of Splice-Altering Variants

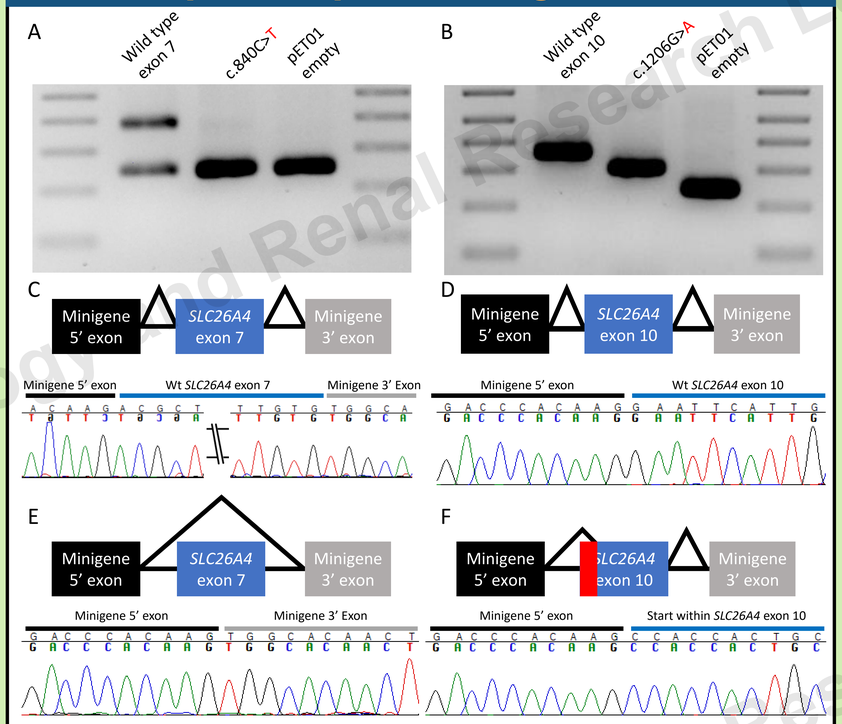

Figure 6: Minigene splicing assays. (A) Gel electrophoresis of wild type exon 7 and the c.840C>T vectors. (B) Gel electrophoresis wild type exon 10 and the c.1206G>A vectors. (C) Schematic drawing and sequencing of impact of the upper band which includes wild type exon 7. (D) Schematic drawing and sequencing of wild type exon 10. (E) Schematic drawing and sequencing of the c.840C>T variant. The c.840C>T variant results in skipping of exon 7. Skipping of exon 7 will preserve the reading frame. (F) Schematic drawings and sequencing of impact of the c.1206G>A variant on splicing. The c.1206G>A variant results in the formation of an alternative splice acceptor 58 bases into exon 10 and a frameshift.

ROC Curves

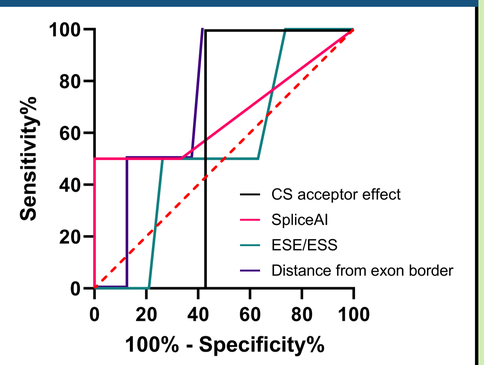

Figure 7: ROC curves of splice predictions by parameter. There were no observed parameters that increase the accuracy of *in silico* predictions.

Conclusions

- **The effect of coding variants (synonymous and missense) on splicing is underappreciated. It is important to assess their effect in clinical diagnostic settings.**
- ***In silico* splice prediction tools are inaccurate and caution should be taken when using them.**

References and Acknowledgments

1. Azaiez H, Yang T, Prasad S, Sorensen JL, Nishimura CJ, Kimberling WJ, Smith RJ. Genotype-phenotype correlations for *SLC26A4*-related deafness. Hum Genet. 2007 Dec;122(5):451-7. PMID: 17690912.
2. Booth KT, Azaiez H, Kahrizi K, Wang D, Zhang Y, Frees K, Nishimura C, Najmabadi H, Smith RJ. Exonic mutations and exon skipping: Lessons learned from *DFNA5*. Hum Mutat. 2018 Mar;39(3):433-440. PMID: 29266521
3. Sloan-Heggen CM, et al. Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. Hum Genet. 2016 Apr;135(4):441–450. PMCID: PMC4796320

This study was supported in part by NIDCDs R01s DC002842, DC012049 and DC017955 and NIGMS T32 GM139776.

Please send questions to Joseph Chin: Joseph-Chin@uiowa.edu
https://morl.lab.uiowa.edu/