INVESTIGATING THE MIS-SPLICING PHENOMENA OF GENES ON EXON-EXON JUNCTIONS AND V5-TAG

DJAN MATTHEW
Supervised by Professor Dong-Yan Jin
School of Biomedical Sciences, LKS Faculty of Medicine, The University of Hong Kong

Research Colloquium for Science Undergraduate Students 2021-22

Poster Number: C5
Name: DJAN Matthew
University No: 3035663459
Major: Molecular Biology and Biotechnology (Intensive)

ABSTRACT

RNA splicing is a highly conserved process in eukaryotic gene expression to transform precursor messenger RNA (pre-mRNA) to mature messenger RNA (mRNA) for translation. With its important role in governing protein expression, any incorrect RNA splicing could compromise the normal protein-encoding potential. Derived from the observed abnormal protein expression of FACI gene, it was found that one exon-exon junction sequence (AG|GTAAG) in FACI open reading frame (ORF) coincided with the splicing donor sequence (AG|GTpuAG, pu=A or G). Such observation led to the search of other genes with exon-exon junction sequences fitting the splicing donor sequences. Through bioinformatics analysis, 16 PFSGs (Prone to False Splicing Genes) were randomly-selected and their expression clones were constructed to further verify the mis-splicing occurrence. Splice-site sequence analysis and immunoblot showed incorrect splicing of all 16 genes, indicating the high mis-splicing risks for PFSGs. Since most of the mis-spliced PFSGs can be successfully expressed, this could greatly affect the accuracy of gain-of-function studies and disrupt the normal protein function.

INTRODUCTION

Eukaryotic mRNA processing involves RNA splicing, a mechanism in gene expression which alters the genetic information by removing introns and ligating exons to form a mature mRNA for further translation (1). This process is mediated by spliceosome, a large ribonucleoprotein complex which recognizes the specific 5′ donor splice site and 3′ acceptor splice site sequences of introns in the precursor messenger RNA. In mammals, the **5′ splice site** consists of 9 nucleotides with amino acid sequence of **YAG|GURAGU ('Y' is pyrimidine, 'R' is adenine or guanine, '|' denotes the splice site)**, whereas the **3′ splice site** consists of **12-nucleotide pyrimidine stretch followed by an AG dinucleotide** (2).

The complexity of mRNA splicing requires high fidelity because if exon ligation or splicing occurs incorrectly, normal protein expression might be compromised. For example, exon-skipping could disrupt translational reading frame, introduce premature stop codon, and produce unstable proteins associated with diseases (3).

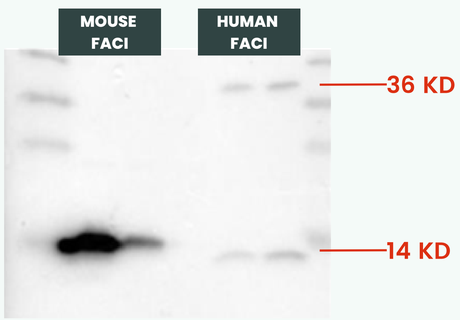
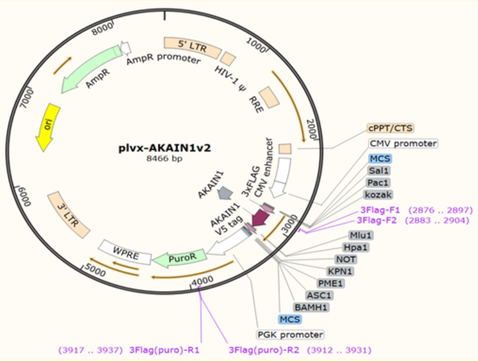


Figure 1. Extra immunoblot of 36KD on Human FACI cloned into inducible lentiviral pCW57-Puro.

With the severity of aberrant RNA splicing, unravelling potential mis-splicing is important to understand its effects on protein expression and explore ways to avoid such occurrence. In previous Professor Jin's lab research about a novel gene called FACI, they observed an extra immunoblot band of 36 KD on human FACI compared to mouse FACI with only 14 KD band (Figure 1). Sequence analysis revealed the **match of exon-exon junction sequence (CAG|GTAAG) in FACI with the splicing donor site "AG|GTpuAG"**, thus contributing to the **generation of mis-spliced human FACI transcript**. They further hypothesized that other genes bearing potential splice sites on the exon-exon junction would also be likely to undergo mis-splicing. Whole genome analysis of potential PFSGs (Prone to False Splicing Genes) was performed by comparing 20,394 human coding genes and 17,056 mouse coding genes. Through bioinformatic strategies, 1099 human PFSGs and 659 mouse PFSGs were narrowed down into 16 genes with potential splice sites on the exon-exon junction. **This research sought to prove the presence of unexpected splice sites on 16 PFSGs mainly through PCR sequencing and immunoblot.**

METHODOLOGY



PLASMID CONSTRUCT

The schematic diagram of one of the PFSGs, AKAIN1V2, which was cloned into pLVX-3xFLAG-V5 vector. 3xFLAG (F1 and F2), AKAIN1V2, and V5-tag were inserted into the vector's multiple cloning site consecutively. Another 3xFLAG (R1 and R2) were inserted into the PuroR sequence.

01 CELL CULTURE AND TRANSFECTION

Transfect 2 μg of lentiviral plasmid containing PFSG gene to HEK293T (Human embryonic kidney cell line) and AML12 (Mouse immortal hepatic cell line) → transduce for 48 hours → puromycine selection to kill lentiviral-negative cells.

02 PROTEIN EXTRACTION AND IMMUNOBLOTTING

Lyse cell or tissue samples with RIPA buffer → SDS-PAGE electrophoresis → gel transfer → incubate membrane with primary anti-flag (M2) antibodies and secondary goat anti-mouse IgG antibodies → visualize with chemiluminescence substrate.

03 RNA EXTRACTION AND RT-PCR

Extract RNA using RNAiso Plus Reagent → incubate with DNaseI for genomic DNA removal → reverse transcription using Transcriptor First Strand cDNA synthesis reagents → amplify RT-PCR products with forward primer on N-terminal flag tag and reverse primer on PuroR.

04 MOLECULAR CLONING AND SEQUENCING

Purify PCR fragments with QIAquick PCR and Gel Cleanup Kit → TA cloning into pGEM®-T Vector → transform to competent cells → blue/white colony selection → sequence positive clones using universal sequencing primer M13F and M13R located on pGEM®-T Vector.

RESULTS

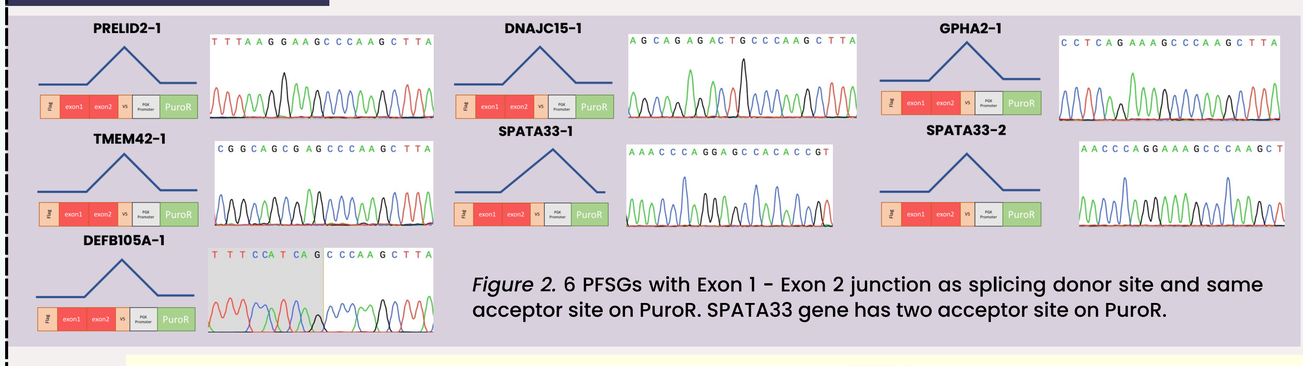


Figure 2. 6 PFSGs with Exon 1 - Exon 2 junction as splicing donor site and same acceptor site on PuroR. SPATA33 gene has two acceptor site on PuroR.

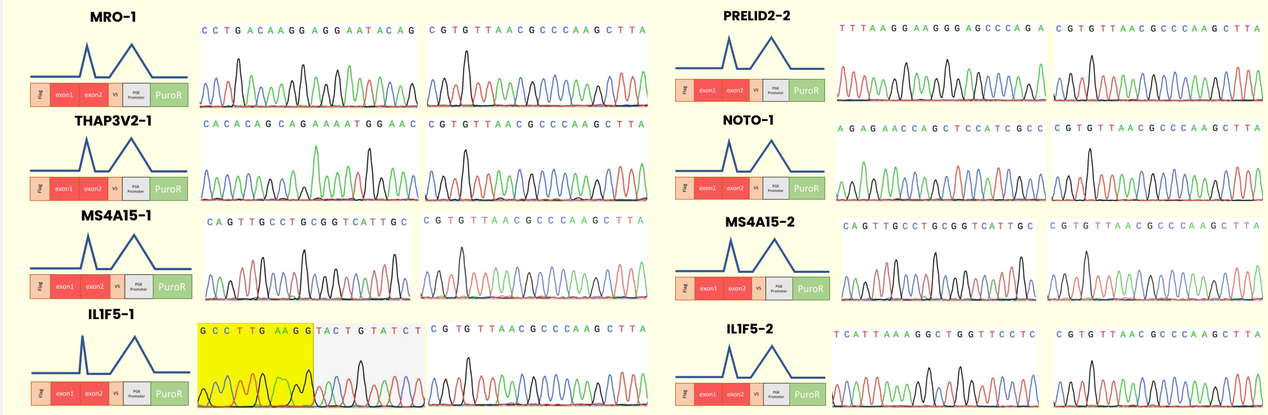


Figure 3. 6 PFSGs with Exon 1 - Exon 2 junction as splicing donor site and acceptor site on their coding sequences.

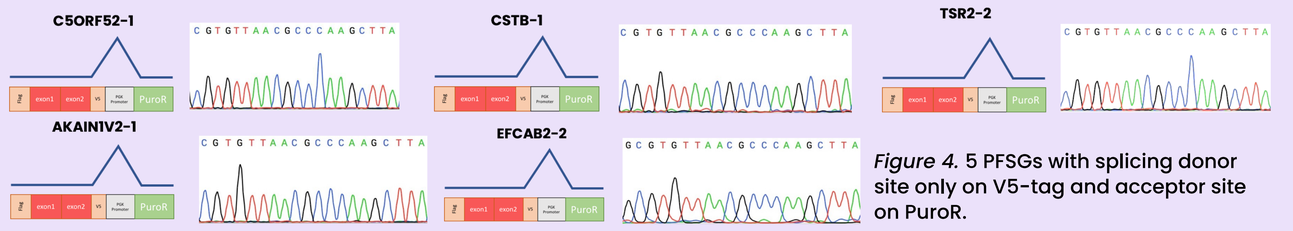


Figure 4. 5 PFSGs with splicing donor site only on V5-tag and acceptor site on PuroR.

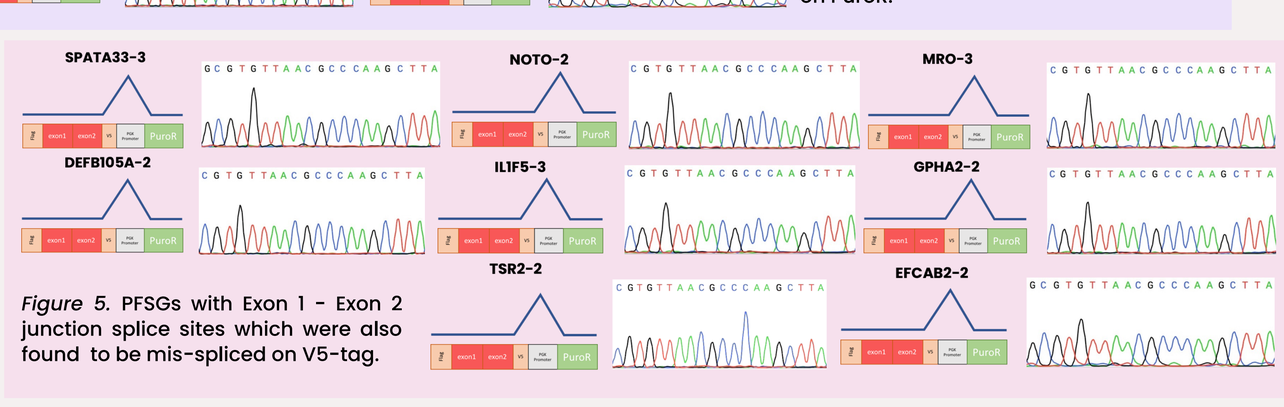


Figure 5. PFSGs with Exon 1 - Exon 2 junction splice sites which were also found to be mis-spliced on V5-tag.

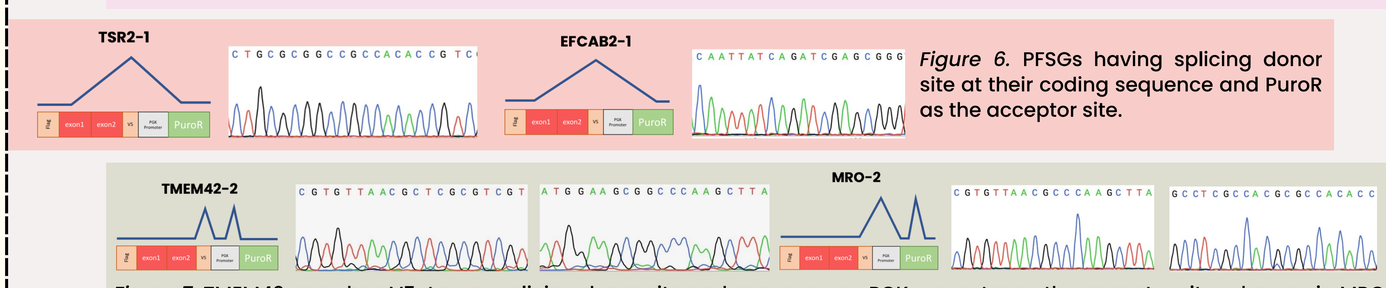


Figure 6. PFSGs having splicing donor site at their coding sequence and PuroR as the acceptor site.

Figure 7. TMEM42 gene has V5-tag as splicing donor site and sequence on PGK promoter as the acceptor site, whereas in MRO gene, mis-splicing occurs randomly on PuroR besides V5-tag.

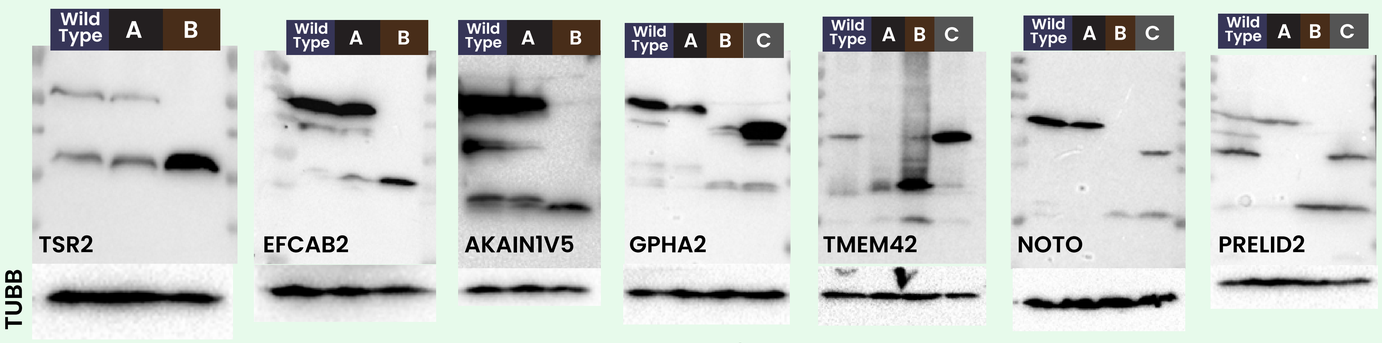


Figure 8. Mutation of the 5′ splicing donor site at exon-exon junction and/or V5-tag of PFSGs. Wild type: unmutated splice sites, A: exon-exon junction mutation, B: exon-exon junction and V5-tag mutation, and C: V5-tag mutation. TUBB (Tubulin Beta Class I) expression served as control.

DISCUSSION

1. Among 16 PFSGs, **11 genes had false splicing on the exon-exon junction**. Within these 11 genes, 6 genes (DNAJC15, DEFB105A, GPHA2, TMEM42, SPATA33, PRELID2) have acceptor site only on PuroR (Figure 2) and 5 genes (MRO, MS4A15, NOTO, THAP3V3, IL1F5) have acceptor site on their coding sequences (Figure 3).
2. **Mis-splicing on V5-tag of the lentiviral plasmid was also found in 14 out of 16 genes**, especially in AKAIN1V2, C5ORF52, CSTB, TSR2, and EFCAB2 genes (Figure 5). V5-tag has **"G|GTAAG" sequence which resembles a potential donor site** → removal of V5-tag **impaired expressed protein visualization by V5-tag antibody.**
3. Silent mutation experiments further verified the mis-splicing of TSR2, EFCAB2, and AKAIN1V2 on V5-tag, whereas GPHA2 and TMEM42 have mis-splicing on the exon-exon junction. NOTO and PRELID2 have splice sites on both exon-exon junction and V5-tag (as demonstrated by immunoblot result in Figure 8).

Overall, mis-splicing is likely to occur during foreign gene delivery if exon-exon junction and/or V5-tag sequence match the 5' splicing donor motif, and the 3' splicing site located on either the ORF of plasmid cassette expression or downstream vector regions. Using lentiviral and retroviral vectors for PFSGs transduction could result in mis-spliced mRNAs as they lack polyA signal for transcription termination.

REFERENCES

1. Jurica, M. S., & Roybal, G. A. (2013). RNA Splicing. In W. J. Lennarz, & M. D. Lane, *Encyclopedia of Biological Chemistry II* (pp. 185-190). Elsevier.
2. Busch, A., & Hertel, K. J. (2012). Evolution of SR Protein and HnRNP Splicing Regulatory Factors. *Wiley Interdisciplinary Reviews RNA*, 3(1), 1-12. doi:10.1002/wrna.100
3. Fackenthal, J. D., & Godley, L. A. (2008). Abberant RNA splicing and its functional consequences in cancer cells. *Disease Models and Mechanisms*, 1(1), 37-42. doi:10.1242/dmm.000331

ACKNOWLEDGEMENT

This research project is supported by the Laidlaw Scholars Programme. I would like to thank Professor Dong-Yan Jin for giving this research opportunity, Dr. Cheng Yun and Ms. Kang Xiao-Zhuo for their assistance throughout the project and providing data collection for analysis.