Name: DJAN MATTHEW
Supervisor: Professor Dong-Yan Jin

Investigating the mis-splicing phenomena of genes on exon junctions and V5-tag

I. Introduction

The Central Dogma of Molecular Biology describes the flow of genetic information in which DNA in the nucleus is transcribed into messenger RNA (mRNA) and subsequently translated into protein in the cytoplasm. Between the steps of transcription and translation, mRNA transcripts are extensively processed to acquire 3' poly-adenine tail and 5' methyl guanine cap before transported to the cytoplasm (Fackenthal & Godley, 2008). The mRNA was believed to be identical copies of DNA until 1977 when Richard Roberts and Phillip Sharps detected shorter cytoplasmic mRNA compared to its longer transcript in the eukaryotic nucleus (Jurica & Roybal, 2013). It was later found that mRNA also undergoes RNA splicing, a process in eukaryotic gene expression which alters the genetic information by removal of non-coding sequences of genes (introns) and ligation of protein-coding sequences (exons) to form a mature mRNA for further translation (Jurica & Roybal, 2013).

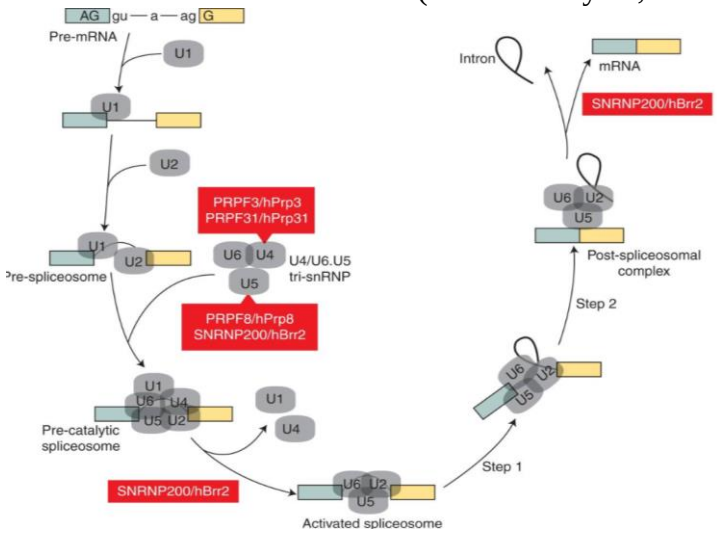


Figure 1. The mechanism of RNA splicing specifying the interaction of snRNPs (U1, U2, U4, U5, U6) and pre-mRNA. Step 1 and step 2 refers to the trans-esterification reaction. From Poulos, M. G., Batra, R., Charizanis, K., & Swanson, M. S. (2011). Developments in RNA Splicing and Disease. *Cold Spring Harbor Perspectives in Biology*, *3*(1), 1-14. doi:10.1101/cshperspect.a000778

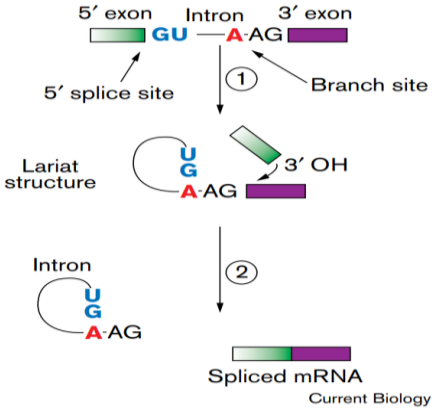


Figure 2. Trans-esterification reaction of mRNA splicing. In the first reaction, the 2'OH of the adenosine (red) at the branch site attacks the 5' splice site (blue) and releases the 5' exon (green). The 5' end of the intron (blue) is joined by a 2' – 5' phosphodiester

bond to the adenosine (red) and forms a lariat structure. In the second reaction, the 3'OH of the 5' exon intermediate (green) attacks the 3' splice site, hence forming spliced mRNA and lariat-shaped intron. From Newman, A. (1998). RNA splicing. *Cell, 9*(25). doi:10.1016/S0960-9822(98)00005-0

RNA splicing is facilitated by spliceosome, a complex molecular machine composed of five small nuclear RNAs (snRNAs) – U1, U2, U4, U5, and U6 and approximately 100 proteins (Wilkinson et al., 2019). Each of these snRNAs binds to a specific set of additional proteins to form small nuclear ribonucleoprotein (snRNP). Such splicing is governed by four main regulatory sequences, including the 5' and 3' splice sites located at the boundaries of exon-intron, the branch site and the polypyrimidine tract (PPT) (Leader et al., 2021). 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 (Busch & Hertel, 2012). The spliceosome carries out splicing through two successive transesterification reactions. Before the first reaction, U1 snRNP is base-paired with the 5' splice site and U2 auxiliary factor (U2AF) binds to the polypyrimidine tract (PPT) which subsequently facilitates the recruitment of U2 snRNP to the branch site and forms pre-spliceosome (Will & Lührmann, 2011). Meanwhile, U4/U6.U5 tri-snRNP is recruited to the complex and generates pre-catalytic complex (Will & Lührmann, 2011). Because of major rearrangements in RNA-RNA and RNA-protein interactions, the U1 and U4 snRNPs destabilize and lead to the activation of spliceosome (Will & Lührmann, 2011). With the aid of DEAH-box helicase Prp28, the activated spliceosome begins the transesterification reaction (Will & Lührmann, 2011). In the first reaction, the 2' hydroxyl of a branch point nucleotide attacks the phosphodiester group at the 5' splice site (Wilkinson et al., 2019). Through this reaction, the 5' exon is cleaved and the 5'-phosphate of the first intron nucleotide forms a lariat structure with the 2' oxygen of the branch site nucleotide (Wilkinson et al., 2019). In the second reaction, the 3' hydroxyl from the 5' exon attacks the last nucleotide of the intron at the 3' splice site, hence the 5' and 3' exons are ligated and the lariat intron is removed (Wilkinson et al., 2019).

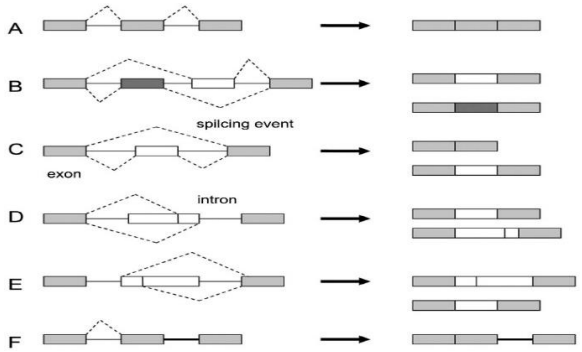


Figure 3. Five main types of alternative splicing events. (A) Constitutive splicing; (B) Mutually exclusive exons; (C) Cassette alternative exon; (D) Alternative 3′ splice site; (E) Alternative 5′ splice site; and (F) Intron retention. From Wang, Y., Liu, J., Huang, B., Xu, Y.-M., Li, J., Huang, L.-F., . . . Wang, X.-Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports, 3*(2), 152-158. doi:10.3892/br.2014.407

Based on deep sequencing and microsequencing analysis, 95 percent of human genes undergo alternative splicing which allows the relatively small amount of protein-coding genes (around 25,000) to generate more than 90,000 different proteins (Wang et al., 2015). While constitutive splicing removes intron and ligates exon in the order in which they appear in a gene, alternative splicing involves skipping or including certain exons which resulted in various forms of mature mRNA (Wang et al., 2015). Such capability allows the same gene sequence to produce different types of protein isoforms, hence enriching the proteome diversity. As shown in figure 3, there are several types of alternative splicing with the cassette alternative exon (exon skipping) as the most prevalent type (around 25 percent) in vertebrates and the intron retention as the rare type of splicing primarily found in the untranslated regions (UTRs) of the human mRNA transcript (Wang et al., 2015). The determination of exons that end up in the mature mRNA during alternative splicing is modulated by the interaction between cis-acting elements and trans-acting elements (Wang et al., 2015). Cis-acting elements include exonic splicing enhancers (ESEs) and intronic splicing enhancers (ISE) which interact with positive trans-acting factors such as SR proteins (serine/arginine-rich family of nuclear phosphoproteins) (Wang et al., 2015). On the other hand, exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs) are bound by negative trans-acting factors such as heterogeneous nuclear ribonucleoproteins (hnRNPs) (Wang et al., 2015). As a result, Cis-acting elements play an important role in constitutive splicing, whereas the silencers are more dominant in alternative splicing (Wang et al., 2015). Besides their role in alternative splicing, cis-acting sequence elements are also found to promote splice-site recognition (Newman, 1998). This was due to the low conserved nucleotide sequence of GU and AG in the intron termini of higher eukaryotes compared to yeast, therefore requiring additional factors to improve the splicing accuracy and avoid unwanted mutation that could lead to catastrophic biological effects (Newman, 1998). For instance, SR protein in eukaryotes enhances the recruitment of U1 snRNP to 5' splice sites or U2AF to the polypyrimidine tract which is crucial to define intron-exon boundaries and initiate the formation of spliceosome (Newman, 1998).

The complexity of mRNA splicing requires high fidelity because if exon ligation or splicing occurs incorrectly, normal protein-encoding potential might be compromised. For example, exon-skipping which occurs on conserved reading frame of triplet codons could result in protein lacking crucial amino acids (Fackenthal & Godley, 2008). More often, exon-skipping would disrupt the translational reading frame and cause premature stop codon (Fackenthal & Godley, 2008). Although 90 percent of these kinds of mRNAs would be degraded through nonsense-mediated mRNA degradation (NMD), some of these truncated products might escape this process and produce unstable proteins (Fackenthal & Godley, 2008). In other case such as non-sense altered splicing (NAS), the tendency of cells to upregulate the expression of splice variants with exon-skipping containing frameshift mutation would result in mRNA defects which might contribute to the development of diseases (Fackenthal & Godley, 2008).

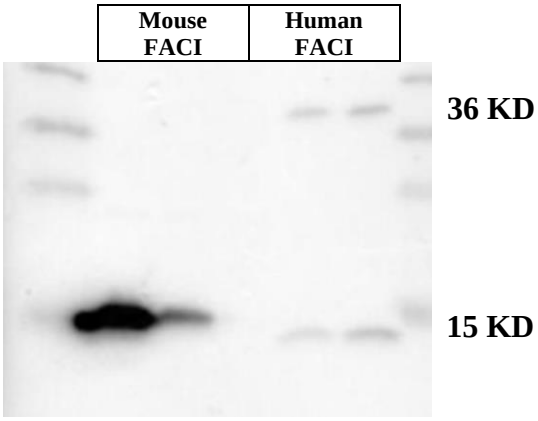


Figure 4. Extra immunoblot of 36 KD on Human FACI cloned into inducible lentiviral vector pCW57-Puro.

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. When my colleagues were investigating 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. At first, the 36 KD band was thought to be due to noncovalent interaction or protein dimerization. However, treating the sample with harsh denaturing and reducing condition did not diminish the 36 KD band. After performing sequence analysis and mutation, they found that the exon-exon junction sequence (CAG|GTAAG) in FACI coincidentally fits the splicing donor site "AG|GTpuAG", hence 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. Since bioinformatics could be used to characterize common sequence features of exons, splice sites, and introns, they conducted whole genome analysis of potential PMSGs (Prone to Mis-Splicing Genes) by comparing 20,394 human coding genes and 17,056 mouse coding genes. Through bioinformatic strategies, they found 1099 human PMSGs and 659 mouse PMSGs, and narrowed the result into 16 genes with potential splice sites on the exon-exon junction. With sufficient information of the genes, this research sought to prove the presence of unexpected splice sites on the 16 genes mainly through PCR sequencing and western blot.

II. Methodology

Touchdown PCR

Touchdown PCR was utilized due to higher specificity and yield of amplified products. The annealing temperature was initially set to be 5^oC - 10^oC higher than the T_m of the primers to allow perfect primer-template hybrids followed by gradual decrease of 1^oC in subsequent cycles until reaching the optimal annealing temperature of 2^oC - 5^oC below the T_m. Such process disfavours primer-dimer binding and non-specific primer-template complexes, hence minimizing undesirable amplification. A 10 μL of PCR reaction consists of 5 μL of Q5 Hot Start High-Fidelity 2X Master Mix, 3 μL of cDNA samples which previously underwent RT-PCR from mRNA samples, 2 μL of H_2O, and 0.6 μL of forward (on N-terminal flag tag) and reverse (on PuroR) primer mix (10 mM).

TA Cloning

Prior to TA Cloning, the PCR products containing amplified cDNAs were run on a 2 percent DNA gel followed by gel purification to isolate and purify desired DNA sample. Since Q5 Polymerase was used in the Touchdown PCR, the PCR products had blunt ends. Therefore, the samples were treated with Taq Polymerase to add adenine residue to the 3'-end of both DNA strands, thus making them suitable for TA cloning. A 10 μL of 3'A addition reaction requires 0.15 – 1.5 pmol of purified PCR products, 0.2 μL of dATP (10 mM), 1 μL of 10x PCR buffer, 0.2 μL of Taq Polymerase, and ddH_2O. The sample was incubated on the PCR machine for 20 minutes at 72ºC. After obtaining adenine at the 3'-end of the strands, the PCR products were ligated with pGEM®-T vector for 2 hours at room temperature. The ligation mixture consists of 0.5 μL of pGEM®-T vector, 1 μL of T4 Ligase, 2 μL of 5x buffer, PCR products with ligation ratio of 3:1, and ddH_2O.

Bacterial Transformation

The ligated samples were transformed with 50 μL of competent cells and incubated on ice for 30 minutes before performing heat shock for 60 seconds at 42ºC. Following the heat shock, the mixture was incubated with 1 mL of LB broth without any antibiotics for 1 hour on a shaker at 37ºC. The mixture was then centrifuged for 5 minutes with the speed of 4,000 rpm. The pellet formed at the bottom of the microcentrifuge tube was resuspended with 200 μL of the supernatant followed by the addition of 16 μL of IPTG and 40 μL of X-gal. Successful transformation was shown by white colonies on the LB plate after overnight incubation, whereas blue colonies indicated the failure of DNA ligation with pGEM®-T vector. The white colonies were sent for sequencing using universal sequencing primer M13F and M13R located on pGEM®-T vector.

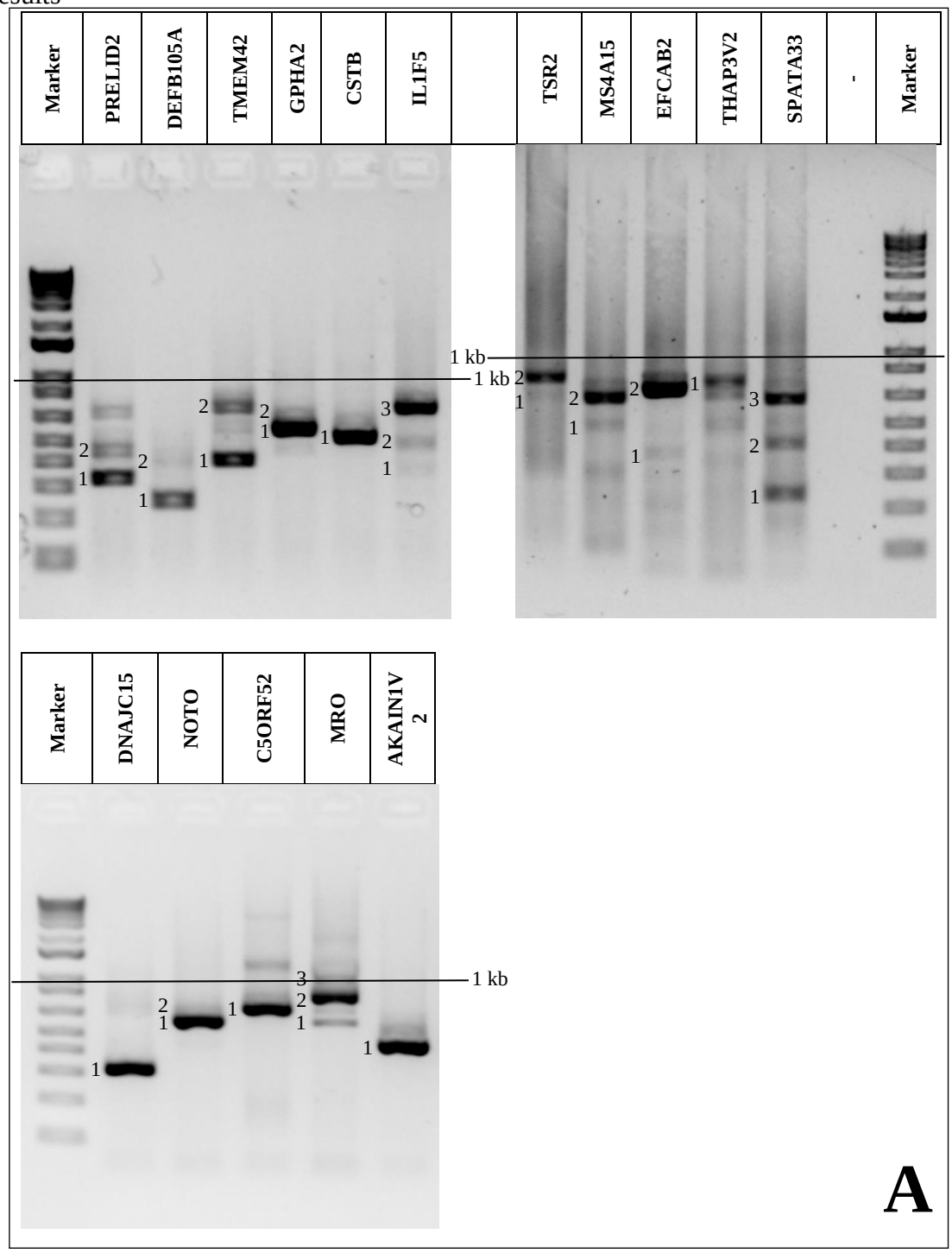
Protein Extraction

The proteins were harvested from HEK293T or AML12 cells transfected 2 days earlier with 2 μg of lentiviral vector containing the investigated genes**.** The cells on the surface of the 6-well culture plate were lysed in RIPA lysis buffer (50 mM and pH 7.4 Tris-HCl, 1 mM EDTA, 150 mM NaCl, 1% NP40, 5 mM NaF, 0.25% Na deoxycholate, and $NaVO_3$) supplemented with protease inhibitor cocktails. For each well, 120 μL of the buffer was added to detach the cells. The culture plate was incubated in 4ºC for 30 minutes before collecting the cell lysate into microcentrifuge tubes. The solution was centrifuged in 4ºC for 15 minutes with the speed of 14,500 rpm. The supernatant was collected for protein analysis.

Western Blot

Prior to SDS-PAGE electrophoresis, 5x SDS-PAGE buffer was added into the samples followed by boiling for 10 minutes at 95ºC. The prepared samples were electrophoresed on a 12% polyacrylamide gel at 120 V. To transfer the proteins from the gel to a PVDF membrane, semi-dry electroblotting was conducted for 30 minutes at 25 V and 1 A. The transferred membrane was blocked in 5 percent slim milk and incubated overnight with primary anti-flag (M2) antibodies followed by secondary goat anti-mouse IgG antibodies for 1 hour, and visualized with Western Bright chemiluminescence substrate.

III. Results


Marker
PRELID2
DEFB105A
TMEM42
GPHA2
CSTB
IL1F5
TSR2
MS4A15
EFCAB2
THAP3V2
SPATA33
-
Marker
1 kb
1 kb
Marker
DNAJC15
NOTO
C5ORF52
MRO
AKAIN1V2
1 kb
A

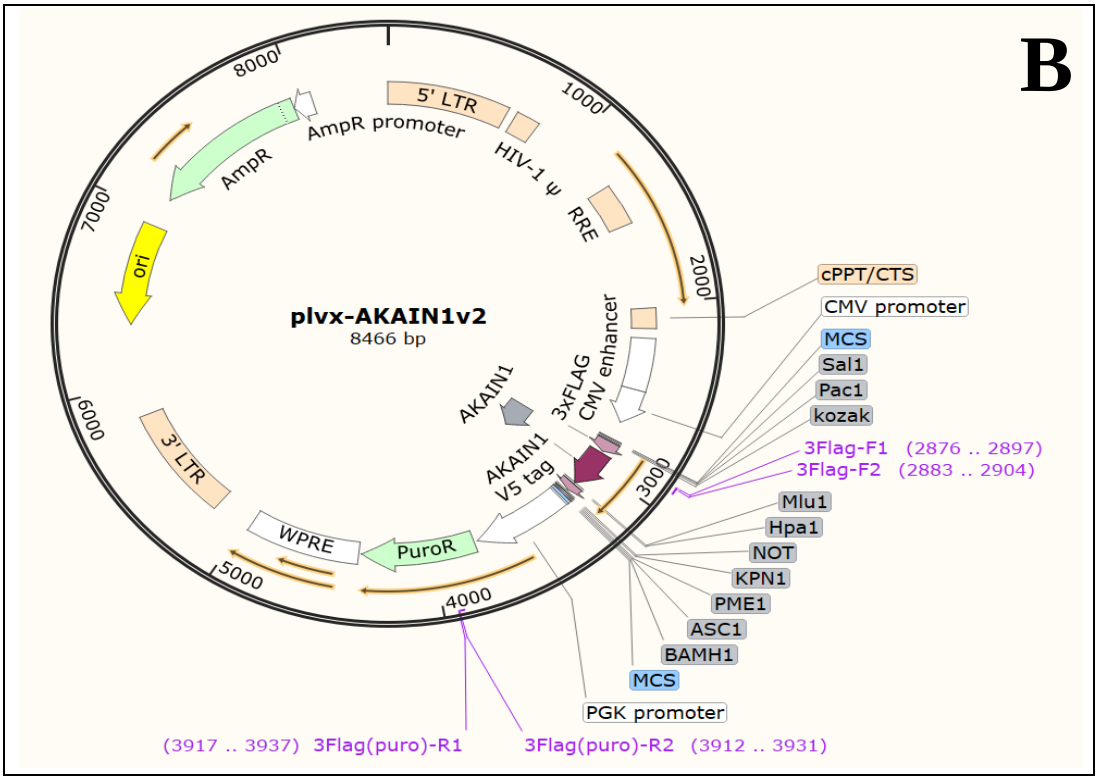
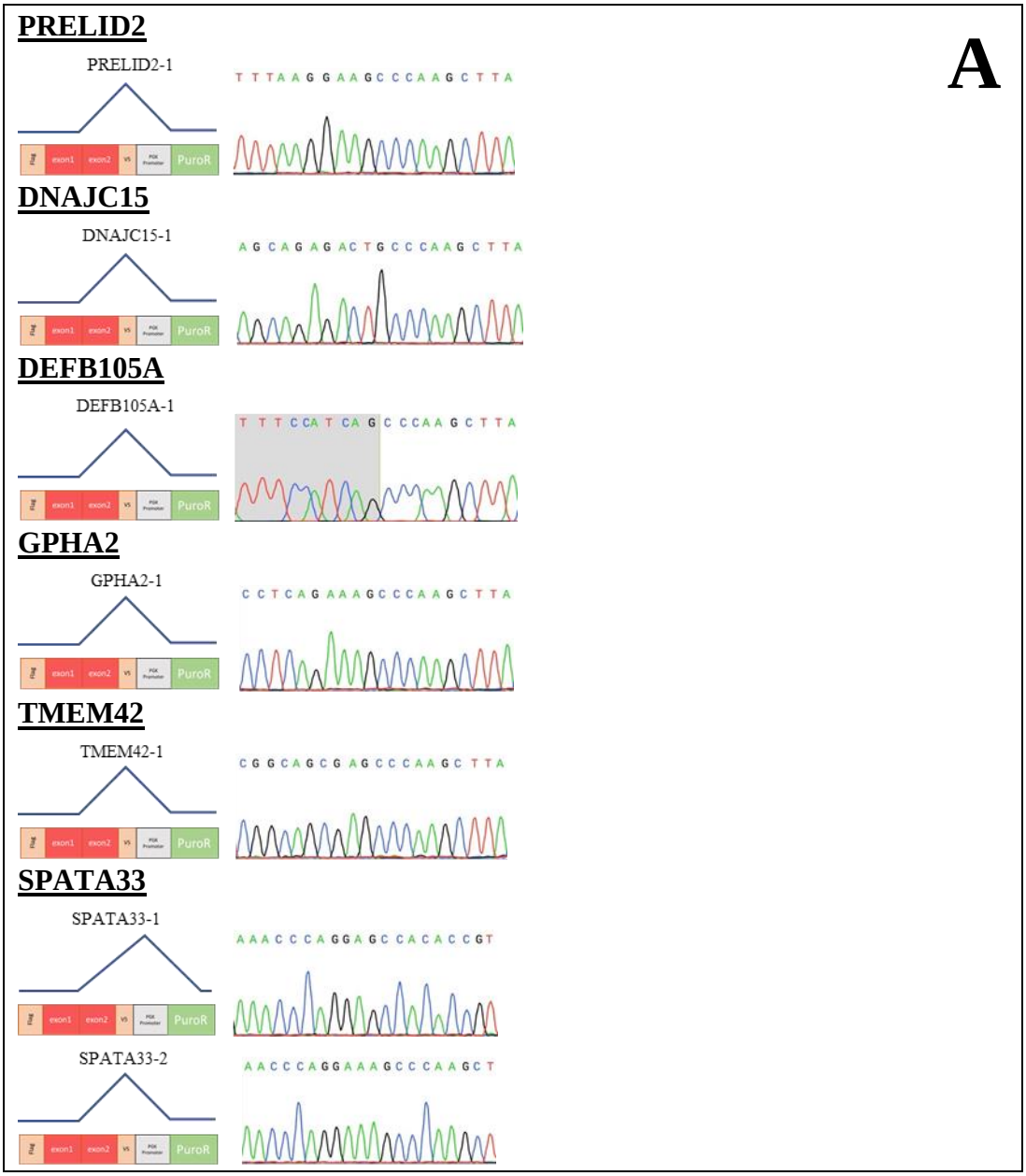
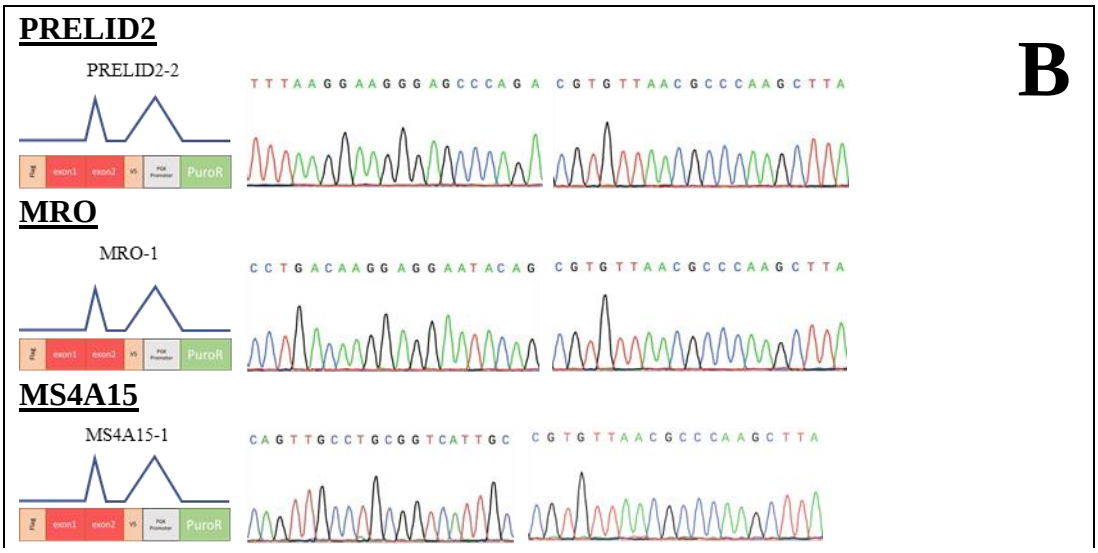


Figure 5. Verification of the false splicing on the 16 PFSGs. (A) The non-splicing PCR product has an expected size of around 1.5 kb, whereas the observed size of less than 1 kb indicated the presence of false splicing. (B) The schematic diagram of one of the PFSGs, AKAIN1V2, which was cloned into pLVX-3xFLAG-V5 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.

With the observed mis-splicing on human FACI gene, it was determined that the presence of 5' splicing site (AG|GTAAG) on the gene's exon-exon junction contributed to such phenomena. This discovery leads to the speculation that other genes having 5' splicing site at their exon-exon junction would also undergo mis-splicing. As the GT subtype accounts for around 99% of 5' splicing sites in human, bioinformatic analysis was conducted on genes having "X|GTAAG" or "X|GTGAG" since these sequences are the most conserved splicing sites. Results showed that 16 genes, 14 genes from human and 2 genes from mouse, have high probability of mis-splicing and proven by the RT-PCR results on figure 5 where the amplified products from mRNA were shorter than expected. For convenience, the 16 genes are termed PFSGs (Pro-False-Splicing Genes).


A
PRELID2
PRELID2-1
T T T A A G G A A G C C C A A G C T T A
DNAJC15
DNAJC15-1
A G C A G A G A C T G C C C A A G C T T A
DEFB105A
DEFB105A-1
T T T C C A T C A G C C C A A G C T T A
GPHA2
GPHA2-1
C C T C A G A A A G C C C A A G C T T A
TMEM42
TMEM42-1
C G G C A G C G A G C C C A A G C T T A
SPATA33
SPATA33-1
A A A C C C A G G A G C C A C A C C G T
SPATA33-2
A A C C C A G G A A A G C C C A A G C T
PuroR


B
PRELID2
PRELID2-2
T T T A A G G A A G G G A G C C C A G A C G T G T T A A C G C C C A A G C T T A
MRO
MRO-1
C C T G A C A A G G A G G A A T A C A G C G T G T T A A C G C C C A A G C T T A
MS4A15
MS4A15-1
C A G T T G C C T G C G G T C A T T G C C G T G T T A A C G C C C A A G C T T A
PuroR

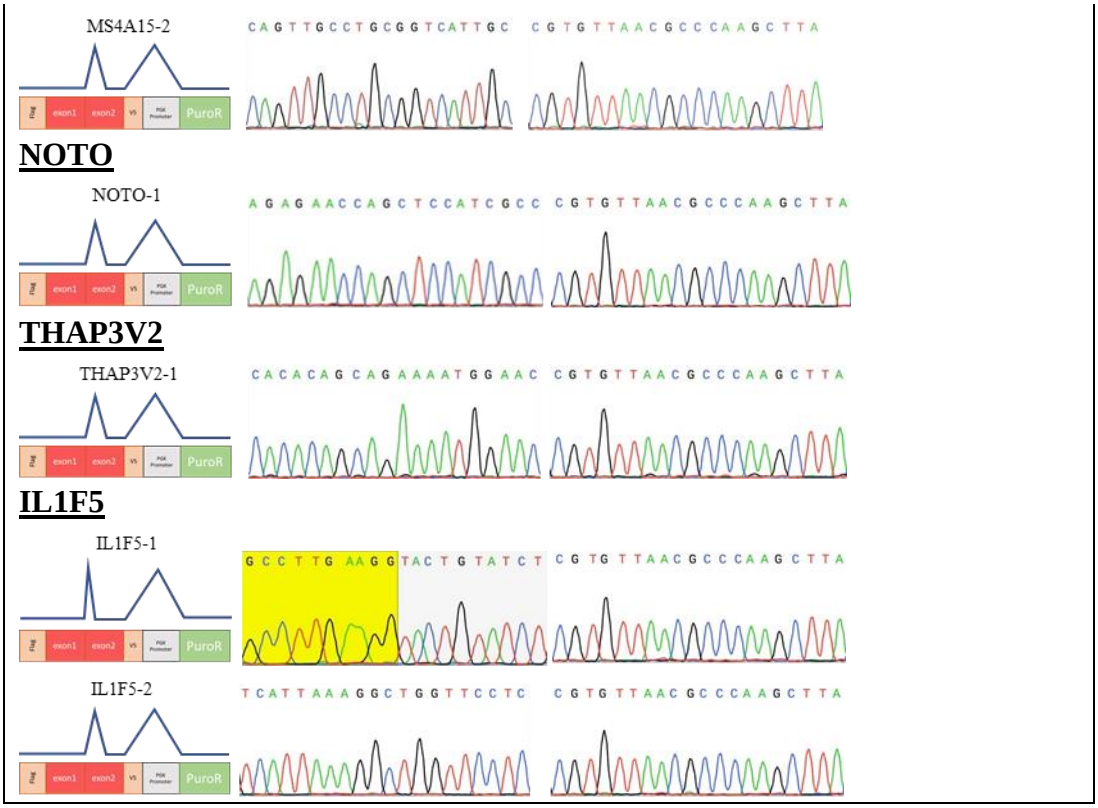
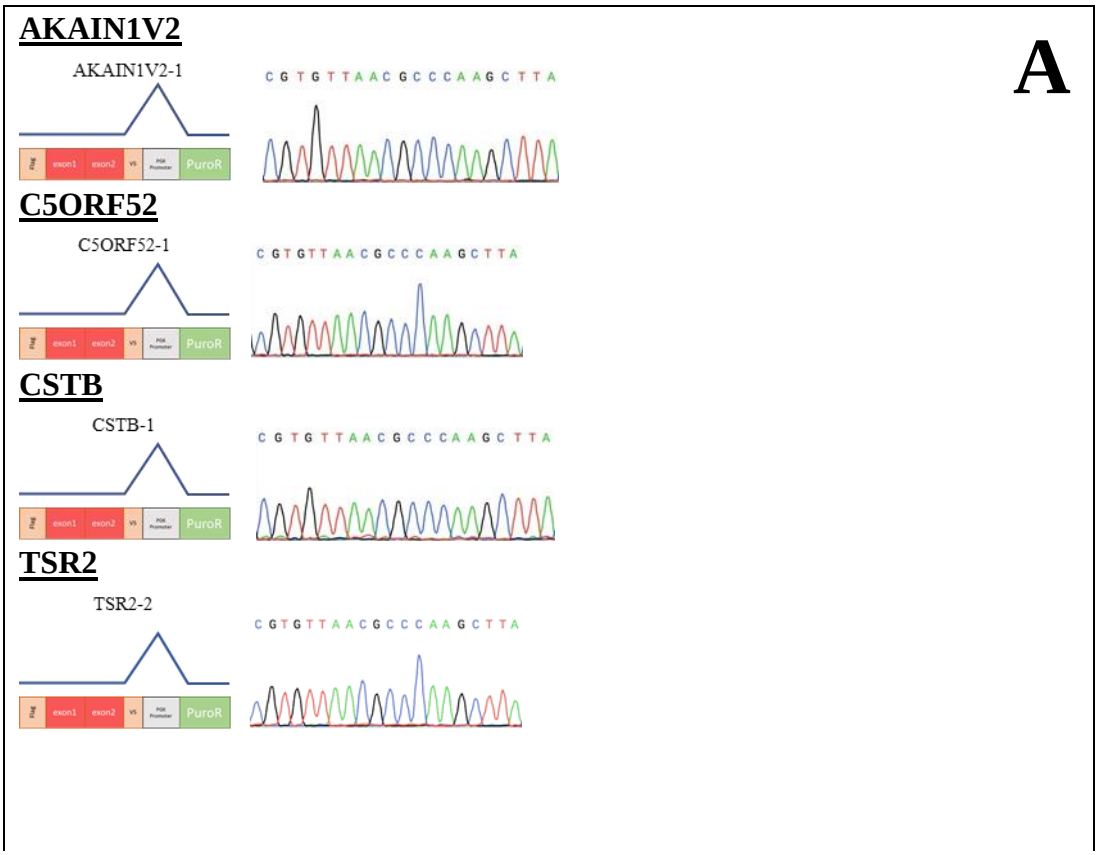
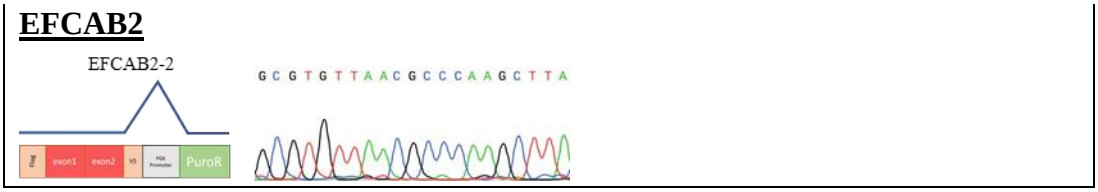
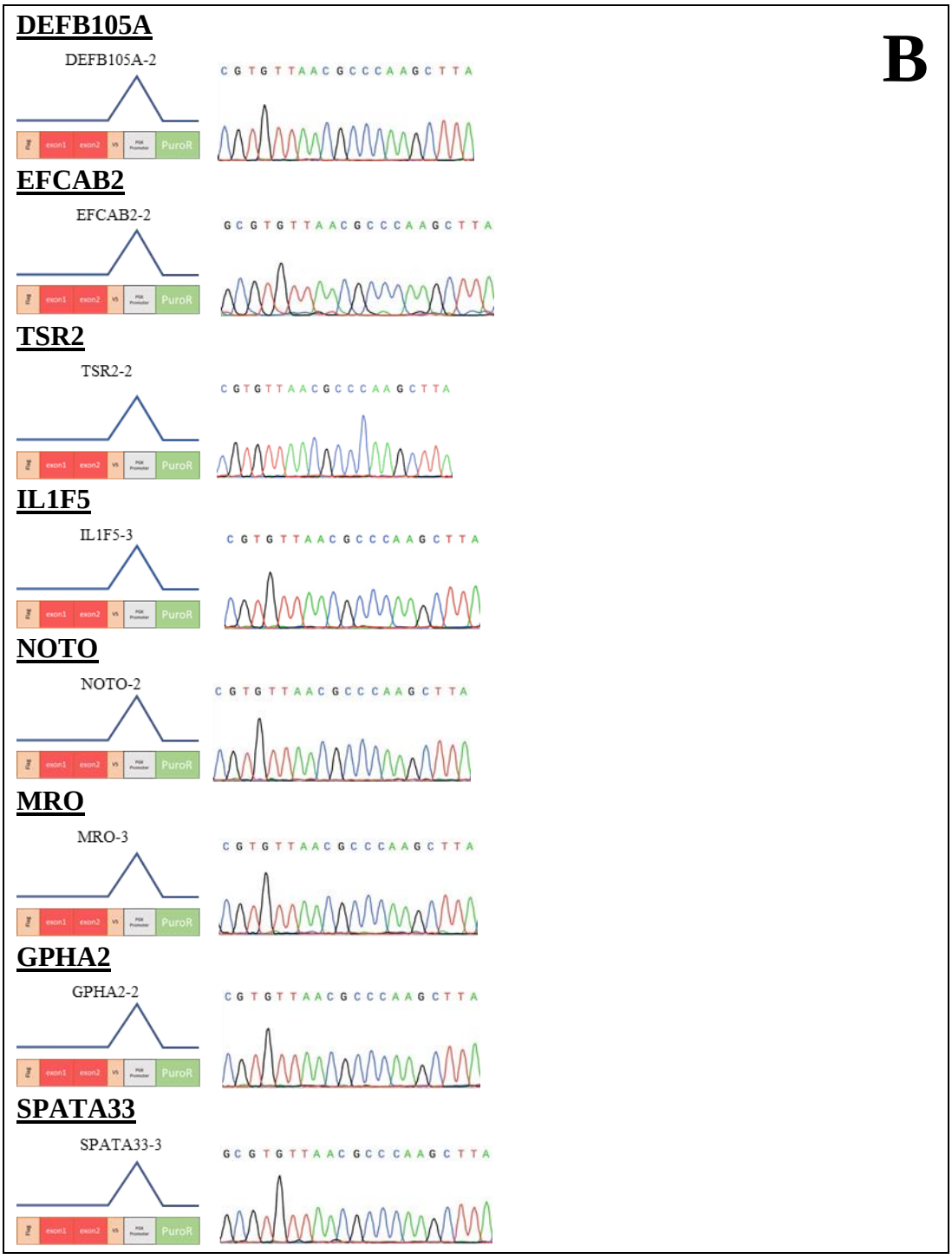


Figure 6. Schematic diagram and chromatography of PFSGs with exon-exon junction as false splicing site. (A) 6 PFSGs with E1-E2 junction as splicing donor site and same acceptor site on PuroR. (B) 6 PFSGs with E1-E2 junction as splicing donor site and acceptor site on their coding sequences.




EFCAB2
EFCAB2-2
G C G T G T T A A C G C C C A A G C T T A
exon1
exon2
PuroR


B
DEFB105A
DEFB105A-2
C G T G T T A A C G C C C A A G C T T A
EFCAB2
EFCAB2-2
G C G T G T T A A C G C C C A A G C T T A
TSR2
TSR2-2
C G T G T T A A C G C C C A A G C T T A
IL1F5
IL1F5-3
C G T G T T A A C G C C C A A G C T T A
NOTO
NOTO-2
C G T G T T A A C G C C C A A G C T T A
MRO
MRO-3
C G T G T T A A C G C C C A A G C T T A
GPHA2
GPHA2-2
C G T G T T A A C G C C C A A G C T T A
SPATA33
SPATA33-3
G C G T G T T A A C G C C C A A G C T T A
exon1
exon2
PuroR

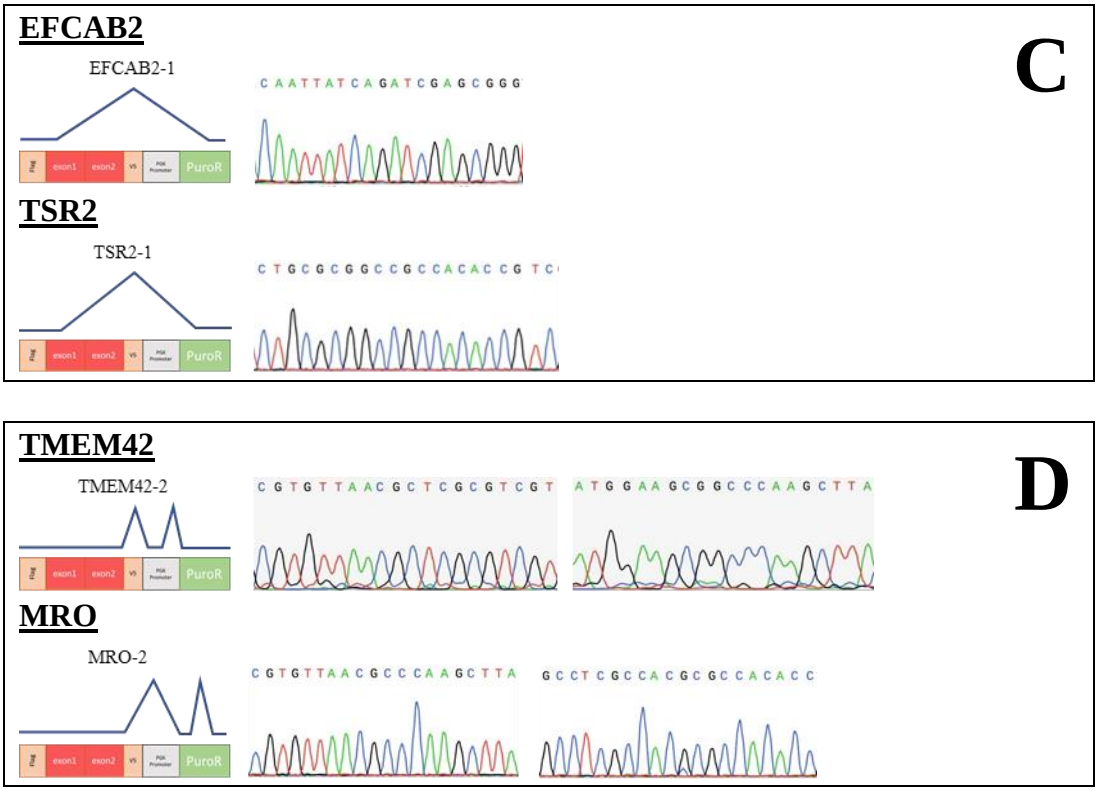
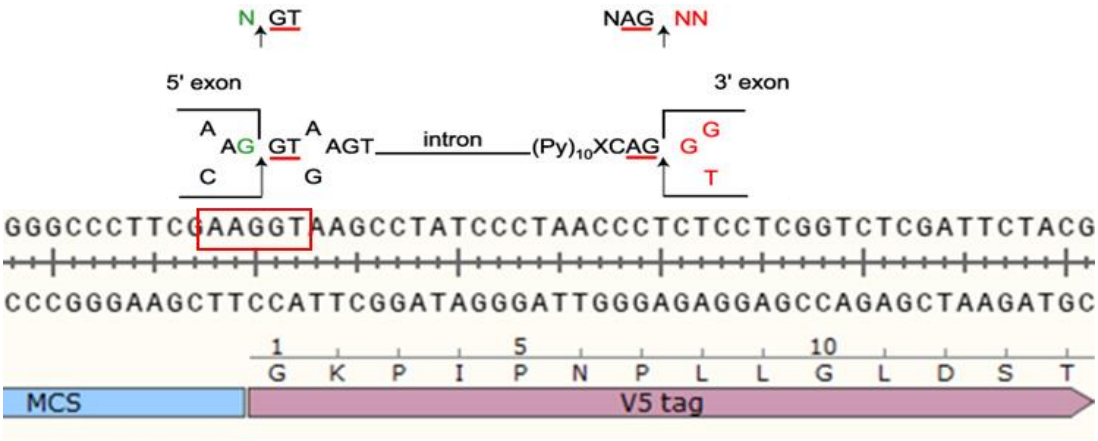


Figure 7. Schematic diagram and chromatography of PFSGs with V5-tag and coding sequence as false splicing site. (A) 5 PFSGs with splicing donor site only on V5-tag and acceptor site on PuroR. (B) PFSGs with E1-E2 junction splice sites which was also found to be mis-spliced on V5-tag. (C) PFSGs having splicing donor site at their coding sequence and PuroR as the acceptor site. (D) TMEM42 has V5-tag as splicing donor site and sequence on PGK promoter as the acceptor site, whereas in MRO, mis-splicing occurs randomly on PuroR besides V5-tag.

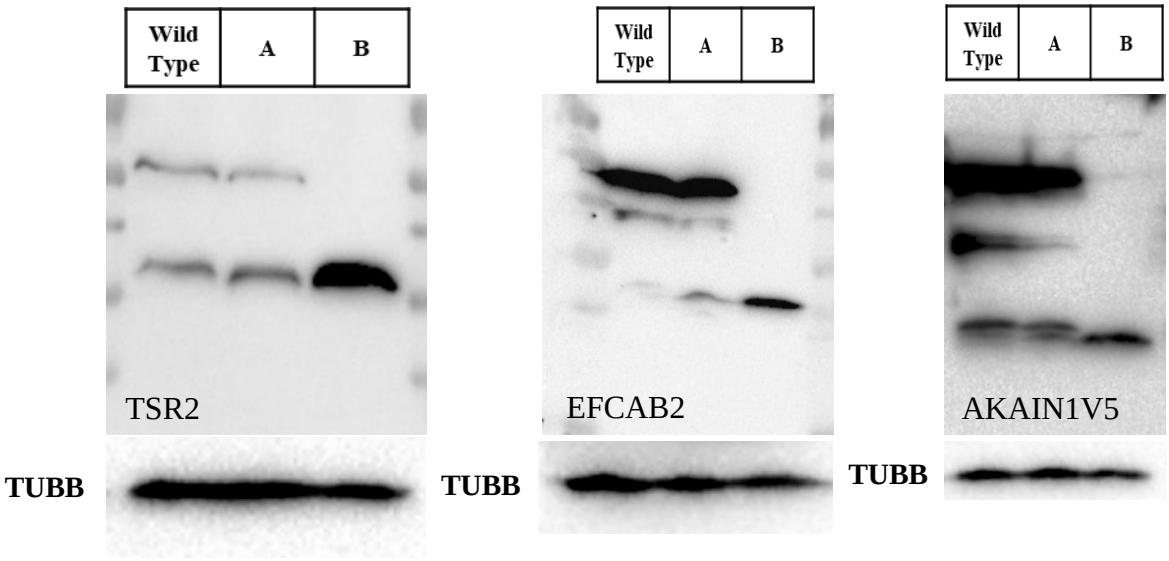
To further analyse the mis-splicing location of PFSGs, TA cloning and gel clean was conducted to purify the products for sequencing. The sequencing results show that among 16 genes, 11 genes had false splicing on the exon-exon junction. The 11 genes were further classified into two groups, with the first group of genes (DNAJC15, DEFB105A, GPHA2, TMEM42, SPATA33, and PRELID2) having acceptor site only on PuroR (Figure 6A), and the second group (MRO, MS4A15, NOTO, THAP3V3, IL1F5 and PRELID2) having another acceptor sites on their coding sequences (Figure 6B). Besides coding sequences, mis-splicing was also found on the V5-tag of the lentiviral plasmid construct. Instead of exon-exon junctions, genes such as AKAIN1V2, C5ORF52, CSTB, TSR2, and EFCAB2 underwent splicing with V5-tag as the splicing donor site and PuroR as the acceptor site (Figure 7A). For TSR2 and EFCAB2, the splicing donor site was present on the coding sequences with PuroR as the acceptor site (Figure 7C). Despite multiple splicing sites for a gene, exon-exon junction splicing was found to be strongly produced as shown by the PCR band results of DNAJC15, DEFB105A, GPHA2, TMEM42, and PRELID2 (Figure 5).



pcDNA3.1/V5-His and pEF6 V5-His

Figure 8. 5' splicing site (AAG|GT) on V5-tag sequence in pcDNA3.1/V5-His and pEF6 V5-His construct.

Based on previous analysis, we also found that V5-tag contributed to the mis-splicing of 14 out of 16 PFSGs. As shown in figure 8, V5-tag has "G|GTAAG" sequence which resembles a potential donor site. Since many research copy the V5-tag sequence from the widely-used commercial vectors (pcDNA 3.1/V5-His, pcDNA4/V5-His and pEF6V5-His) in designing V5-tagged gene expression plasmid, this may increase the risks of gene mis-splicing during expression. Nevertheless, further investigation using non PFSGs should be explored to see if V5-tag splicing could still happen.



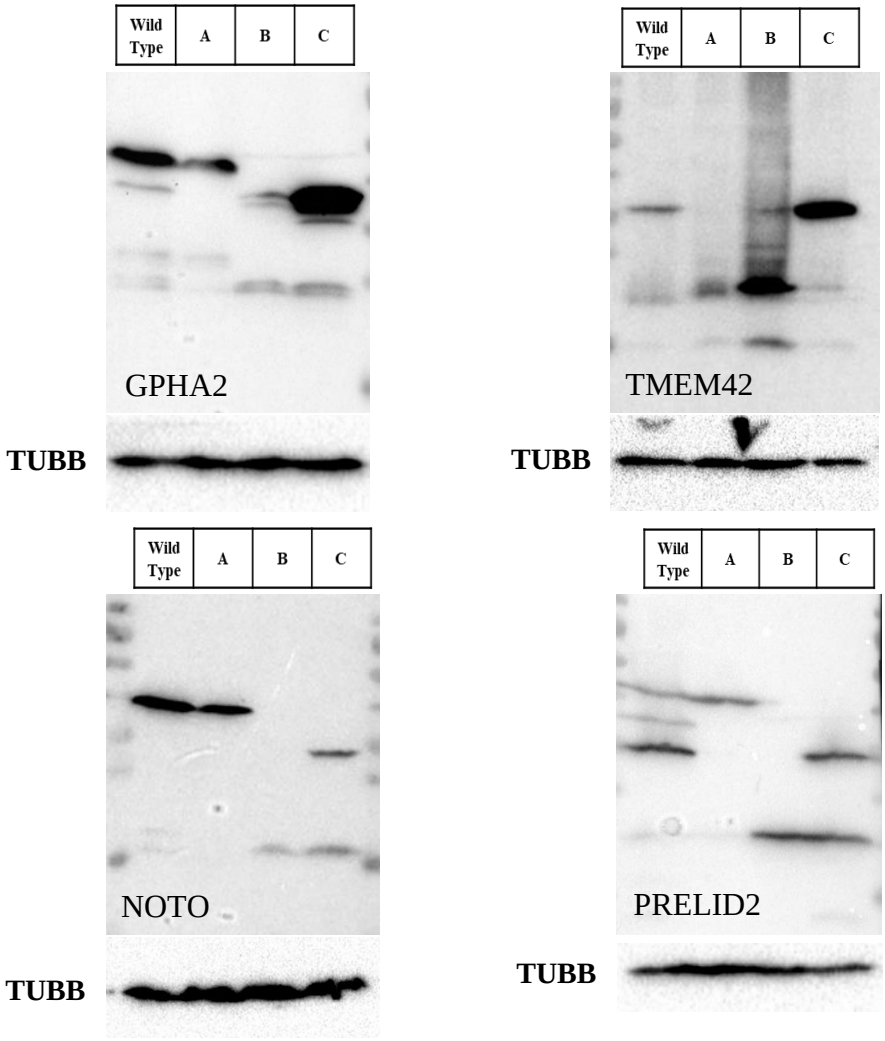


Figure 8. Mutation of the 5' splicing donor site at exon-exon junction and/or V5-tag of the genes. Wild type represents unmutated splice sites, A represents exon-exon junction mutation, B represents exon-exon junction and V5-tag mutation, and C represents V5-tag mutation. To ensure the same quantity of different samples' protein expression in the transfected cell lines, TUBB (Tubulin Beta Class I) expression served as the control.

To validate the mis-splicing phenomena of PFSGs, we performed silent mutation towards the potential splicing donor site on the exon-exon junction and/or V5-tag sequence of the genes followed by transfection into HEK293T or AML12 cells and subsequently protein extraction. Based on the immunoblot result, mutation of splicing donor site at exon-exon junction in TSR2, EFCAB2 and AKAIN1V2 genes produced the exact same size and number of protein bands as the wild type, whereas mutation on both exon-exon junction and V5-tag created a single protein band. This means that the mis-splicing of these genes happened on the V5-tag and the mutation on V5-tag sequence successfully inhibited the occurrence of mis-splicing. In GPHA2 and TMEM42, V5-tag mutation did not remove the splicing site as shown with several bands matching the wild type. Meanwhile, a single protein band was observed on the exon-exon junction mutation which verified the splice site on these genes' exon junction. For NOTO and PRELID2, mutating either

exon-exon junction or V5-tag mutation only deleted partial splice sites as shown by different size of bands between the two mutations corresponding to the wild type. However, the mutation on exon-exon junction and V5-tag produced a single band of protein, hence indicating the presence of splice sites on both location for these two genes.

Overall, our findings indicated that mis-splicing is likely to occur during foreign gene delivery or overexpression of genes if exon-exon junction has sequence which matches the 5' splicing donor motif, and the 3' splicing site is located on either the open reading frame regions of the plasmid cassette expression or the downstream vector regions that contain 3' splicing acceptor motif. Since we utilized lentiviral and retroviral vectors, the transgene expression cassettes do not have polyadenylation signals for transcription termination. Hence, using these types of vectors for the transduction of PFSGs could result in mis-spliced mRNAs and mis-produced proteins. Although mis-splicing could occur on random sequences, the prevalence is low compared to exon-exon junction due to the lack of regulatory elements involved in spliceosome recruitment.

IV. Discussion

In this study, mis-splicing phenomena was found on PFSGs having splicing donor sites on the exon-exon junction. Since most of the mis-spliced mRNAs can be successfully translated into proteins and not degraded, this could raise an issue towards the normal function of target proteins. For instance, high expression level of mis-produced proteins could supress the expression of normal proteins. Furthermore, as these proteins share the same sequences as the target proteins, it is likely that the mis-produced proteins affect the normal function of the target proteins.

Besides exon-exon junction, false splicing was also observed on V5-tag sequence having potential splicing donor site of "XXG|GTAAG". This has resulted in the mis-splicing occurrence on 14 out of 16 PFSGs. The removal of V5-tag during mis-splicing could be detrimental as the protein produced from the transgene expression cassette cannot be visualized by V5-tag antibody. However, further research should be conducted to see if mis-splicing also happens on V5-tagged non PFSGs.

To ensure correct gene delivery and expression, performing synonymous mutation of the false splicing sites on the exon-exon junction and V5-tag could help to minimize mis-splicing. However, there are several aspects which should be considered beforehand, including the time and cost of mapping potential splicing donor sites on PFSGs as well as codon usage bias which may affect translation efficiencies and results.

The findings of this research serve a great purpose to discover the mis-splicing phenomena in PFSGs as well as V5-tag which are widely used in biomedical research. With mis-produced proteins being expressed, this could have a huge impact on the accuracy of many research outcomes focusing on gain-of-function studies. Hence, before designing experiments related with the usage of cDNAs, it is better to perform empirical test to determine if potential splice sites are present on the genes and downstream sequences of the plasmid construct so that mis-splicing could be avoided.

V. References

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

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

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.

Leader, Y., Maor, G. L., Sorek, M., Shayevitch, R., Hussein, M., Hameiri, O., . . . Ast, G. (2021). The upstream 5′ splice site remains associated to the transcription machinery during intron synthesis. *Nature Communications*, 4545. doi:10.1038/s41467-021-24774-6

Newman, A. (1998). RNA splicing. *Cell, 9*(25). doi:10.1016/S0960-9822(98)00005-0

Wang, Y., Liu, J., Huang, B., Xu, Y.-M., Li, J., Huang, L.-F., . . . Wang, X.-Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports, 3*(2), 152-158. doi:10.3892/br.2014.407

Wilkinson, M. E., Charenton, C., & Nagai, K. (2020). RNA Splicing by the Spliceosome. *Annual Review of Biochemistry, 89*, 1-30. doi:10.1146/annurev-biochem-091719-064225

Will, C. L., & Lührmann, R. (2011). Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology, 3*(7), 1-23. doi:10.1101/cshperspect.a003707