

Table S1: Annotations made on transcript introns.

Annotation	Details
Candidate ESE Count	Counts the total number of candidate Exonic Splicing Enhancer hexamers in exon sequences.
G triplet count	Counts the number of "GGG" sequences - a putative intronic splicing enhancer (ISE) - in an intron.
G triplet pairs count	Counts the number of "GGG" sequence pairs of 0 to 4 nucleotides apart in an intron.
Polypyrimidine Tract	Records the sequence, length and percentage of pyrimidines between 3' intron end and the Branch Point.
GC %	Records the percentage of nucleotides that are C or G in an intron.
5' Splice Site	Records the 6 base sequence at the 5' end of an intron and whether it matches the consensus GTRAGT.
3' Splice Site	Records the 3 base sequence at the 3' end of an intron and whether it matches the consensus YAG.
Branch Point	Records, where matching consensus YTN <u>A</u> Y, a sequence of 5 nucleotides between 21 and 34 bases of the 3' intron end.