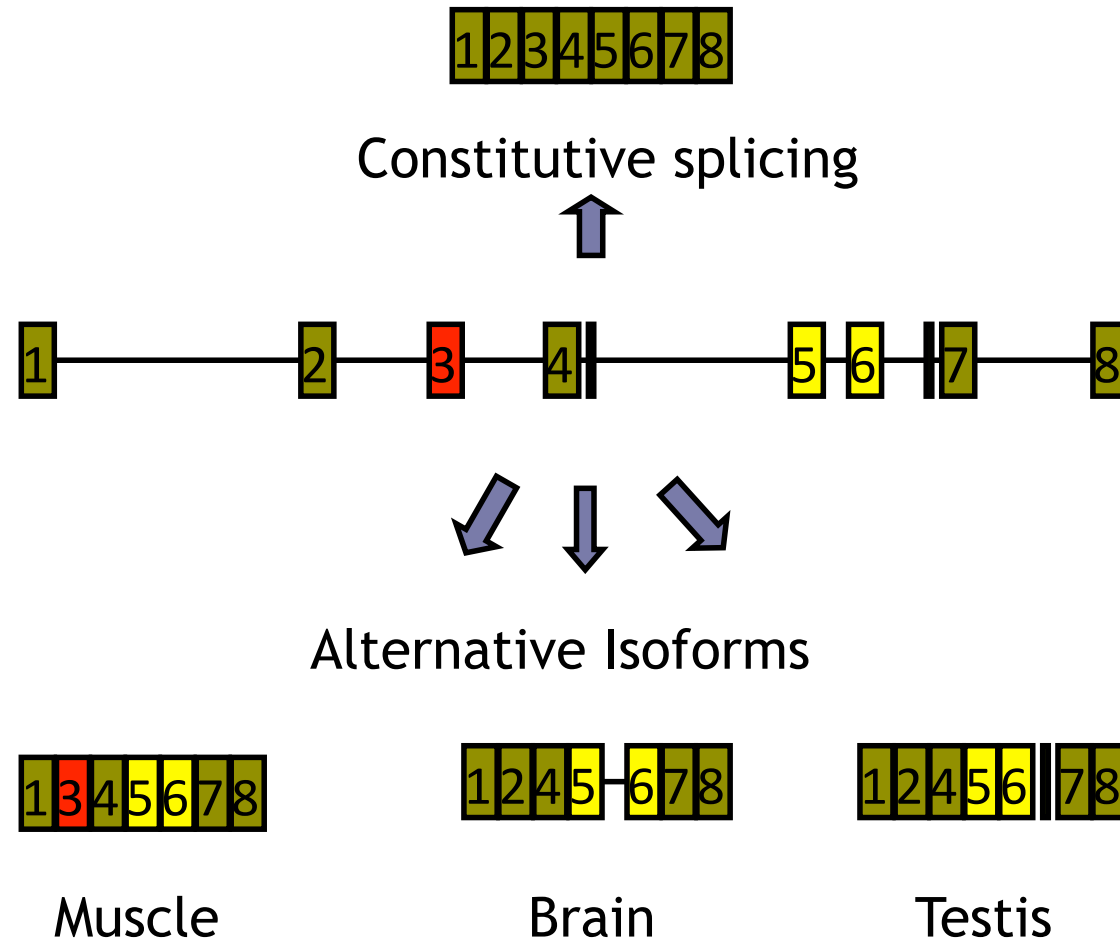


Visualization of the Effect of Unclassified Variant Sequences on Exonic ESE/ESS Content

Dr Sushma Grellscheid
Institute of Genetic Medicine
Newcastle University

Constitutive and Alternative Splicing



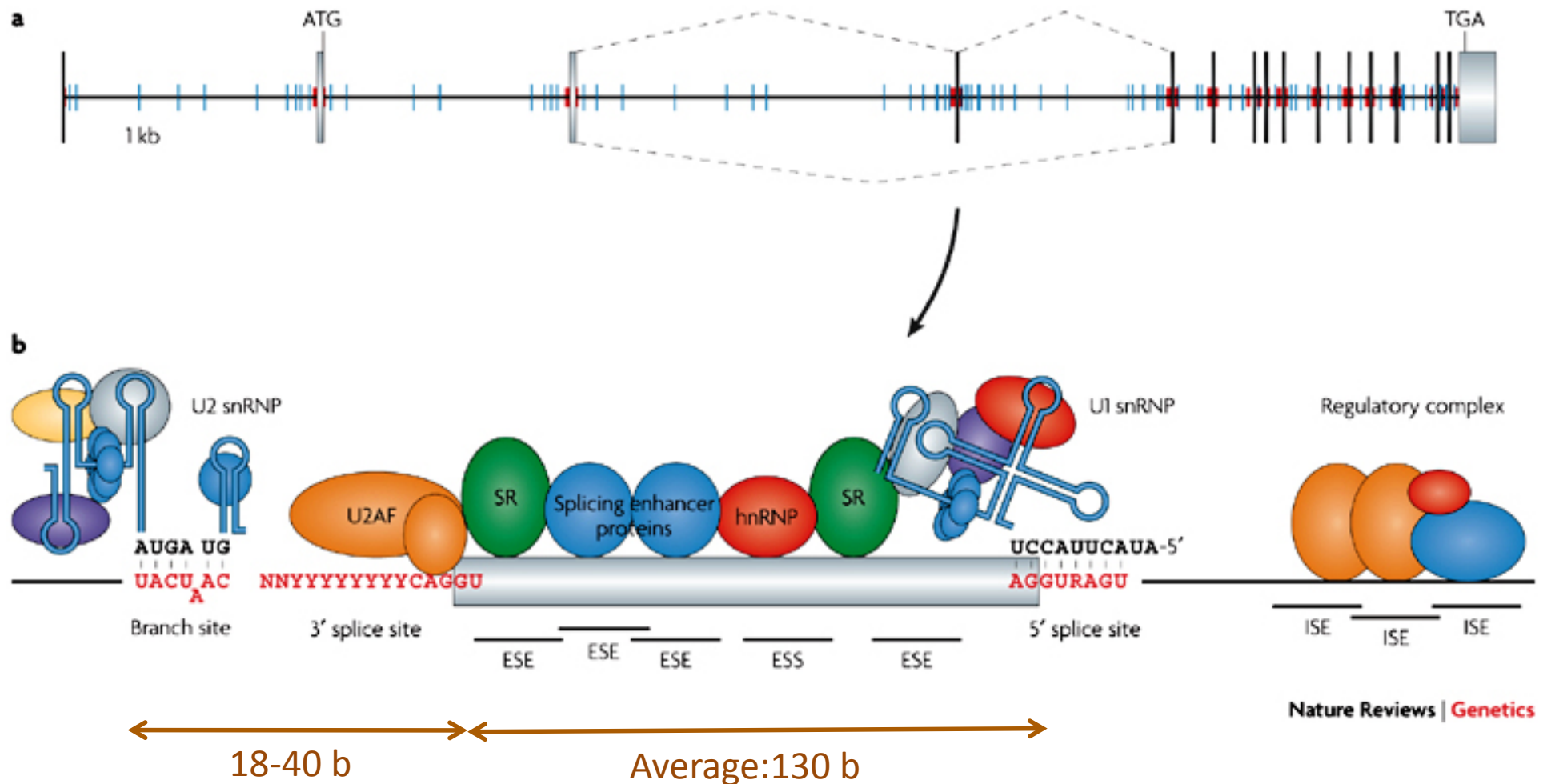
Adapted from Elliott and Grellscheid, 2005

Frequency of splicing defects in some common human genetic disorders

Gene	Disease/phenotype	Splicing/total mutation* (%)
<i>ATM</i>	Ataxia telangectasia	18
<i>BRCA1</i>	Breast cancer predisposition	9
<i>CADM</i>	Medium chain acyl CoA dehydrogenase deficiency	10
<i>CFTR</i>	Cystic fibrosis	14
<i>DMD</i>	Duchenne muscular dystrophy	9
<i>HBA1/2</i>	Blood disorders (thalassaemias, anaemia etc)	3
<i>HBB</i>	Blood disorders (thalassaemias, anaemia etc)	10
<i>HPRT</i>	Hypoxanthine-guanine phosphoribosyltransferase 1 deficiency	15
<i>IKBKAP</i>	Dysautonomia, familial	33
<i>MAPT</i>	Frontotemporal dementia and Parkinsonism	33
<i>MLH1</i>	Colorectal cancer	18
<i>MSH2</i>	Colorectal cancer	9
<i>NF1</i>	Neurofibromatosis type 1	19
<i>NF2</i>	Neurofibromatosis type 2	22
<i>RHO</i>	Retinitis pigmentosa	3
<i>SMN1/2</i>	Spinal muscular atrophy	4
<i>WT1</i>	Wilms tumour	11

*Data calculated from the public Human Gene Mutation Database (24/11/2008).

Splicing Mechanism



Wang & Cooper, Nature Reviews Genetics 8, 749-761

Splicing Mechanism

Splice site sequences are bound by core splicing factors:

U1, U2 snRNPs, BBP, and U2AF65/35

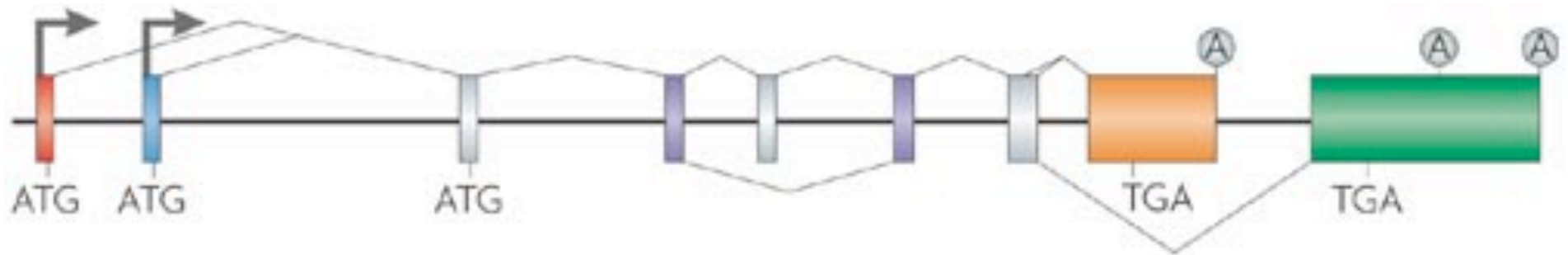
ESE and ESS sequences are bound by SR proteins and hnRNP family proteins.

Regulatory sequences are very degenerate:

GAAGAA: ideal binding site
GATAGA: also binds functionally

Human Genes are Complex

Pathogenic mutations occur in coding as well as regulatory regions

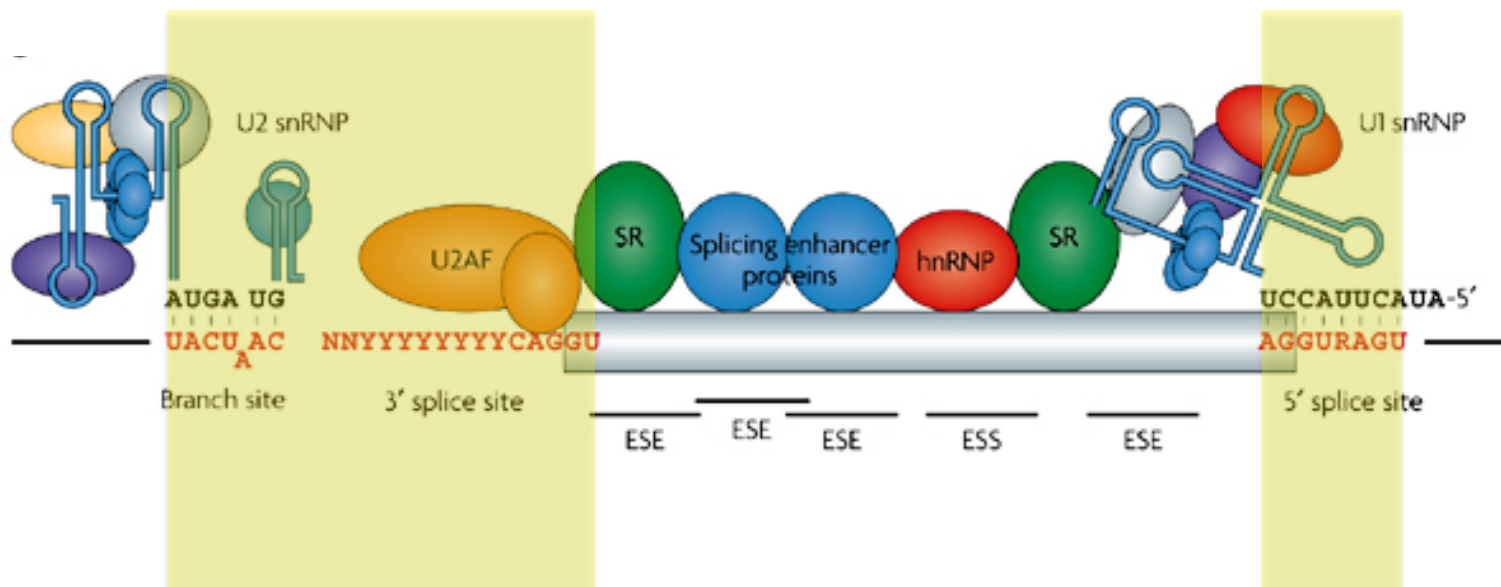


Considering mutations from an RNA processing perspective.

Splicing Mutations-1

Splice-site Mutations

Mutations in acceptor (3'ss) and donor (5'ss) sites
Well characterised matrices exist.
Not the focus of this talk.

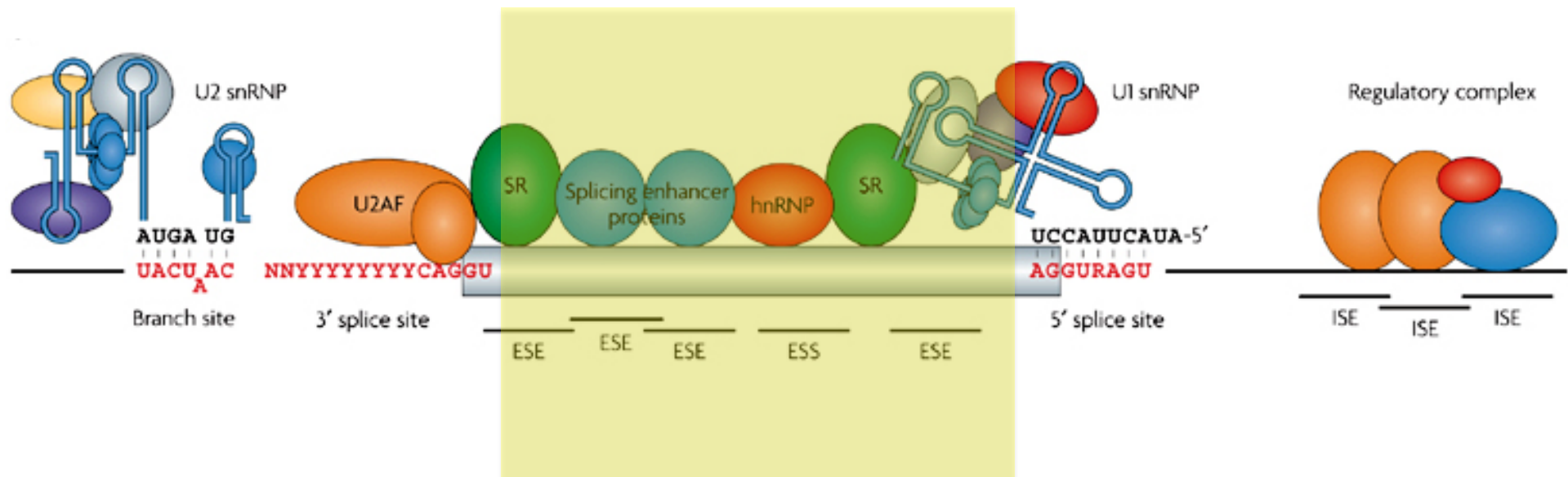


Splicing Mutations-2

Regulatory Element Mutations

ESE or ESS mutations: Unclassified variants or often (mis-) classified as synonymous or missense mutations

Numerous ESE/ESS predictive tools, which one to use?!



Identification of ESE and ESS binding sites

Molecular Methods

SELEX: RNA-protein binding studies

Functional splicing assays to identify sequences that enhance or silence splicing.

Example: ESE-Finder

Bioinformatic Methods

Identification of 6-8 mers over or under represented in exons versus Introns.

Example: RescueESE, PESE, Goren ESRs,

Challenges:

1. Motifs display little overlap between methods, and only partial overlap within methods.
2. Splicing is a binary decision- how does the cell integrate the different facets of information?

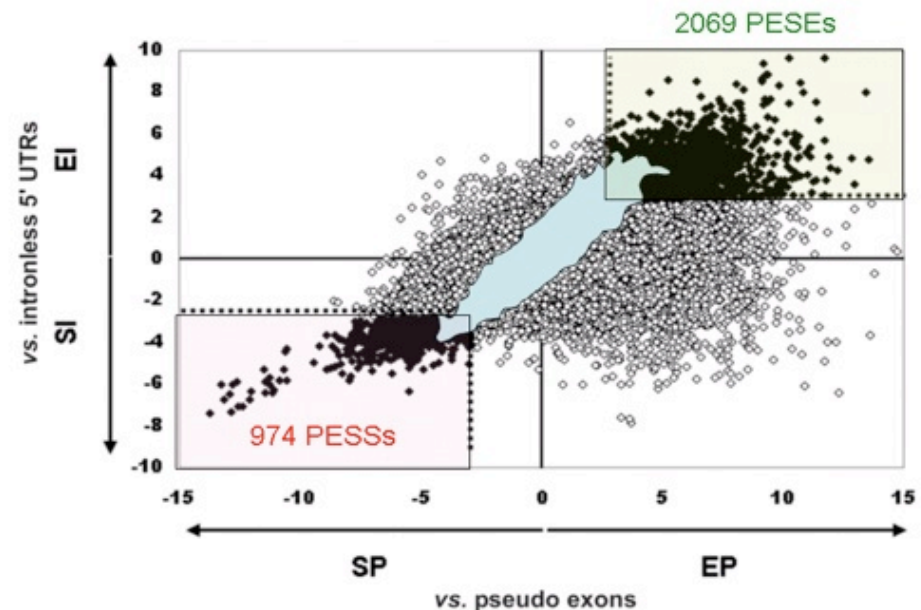
Bioinformatic identification of Exonic Regulatory Sequences

Methods comparing the frequency of all possible 'x-mers' between different populations:

Identified hexamers that are enriched in:

- i) exons vs introns
- ii) exons with weak vs. strong splice sites (RescueESE, Fairbrother et al)
- iii) Conserved exons (ESRs, Goren et al)
- iv) Removal of coding bias: (PESEs, Zhang and Chasin)

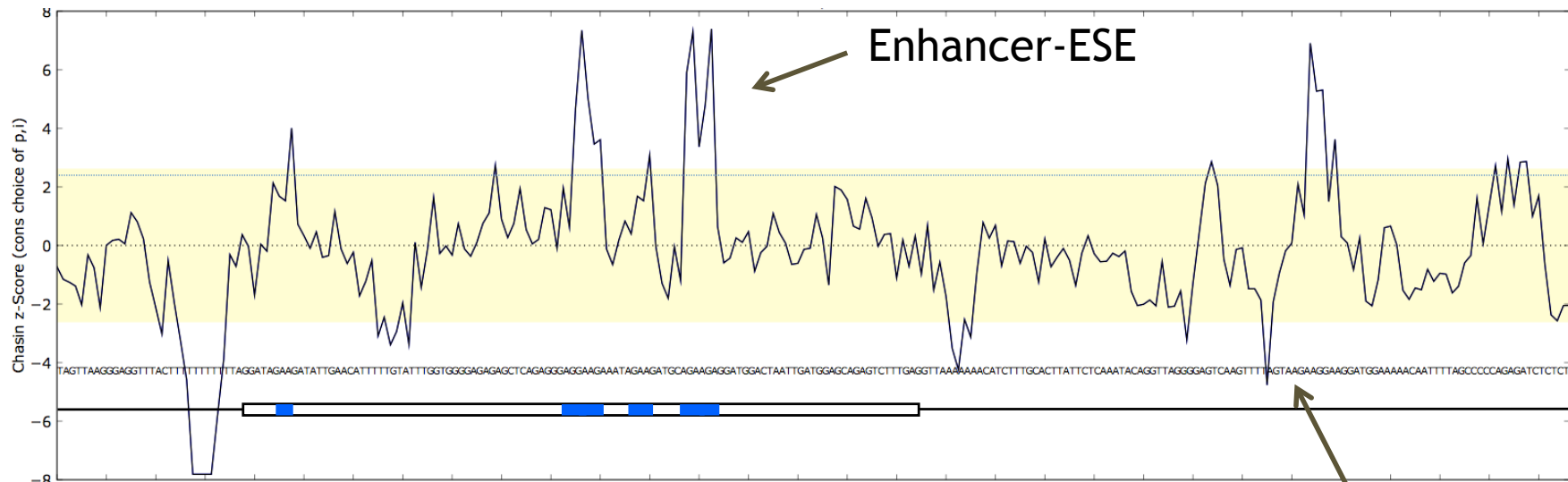
- a. Non-coding real exons vs adjacent pseudo exons.
- b. Non-coding real exons vs UTR of intronless genes.



EP = positive z-score of non-coding exons vs. pseudo exons
SP = negative z-score of non-coding exons vs. Pseudo exons (abs. value)
EI = positive z-score of non-coding exons vs. 5' UTR of intronless genes
SI = negative z-score of non-coding exons vs. 5' UTR of intronless genes (abs. value)

Exon analyser (and mutagenesis designer)

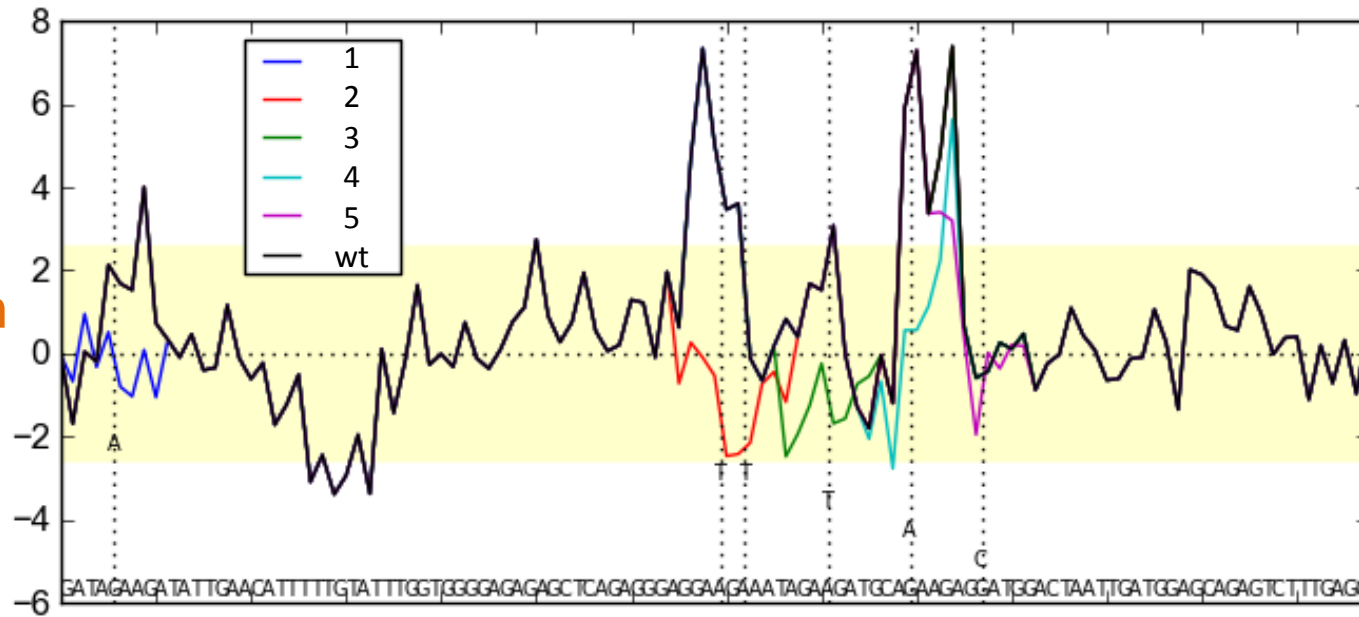
In-house software written in Python



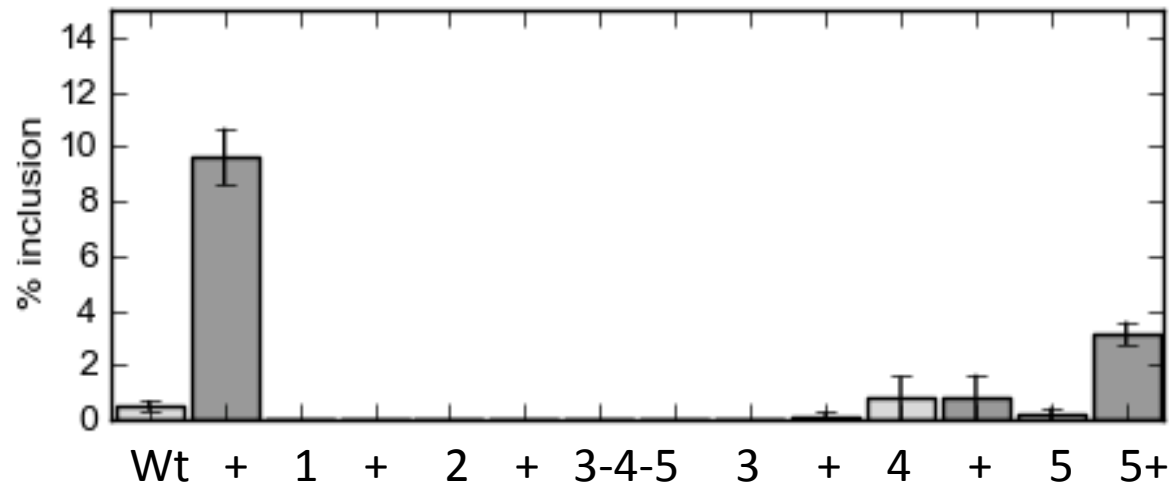
- **Input:**
sequence in fasta format.
- **Output:**
Plots ESE/ESS score per 8 nt in a sliding window.
Interactive for mutagenesis.

Validation-Example 1

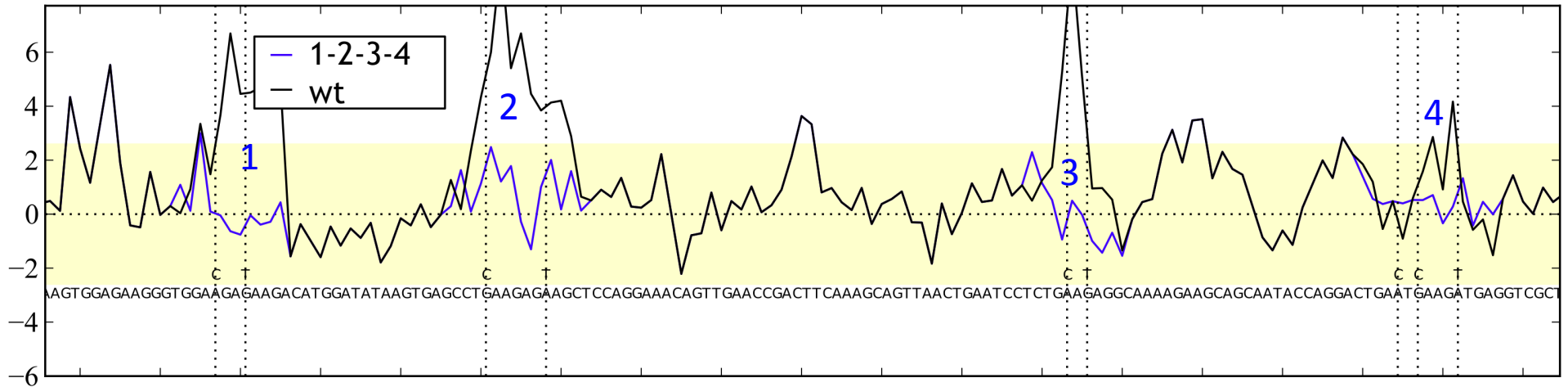
Prediction



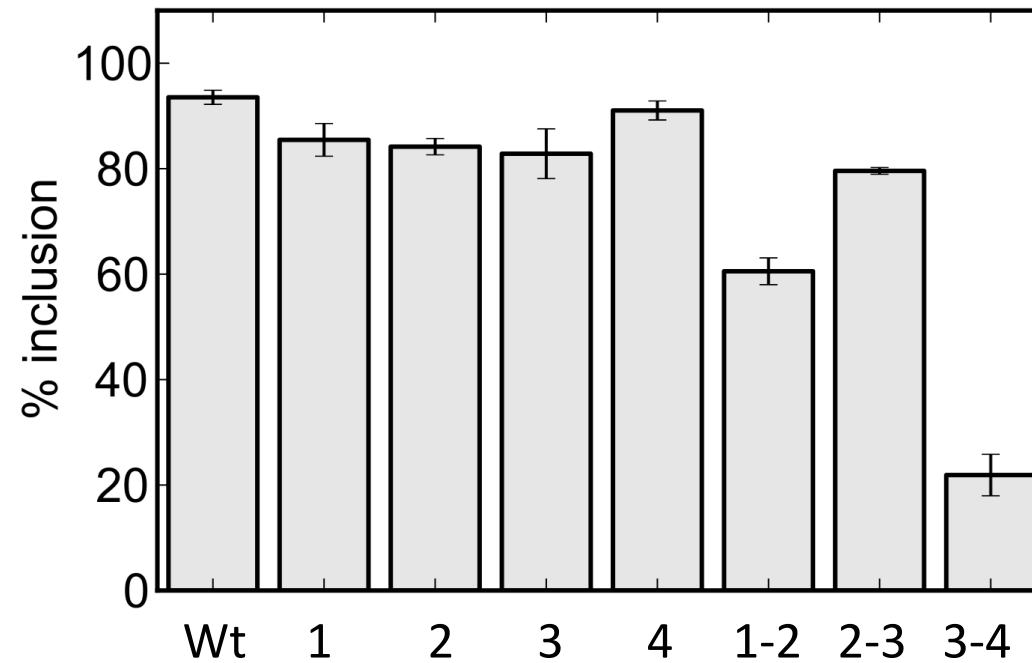
Minigene Validation



Validation-Example 2



Prediction



Minigene
Validation

Summary

Splicing defects are a common genetic basis of disease.

These may result from mutations in the splice sites or in Enhancer or Silencer sequences.

Various online methods are available to evaluate exonic mutations including ESEFinder, RescueESE, Scroogle. The results of the various methods do not correlate very well.

We have developed a software that uses the PESE/PESS matrix, and plots ESE/ESS for a given exon and is interactive to aid in designing mutagenesis.

Method validates very well in a small dataset but needs further validation.

Future Work

Seeking collaborator to carry out validation of this method.

RT-PCRs from patient RNAs with exonic mutations
(Unclassified Variants, Synonymous, mis-sense).

Comparison between predicted outcome and experimental validation.



Institute of Genetic Medicine

Prof Sir John Burn
Prof David Elliott
Dr David Bourne

welcometrust