HExoSplice: a new software based on overlapping hexamer scores for prediction and stratification of exonic variants altering splicing regulation of human genes

Arnaud M. Lefebvre [1, 2, 4], Alexandra Martins [1, 3, 4], Karim Labrèche [1, 5], Vivien Deshaies [1, 5], Alan Lahure [1, 5], Pascaline Gaildrat [1, 3, 4] and Hélène Dauchel [1, 2]

[1]Normandie Université, France,
[2]LITIS EA 4108 Computer science, information processing and systems laboratory, University of Rouen,
[3]INSERM U1079 Genetics of cancer and neuropsychiatric diseases, University of Rouen,
[4]Institute for Research and Innovation in Biomedicine (IRIB), Normandy, [5]Bioinformatics Master, University of Rouen

Sequencing exonic regions in disease-associated genes became a common practice for mutation screening in medical genetics. Besides their potential impact at the protein level, **exonic sequence variations** can induce **aberrant splicing through disruption of Exonic Splicing Regulator elements (ESRs)**, leading to potential pathological effects. **HExoSplice implements a new scoring method** for the prediction of the splicing regulatory mutations through an effective and user-friendly web interface. HExoSplice could help to quickly prioritize a restricted number of variations before performing further investigations. This approach could be particularly relevant for the evaluation of the so called Variations of Unknown Significance (VUS). More generally, HExoSplice could contribute to the annotation and filtering strategies of genetics variations identified by whole or targeted exome sequencing.

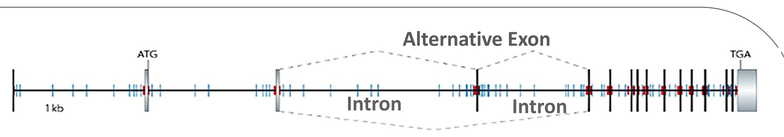
HExoSplice

Home | Run HExoSplice

ESRs and Splicing Regulation

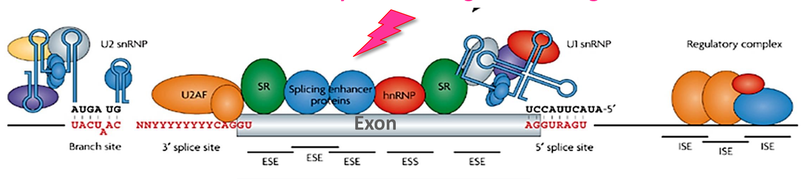
Exonic Splicing Regulator elements (ESRs) are described as 6-8 nucleotide motifs that regulate constitutive and alternative pre-mRNA splicing by recruiting trans-acting factors. Exonic mutations can disrupt ESRs and induce aberrant splicing.

A pre-mRNA : cis-elements within and around an alternative exon are required for its recognition and regulation.



Mutations in the pre-mRNA can decrease the efficiency of exon recognition and regulation

Spliceosome components : **Consensus 5′ and 3′ splice-sites sequences at the intron–exon boundaries and branch site** serve as binding sites for spliceosome components.



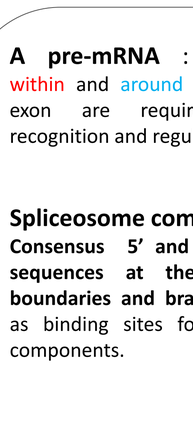
ESRs : ESEs and ESSs

Adapted from Wang & Cooper, Nature Reviews Genetics , 2007

Exons and introns contain diverse sets of enhancer and suppressor elements that refine bone fide exon recognition: **Exon splicing enhancers (ESEs)** recruit and stabilize binding of spliceosome components. **Exon splicing suppressors (ESSs)** repress exon usage. Some intronic splicing enhancers (ISEs) bind auxiliary splicing factors that are not normally associated with the spliceosome to regulate alternative splicing.

Scoring Method : overlapping hexamer-based score

HExoSplice takes advantage of two previously published studies [1, 2]. Firstly, it integrates individual scores of all possible RNA hexamers (4096) ranking their potential functions as ESRs. Secondly, HExoSplice implements the computation of a new promising global quantitative predictive value (*Total ESRseq score change*) for any given exonic single nucleotide substitution to alter potential splicing regulatory signals [1, 2].



List of ESE/ESS score for all 4096 hexamers [1]. E, positive ESRseq score; S, negative ESRseq score; N, null ESRseq score.

WT exonic sequence

Exonic variant ex : C>G

A given exonic variant in a given WT exonic sequence

The scoring method : calculation of the **net effect of the six hexamers overlapping the nucleotide substitution** (11 nucleotide stretch with 5 nucleotides upstream/downstream the position), both in the variant and wild type (WT) contexts, and then computing of the **differential value of *Total ESRseq score change*** (variant versus WT) [2].

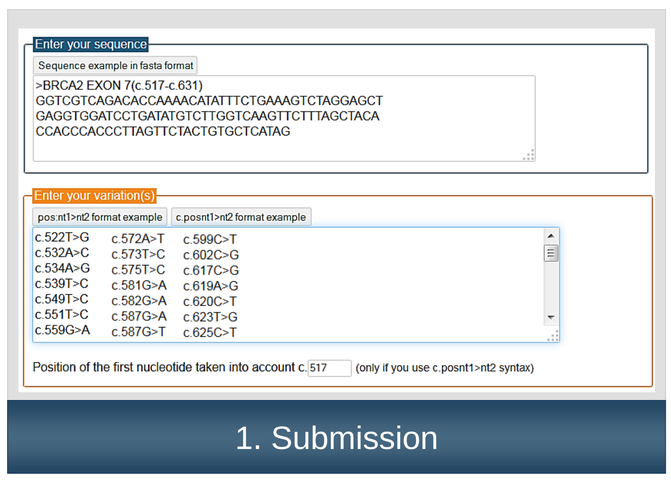
Total ESRseq score change values for stratification of exonic variant as potential splicing regulatory mutations

A positive *Total ESRseq score change* is interpreted as a potential increase in exon inclusion, while a *negative Total ESRseq score change* is interpreted as a potential increase in exon skipping.

References

1. Ke S, Shang S, Kalachikov SM, Morozova I, Yu L, Russo JJ, Ju J, Chasin LA: **Quantitative evaluation of all hexamers as exonic splicing elements**. Genome Res 2011, 21(8):1360-1374. PMID:21659425
2. Di Giacomo D, Gaildrat P, Abuli A, Abdat J, Frébourg T, Tosi M, Martins A: **Functional analysis of a large set of BRCA2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements**. Hum Mutat 2013, 34(11):1547-1557. PMID:23983145

Web interface



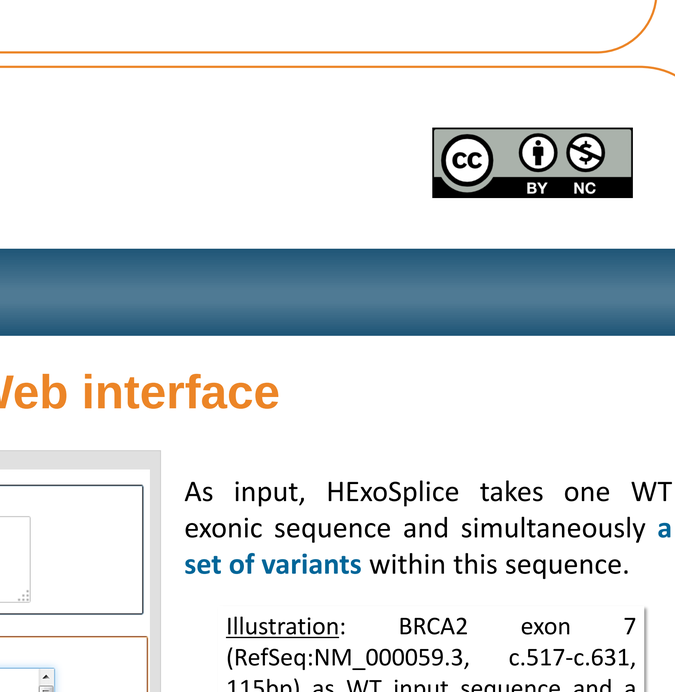
1. Submission

As input, HExoSplice takes one WT exonic sequence and simultaneously **a set of variants** within this sequence.

Illustration: BRCA2 exon 7 (RefSeq:NM_000059.3, c.517-c.631, 115bp) as WT input sequence and a dataset of 22 BRCA2 exon 7 single nucleotide substitutions (as described in [2]) in c. format.

After immediate **batch computing**, the output figures generated by HExoSplice allow a **quick inspection of the full dataset**. The results are displayed as **predictive maps of potential ESRs (a) and variation-induced alterations (b)**.

Illustration : Maps indicate positive values for 11 variations, predicting a potential increase in exon inclusion, and negative values for 11 other variations predicting a potential increase in exon skipping.



2. Output : full dataset

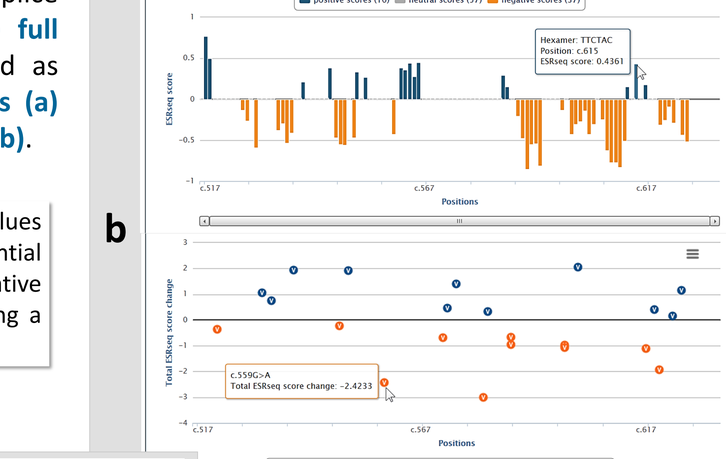
2. Output : full dataset

HExoSplice displays **comprehensive score tables which allow ranking** .

Illustration: As an example of prioritization of the 22 variations within the exon, the sorting was obtained by clicking on the column header Total ESRseq scores change.

Barplot representations (a) and table overview (b) of **ESRseq scores for the six overlapping hexamer** concerned by the variant position

Illustration: c.617C>G exhibiting a negative Total ESRseq score change (-1.1153) as net effect, the barplots show (i) a disruption of three potential ESEs (ii) the creation of a potential novel ESS (iii) and the strengthening of a potential ESS.



3. Output : variation detail