Splicy: a web-based tool for the prediction of possible alternative splicing events from Affymetrix probeset data

Rambaldi D (1), Felice B(1), Praz V (2), Bucher P (2), Guffanti A (1)

(1) The IFOM-IEO Campus, Via Adamello, 16 - 20139 Milano, Italy
(2) ISREC, Ch. des Boveresses 155, Epalinges (Switzerland)

Motivation

The Affymetrix technology is nowadays a well-established method for the detection of gene expression profiles in cancer research studies. As an example, a query with the keyword 'cancer' of the NetAffx Scientific Publications database (http://www.affymetrix.com/community/index.affx) results in the retrieval of 672 publications from 1996. However, changes in gene expression levels are not the only aspect of importance in the link between genes and cancer. The existence of gene isoforms specifically linked with cancer or apoptosis is increasingly found in literature (1,2).

Methods

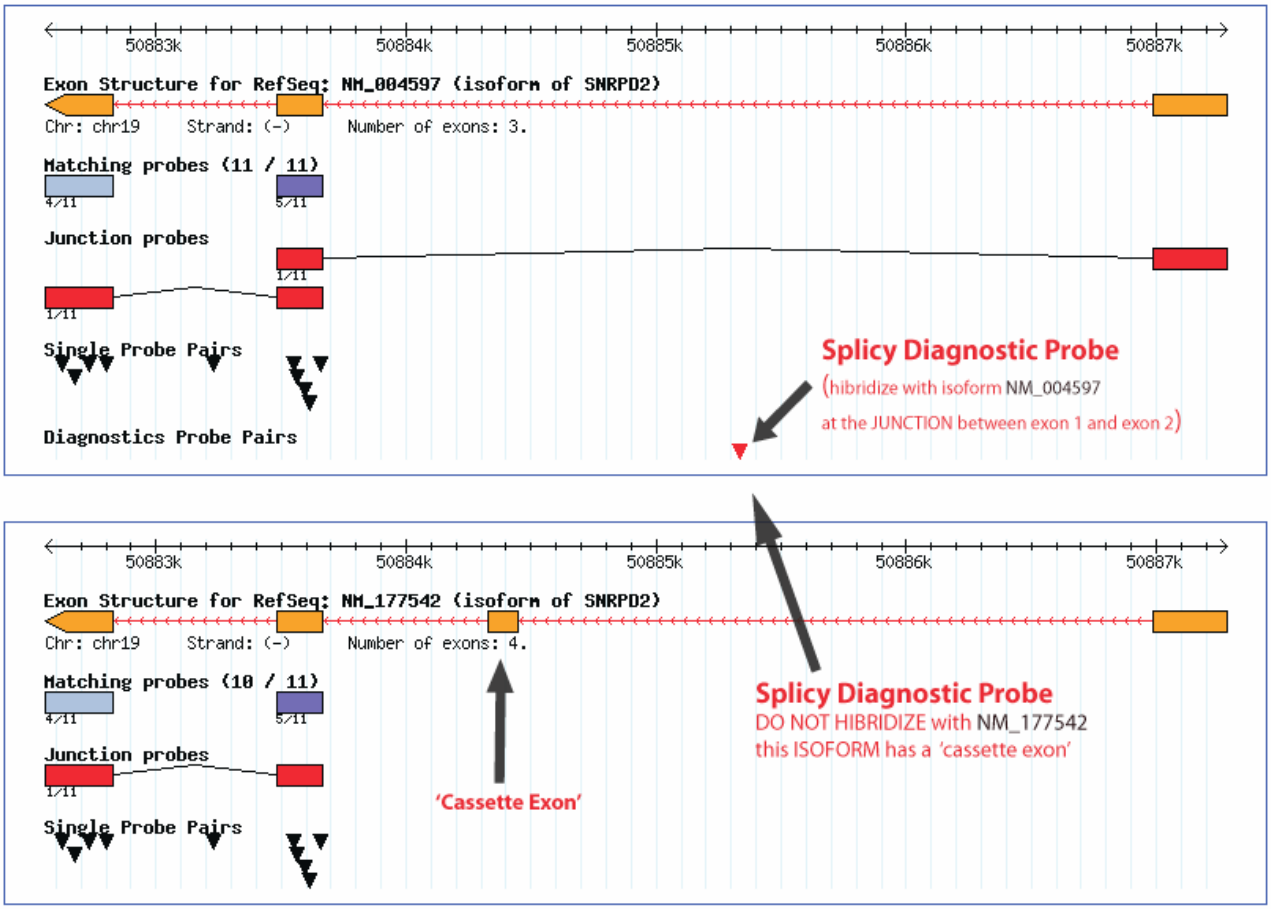
We present here a web-based software tool, Splicy (http://bio.ifom-ieo-campus.it/splicy/), whose primary task is to retrieve data on the mapping of Affymetrix probes to single exons of gene transcripts and display graphically this information on all available transcripts (both RefSeq and cDNAs). The program accepts in input a list of Affymetrix probesets and produces a series of graphical displays, each relative to a transcript associated with the gene targeted by a given probe. Each graphic reports the exon structure for a given transcript, a gliph which evidentiates the exons which contain matching probes, a gliph evidentiating the probes which are at the boundary of two exons, and a series of triangular glyphs evidentiating each single probe on the matching exon. If a given probe belongs to an exon which is skipped in a different transcript belonging to the same gene, it is tagged as a possible "splice diagnostic probe" and marked red. The idea is that a given probeset containing a 'diagnostic' probe will behave differently in the hybridization process, according to the transcript variant which is present in the hybridization mixture. The information on the transcript-by-transcript and exon-by-exon mapping can be retrieved both graphically and in the form of tab-separated files. Other features which can be retrieved from the graphic are the annotation table for each probeset, the Probeset design, RefSeq targets, the complete list of the Probeset probe pairs, the coordinates of the Genome alignments, further notes and links on additional transcript/gene annotation, GO functional classification and a direct link to the corresponding Entrez GENE entry. The mapping of single probes to RefSeq or EMBL cDNAs derives from the ISREC mapping tables which are the basis of the CleanEx Expression Reference Database Project (http://www.cleanex.isb-sib.ch/). We currently maintain mappings on the most popular human and mouse Affymetrix chips, and Splicy can be queried for matches with human and mouse RefSeq or EMBL cDNAs.

Results

We think that Splicy will be useful for giving to the researcher interested in transcriptome diversity a clearer idea of the possible transcript variants linked with a given gene and an additional key of interpretation of microarray experiment data. Splicy is publicly available from the web link http://bio.ifom-ieo-campus.it/splicy/ and has been realized thanks to part of a bioinformatics grant from the Italian Cancer Research Association.

Availability: http://bio.ifom-ieo-campus.it/splicy/

Contact email: davide.rambaldi@ifom-ieo-campus.it


50883k
50884k
50885k
50886k
50887k
Exon Structure for RefSeq: NM_004597 (isoform of SNRPD2)
Chr: chr19
Strand: (-)
Number of exons: 3.
Matching probes (11 / 11)
4/11
5/11
Junction probes
1/11
1/11
Single Probe Pairs
Diagnostics Probe Pairs
Splicy Diagnostic Probe
(hibridize with isoform NM_004597
at the JUNCTION between exon 1 and exon 2)
50883k
50884k
50885k
50886k
50887k
Exon Structure for RefSeq: NM_177542 (isoform of SNRPD2)
Chr: chr19
Strand: (-)
Number of exons: 4.
Matching probes (10 / 11)
4/11
5/11
Junction probes
1/11
Single Probe Pairs
'Cassette Exon'
Splicy Diagnostic Probe
DO NOT HIBRIDIZE with NM_177542
this ISOFORM has a 'cassette exon'