BioSystems 114 (2013) 31–38

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Review Article

Alternative splicing of mutually exclusive exons—A review

Martin Pohl[a,*], Ralf H. Bortfeldt[b], Konrad Grützmann[a], Stefan Schuster[a]

[a] Department of Bioinformatics, Friedrich Schiller University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany
[b] Department of Crop and Animal Sciences, Humboldt University Berlin, Invalidenstrasse 42, 10115 Berlin, Germany

ARTICLE INFO

Article history:
Received 30 May 2013
Accepted 3 July 2013

Keywords:
Alternative splicing
Exon clusters
Mutually exclusive exons
Non-sense mediated decay
Splicing mechanisms

ABSTRACT

Alternative splicing (AS) of pre-mRNAs in higher eukaryotes and several viruses is one major source of protein diversity. Usually, the following major subtypes of AS are distinguished: exon skipping, intron retention, and alternative 3′ and 5′ splice sites. Moreover, mutually exclusive exons (MXEs) represent a rare subtype. In the splicing of MXEs, two (or more) splicing events are not independent anymore, but are executed or disabled in a coordinated manner. In this review, several bioinformatics approaches for analyzing MXEs are presented and discussed. In particular, we revisit suitable definitions and nomenclatures, and bioinformatics tools for finding MXEs, adjacent and non-adjacent MXEs, clustered and grouped MXEs. Moreover, the molecular mechanisms for splicing MXEs proposed in the literature are reviewed and discussed.

Contents

1. Introduction

1.1. Molecular biology background

In addition to the genetic code, several other codes are used by the living cell at the molecular level, for example, the calcium oscillation code and the code used for signaling among plants by volatile chemicals. In eukaryotes one of these is the splicing code, by which the cell decides which sequence parts are finally used (Choudhary and Krithivasan, 2007; Barbieri, 2008; Barash et al., 2010; Reddy et al., 2012).

In the post-genomic era, alternative splicing (AS) of pre-mRNAs in higher eukaryotes got in the focus of research as one major source of protein diversity (Black, 2000; Graveley, 2001; Kim et al., 2008; Nilsen and Graveley, 2010; Chen et al., 2012a). AS was discovered in adenoviruses (Berget et al., 1977) and also occurs in several other viruses such as cytomegalovirus (Gatherer et al., 2011). Protein variability contributes to a high complexity of higher eukaryotes while keeping the numbers of genes relatively low. AS is a means to change proteins, in dependence on gender, developmental stage or environmental conditions and can affect binding

Abbreviations: AS, alternative splicing or alternatively spliced; MXE, mutually exclusive exon; EST, expressed sequence tag; NMD, nonsense mediated mRNA decay; ORF, open reading frame; *Dscam*, *Drosophila* Down Syndrome Cell Adhesion Molecule.

* Corresponding author. Tel.: +49 3641 949581.
E-mail addresses: m.pohl@uni-jena.de (M. Pohl), ralf.bortfeldt@agrar.hu-berlin.de (R.H. Bortfeldt), konrad.g@uni-jena.de (K. Grützmann), stefan.schu@uni-jena.de (S. Schuster).

0303-2647/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biosystems.2013.07.003

properties, intracellular localization, enzymatic activity and many more properties of proteins (Stamm et al., 2005; Yap and Makeyev, 2013). Estimations raised from one third up to 95% of human genes affected by AS, with other mammals showing similar high AS levels (Florea, 2006; Pan et al., 2008; Wang et al., 2008). Alternative splicing and splicing in general is a major problem in gene finding in eukaryotes because it may disrupt ORFs (Pohl et al., 2012).

The potential for variability is enormous. For instance, the human calcium-activated potassium channel subunit alpha-1 gene and the three neurexin genes could potentially generate 500 and more than 2000 different protein isoforms, respectively, by different ways of splicing (Black, 1998; Tabuchi and Südhof, 2002). The *Drosophila* Down Syndrome Cell Adhesion Molecule gene (*Dscam*) has several sets of cassette exons with one of them involving 48 alternative exons among which one is selected (Graveley, 2005; Anastassiou et al., 2006; Meijers et al., 2007; Olson et al., 2007; Hemani and Soller, 2012; Wang et al., 2012). This leads to 38,016 theoretical splicing variants.

AS is thought to lower selective pressure on gene sequences allowing a higher trial and error rate by mutations in one of the isoforms without compromising the acquired functionality of the other isoform (Boué et al., 2003; Chen et al., 2006; Noh et al., 2006). The apparent evolutionary advantages of AS require, however, significant energetic and metabolic costs because the spliceosome, which performs the splicing reaction, is a large complex of proteins and RNA including up to several hundreds of constituents (Jurica and Moore, 2003; Kielbassa et al., 2009; Bortfeldt et al., 2010; Hoskins et al., 2011). Given the enormous effort to assemble such complicated molecular machinery it can be assumed that the benefit of transcript flexibility outweighs the biochemical costs. In contrast, some organisms such as many plants, seem to have achieved their level of protein variability mainly by gene duplications i.e., an increase in genome length (Kopelman et al., 2005).

The ability to cope with stress is widely enhanced via transcriptome plasticity (Mastrangelo et al., 2012). Moreover, the involvement and prevalence of AS in many diseases is becoming increasingly clear. Hence, protein variability as generated by alternative splicing is of great medical and biotechnological importance because different isoforms are often associated with diseases such as cancer (Hernandez-Lopez and Graham, 2012) or with the distinction between intracellular and extracellular enzymes (Andreassi and Riccio, 2009). This renders AS and its regulation a potential therapeutic target (Mount and Pandey, 2005; Garcia-Blanco, 2006; He et al., 2009; Tazi et al., 2009, 2010; Douglas and Wood, 2011; Germann et al., 2012; Hernandez-Lopez and Graham, 2012; Sanchez-Pla et al., 2012).

Several attempts for general AS annotations have been presented (Xing et al., 2004; Nagasaki et al., 2006; Sammeth et al., 2008; Kroll et al., 2012). Among the well-known subtypes of AS are exon skipping (Sorek et al., 2004b), intron retention (Wang et al., 2006), alternative 5′ splice sites (Dou et al., 2006; Bortfeldt et al., 2008; Hiller and Platzer, 2008), alternative 3′ splice sites (Bortfeldt et al., 2008; Hiller and Platzer, 2008). A less abundant subtype of AS is represented by mutually exclusive exon (MXE) splicing.

MXEs are characterized by splicing of exons in a coordinated manner such that two or more splicing events are not independent. As the name “mutually exclusive” indicates, exactly one out of two exons (or one group out of two exon groups) is retained, while the other one is spliced out. Sammeth (2009) applies the term in a less strict way, allowing the case that none or all of the exons under consideration are retained. In contrast to other variants of alternative splicing, mutually exclusive splicing can leave the size of the protein unchanged provided that the exchanged sequence is of the same length and does not introduce a premature stop codon. Depending on the similarity of exchanged exon sequences, minor changes as in subtle alternative 5′ and 3′ splicing events or major changes of whole protein domains as in exon skipping are possible. In case of minor protein sequence changes, MXEs may provide an advantage to many types of proteins, such as ion channels, because the spatial structure is preserved, while the protein exhibits an altered function (Birzele et al., 2008a). Interestingly, another RNA processing mechanism, RNA editing, can also occur in a mutually exclusive manner as shown for the *TPH2* gene (Grohmann et al., 2010) resulting in a similar effect as mutually exclusive exon splicing.

A common assumption is that MXEs have originated from exon duplication and, hence, are highly similar (Letunic et al., 2002; Copley, 2004; Sorek, 2009; Pillmann et al., 2011). Accordingly, some authors (Stephan et al., 2007; Pillmann et al., 2011) define MXEs based on similar length and sequence. In our opinion, these criteria are not necessary. The term “mutually exclusive” only implies that exons do not occur together but does not refer to length, sequence or exon numbers. In general, also a group (cluster) of exons can be mutually exclusive with respect to another group (cluster) of exons. Such cases should be distinguished from exon cassettes where exactly one out of several exons is retained in the mature transcript, such as in the *Dscam* gene in *Drosophila*. However, the terminology is not used consistently among researchers, MXE were previously also termed as “exon clusters” (Pillmann et al., 2011) or “cassette exons” (Stephan et al., 2007).

MXEs turned out to be very promising candidates for generation of highly diverse but specific processes (Anastassiou et al., 2006; Soom et al., 2008). The alternative selection of exons enables the encoding of a whole class of proteins with similar scaffold and similar length but with highly specific functionality. Beside the above-mentioned *Drosophila Dscam* gene, examples of biological relevance are provided by the voltage dependence of ion channels (Soom et al., 2008) and calcium sensitivity of muscle proteins in higher animals (Waites et al., 1992). Like other AS types, MXEs proved to be of medical relevance, e.g., at regulation of expression levels of the mammalian pyruvate kinase M isoforms (Chacko and Ranganathan, 2009b; Chen et al., 2012b). Examples of MXEs have been described in human (Soom et al., 2008), mouse (Chacko and Ranganathan, 2009a), rat (Gustafson et al., 1993), chicken (Waites et al., 1992; Chacko and Ranganathan, 2009a), cow (Chacko and Ranganathan, 2009b), nematode (Johnson et al., 2003) and other species.

1.2. Bioinformatics resources for analyzing MXE splicing

As biochemical analyses are expensive and time consuming, computational approaches have attracted an ever increasing interest. Accordingly, AS is an important topic in bioinformatics (Dou et al., 2006; Zavolan and van Nimwegen, 2006; Hiller et al., 2007; Bortfeldt et al., 2008; Hiller and Platzer, 2008; Sammeth et al., 2008; Busch and Hertel, 2012; Chen et al., 2012a; Sanchez-Pla et al., 2012). To date many resources on AS emerged thanks to the growing amount of sequence and alignment data, in spite of incompleteness and considerable noise within the data (Black, 2003; Lareau et al., 2004; Chen et al., 2012a). Relevant databases that emerged in the context of MXE are MAASE (Zheng et al., 2005), HOLLYWOOD (Holste et al., 2006), ASAP II (Kim et al., 2007), ECgene (Lee et al., 2007), Ensembl (including former ASD/ATD/ASTD/AEdb projects (Koscielny et al., 2009), SPLOOCE (Kroll et al., 2012).

Also, the assembly of the spliceosome has been described by bioinformatics approaches (Kielbassa et al., 2009; Bortfeldt et al., 2010; Hoskins et al., 2011). Different types of the spliceosome were suggested to produce MXE splicing patterns (see Section 4). Beside the major spliceosome, a minor spliceosome can process splice sites that have distinct consensus sequences and are incompatible with the major spliceosome (Will and Lührmann, 2005).

In this review, we discuss several bioinformatics approaches for analyzing MXE splicing. In particular, we will focus on appropriate