

Mechanistic insights into mutually exclusive splicing in *dynammin 1*

Mikita Suyama^{1,2}¹Medical Institute of Bioregulation, Kyushu University and ²Japan Science and Technology Agency, CREST, Higashi-ku, Fukuoka 812-8582, Japan Associate Editor: Martin Bishop

ABSTRACT

Summary: Mutually exclusive splicing is a strictly regulated pattern of alternative splicing. A specific group of mutually exclusive splicing events has been shown to be regulated by the formation of specific RNA secondary structures. This type of regulation has been shown to exist only in arthropods. The present study involved a detailed sequence analysis of human gene structures that undergo mutually exclusive splicing, which showed that this type of regulation may also occur in *dynammin 1* in mammals. A phylogenetic analysis revealed that the *dynammin 1* orthologs in invertebrates did not share the same sequence features, which suggests that the regulatory mechanism has independently evolved in the mammalian lineage. Therefore, the emergence of this elaborate mechanism for mutually exclusive splicing may be attributable to mechanistic convergence.

Contact: mikita@bioreg.kyushu-u.ac.jp

Supplementary information: Supplementary data are available at Bioinformatics online.

Received on May 7, 2013; revised on June 12, 2013; accepted on June 19, 2013

1 INTRODUCTION

The alternative splicing of pre-mRNA is regulated in a spatiotemporal manner and is essential for generating proteomic diversity. Mutually exclusive splicing is an alternative splicing pattern where only a single exon is precisely selected from a cluster of candidate exons. Several different mechanisms have been proposed for mutually exclusive splicing (Smith, 2005). For example, an elegant model that involves specific RNA secondary structures in introns was first proposed for the exon 6 cluster of *Dscam* in *Drosophila* species (Anastassiou *et al.*, 2006; Graveley, 2005). In this model, the formation of RNA secondary structures between a docking site and selector sequences in pre-mRNA allows a single exon to be selected, although other cases that can be explained using this model were not observed for several years.

A recent study by Yang *et al.* (2011) expanded the applicability of this model to other genes in insects, such as *14-3-3ξ* and *myosin heavy chain*. In *Dscam*, it appeared that clusters of mutually exclusive exons other than exon 6 could also be explained using the docking site-selector sequence interaction model. They not only predicted the intronic elements responsible for mutually exclusive splicing but also experimentally demonstrated that the elements were directly involved with the precise selection of the exon (Yang *et al.*, 2011). At present, this mechanism is only known to exist in arthropods (Brites *et al.*, 2008), and it is uncertain whether it applies to genes in other clades, including mammals.

In this report, a genome-wide analysis of cases of mutually exclusive splicing was performed by scanning the structures of human transcripts to gain mechanistic insights into the regulation of mutually exclusive splicing. The RNA base-pairing potential for pre-mRNA sequences that underwent mutually exclusive splicing was calculated, and a characteristic sequence pattern that fitted well with the docking site-selector sequence pairing model was observed in the sequence of *dynammin 1* (*dnm1*). The applicability of this model was supported by evidence, such as the conservation of sequence patterns in other mammalian species. The discovery that the pre-mRNA sequence of *dnm1* in humans has sequence features similar to that of the alternative splicing patterns observed in insects greatly expands the applicability of the docking site-selector sequence pairing model to bilaterian animals.

2 METHODS

2.1 Gene and genome sequences

The genomic sequences and mammalian genome alignments were downloaded from the UCSC genome browser (<http://genome.ucsc.edu>) (Dreszer *et al.*, 2012). The mammalian genome sequences used in the analysis were as follows (the release date and assembly identifiers are indicated in parentheses): human (March 2006, hg18), orangutan (July 2007, ponAbe2), marmoset (June 2007, calJac1), mouse lemur (June 2003, micMur1), rat (November 2004, rn4), kangaroo rat (July 2008, dipOrd1), cow (October 2007, bosTau4), horse (September 2007, equCab2), dog (May 2005, canFam2), megabat (July 2008, pteVam1), rock hyrax (July 2008, proCap1), tenrec (July 2005, echTel1), armadillo (July 2008, dasNov2), and opossum (January 2006, monDom4). The conservation scores for the mammalian genome alignments were calculated using the phastCons program (Siepel *et al.*, 2005) and also downloaded from the UCSC genome browser (Dreszer *et al.*, 2012).

The genomic coordinates of human transcripts were downloaded from the UCSC genome browser (Dreszer *et al.*, 2012). There were 66 803 transcripts ('knownGene' in UCSC genome browser), which corresponded to 26 570 distinct genes.

2.2 RNA secondary structure analysis

The base-pairings between the surrounding introns of mutually exclusive exons in pre-mRNA sequences were calculated using the RNA duplex program in the ViennaRNA package (Hofacker *et al.*, 1994) (see Supplementary Materials for detailed procedure). The significance levels of the base-pairing energy values were evaluated by comparing them with the background energy distribution, which was obtained by applying the same program to randomly shuffled versions of the original sequences. This procedure was used to exclude base-pairing with low energy values because of biased nucleotide compositions.

2.3 Molecular phylogeny

The amino acid sequences of dynamins were downloaded from National Center for Biotechnology Information (Sayers *et al.*, 2012) and aligned using the MAFFT program (Katoh *et al.*, 2005) with the default parameters. A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei, 1987) with the JTT model (Jones *et al.*, 1992) for amino acid substitutions. Bootstrap values were calculated on the basis of 100 resamplings of the original alignment data.

3 RESULTS AND DISCUSSION

3.1 Identification of mutually exclusive splicing in the human transcriptome

For each gene, the structures of the transcripts were compared with each other to identify mutually exclusive splicing. The detailed procedures were as follows: (i) For a pair of transcripts that belong to the same gene, candidate mutually exclusive exons were observed by comparing the coordinates of the exons. If the upstream and downstream constitutive exons had the exact same coordinates and the coordinates of the alternatively spliced exon located between the constitutive exons were different in each transcript, these exons were marked as candidate mutually exclusive exons. (ii) After finding the candidate mutually exclusive exons, the structures of all the remaining transcripts belonging to the gene were analyzed, and candidates were rejected if a transcript skipped or included both candidate mutually exclusive exons. If the candidate exons were not rejected during comparisons with all the remaining transcripts, i.e. all the transcripts in a locus have either one of the candidate mutually exclusive exons, these exons were stored in the final list of mutually exclusive

exons. (iii) This procedure was repeated for all possible combinations of transcripts to exhaustively search for examples of mutually exclusive splicing.

This procedure was applied to 66 803 transcripts corresponding to 26 570 genes, and 118 examples of mutually exclusive splicing were identified (Supplementary Table S1). These examples were visually confirmed using the UCSC genome browser (Dreszer *et al.*, 2012), and a clear pattern of mutually exclusive splicing was observed, which was retained after comparing the EST sequences mapped on these loci. The total number of the identified examples is probably a lower limit because rather stringent criteria were adopted for further analyses of the regulatory mechanisms.

3.2 The docking site-selector sequence interaction model can be applied to *dnm1*

To identify examples that may be explained by the docking site-selector sequence pairing model, the 118 examples of mutually exclusive splicing were investigated in detail by measuring their potential to form base-pairings with the surrounding introns in the mutually exclusive exons. A striking example was observed in *dnm1* where a pair of complementary elements existed in two intronic regions, i.e. a potential selector sequence in the downstream region of exon 10a and a potential docking site in the proximal upstream region of exon 11 (Fig. 1A). The positions of these elements relative to the mutually exclusive exons and neighboring constitutive exons exactly matched those observed in the *14-3-3 ξ* gene in insects (see Fig. 1 in Yang *et al.*, 2011). These two elements were located in highly conserved regions (Fig. 1A and B), suggesting that selection pressures were acting on these

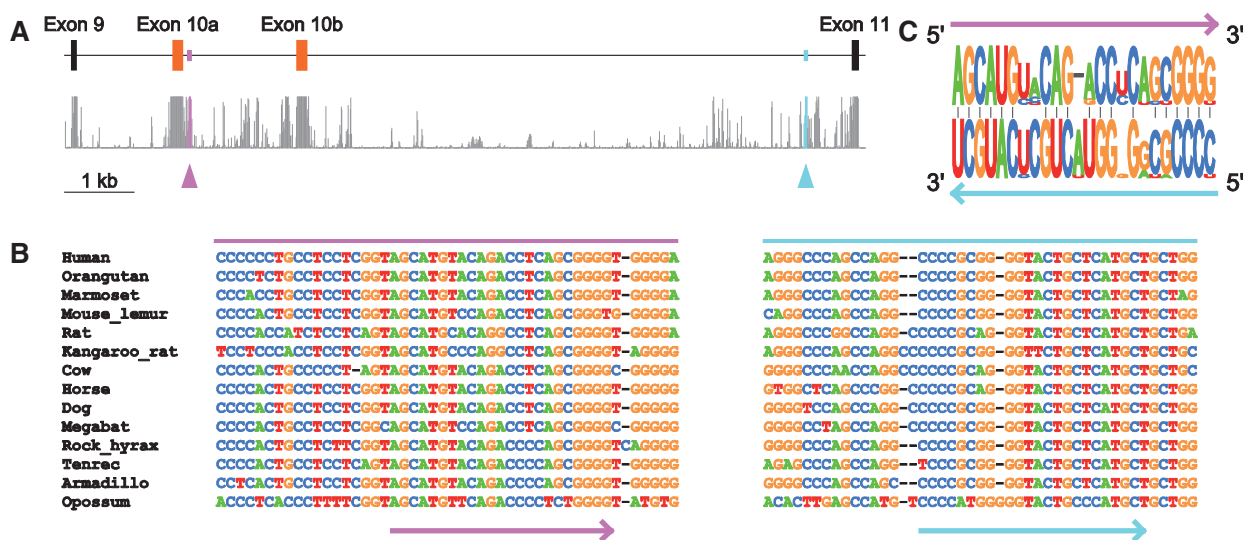


Fig. 1. Conserved complementary sequence elements in the human *dnm1* gene. (A) The partial gene structure around the mutually exclusive exons. Mutually exclusive and constitutive exons are shown in orange and black boxes, respectively. Conservation in mammalian genomes downloaded from the UCSC genome browser (Dreszer *et al.*, 2012) is shown under the gene structure. The two regions that contained the complementary sequence elements are shown in pink (the selector sequence) and cyan (the docking site). These regions are also indicated by colored triangles in the conservation graph. The partial gene structure is drawn to scale. The scale is shown at the lower left of this panel. (B) Genome sequence alignments of the regions that contained the complementary sequence elements for selected mammalian species. The colored bars at the top of the alignments correspond to the gene structure in panel (A). The complementary sequence elements are indicated by arrows at the bottom of each alignment. (C) The complementary sequence elements based on their predicted RNA base-pairing

sites. Moreover, both elements can form significantly stable base-pairing ($\Delta G = -30.8$ kcal/mol; z score = -3.34 , calculated from the background energy distribution) (Fig. 1C). Some sites that were not strictly conserved were not complementary with each other when the elements formed a specific RNA secondary structure, which also supports the functional importance of these elements during base-pairing. Overall, these features strongly suggest that RNA pairing between the two conserved elements observed in *dnm1* in mammals directs the choice of the mutually exclusive exons. This is the first clear evidence of the existence of this mechanism in mammals, which suggests that this is a widely used mechanism for mutually exclusive splicing in bilaterian animals.

3.3 Proposed model for the regulation of mutually exclusive splicing in *dnm1*

Based on the sequence features of *dnm1*, a model is proposed for the regulation of mutually exclusive splicing, which is analogous to that observed in mutually exclusive splicing events in insects (Anastassiou *et al.*, 2006; Graveley, 2005; Yang *et al.*, 2011) (Fig. 2). According to this model, mutually exclusive splicing is accomplished by competition between the formation of a specific RNA secondary structure via the complementary sequence elements and binding of a repressor protein (Fig. 2). However, what is the identity of the trans-factor that acts as a repressor during the regulation of mutually exclusive splicing? In case of the *Dscam* exon 6 cluster in *Drosophila*, Hrp36 is the factor that ensures fidelity of the mutually exclusive choice of exon 6 (Olson *et al.*, 2007). Thus, the corresponding human homolog, hnRNP A1 (Zu *et al.*, 1998), which is known to repress splicing by binding to pre-mRNAs (Smith and Valcarcel, 2000), may be the trans-factor involved with regulation of mutually exclusive splicing in *dnm1*. Another possible factor may be the Fox-1/Fox-2 splicing regulator, which is known to act as an activator and repressor during exon skipping (Zhang *et al.*, 2008). The 5'-residues of the selector sequence (AGCAUG) were highly similar to the Fox protein recognition sequence (UGCAUG) (Underwood *et al.*, 2005), and binding to the site may directly prevent the pre-mRNA from forming specific base-pairing, thereby inactivating

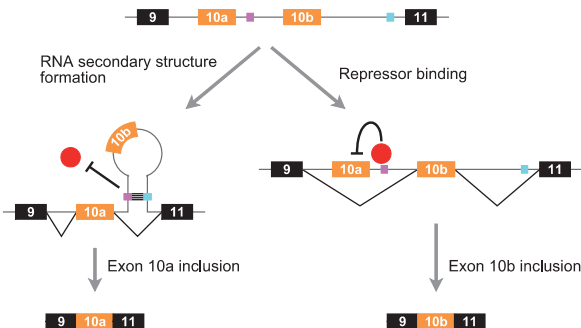


Fig. 2. Proposed model for the regulation of mutually exclusive splicing. Mutually exclusive and constitutive exons are shown in orange and black boxes, respectively, along with their exon numbers. The complementary sequence elements are shown in pink (the selector sequence) and cyan (the docking site). The filled red circle indicates the hypothetical repressor involved with the regulation of mutually exclusive splicing

exon 10a inclusion. This is also supported by the fact that *dnm1* and *fox-1/fox-2* are both highly expressed in the nervous system (Powell and Robinson, 1995; Underwood *et al.*, 2005).

The involvement of Fox proteins in mutually exclusive exon selection in *dnm1* was analyzed further using the CLIP-seq data obtained for Fox-2 (Yeo *et al.*, 2009). CLIP-seq is used to identify transcriptome-wide interactions among RNA-binding proteins and targets RNAs using cross-linking immunoprecipitation coupled with massively parallel sequencing (Yeo *et al.*, 2009). The Fox-2 CLIP-seq data were downloaded from the UCSC genome browser (the 'FOX2 CLIP-seq' track) (Dreszer *et al.*, 2012) and analyzed by counting the sequence reads mapped on each gene locus. Among the total of 4418213 CLIP-seq reads, 825 reads mapped on the *dnm1* locus, and *dnm1* was ranked among the top 3% of the total genes (729th of 26570 genes; $P < 0.0001$, Poisson distribution) (Supplementary Fig. S1). This also supported the involvement of Fox proteins in the mutually exclusive selection of the exons in *dnm1*, and it also reinforced the proposed model, which is explained by the competition between the formation of specific base-pairing and binding of a splicing regulator.

3.4 Molecular phylogenetic analysis indicates that the regulatory elements are specific invention in mammalian lineage

To obtain insights into the molecular evolutionary process underlying the emergence of this elaborate mechanism for mutually exclusive splicing, a phylogenetic analysis was performed using homologous sequences from several species. Each mammalian species had three closely related dynamins (*dnm1*, *dnm2* and *dnm3*), whereas *Caenorhabditis elegans* and *Drosophila* each appeared to have only one dynamin (van der Bliek, 1999). The phylogenetic relationships among the dynamins indicated that the expansion of *dynammin* paralogs occurred in the common ancestor of mammals (Fig. 3). Pairs of mutually exclusive exons were observed only in mammalian *dnm1* and *dnm2*, and were absent from all other dynamins (Boumil *et al.*, 2010). According to the phylogenetic tree, *dnm2* and *dnm3* were closer to each other than to *dnm1*. This suggests that exon duplication occurred in the common ancestor of mammals, and then *dnm3*

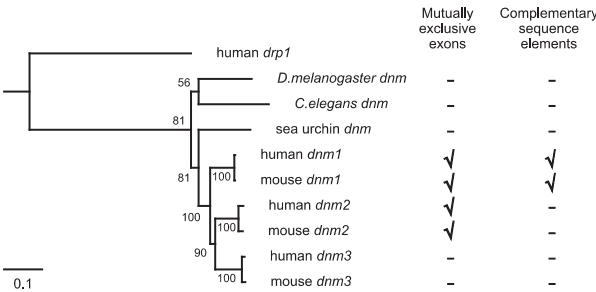


Fig. 3. Phylogenetic tree of *dynammin* genes. The tree was constructed using the neighbor-joining method (Saitou and Nei, 1987). Human *dynammin-related protein 1* (*drp1*) was used as an outgroup. The numbers at the internal nodes indicate bootstrap values based on 100 replicates. The existence of mutually exclusive exons and complementary sequence elements is indicated on the rightside of the tree

lost one of the duplicated exons; otherwise, multiple independent duplication would be required to explain the observed presence/absence pattern in the mutually exclusive exons. The complementary sequence elements were observed only in *dnm1* in mammals. This can be explained by the loss of sequence elements in *dnm2* or by the emergence in *dnm1*. Both possibilities suggest that these sequence elements independently evolved in the mammalian lineage and not in the bilaterian ancestor.

4 CONCLUSIONS

Recently, it has been reported that long-range RNA structures are important for the regulation of splicing in mammals and that a mechanism based on docking site-selector sequence pairing may be present in mammalian genes (Pervouchine *et al.*, 2012). To the best of my knowledge, this is the first example of a possible regulatory mechanism facilitated by docking site-selector sequence pairing during the precise selection of mutually exclusive exons in a mammalian lineage. A direct experimental proof is still required, but the lines of evidence presented in this study suggest that the regulatory mechanism responsible for mutually exclusive splicing is also present in *dnm1* in mammals. This mechanism was first proposed in the exon 6 cluster of *Dscam* in insects (Anastassiou *et al.*, 2006; Graveley, 2005). A molecular evolutionary study of *Dscam* homologs revealed that the regulatory mechanism exists only in arthropods, and that this was not because of its loss in other bilaterian species but because it had independently evolved in arthropods (Brites *et al.*, 2008), which is the phylogenetic pattern opposite to that proposed for the dynamins presented in this study. Thus, although mutually exclusive splicing is regulated the same way in *dnm1* in mammals and *Dscam* in arthropods, neither gene had a common ancestral gene with the same regulatory mechanism, which suggests that this type of regulation may be a result of mechanistic conversion. Convergent evolution of gene structures has often been observed in mutually exclusive exons (Copley, 2004). The example presented here indicates that convergence also occurs at the mechanistic level of mutually exclusive splicing where distinct genes in different clades have independently evolved the same mechanism of regulation to achieve the precise selection of mutually exclusive exons.

ACKNOWLEDGEMENT

The author thanks Tetsuya Sato for critical comments regarding the manuscript.

Funding: Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan

(21510215 and 22132005); and Kyushu University Interdisciplinary Programs in Education and Projects in Research Development (P&P).

Conflict of Interest: none declared.

REFERENCES

- Anastassiou,D. *et al.* (2006) Variable window binding for mutually exclusive alternative splicing. *Genome Biol.*, **7**, R2.
- Boumil,R.M. *et al.* (2010) A missense mutation in a highly conserved alternate exon of dynamin-1 causes epilepsy in fitful mice. *PLoS Genet.*, **6**, e1001046.
- Brites,D. *et al.* (2008) The Dscam homologue of the crustacean *Daphnia* is diversified by alternative splicing like in insects. *Mol. Biol. Evol.*, **25**, 1429–1439.
- Copley,R.R. (2004) Evolutionary convergence of alternative splicing in ion channels. *Trends Genet.*, **20**, 171–176.
- Dreszer,T.R. *et al.* (2012) The UCSC Genome Browser database: extensions and updates 2011. *Nucleic Acids Res.*, **40**, D918–D923.
- Graveley,B.R. (2005) Mutually exclusive splicing of the insect *Dscam* pre-mRNA directed by competing intronic RNA secondary structures. *Cell*, **123**, 65–73.
- Hofacker,I.L. *et al.* (1994) Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.*, **125**, 167–188.
- Jones,D.T. *et al.* (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.*, **8**, 275–282.
- Katoh,K. *et al.* (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.*, **33**, 511–518.
- Olson,S. *et al.* (2007) A regulator of *Dscam* mutually exclusive splicing fidelity. *Nat. Struct. Mol. Biol.*, **14**, 1134–1140.
- Pervouchine,D.D. *et al.* (2012) Evidence for widespread association of mammalian splicing and conserved long-range RNA structures. *RNA*, **18**, 1–15.
- Powell,K.A. and Robinson,P.J. (1995) Dephosphin/dynamin is a neuronal phosphoprotein concentrated in nerve terminals: evidence from rat cerebellum. *Neuroscience*, **64**, 821–833.
- Saitou,N. and Nei,M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- Sayers,E.W. *et al.* (2012) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **40**, D13–D25.
- Siepel,A. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.
- Smith,C.W. (2005) Alternative splicing—when two's a crowd. *Cell*, **123**, 1–3.
- Smith,C.W. and Valcarcel,J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.*, **25**, 381–388.
- Underwood,J.G. *et al.* (2005) Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol. Cell. Biol.*, **25**, 10005–10016.
- van der Bliek,A.M. (1999) Functional diversity in the dynamin family. *Trends Cell Biol.*, **9**, 96–102.
- Yang,Y. *et al.* (2011) RNA secondary structure in mutually exclusive splicing. *Nat. Struct. Mol. Biol.*, **18**, 159–168.
- Yeo,G.W. *et al.* (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.*, **16**, 130–137.
- Zhang,C. *et al.* (2008) Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev.*, **22**, 2550–2563.
- Zu,K. *et al.* (1998) Separable roles in vivo for the two RNA binding domains of *Drosophila* A1-hnRNP homolog. *RNA*, **4**, 1585–1598.