Published online 8 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1* **1–10**
doi:10.1093/nar/gkr648

SURVEY AND SUMMARY

Regulated functional alternative splicing in *Drosophila*

Julian P. Venables, Jamal Tazi and François Juge*

Université Montpellier 2, UMR 5535, Institut de Génétique Moléculaire de Montpellier, CNRS, 1919 Route de Mende, 34293 Montpellier cedex 05, France

Received April 3, 2011; Revised July 4, 2011; Accepted July 24, 2011

ABSTRACT

Alternative splicing expands the coding capacity of metazoan genes, and it was largely genetic studies in the fruit-fly *Drosophila melanogaster* that established the principle that regulated alternative splicing results in tissue- and stage-specific protein isoforms with different functions in development. Alternative splicing is particularly prominent in germ cells, muscle and the central nervous system where it modulates the expression of various proteins including cell-surface molecules and transcription factors. Studies in flies have given us numerous insights into alternative splicing in terms of upstream regulation, the exquisite diversity of their forms and the key differential cellular functions of alternatively spliced gene products. The current inundation of transcriptome sequencing data from *Drosophila* provides an unprecedented opportunity to gain a comprehensive view of alternative splicing.

INTRODUCTION

Alternative splicing is a means by which an organism can create multiple mRNAs from a single gene; the connection of different parts of a pre-mRNA leads to different mRNAs. Alternative splicing can affect 5′- and 3′-untranslated regions of mRNAs and consequently modulate translation, stability or localization of mRNA. Also, by modulating inclusion and exclusion of coding exons, this process dramatically increases overall protein diversity and is thought to account for some of the complexity of higher organisms such as vertebrates and equally that of complex tissues such as the brain. Most importantly, alternative splicing can lead to the expression of functionally different proteins from a single gene in a temporal- or tissue-specific manner. About 95% of human genes produce alternatively spliced mRNAs (1,2) but the functional consequences of these alternatilve splicing events are mostly unknown. Indeed, it is not obvious that different protein isoforms produced by alternative splicing necessarily have different functions. The answer to this important question requires extensive *in vivo* studies.

Drosophila is the geneticist's organism of choice because of easy manipulation, quick generation times, extensive collections of mutants and the numerous tools available for manipulating gene expression. Pioneer studies in *Drosophila* have led to the identification of regulatory sequences in the pre-mRNA required for alternative splicing regulation as well as the splicing regulators that bind to them. These discoveries have paved the way for alternative splicing regulation studies in other organisms. *Drosophila* uses every alternative splicing strategy imaginable with an elegance and complexity that often eclipses mammals. To summarize these findings, we highlight several examples of alternative splicing events illustrating the diversity of alternative splicing strategies used by *Drosophila*. The examples of alternatively spliced genes are organized into different sections according to the localization of their encoded proteins in Figure 1. For easy reference, the examples are discussed in the same order in the text along with their diverse biological implications, where known. The subsequent sections highlight common themes in the regulation and evolutionary aspects of tissue-specific alternative splicing and the final section assesses the implication of new technologies for our understanding of the systems biology of alternative splicing and the crucial role *Drosophila* studies will play. We refer throughout to the current names of genes in Flybase (http://flybase.org/) as these may differ from those in their original references.

SEX-SPECIFIC SPLICING

The concept of alternative splicing as a major controller of gene expression came largely from work in the 1990s on the *Drosophila* sex-determination pathway. Sex determination in *Drosophila* is controlled by a cascade of splicing factors that are themselves alternatively spliced, ultimately leading to the sex-specific expression of two different

*To whom correspondence should be addressed. Tel: +33 (0) 4 34 35 96 85; Fax: +33 (0) 4 34 35 96 34; Email: francois.juge@igmm.cnrs.fr

variants of the Doublesex (Dsx) transcription factor. These proteins are differentially expressed throughout male and female bodies to control characteristically different sets of genes (Figure 1A). The key splicing differences between males and females are set up by the expression of the Sex-lethal (Sxl) protein in females. The early expression of Sxl in female embryos, in response to X chromosome number, also allows Sxl to auto-regulate its own productive splicing in females and thus dictate female development [reviewed in ref. (3)]. Female Sxl protein represses the inclusion of an alternative 'poison exon' containing a stop codon (exon 3) in its own mRNA, allowing the expression of full-length protein. In males, the poison exon is spliced into the mRNA and a truncated Sxl protein is produced. Conclusively, ectopic expression of the female form of Sxl in male flies causes the female splicing pattern of endogenous *Sxl* (4). Sxl governs multiple aspects of female differentiation. For example, Sxl represses the expression of the Male-specific lethal (Msl2) protein, which is involved in X-dosage compensation in males, by inhibiting both splicing and translation of the *msl2* transcript in females (5). Downstream in the sex-determination splicing cascade, Sxl interferes with the default proximal 3′-splice site in the *transformer* (*tra*) gene in females thus leading to functional Tra protein in females only (6).

Tra is also a pre-mRNA splicing factor controlling key alternative splicing events in females further downstream in the sex-determination cascade and it was an early example of combinatorial interplay between RNA-binding proteins controlling alternative splicing. Here, Tra binds in a complex with the Transformer 2 (Tra2) protein and the SR protein RBP1 on the 13-nt repeat sequences in the female-specific exon of the *doublesex* (*dsx*) gene, and with SRp30 on the adjacent purine-rich element (7). This leads to male and female splice forms of the Dsx transcription factor with their own different sets of target genes that drive sex-specific gene expression. The splicing differences in the sex-determination pathway are found in every cell of the male and female bodies; there are also sex-specific splicing differences that appear only in the brain and in germ cells, and some are controlled by other proteins than the known sex-determination genes (8).

In the brain, another gene downstream of Tra represents one of the most striking demonstrations of the function of alternative splicing. The *fruitless* locus (*fru*) generates a complex set of mRNAs through the use of alternative promoters and alternative splicing at both the 5′- and 3′-ends of the transcripts. One set of transcripts, arising from the P1 promoter, are spliced differently in males and females in order to express a specific isoform of Fru only in males (9). The protein is not expressed in females because female-specific splicing uses a downstream 5′-splice site that interrupts the reading frame of the protein (10). The female-specific 5′-splice site depends for its use on an enhancer complex containing Tra/Tra2 which assembles on the RNA 40–240 nt upstream of the female-specific splice site (11). Remarkably, male splicing of *fru* is necessary for male courtship behaviour and it is sufficient to generate male behaviour in otherwise wild-type females (9).

Although Tra2 is ubiquitously expressed, it is most highly expressed in testis where its high level leads to direct repression of removal of an intron in its own message to negatively auto-regulate its expression (Figure 1B) (12,13). Limiting the overabundance of Tra2 by this mechanism would appear to be important as increasing gene dosage of Tra2 in male flies had an increasingly severe effect on male fertility (14). In another example of testis-specific splicing, TAF1 (TBP-associated factor 1) needs two AT hooks to bind DNA, and alternative splicing of one of the hook regions comes from a 33 amino acid cassette exon that is included at much higher levels in testis than in any other tissue (15). Tra2 is likely involved in this testis-specific splicing event since Tra2 level affects *Taf1* splicing in S2 cells (16). Splicing in the testis is also controlled, in part, by a testis-specific homologue of the U2AF large subunit (17).

A study that demonstrates the power of *Drosophila* genetics to decipher the function and mechanism of alternative splicing involved the *ovarian tumour* (*otu*) gene, which is required for female germ-line differentiation. *otu* has a 126-base in-frame alternative exon that encodes a Tudor domain, which is included in early stages of oogenesis but excluded in late stages. Inclusion of this alternative exon requires the activity of the pUf68 (poly U binding factor 68 kDa a.k.a. half pint) splicing factor since *pUf68* mutant flies specifically omit the exon. Remarkably, *pUf68* mutant ovaries display ovarian phenotypes that can be rescued by correctly spliced *otu*, revealing a critical role of the long Otu isoform in oogenesis (18).

ALTERNATIVE SPLICING IN MUSCLE

Some of the first reports of regulated alternative splicing discovered, which probably have important implications for cellular function, came in four *Drosophila* myofibrillar genes (Figure 1C). Many alternatively spliced *Drosophila* genes have an elegant organization of mutually exclusively used exons that provide subtly different coding regions to genes. For example, the myosin heavy-chain isoforms are encoded by a single gene (*Mhc*), which has in-frame alternative cassette exons in five different regions (exons 3, 7, 9, 11 and 15) and a cassette exon (number 18) that encodes an alternative C-terminus. Thus, by multiplying these independent choices, it is apparent that there are potentially 480 different myosin isoforms in *Drosophila*. An early study documented many developmental splicing changes in *Mhc* in different kinds of muscle (19). For example, exclusive use of exon 15a was observed in thorax (20) and more recently, widespread regulation of the *Mhc* exons has been shown in 33 different muscle types from embryos (21). The major species were the 7(a)9(b-c)11(c-d)15(b)18(b) form in slow contractile muscle and 7(d)9(a)11(b)15(a)18(a) in fast contractile muscle (Figure 1C). Therefore, different muscle types express specific Mhc isoforms through specific combinations of cassette exons, suggesting regulated fine control of the contractile properties of each kind of muscle. Indeed, swapping the converter domain (encoded by the exon

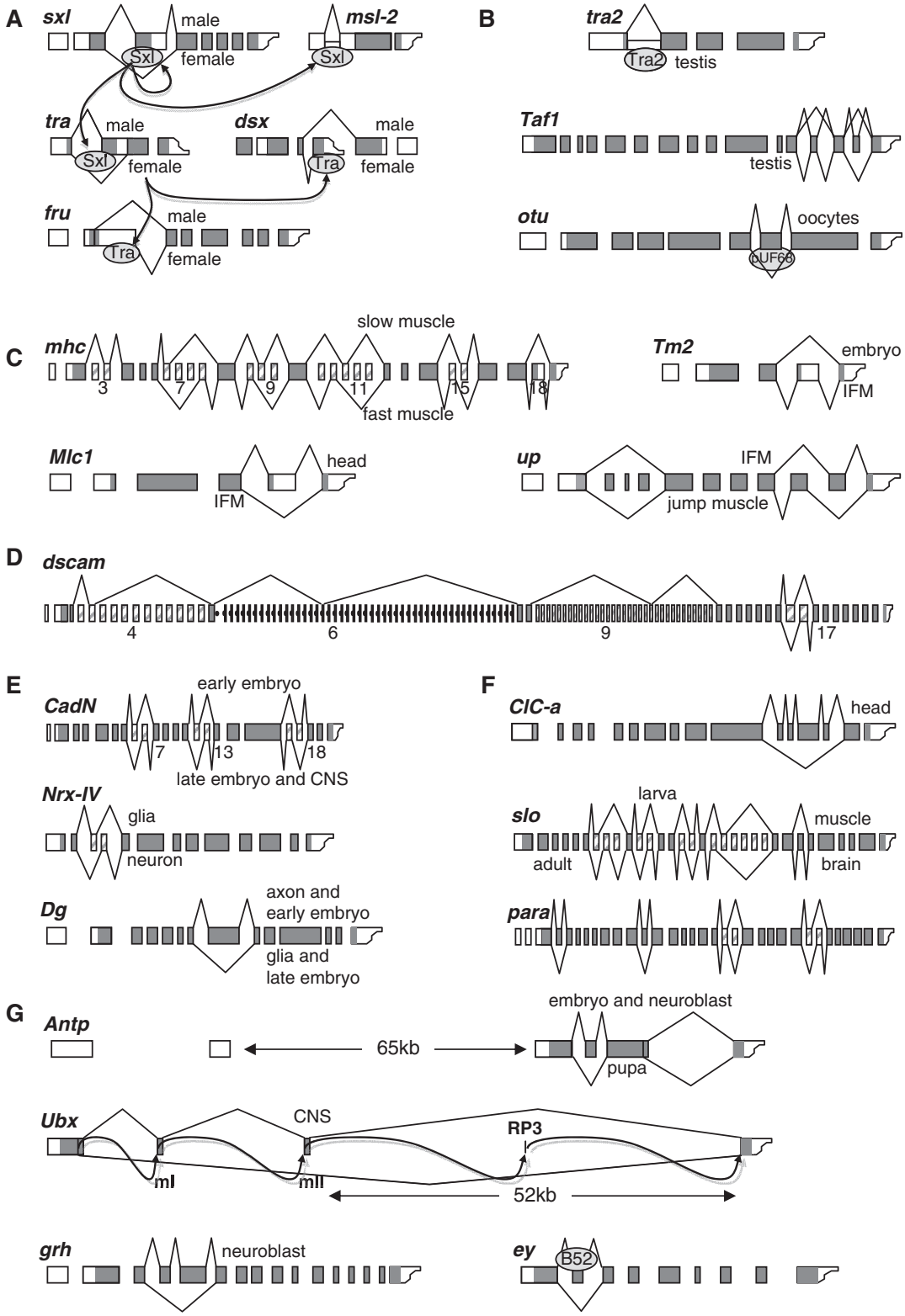

Figure 1. Examples of alternative splicing patterns in *Drosophila* genes. Schematics illustrating regulated alternative splicing patterns of several *Drosophila* genes. Exons and intervening intron size differences have been minimized to standardize formatting; therefore, genes are not depicted to scale. For simplicity, alternative 5′- and 3′-terminal exons are not shown. Exons are shown as boxes, coding regions are grey tint and mutually exclusive alternative exons are crosshatched. Splice forms are labelled according to the stage or cell type they are preferentially expressed in. (**A**) The autosomal sex-specific splicing cascade. (**B**) Germ cell-specific splicing. (**C**) Muscle-specific splicing. IFM = indirect flight muscle. (**D**) Alternative splicing of *Dscam*. Within the cluster of exon 6, 'acceptor' and 'docking' sites are shown as black dots (see the text). (**E**) Alternative splicing in genes encoding cell surface molecules (**F**) Alternative splicing in genes encoding ion channels. (**G**) Alternative splicing of transcription factor genes. In *Ubx*, RP3 represents a ratcheting point used for recursive splicing.

11 cluster) from flight muscle to an embryonic exon influenced the kinetic properties of the muscle fibres (22).

In addition to Mhc, there are indirect flight muscle (IFM)-specific isoforms of myosin alkali light chain 1, Tropomyosin 2 and Troponin-T. The *Myosin alkali light chain 1* gene (*Mlc1*) is spliced whereby the long penultimate exon is skipped only in the IFM giving rise to alternative 12 residue C-termini (23). The intron upstream from the alternative exon has no polypyrimidine tract, which implies that special splicing factors might be needed for its recognition. Tropomyosin 2 (*Tm2*) has a 429 base cassette, penultimate exon, used differentially in embryo and in thoracic IFM giving rise to two highly homologous alternative C-termini of 27 amino acids. The two types of Tropomyosin correlate with different muscle types and filament organization (24). In both *Mlc1* and *Tm2*, inclusion of the alternative exon could open the transcript up to potential nonsense-mediated decay, which might be an additional factor in their regulation. A further noticeable example of muscle-specific splicing is found in the *upheld* (*up*) gene encoding Troponin-T. *up* has two mutually exclusive exons of identical size encoding part of the C-terminal domain. The 10b exon is used in jump muscle whereas exon 10a is almost entirely selected in IFM (25). Moreover, in the N-terminus of Troponin-T, three micro-exons, one being just 3 nt in length, are collectively skipped in IFM and jump muscle (26).

NEURAL ALTERNATIVE SPLICING

Splice form diversity is most extensive in the mammalian brain, and in *Drosophila* too there are many examples of neural-specific alternative splicing events including, most notably, cell-surface molecules and transcription factors (Figure 1D–G). The most exquisite example of alternative splicing controls the cell-surface diversity of the Down syndrome cell adhesion molecule (Dscam) (Figure 1D). *Dscam* is organized into four clusters of 12, 48, 33 and 2 mutually exclusive exons, thus potentially encoding over 38 000 different proteins (27). The molecular bar code provided by alternatively spliced Dscam isoforms on the cell surface impedes self-crossing of axons and is essential for nervous system development, as artificial restriction of *Dscam* alternative splicing diversity resulted in severe disorganization of neural circuitry (28). Exon 4 variants are all used in similar proportions in brains, where neurone density is highest, whereas the second exon of the exon 4 array populates most of the transcripts in wings and legs (29).

Among the 48 exons of the *Dscam* exon 6 array, only one is used at a time. The heterogeneous nuclear ribonucleoprotein (hnRNP) hrp36 is a major player in preventing inappropriate inclusion of exon 6 cassettes, as depletion of this protein in S2 cells resulted in tandem inclusion of multiple exon 6 variants in the mature *Dscam* RNA. This effect was abrogated in the absence of the 'SR protein' B52 implying that a balance of these splicing factors controls the exon 6 array (30). However, there needs to be an additional, failsafe, mechanism to ensure that only one exon from the region is selected. It was observed that a sequence upstream of each exon in the exon 6 array, termed the 'selector sequence', is complementary to a putative 'docking site' sequence downstream of exon 5. It therefore seems likely that complementarity between the docking site and the selected exon's selector sequence confers mutual exclusivity (31). Indeed, removal of the docking site in a *Dscam* minigene led to widespread exon skipping and reduced the alternative splicing diversity by favouring predominant selection of the first exon of the array (32). There is now evidence that other arrays of multiple mutually exclusive exons in *Drosophila* (e.g. *Mhc* discussed above) may be similarly controlled, confirming the importance of the RNA structure in the control of alternative splicing (33).

Other alternative splicing events affecting the neuronal cell surface include Cadherin-N (CadN). *CadN* exons 7, 13 and 18 each come from a choice of a pair of exons and their splicing is temporally regulated during the first 24 h of embryonic development, with all three variable regions shifting in concert from the 7a13b18b form to the 7b13a18a form (34,35) (Figure 1E). Neurexin-IV (Nrx-IV) is a component of septate junctions in glial cells, and mutually exclusive alternative exons in the extracellular domain are differentially incorporated between glial cells and neurons. Utilization of exon 4 predominates in neurons and increases Nrx-IV binding affinity for the glial cell surface protein Wrapper (36). Glial- versus neuronal-specific alternative splicing is also seen in the *dystroglycan* (*Dg*) gene. Dystroglycan links the extracellular matrix to the cytoskeleton in a variety of cell types and its mucin-like domains, which bind to Laminin, are encoded by a large 795-base alternatively spliced exon. The long form is expressed on axons while the short form, lacking this region, is found on glial cells. There is also a shift to the short form during embryogenesis (Figures 1E and 2A) (37). Further studies are required to elucidate the specific function of the short form of Dystroglycan lacking the mucin-like domains.

Several ion channels are alternatively spliced in the brain (Figure 1F). For example, the *Chloride Channel-a* (*ClC-a*) gene has a 127 amino acid insertion in its intracellular region that is thought to be a regulatory site for channel gating and trafficking. The alternative splice consists of four adjacent exons that are only included in *Drosophila* heads (38). The *slowpoke* (*slo*) calcium-activated potassium channel is important for neuronal firing, and gene-product diversity is provided by 5 different promoters and 14 alternative exons in 5 different parts of its extracellular domain. The first alternative cassette encodes the mouth of the ion pore, and alternative exon selections in the first and third regions both affect calcium ion sensitivity of the channel. Coordinated shifting of all four alternatively spliced regions was seen between larvae and adults. There is also higher inclusion of the final cassette exon in brain than in other tissues (39). The *paralytic* gene (*para*) encodes a voltage-gated sodium channel, and it has several regions of alternative splicing which are differentially regulated. One alternative cassette exon encodes 22 amino acids in the first intracellular loop with a PKA phosphorylation site, and some individual neurons express only the long form while others only

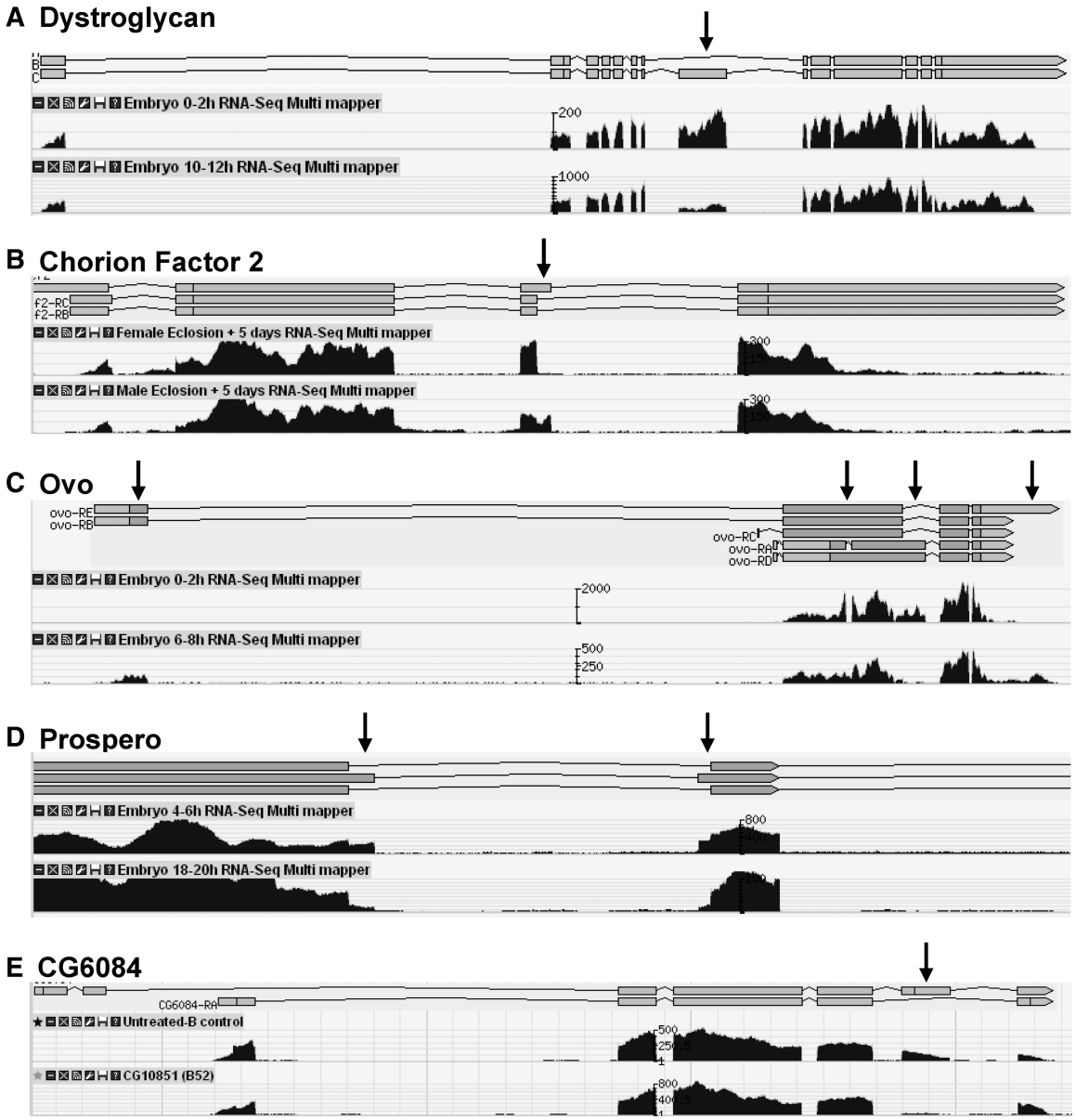

Figure 2. Visualization of alternative splicing events by RNA-Seq. All data are taken from the modENCODE website http://www.modencode.org/ (71). Different known transcripts of each gene are shown with exons as boxes and introns as lines between. RNA-Seq short reads from different tissues, stages or conditions, are piled below. Arrows point to regions where the number of reads changes relative to the gene as a whole. (**A**) Maternal *Dystroglycan* mRNA detected in early embryos includes an alternative exon, which is skipped later on during embryogenesis (see the text and Figure 1E). (**B**) Sex-specific splicing of the zinc finger transcription factor Chorion factor 2. The male-specific splice form has two extra zinc fingers due to the use of a downstream 5'-splice site (72). (**C**) In the *Ovo* gene, encoding a zinc finger transcription factor, inclusion of the alternatively spliced region correlates with the use of the proximal promoter and polyadenylation sites (59). (**D**) The so-called Twintron of the *prospero* gene has a seemingly ordinary intron in early embryogenesis but a switch occurs to a different set of borders, which are spliced by the minor U12 spliceosome (73). This change may be caused by downregulation of the heterogeneous ribonucleoproteins Hrp38 and Hrp36 (74). (**E**) The *CG6084* gene was shown to be alternatively spliced in cells in which the *B52* gene was knocked down, by the use of splicing-specific microarrays (55).

have the short form and this correlates with sodium currents in those individual cells (40). Therefore, alternative splicing of *para* (like *dscam*, see above) appears to differentiate individual neurons from one another with implications for CNS complexity. Recent studies have revealed an extreme complexity of *para* isoforms and different functions for the many different variants (41,42).

ALTERNATIVE SPLICING OF TRANSCRIPTION FACTOR GENES

In addition to the Dsx transcription factor responsible for sex-specific differentiation (see above), other transcription factors are also alternatively spliced in *Drosophila*. Homeotic genes encode transcription factors involved in the specification of cell identities according to their

position along the antero-posterior axis. Among them, Antennapedia (*Antp*) and Ultrabithorax (*Ubx*) are encoded by large genes, spanning of the order of 100 kb, which both have small alternatively spliced exons (Figure 1G). *Antp* has two regions of alternative splicing directly upstream from the region encoding the homeodomain of the protein. Thus, four alternative splice forms can be made due to an alternative exon of 13 codons and alternative 5′-splices sites of the downstream exon that are four codons apart. The form that omits both regions is abundant in early embryogenesis and the longer form progressively accumulates in late embryos and subsequent stages (43). There are also variations in isoforms ratio between adult tissues (44).

Ubx is alternatively spliced to produce six mRNAs, by the use of two 5′-splice sites (*a* and *b*) at the end of exon 1 and by inclusion of two alternative micro-exons (mI and mII). The 34 amino acids encoded by the two alternative exons form a linker region that modulates Ubx's DNA-binding activity and its capacity to activate the *dpp* target gene (45) or to repress *Distalless* gene expression (46) in transgenic assays. Skipping of mI and mII alternative exons requires the removal of 74 kb of intronic sequence. Analysis of this splicing event led to the discovery of a new concept in splicing research, namely 'recursive splicing', whereby portions of a large intron can be spliced out sequentially. Here, splicing of mI and mII to the 5′ *a* site re-creates a 5′-splice site each time that can then be re-used for splicing out the remaining downstream part of the intron (Figure 1G) (47). In a separate study, bioinformatic analysis revealed the widespread presence of 'zero-length exons' (i.e. juxtaposed 3′- and 5′-splice sites of the form $(Y)_n$NCAG/GUAAGU, where '/' marks the splice junction) within large introns that can be used as 'ratcheting points (or RP sites) for recursive splicing (48). Experimental validation confirmed the use of several RP sites *in vivo*, such as the RP3 site found in the 52 kb third intron of *Ubx*, indicating that RP sites may facilitate splicing of large *Drosophila* introns by sequentially truncating them during transcription (48). RP sites are also significantly enriched in other invertebrates' large introns but not in vertebrate ones indicating that recursive splicing may not be used in vertebrates (49). Finally, two adjacent exons encoding a total of 270 amino acids in the Grainyhead (*grh*) transcription factor are specifically incorporated in neuroblasts and may function by altering its dimerization properties (50).

CONTROL OF ALTERNATIVE SPLICING

While the factors that control alternative splicing in the sex-differentiation pathway are well known (Figure 1A), relatively little is known as yet about what controls the other diverse splicing events that have been observed. For the most part, alternative splicing has been considered to result from the independent recruitment to the pre-mRNA of RNA-binding proteins. RNA interference was used to knockdown 250 *Drosophila* RNA-binding proteins in S2 cells, and reverse-transcribed RNA samples were tested by PCRs targeting 19 alternative splicing events in three genes; 47 of the RNA-binding proteins were thus identified as direct or indirect controllers of one or more alternative splice events (51). Among these were known alternative splicing factors (i.e. proteins that are not essential for splicing generally but have substrate-specific effects) and other proteins that had previously been considered as constitutive- rather than alternative-splicing factors. Both alternative and constitutive splicing factors were found to regulate *Taf1* alternative splicing (Figure 1B) in response to ATR pathway activation using the same strategy (16). These substrate-specific effects of knockdowns of core spliceosomal components are likely of relevance to human diseases involving constitutive splicing factors. The identification of multiple splicing factors controlling individual alternative splice events also suggested combinatorial action of factors. For example, Pasilla and Mub coordinately regulate splicing of *Dscam* exon 4.2, whereas Rox8/dTIAR and B52 might coordinately regulate an alternative splice site in *dADAR* (51).

B52 is a ubiquitously expressed member of the 'SR protein' family of splicing factors. Several strategies have been used to identify alternative splicing events regulated by B52, such as genomic Systematic Evolution of Ligands by EXponential enrichment (SELEX) (52) or immunoprecipitation of RNA bound by a tagged-B52 overexpressed in specific tissues (53). Analysing the phenotypes induced by B52 overexpression also gave insights into the developmental processes modulated by this protein *in vivo* and therefore implicated potential target genes. For example, B52 overexpression during early eye development led to flies harbouring partially to totally reduced eyes, a phenotype that resembles the *eyeless (ey)* mutant phenotype. It turned out that *ey* is indeed a direct target gene of B52; B52 binds to and promotes the inclusion of an alternative exon, which inserts 60 amino acid residues upstream of the paired domain (Figure 1G). The Ey isoform including the alternative exon has lower affinity for its DNA targets *in vitro* and is less potent in driving eye development *in vivo* (54).

To systematically study the targets of SR proteins, Blanchette and co-workers knocked down B52 and dASF/SF2 in S2 cells and monitored splicing of 8315 alternative splicing events in 2797 genes by the use of splicing-specific microarrays. The knockdowns affected 107 and 319 alternative splicing events, respectively (55). The same group also knocked down four hnRNPs. In parallel to their splicing analysis, a global analysis of pre-mRNA regions bound by these proteins was performed using RNA-immunoprecipitation followed by hybridization to tiling arrays. All four hnRNPs affected a similar number of splicing events, which overlapped significantly with the targets they bound; however, there were interesting differences. For example, hrp36 and hrp38 had significant overlap in their splicing targets, suggesting a redundant function *in vivo*, and bound to introns significantly more than hrp48, which bound mostly to exons (56). The binding patterns of hnRNPs to genes offered insights into the different mechanisms of hnRNP action, including the 'looping out' model whereby hnRNP binding at either end of an intron brings splice sites

closer together, and also the nucleation model whereby specific binding to RNA by an hnRNP is followed by multiple binding and spreading of hnRNPs over a large region (56). Recently, the targets of Pasilla (PS), the homologue of mammalian NOVA1/2 splicing factors, have been investigated by new generation full transcriptome sequencing (RNA-Seq) after depletion of PS by RNAi in S2 cells (57). This approach identified 405 alternative splicing events in 323 genes. Interestingly many splicing events modified by PS depletion were previously unknown; this discovery of previously unannotated splice junctions highlights an important advantage of RNA-Seq over pre-designed microarray methods.

Despite the valuable mechanistic insights these studies have provided, simple up- or downregulation of splicing factors in cells cannot fully uncover the complex network underlying alternative splicing, as the activity of splicing factors is controlled by the tissue-specificity of upstream signals at different levels. For example, the Tra, Tra2 and RBP1 splicing factors, involved in *dsx* alternative splicing, are phosphorylated by the Darkener of apricot (Doa) kinase, and in a sensitized genetic background for *tra* or *tra2*, *Doa* mutants females display altered splicing of *dsx* and an intersex phenotype (58). Tra2 protein level can also be controlled by proteasomal degradation as shown in S2 cells in response to ATR signalling pathway activation (16).

Many of the genes with multiple regions of alternative splicing highlighted above have correlated alternative splices between different regions within the gene, that is to say the alternative splicing choice in one region is often linked to alternative splicing in the other regions and not independent. Examples include *Mhc*, *para*, *slow* and *CadN*. In other genes like *up*, *ClC-a* and *grh,* several adjacent exons are collectively skipped or included (Figure 1). In addition, there is sometimes a clear correlation between alternative splicing and specific promoter and polyadenylation site use as seen in the *ovo* gene (59) (Figure 2C). Further studies are required to establish if these linked changes within a gene are due to coordinated action of *trans*-acting factors or whether they are controlled at the level of transcription. The prevailing theory is that fast transcription favours alternative exon skipping and slow transcription allows time for alternative exon selection (60). Co-transcriptional effects are likely to play an especially important role in genes with long introns that can take many minutes to be transcribed (e.g. *Antp* and *Ubx*, Figure 1G).

CONSERVATION OF REGULATED FUNCTIONAL ALTERNATIVE SPLICING

There are big differences in gene structure between *Drosophila* and mammals. For example, in humans introns are usually much bigger than exons (87% of introns are >250 nt), whereas in *Drosophila* the reverse is true (66% of introns are <250 nt) (61). There are also several unique facets of splicing in *Drosophila* that are not found in mammals. For example, multiple variants of mutually exclusive exons and recursive splicing are not found in mammals. Another non-mammalian phenomenon is that of *trans*-splicing, which is widespread in *Caenorhabditis elegans* and also found in some *Drosophila* genes (62). However, within mammals on the one hand and insects on the other, there is extensive conservation of alternative splicing patterns.

Several examples of regulated alternative splicing are highly conserved among insects. For example, Cadherin-N has a broadly conserved exon structure across 300 million years of arthropod evolution with the orthologues of the 'a' and 'b' exons in other species being far more conserved than the paralogous 'a' and 'b' exons within the *Drosophila CadN* gene. Importantly, the central pair of mutually exclusive exons is differentially expressed in embryonic mesoderm and neurons in both *Drosophila* and beetles (35). However, several attempts to discern the functional differences between *CadN* isoforms have failed, but clearly the conservation of the regulation entails an important functional difference. Similarly, in the sex-determination cascade, *dsx* and *fru* splicing are controlled by Tra in diverse insects and this is the subject of recent reviews (63,64). Conservation of sex-specific splicing has led to the development of bio-control technologies that induce the sex-specific gene expression of toxic genes in insect agricultural pest species. For example, in the Mediterranean fruit-fly *Ceratitis capitata*, it has been possible to exploit female-specific splicing controlled by endogenous Tra/Tra2 to engineer sterilization genes into females (65). Alternative splicing is also currently being used to engineer flightless female mosquitoes, thus allowing release of sterile males to control Dengue virus spread (66).

Spliceosome dynamics and composition are very similar between human and *Drosophila* (67) and most known mammalian splicing factors have orthologues in flies (51). These include members of the SR protein family [e.g. Srp55/B52 (53–55)], the hnRNP/hrp family (8,30,55,56) and specific alternative splicing factors such as Tra2 (7,14,58), TIAR/Rox8 (5,51), PUF60/pUf68 (18), Fox proteins (68) and Nova/Pasilla (57,69). A recent study compared the action of homologues of the mammalian brain-specific splicing factors of the Nova family from six different deuterostome and protostome species. This showed, consistent with an earlier study (57), that the mechanism of Nova/Pasilla proteins in enhancer- and inhibitor-dependent splicing was conserved across the animal kingdom, but in contrast, that the exons that Nova controls are only conserved within the vertebrates. Consistent with this, Nova expression is specific to the nervous system in chordates, but not in flies (69). The synopsis of these findings is that splicing factors and their action are conserved across the animal kingdom but their location, targets and therefore their developmental functions have changed. This means that studies in flies are highly relevant to the mechanisms of alternative splicing in humans and that we need to keep flies in the bigger picture to understand the function of alternative splicing in development and the way that evolution adapts it to its needs.

PERSPECTIVE

Recently, two RNA-Seq studies have been published characterizing the full transcriptome of multiple embryonic, larval, pupal and adult time-points for both *Drosophila* sexes (34,70). A full analysis of these new data over the coming years will likely greatly impact our understanding of the diversity of alternative splicing, its mechanisms and its biological impact. For example, the study of Graveley *et al*. (34) reports 11 457 intron retention events, 3548 alternative 3′-splices sites, 3457 alternative 5′-splice sites, 1844 cassette exons, 831 multiple cassettes events and 145 mutually exclusive splicing events that shift significantly between one cell type and another. These data broadly confirm past findings (Figure 2) but also identify a huge number of new potential regulated functional alternative splice events. In addition, to gain general insights into alternative splicing regulators' targets, 26 RNA-binding proteins have been knocked down by RNAi in S2 cells and analysed by RNA-Seq and also made available through modencode (Figure 2E). When fully assimilated, this information will certainly reveal further insights into the systems biology and combinatorial control of alternative splicing.

One way of understanding this new mountain of RNA-Seq data is to group alternative splice events that shift with similar time courses in development (34). Gene ontology classifications can also give indications of what the function of alternative splicing programmes are. Similarly, bioinformatic searches in the sequences of co-regulated genes can identify RNA target motifs and hence what splicing factors might control these changes. Finally, one important way to distill the data will be to compare alternative splicing in *Drosophila* with RNA-Seq data from multiple tissues of other arthropods as these become available, on the assumption that most important regulated and functional alternative splicing events will be conserved to some extent. In summary, as with classical genetics, the easy manipulation of *Drosophila* has led to the foundation of our view of how alternative splicing is used to control development of higher organisms. Studies in this model organism will likely continue to give key insights into the complex questions now facing the splicing field, such as how organisms exploit combinatorial control by multiple splicing factors and what are the functions of programs of multiple co-regulated alternatively spliced genes.

ACKNOWLEDGEMENTS

Thanks to Philippe Fort for useful discussions and comments on the manuscript.

FUNDING

J.P.V. and J.T. were supported by the European Alternative Splicing Network of Excellence (EURASNET, FP6 life sciences, genomics and biotechnology for health) and a grant from Canceropôle and l'INCa 2009-1-RT-10-CNRS13-1. J.T. was supported by Institut Universitaire de France as a senior member. Funding for open access charge: Agence Nationale de la Recherche (ANR – 05 –BLAN – 0261 – 01); European Alternative Splicing Network of Excellence (EURASNET, FP6 life sciences, genomics and biotechnology for health).

Conflict of interest statement. None declared.

REFERENCES

1. Pan,Q., Shai,O., Lee,L.J., Frey,B.J. and Blencowe,B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–1415.
2. Wang,E.T., Sandberg,R., Luo,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P. and Burge,C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.
3. Salz,H.K. and Erickson,J.W. (2010) Sex determination in Drosophila: the view from the top. *Fly*, **4**, 60–70.
4. Bell,L.R., Horabin,J.I., Schedl,P. and Cline,T.W. (1991) Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. *Cell*, **65**, 229–239.
5. Forch,P., Merendino,L., Martinez,C. and Valcarcel,J. (2001) Modulation of msl-2 5' splice site recognition by Sex-lethal. *RNA*, **7**, 1185–1191.
6. Valcarcel,J., Singh,R., Zamore,P.D. and Green,M.R. (1993) The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature*, **362**, 171–175.
7. Lynch,K.W. and Maniatis,T. (1996) Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila doublesex* splicing enhancer. *Genes. Dev.*, **10**, 2089–2101.
8. Hartmann,B., Castelo,R., Minana,B., Peden,E., Blanchette,M., Rio,D.C., Singh,R. and Valcarcel,J. (2011) Distinct regulatory programs establish widespread sex-specific alternative splicing in Drosophila melanogaster. *RNA*, **17**, 453–468.
9. Demir,E. and Dickson,B.J. (2005) Fruitless splicing specifies male courtship behavior in Drosophila. *Cell*, **121**, 785–794.
10. Ryner,L.C., Goodwin,S.F., Castrillon,D.H., Anand,A., Villella,A., Baker,B.S., Hall,J.C., Taylor,B.J. and Wasserman,S.A. (1996) Control of male sexual behavior and sexual orientation in Drosophila by the fruitless gene. *Cell*, **87**, 1079–1089.
11. Heinrichs,V., Ryner,L.C. and Baker,B.S. (1998) Regulation of sex-specific selection of fruitless 5' splice sites by transformer and transformer-2. *Mol. Cell. Biol.*, **18**, 450–458.
12. Chandler,D.S., Qi,J. and Mattox,W. (2003) Direct repression of splicing by transformer-2. *Mol. Cell. Biol.*, **23**, 5174–5185.
13. Mattox,W. and Baker,B.S. (1991) Autoregulation of the splicing of transcripts from the transformer-2 gene of Drosophila. *Genes. Dev.*, **5**, 786–796.
14. McGuffin,M.E., Chandler,D., Somaiya,D., Dauwalder,B. and Mattox,W. (1998) Autoregulation of transformer-2 alternative splicing is necessary for normal male fertility in Drosophila. *Genetics*, **149**, 1477–1486.
15. Katzenberger,R.J., Marengo,M.S. and Wassarman,D.A. (2006) ATM and ATR pathways signal alternative splicing of Drosophila TAF1 pre-mRNA in response to DNA damage. *Mol. Cell. Biol.*, **26**, 9256–9267.
16. Katzenberger,R.J., Marengo,M.S. and Wassarman,D.A. (2009) Control of alternative splicing by signal-dependent degradation of Splicing-regulatory Proteins. *J. Biol. Chem.*, **284**, 10737–10746.
17. Taliaferro,J.M., Alvarez,N., Green,R.E., Blanchette,M. and Rio,D.C. (2011) Evolution of a tissue-specific splicing network. *Genes. Dev.*, **25**, 608–620.
18. Van Buskirk,C. and Schupbach,T. (2002) Half pint regulates alternative splice site selection in Drosophila. *Dev. Cell.*, **2**, 343–353.
19. Hastings,G.A. and Emerson,C.P. Jr (1991) Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of Drosophila. *J. Cell. Biol.*, **114**, 263–276.

20. George,E.L., Ober,M.B. and Emerson,C.P. Jr (1989) Functional domains of the Drosophila melanogaster muscle myosin heavy-chain gene are encoded by alternatively spliced exons. *Mol. Cell. Biol.*, **9**, 2957–2974.
21. Zhang,S. and Bernstein,S.I. (2001) Spatially and temporally regulated expression of myosin heavy chain alternative exons during Drosophila embryogenesis. *Mech. Dev.*, **101**, 35–45.
22. Swank,D.M., Knowles,A.F., Suggs,J.A., Sarsoza,F., Lee,A., Maughan,D.W. and Bernstein,S.I. (2002) The myosin converter domain modulates muscle performance. *Nat. Cell. Biol.*, **4**, 312–316.
23. Falkenthal,S., Graham,M. and Wilkinson,J. (1987) The indirect flight muscle of Drosophila accumulates a unique myosin alkali light chain isoform. *Dev. Biol.*, **121**, 263–272.
24. Basi,G.S., Boardman,M. and Storti,R.V. (1984) Alternative splicing of a Drosophila tropomyosin gene generates muscle tropomyosin isoforms with different carboxy-terminal ends. *Mol. Cell. Biol.*, **4**, 2828–2836.
25. Nongthomba,U., Ansari,M., Thimmaiya,D., Stark,M. and Sparrow,J. (2007) Aberrant splicing of an alternative exon in the Drosophila troponin-T gene affects flight muscle development. *Genetics*, **177**, 295–306.
26. Benoist,P., Mas,J.A., Marco,R. and Cervera,M. (1998) Differential muscle-type expression of the Drosophila troponin T gene. A 3-base pair microexon is involved in visceral and adult hypodermic muscle specification. *J. Biol. Chem.*, **273**, 7538–7546.
27. Schmucker,D., Clemens,J.C., Shu,H., Worby,C.A., Xiao,J., Muda,M., Dixon,J.E. and Zipursky,S.L. (2000) Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell*, **101**, 671–684.
28. Hattori,D., Demir,E., Kim,H.W., Viragh,E., Zipursky,S.L. and Dickson,B.J. (2007) Dscam diversity is essential for neuronal wiring and self-recognition. *Nature*, **449**, 223–227.
29. Celotto,A.M. and Graveley,B.R. (2001) Alternative splicing of the Drosophila Dscam pre-mRNA is both temporally and spatially regulated. *Genetics*, **159**, 599–608.
30. Olson,S., Blanchette,M., Park,J., Savva,Y., Yeo,G.W., Yeakley,J.M., Rio,D.C. and Graveley,B.R. (2007) A regulator of Dscam mutually exclusive splicing fidelity. *Nat. Struct. Mol. Biol.*, **14**, 1134–1140.
31. Graveley,B.R. (2005) Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell*, **123**, 65–73.
32. May,G.E., Olson,S., McManus,C.J. and Graveley,B.R. (2011) Competing RNA secondary structures are required for mutually exclusive splicing of the Dscam exon 6 cluster. *RNA*, **17**, 222–229.
33. Yang,Y., Zhan,L., Zhang,W., Sun,F., Wang,W., Tian,N., Bi,J., Wang,H., Shi,D., Jiang,Y. *et al.* (2011) RNA secondary structure in mutually exclusive splicing. *Nat. Struct. Mol. Biol.*, **18**, 159–168.
34. Graveley,B.R., Brooks,A.N., Carlson,J.W., Duff,M.O., Landolin,J.M., Yang,L., Artieri,C.G., Van Baren,M.J., Boley,N., Booth,B.W. *et al.* (2011) The developmental transcriptome of Drosophila melanogaster. *Nature*, **471**, 473–479.
35. Hsu,S.N., Yonekura,S., Ting,C.Y., Robertson,H.M., Iwai,Y., Uemura,T., Lee,C.H. and Chiba,A. (2009) Conserved alternative splicing and expression patterns of arthropod N-cadherin. *PLoS. Genet.*, **5**, e1000441.
36. Stork,T., Thomas,S., Rodrigues,F., Silies,M., Naffin,E., Wenderdel,S. and Klambt,C. (2009) Drosophila Neurexin IV stabilizes neuron–glia interactions at the CNS midline by binding to Wrapper. *Development*, **136**, 1251–1261.
37. Schneider,M. and Baumgartner,S. (2008) Differential expression of Dystroglycan-spliceforms with and without the mucin-like domain during Drosophila embryogenesis. *Fly*, **2**, 29–35.
38. Flores,C.A., Niemeyer,M.I., Sepulveda,F.V. and Cid,L.P. (2006) Two splice variants derived from a Drosophila melanogaster candidate ClC gene generate ClC-2-type Cl- channels. *Mol. Membr. Biol.*, **23**, 149–156.
39. Yu,J.Y., Upadhyaya,A.B. and Atkinson,N.S. (2006) Tissue-specific alternative splicing of BK channel transcripts in Drosophila. *Genes. Brain Behav.*, **5**, 329–339.
40. O'Dowd,D.K., Gee,J.R. and Smith,M.A. (1995) Sodium current density correlates with expression of specific alternatively spliced sodium channel mRNAs in single neurons. *J. Neurosci.*, **15**, 4005–4012.
41. Lin,W.H., Wright,D.E., Muraro,N.I. and Baines,R.A. (2009) Alternative splicing in the voltage-gated sodium channel DmNav regulates activation, inactivation, and persistent current. *J. Neurophysiol.*, **102**, 1994–2006.
42. Olson,R.O., Liu,Z., Nomura,Y., Song,W. and Dong,K. (2008) Molecular and functional characterization of voltage-gated sodium channel variants from Drosophila melanogaster. *Insect Biochem. Mol. Biol.*, **38**, 604–610.
43. Bermingham,J.R. Jr and Scott,M.P. (1988) Developmentally regulated alternative splicing of transcripts from the Drosophila homeotic gene Antennapedia can produce four different proteins. *EMBO J.*, **7**, 3211–3222.
44. Stroeher,V.L., Gaiser,J.C. and Garber,R.L. (1988) Alternative RNA splicing that is spatially regulated: generation of transcripts from the Antennapedia gene of Drosophila melanogaster with different protein-coding regions. *Mol. Cell. Biol.*, **8**, 4143–4154.
45. Reed,H.C., Hoare,T., Thomsen,S., Weaver,T.A., White,R.A., Akam,M. and Alonso,C.R. (2010) Alternative splicing modulates Ubx protein function in Drosophila melanogaster. *Genetics*, **184**, 745–758.
46. Gebelein,B., Culi,J., Ryoo,H.D., Zhang,W. and Mann,R.S. (2002) Specificity of distalless repression and limb primordia development by abdominal Hox proteins. *Dev. Cell.*, **3**, 487–498.
47. Hatton,A.R., Subramaniam,V. and Lopez,A.J. (1998) Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon–exon junctions. *Mol. Cell.*, **2**, 787–796.
48. Burnette,J.M., Miyamoto-Sato,E., Schaub,M.A., Conklin,J. and Lopez,A.J. (2005) Subdivision of large introns in Drosophila by recursive splicing at nonexonic elements. *Genetics*, **170**, 661–674.
49. Shepard,S., McCreary,M. and Fedorov,A. (2009) The peculiarities of large intron splicing in animals. *PLoS One*, **4**, e7853.
50. Uv,A.E., Harrison,E.J. and Bray,S.J. (1997) Tissue-specific splicing and functions of the Drosophila transcription factor grainyhead. *Mol. Cell. Biol.*, **17**, 6727–6735.
51. Park,J.W., Parisky,K., Celotto,A.M., Reenan,R.A. and Graveley,B.R. (2004) Identification of alternative splicing regulators by RNA interference in Drosophila. *Proc. Natl Acad. Sci. USA*, **101**, 15974–15979.
52. Kim,S., Shi,H., Lee,D.K. and Lis,J.T. (2003) Specific SR protein-dependent splicing substrates identified through genomic SELEX. *Nucleic. Acids Res.*, **31**, 1955–1961.
53. Gabut,M., Dejardin,J., Tazi,J. and Soret,J. (2007) The SR family proteins B52 and dASF/SF2 modulate development of the Drosophila visual system by regulating specific RNA targets. *Mol. Cell. Biol.*, **27**, 3087–3097.
54. Fic,W., Juge,F., Soret,J. and Tazi,J. (2007) Eye Development under the control of SRp55/B52-mediated alternative splicing of eyeless. *PLoS ONE*, **2**, e253.
55. Blanchette,M., Green,R.E., Brenner,S.E. and Rio,D.C. (2005) Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. *Genes. Dev.*, **19**, 1306–1314.
56. Blanchette,M., Green,R.E., MacArthur,S., Brooks,A.N., Brenner,S.E., Eisen,M.B. and Rio,D.C. (2009) Genome-wide analysis of alternative pre-mRNA splicing and RNA-binding specificities of the Drosophila hnRNP A/B family members. *Mol. Cell.*, **33**, 438–449.
57. Brooks,A.N., Yang,L., Duff,M.O., Hansen,K.D., Park,J.W., Dudoit,S., Brenner,S.E. and Graveley,B.R. (2011) Conservation of an RNA regulatory map between Drosophila and mammals. *Genome. Res.*, **21**, 193–202.
58. Du,C., McGuffin,M.E., Dauwalder,B., Rabinow,L. and Mattox,W. (1998) Protein phosphorylation plays an essential role in the regulation of alternative splicing and sex determination in Drosophila. *Mol. Cell.*, **2**, 741–750.
59. Salles,C., Mevel-Ninio,M., Vincent,A. and Payre,F. (2002) A germline-specific splicing generates an extended ovo protein isoform required for Drosophila oogenesis. *Dev. Biol.*, **246**, 366–376.
60. de la Mata,M., Alonso,C.R., Kadener,S., Fededa,J.P., Blaustein,M., Pelisch,F., Cramer,P., Bentley,D. and

Kornblihtt,A.R. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell.*, **12**, 525–532.
61. Fox-Walsh,K.L., Dou,Y., Lam,B.J., Hung,S.P., Baldi,P.F. and Hertel,K.J. (2005) The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proc. Natl Acad. Sci. USA*, **102**, 16176–16181.
62. McManus,C.J., Duff,M.O., Eipper-Mains,J. and Graveley,B.R. (2010) Global analysis of trans-splicing in Drosophila. *Proc. Natl Acad. Sci. USA*.
63. Salvemini,M., Polito,C. and Saccone,G. (2010) Fruitless alternative splicing and sex behaviour in insects: an ancient and unforgettable love story? *J. Genet.*, **89**, 287–299.
64. Shukla,J.N. and Nagaraju,J. (2010) Doublesex: a conserved downstream gene controlled by diverse upstream regulators. *J. Genet.*, **89**, 341–356.
65. Fu,G., Condon,K.C., Epton,M.J., Gong,P., Jin,L., Condon,G.C., Morrison,N.I., Dafa'alla,T.H. and Alphey,L. (2007) Female-specific insect lethality engineered using alternative splicing. *Nat. Biotechnol.*, **25**, 353–357.
66. Fu,G., Lees,R.S., Nimmo,D., Aw,D., Jin,L., Gray,P., Berendonk,T.U., White-Cooper,H., Scaife,S., Kim Phuc,H. *et al.* (2010) Female-specific flightless phenotype for mosquito control. *Proc. Natl Acad. Sci. USA*, **107**, 4550–4554.
67. Herold,N., Will,C.L., Wolf,E., Kastner,B., Urlaub,H. and Luhrmann,R. (2009) Conservation of the protein composition and electron microscopy structure of Drosophila melanogaster and human spliceosomal complexes. *Mol. Cell. Biol.*, **29**, 281–301.
68. Damianov,A. and Black,D.L. (2010) Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA*, **16**, 405–416.
69. Irimia,M., Denuc,A., Burguera,D., Somorjai,I., Martin-Duran,J.M., Genikhovich,G., Jimenez-Delgado,S., Technau,U., Roy,S.W., Marfany,G. *et al.* (2011) Stepwise assembly of the Nova-regulated alternative splicing network in the vertebrate brain. *Proc. Natl Acad. Sci. USA*, **108**, 5319–5324.
70. Daines,B., Wang,H., Wang,L., Li,Y., Han,Y., Emmert,D., Gelbart,W., Wang,X., Li,W., Gibbs,R. *et al.* (2011) The Drosophila melanogaster transcriptome by paired-end RNA sequencing. *Genome. Res.*, **21**, 315–324.
71. Celniker,S.E., Dillon,L.A., Gerstein,M.B., Gunsalus,K.C., Henikoff,S., Karpen,G.H., Kellis,M., Lai,E.C., Lieb,J.D., MacAlpine,D.M. *et al.* (2009) Unlocking the secrets of the genome. *Nature*, **459**, 927–930.
72. Hsu,T., Gogos,J.A., Kirsh,S.A. and Kafatos,F.C. (1992) Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene. *Science*, **257**, 1946–1950.
73. Scamborova,P., Wong,A. and Steitz,J.A. (2004) An intronic enhancer regulates splicing of the twintron of Drosophila melanogaster prospero pre-mRNA by two different spliceosomes. *Mol. Cell. Biol.*, **24**, 1855–1869.
74. Borah,S., Wong,A.C. and Steitz,J.A. (2009) Drosophila hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron. *Proc. Natl Acad. Sci. USA*, **106**, 2577–2582.