

Alternative RNA splicing regulation in the testis

David J Elliott and Sushma N Grellscheid

Institute of Human Genetics, University of Newcastle, International Centre for Life, Central Parkway, Newcastle NE1 3BZ, UK

Correspondence should be addressed to D J Elliott; Email: david.elliott@ncl.ac.uk

Abstract

Alternative splicing regulation has been shown to be critically important for several developmental pathways. It is particularly prevalent in the testis, which is the site of an extensive adult developmental programme. Alternative splicing is controlled by a splicing code, in which transcripts respond to subtle cell type-specific variations in positive and negative trans-acting RNA-binding proteins according to their unique set of binding sites for these proteins. Because of their unique combinations of cis-acting sequence elements, specific transcripts are able to respond individually to this code. In this review, we discuss how this code may be deciphered in germ cells to mediate a splicing response.

Reproduction (2006) **132** 811–819

Alternative splicing is very frequent in the testis

In 1993, the Nobel Prize in Physiology or Medicine was awarded to Phil Sharp and Richard Roberts for the discovery that many protein-coding genes in eukaryotes are split into exons and introns (Fig. 1). Exons comprise the sequences which are included in the mature mRNA, which is exported into the cytoplasm (i.e. coding information, and 5' and 3' UTRs), while introns are removed in the nucleus before the RNA leaves. This means that most eukaryotic genes are much longer than their mature transcripts. In humans, the average gene length is ~27 kb, while the average mature transcript which is exported to the cytoplasm is reduced in size to ~1.3 kb (Lander *et al.* 2001, Venter *et al.* 2001). The implication of these numbers is that introns are usually much longer than exons. Splicing is carried out in the nucleus by a large macromolecular complex called the spliceosome, which consists of both proteins and small nuclear RNAs. Although spliceosomes might be pre-assembled in some cases, they are thought to largely assemble *de novo* on transcripts (Gornemann *et al.* 2005, Lacadie & Rosbash 2005, Azubel *et al.* 2006).

While splicing is an essential step in gene expression, alternative splicing pathways provide a versatile means of regulating gene expression using different combinations of exons from the same primary transcript (often called the pre-mRNA), resulting in the generation of different mature transcripts (reviewed by Matlin *et al.* 2005). Hence, while a gene may be widely or ubiquitously transcribed into the same pre-mRNA,

variant transcripts from this gene might be constructed either within the same cell, or between different tissues and cell types (summarised in Fig. 1a). Exons might be skipped or included in these variant transcripts (these are called 'cassette exons'), or modified (using alternative 5' or 3' splice sites). This area of research is presently very topical for three reasons. First, alternative splicing vastly extends the coding capacity of the human genome, enabling many more proteins to be made than the approximately 30 000 full-length genes which are apparently encoded (Stamm 2002). Secondly, it establishes a new level at which gene expression can be controlled. One aspect of this might be the splicing of exons which contain premature termination codons (PTC). This can lead to subsequent transcript degradation by the nonsense-mediated decay pathway, although recent data suggest this might be quantitatively less important than previously thought for regulating global transcript levels (Pan *et al.* 2006). A third reason is the sheer numbers involved. Even conservative estimates suggest that about 70% of all human genes might be alternatively spliced, and up to half of these alter the reading frame of the transcript (Johnson *et al.* 2003).

Because of the short relative length of most alternative exons compared with the full transcript length, alternatively spliced transcripts are usually not visible on northern blots. More often they are identified bioinformatically from cDNA or expressed sequence tag (EST) comparisons, and then confirmed empirically by RT-PCR using primers which bind to constitutive exons flanking the alternative exon (Fig. 1c). This results in a larger

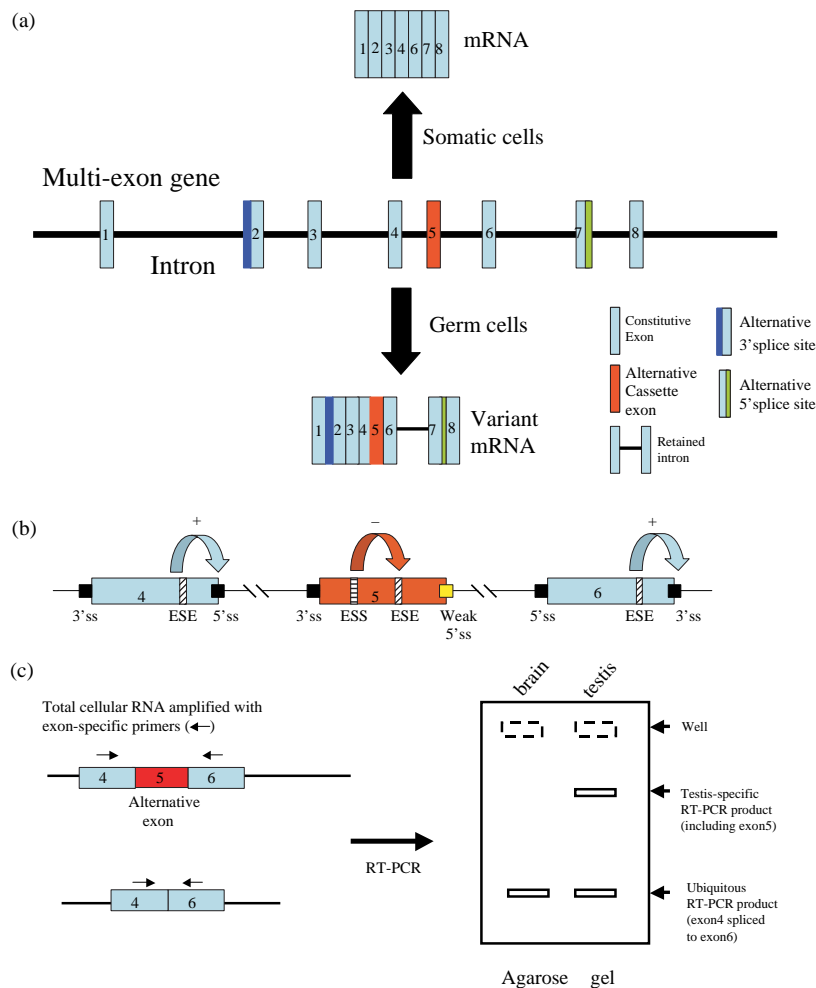


Figure 1 Alternative splicing. (a) Most eukaryotic genes are split into exons (coloured boxes) joined by non-coding introns (horizontal lines). A cartoon of a typical 8 exon gene is shown, with the constitutive exons spliced together (top) and possible alternative splices included (bottom). Constitutive exons (light blue) are spliced together in all tissues, while alternative exons (red: full-length alternatively spliced exons are called cassette exons) are only spliced into transcripts in some tissues (e.g. testis) or cell types (e.g. germ cells). Other alternative splicing events use alternative 5' and 3' splice sites on the constitutive exons (making these exons longer or shorter), or retain introns in the final transcript. More detail of exons 4–6 is shown in (b), in which the constitutive exons have strong splice sites (black boxes), and the alternative exon has a weak 5' splice site (yellow box). All the three exons contain enhancer sequences (ESE) which assist their recognition by the spliceosome, while the alternative exon 5 also contains a silencer (ESS) which acts to block the action of the ESE, possibly by binding a negative regulator. (c) Alternatively spliced exons are frequently detected by reverse transcription-PCR (RT-PCR) on total RNA using primers specific for flanking exons. In this case, the alternative exon 5 is spliced into the transcript in the testis but not in the brain, so on agarose gel electrophoresis there will be a larger size product identified in the testis.

RT-PCR product, which can be resolved by gel electrophoresis where there is an additional exon, and a corresponding difference in size if an alternative splice site is used. Bioinformatic comparisons have shown that alternative splicing is particularly prevalent in the testis (Yeo *et al.* 2004). Like in other tissues, testis-specific splicing includes examples of cassette exons which are included or repressed in the testis, alternative 5' or 3' splice sites which are used in the testis compared with other tissues, and intron retention (in this case exons are not spliced together, but joined by the intervening intron which is not removed). Thirty percent of human ESTs (expressed sequence tags, which correspond to partial cDNA sequences) in the testis were recorded to undergo

alternative splicing, a frequency just less than the brain. The testis was also the second most abundant anatomic site for cassette exons, the most frequent form of alternative splicing (Yeo *et al.* 2004).

Why is alternative splicing so abundant in the testis? A relatively unique feature of the testis is that it is the site of an extensive adult developmental process. Male germ cell development continues lifelong, and the pathway involves the coordinated division and differentiation of huge numbers of cells. An adult human testis produces 10^8 sperm/day (Gilbert 2000). Consistent with this explanation, alternative splicing is also very abundant in the thymus, where it plays a key role in adult T-cell development and differentiation (Lynch 2004).

Alternative splicing is likely to be important in many if not all developmental pathways and has been proven to be essential in the mouse for normal cardiac, neural and thymus function (Wang *et al.* 2001, Ding *et al.* 2004, Xu *et al.* 2005). Alternative splicing is probably very important throughout germ cell development, but is known to play a critical role in transcriptional re-programming after meiosis where it converts the transcription factor cAMP responsive element modulator (CREM) from an antagonist to a potent activator required for the transcription from an array of promoters in round spermatids (Foulkes *et al.* 1992). However, other factors may also be important in driving the volume of alternative splicing in the testis, since alternative splicing is also particularly high in the adult liver which is terminally differentiated and not particularly complex in cell type content (Yeo *et al.* 2004). A further possibility for explaining the high levels of alternative splicing in the testis is that although specific examples of alternative splicing might be crucial, there might be a higher level of 'background' noise of alternative splicing errors in the testis. This latter hypothesis was suggested by the observation that there is an increased frequency of species-specific splicing events in the testis (not conserved between mouse and human) compared with the brain (Kan *et al.* 2005). This might perhaps be due to either increases in cell proliferation, decreases in quality control or global fluctuations in splicing regulators (particularly down-regulation of heterogeneous ribonucleoprotein A1 (hnRNPA1), as discussed below). Like other conserved features, alternative splices which are shared between humans and mouse might have been maintained since they play an important function.

The generation of novel and potentially tissue/cell type-specific proteins together with new levels in gene expression control are likely to have made an important contribution to the development of sophisticated and rapidly evolving metazoan tissues. Reproductive systems are highly evolving between species, including between the mouse and the human (Wyckoff *et al.* 2000), and an increase in alternative splicing might be one of the mechanisms driving this. Alternative splicing is

particularly compatible with evolutionary experimentation since it provides a possible escape route for the testing of evolutionary variants. Most alternative splices produce the original exon combination in parallel with the alternative splice and so enable new variants to be made, while the original gene is left intact to provide normal gene function. Like the reproductive system, the human brain is also evolving rapidly in the primate lineage.

Alternative exon recognition

How exons are recognised by the splicing machinery lies at the root of alternative splicing. This initial recognition is largely by a process of 'exon definition'. The first interactions between spliceosome components and pre-mRNAs take place across individual exons, and only later across introns between separate exons which are destined to be spliced together. Constitutive exons typically have 'strong' splice sites which conform to the ideal consensus (see Box 1), while alternatively spliced exons often tend to have splice sites which differ from the consensus and so are sub-optimally recognised (a weak 5' ss is shown as a yellow box in Fig. 1b). Alternatively spliced exons might also be relatively short (this makes the protein-protein interactions across these exons during the initial steps of exon definition spatially more difficult to achieve). Another important factor can be secondary structure of RNA. RNA is usually thought of as single-stranded, but forms extensive intramolecular base pairing, which can occlude splice sites. The result of these shortcomings are that weak alternatively spliced exons need additional help to be recognised, and it is this 'help' which provides the mechanistic base of alternative splicing.

Two further types of *cis*-acting sequence which provide this help and control the efficiency of exon recognition are called enhancers and silencers (Box 2). Exon inclusion depends on the frequently complex arrangement of these positive and negative control elements in the transcript, and different transcripts will respond differently to the same cellular splicing code

Box 1 Splice sites are recognised and joined together by the spliceosome. Exons are flanked by short conserved sequences called splice sites which are recognised and spliced together by the spliceosome (Fig. 1b). Each exon has a 5' splice site (5'ss) and a 3' splice site (3'ss). Slightly confusingly, these are labelled relative to the intron being removed rather than the exon being spliced: hence the 3' splice site is within the intron just upstream of the exon to be spliced, and the 5' splice site is just downstream and overlapping the exon to be spliced. Strong splice sites are efficiently and quantitatively recognised by the spliceosome (black 5'ss and 3'ss boxes in Fig. 1b). This consensus is based on the degree of partial bp complementarity between snRNA components of the snRNPs and the splice sites. Strong splice sites have more complementarity with the snRNA components of the spliceosome. In Fig. 1b, exon 5 has a weak (poorly complementary to U1 snRNA) 5' splice site (shown as a yellow box). This results in exon 5 normally being skipped, but included in some tissues with appropriate *trans*-acting factors which can enable this exon to be recognised. This situation is analogous to the testis-specific splicing of an exon in the HIPK3 transcript discussed in the text, in which the 5' splice site is weak.

Box 2 Splicing enhancers and silencers moderate exon recognition by the spliceosome. These are short nucleotide sequences which bind splicing activator proteins (e.g. Tra2 or SR proteins) or repressors (e.g. hnRNPA1) respectively. They can be either located within exons (exonic splicing enhancers abbreviated to ESEs or exonic splicing silencers abbreviated to ESSs) or within introns (intronic splicing enhancers abbreviated to ISE or intronic splicing silencers abbreviated to ISS).

based on their unique combination of these *cis*-acting sites. Hence, combinatorial control of splicing through an interplay of *cis*- and *trans*-acting factors provides an extremely flexible system for the coordinate and differential regulation of all intron-containing genes in the human genome. Splicing enhancers are also important for the splicing of constitutive exons (Fig. 1b), and help the splicing machinery identify these exons from potential 'decoy' sequences which might appear to the casual observer as exons with reasonably strong splice sites (Grellscheid & Smith 2006). In fact, most exons can be considered to be on a spectrum ranging from very weak to very strong, with multiple factors positively and negatively contributing, and the balance of these forces determining whether the exon is included or not.

Another factor impinging on alternative splicing in the testis is transcription. The rate of transcription elongation can have a strong effect on splice site choice since it controls the rate at which competing splice sites are synthesised and so become available to the splicing machinery (Batsche *et al.* 2006). Overall genomic levels of transcription are very high in meiosis, which is possibly linked to DNA repair (Hackstein *et al.* 2000). Following meiosis in the round spermatid a number of subunits of RNA polymerase II are found at particularly high concentration, where they contribute to selection and frequent use of promoters of an array of developmentally important genes (Schmidt & Schibler 1995, Schmidt 1996). This affects the transcriptional strength and rates of elongation of specific genes, and so likely also impacts on their alternative splicing pathways. Patterns of 3' end formation are also somewhat unusual in the testis, and frequently the canonical AAUAAA polyadenylation signal is not used (Wallace *et al.* 1999).

The germ cell splicing code

An important current paradigm for understanding alternative splicing in the testis is that of the cellular splicing code in which the cell type-specific complement of *trans*-acting RNA-binding proteins controls the output of spliced transcripts made (Fu 2004, Matlin *et al.* 2005). Splicing regulation involves combinatorial control, in which SR (serine/arginine-rich) proteins, heterogeneous ribonucleoproteins (hnRNPs) and other splicing regulators combine with transcript-specific collections of splicing enhancer and repressor sequences in the nascent RNA as it emerges from RNA polymerase II.

These sequences act as punctuation marks which are recognised by the spliceosome and determine splicing decisions. RNA is likely to be non-specifically associated with hnRNP proteins immediately after transcription (defined in *in vitro* splicing reactions as the H complex; Jurica & Moore 2003). Subsequently, these non-specific associations might be replaced by factors recognising the 5' and 3' splice sites, and splicing enhancers/silencers leading to the formation of a splicing-competent E complex. While very early complexes have been more biochemically intractable than the later splicing complexes, H complex protein components might have a significant effect on alternative splicing choices through recruitment or exclusion of key splicing factors (Black 2003). *In vivo*, early splicing complexes and biochemical splicing events are likely to be associated with nascent transcripts attached to chromatin (Swinburne *et al.* 2006).

Biochemical and proteomic analyses of splicing competent extracts made from transformed cell lines have identified many protein components of 'standard spliceosomes' (Zhou *et al.* 2002, Jurica & Moore 2003). These essential components are considered to be general splicing factors, although in fact their levels of expression are frequently differentially regulated between tissues at the RNA level (<http://genome.ucsc.edu/cgi-bin/hgNear>). These variations are likely to impact on the regulation of alternative splicing: in *Drosophila* development, even quite subtle differences in the expression of these general factors correlate with significant changes in cellular splicing patterns (Park *et al.* 2004). These patterns of variation may be adjusted/amplified at the protein level through differential translation of these mRNAs. Translation can be differentially regulated and this form of regulation is particularly common in the testis (Iguchi *et al.* 2006).

Substantial changes in the expression of both ubiquitous, tissue and cell type-specific nuclear RNA-binding proteins take place during germ cell development. The otherwise ubiquitously expressed hnRNPA1 is expressed in some spermatogonia and Sertoli cells, but reduced in meiotic cells (Fig. 2; Kamma *et al.* 1995). hnRNPA1 is normally one of the most abundant proteins in the nucleus, where it operates as a potent splicing repressor. Systematic evolution of ligands by exponential enrichment (SELEX) experiments have identified the sequence UAGGG(A/U) as a high affinity hnRNPA1-binding site, although hnRNPA1 might bind RNA more generally at low affinity and play a packaging role in H complexes

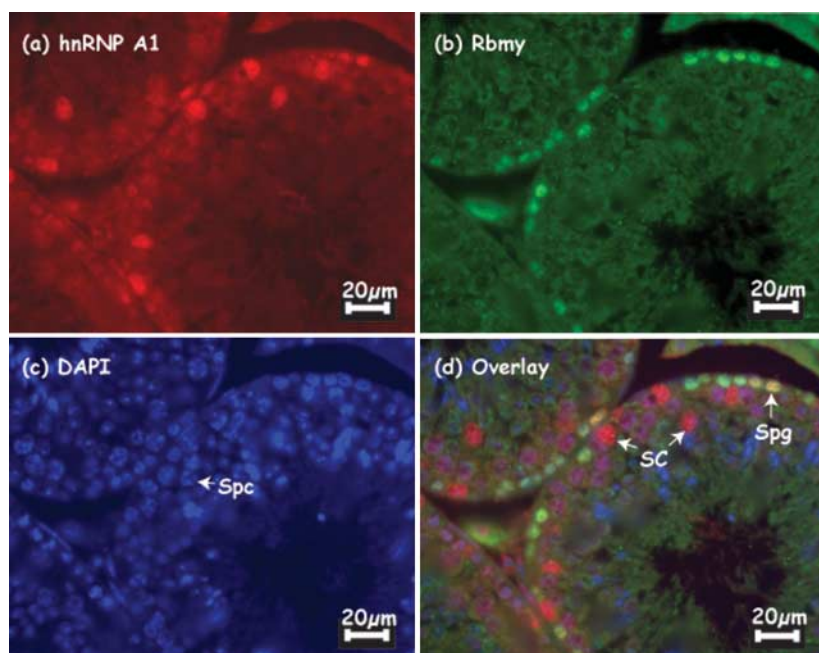


Figure 2 The global splicing repressor hnRNP A1 is down-regulated over germ cell development. A section of mouse testis is shown, stained for (a) hnRNP A1 (indirect immunofluorescence with a MAB specific for hnRNP A1. Only some spermatogonia and Sertoli cell nuclei are positive for hnRNP A1 expression, while meiotic germ cells and post-meiotic stages are negative); (b) Rbm (spermatogonia show nuclear staining. The arrowed nucleus expresses Rbm and hnRNP A1, although some adjacent cells only express detectable Rbm and no hnRNP A1); (c) DAPI (4'-6 diamidino-2-phenylindole DNA staining to show the position of all cell nuclei) and (d) an overlay of these pictures. Abbreviations used: Spg, spermatogonium; SC, Sertoli cells; Spc, spermatocytes.

(Burd & Dreyfuss 1994). Consistent with a role for hnRNP A1 in regulating testis-specific splicing events, the slightly shorter hnRNP A1-binding sequence UAGG has been found to be enriched in cassette exons from human testis (Yeo *et al.* 2004), suggesting hnRNP A1 might play a role in regulating these exons. hnRNP A1 functions by occluding binding sites for splicing activators and multimerises through its glycine-rich C-terminal domain, so binding of one hnRNP A1 molecule to a single high affinity target site can result in oligomerisation of multiple molecules along the transcript resulting in a zone of repression on the target RNA, getting in the way of other potential protein-binding sites (Matlin *et al.* 2005). hnRNP A1 proteins bound to spatially separate sites can also interact with each other, resulting in looping out of regulated exons. hnRNP A1 also plays roles in transcriptional regulation at promoters and in telomere maintenance (Xia 2005, Zhang *et al.* 2006). Since spermatogonia include the stem cells within the testis (as a sub-population of different types of germ cells which are able to divide or proliferate according to their cellular programme), the hnRNP A1 down-regulation after this stage may be associated with the absence of telomere maintenance in differentiating cells. However Sertoli cells, which are also terminally differentiated, do express hnRNP A1 (Fig. 2; Kamma *et al.* 1995).

A number of nuclear RNA-binding proteins are primarily or exclusively expressed in the testis, including

RBM Y (an acronym of RNA-binding motif encoded by the Y chromosome), hnRNP G-T (heterogeneous ribonucleoprotein G-testis) and T-STAR (testis-signal transduction and RNA processing). RBMY was originally identified as a gene within a region called Azoospermia factor b microdeleted from the Y chromosome in some infertile men, suggesting that pre-mRNA splicing pathways might be disrupted in these men (reviewed by Elliott 2000). RBMY is only expressed in germ cells, but is related to a more ubiquitously expressed gene called RBMX on the X chromosome. RBMX is expressed ubiquitously, and encodes a nuclear RNA-binding protein called hnRNP G (a member of the group of heterogeneous ribonucleoproteins) ((Mazeyrat *et al.* 1999), reviewed in (Elliott 2004)). Although the precise RNA target sequences of RBMY or hnRNP G have not yet been reported, both proteins can interact with splicing factors in the SR and SR-related groups, and influence pre-mRNA splicing patterns. The fact that RBMY seems to be conserved on the Y chromosome in all mammals is an indication that it likely performs an important function: while once it was approximately equal in size to the extant X chromosome, the Y chromosome has shrunk to its present size by getting rid of unnecessary genes. A third homologous gene called hnRNP G-T which is most closely related to RBMX is expressed exclusively in the testis (Elliott *et al.* 2000). hnRNP G-T arose by retrotransposition from RBMX before

divergence of the mouse and the human lineages, but unlike most retrogenes has retained its ability to encode a functional protein, suggesting it is under selective pressure. hnRNPG-T is specifically expressed in meiotic prophase, when the X and Y chromosomes are inactivated in a structure called the XY body, and then at a slightly lower level after meiosis. Meiotic prophase lasts several days in the mouse, and hnRNPG-T most likely provides replacement gene function for RBMX during this time (RBMX is expressed ubiquitously apart from during meiotic prophase, and so is likely to have an essential housekeeping role).

Both RBMY and hnRNPG-T interact with two other proteins, Sam68 (Src associated in mitosis) and T-STAR (Elliott 2004, Venables *et al.* 1999). These are members of the STAR family of splicing regulators, which are thought to link splicing regulation with signal transduction: as well as KH (hnRNPK homology) domains which mediate protein–protein and protein–RNA interactions, they also contain SH2 and SH3 domains which are frequently found in signalling molecules. Sam68 is ubiquitously expressed, while T-STAR is mainly expressed in the testis in the adult (Venables *et al.* 2004) but also in the developing nervous system (<http://expression.gnf.org/cgi-bin/index.cgi>). Both T-STAR and Sam68 can potentially activate splicing of the alternatively spliced exon V5 in transcripts encoded by the CD44 minigene. The mechanism for this might include interactions with chromatin remodelling factors which slow transcriptional elongation, thereby enabling selection of the weaker CD44 exon V5 splice sites (Batsche *et al.* 2006).

Mechanisms of testis-specific splicing: HIPK3 – a model for testis-specific alternative splicing

How might these changes in the germ cell splicing code induce testis-specific splicing patterns? The molecular mechanisms controlling testis-specific splicing of most transcripts, including CREM, are not well understood. To remedy this, we examined the molecular mechanisms regulating the splicing of a testis-specific exon which we discovered in the homeodomain interacting protein kinase 3 (HIPK3) transcript (Venables *et al.* 2005). HIPK3 encodes a protein which is normally targeted to PML bodies within the cell nucleus, sites which are associated with phosphorylation of p53 and the control of apoptosis (Dellaire & Bazett-Jones 2004). An alternatively spliced cassette exon 'T' (for 'testis') is included between exons 3 and 4 in the human testis in some transcripts, but this is skipped in every other tissue (i.e. exons 3 and 4 are directly joined, skipping exon 'T', Fig. 3a). Since it has previously been difficult to analyse mechanisms of human germ cell splicing because of the difficulty in getting fresh human germ cells and lack of a human transformed

germ cell line, we developed a FACS-based assay which allows us to make germ cell-specific changes to the cellular splicing code in somatic cell lines, and then monitor the resulting splicing response of endogenous transcripts (Venables *et al.* 2005).

There are two questions to be answered regarding the regulation of the HIPK3 'T' exon: why is it repressed in somatic cells (i.e. elsewhere in the body), but included in the testis? A critical factor preventing the HIPK3 'T' exon being normally recognised by the splicing machinery outside the testis is likely that it has a weak 5' splice site (Fig. 3b). Just downstream of this 5' splice site is a potentially stronger site which is not used, possibly since it overlaps with a quartet of G residues. Such an arrangement of G residues is very infrequently found associated with 5' splice sites, but where it is found is often associated with exon skipping by creating a high affinity hnRNPA1 site (Han *et al.* 2005). Within the HIPK3 'T' exon is a purine-rich sequence element. In nuclear extracts this binds the positive splicing regulators Tra2 β (transformer 2), ASF/SF2 (alternative splicing factor/splicing factor 2) and SRp40 (serine/arginine rich protein 40), and the negative splicing regulators hnRNPA1 and hnRNPH (predicted binding sites for these splicing regulators are shown in Fig. 3b). Consistent with binding both positive and negative splicing regulators binding the purine-rich sequence, mutagenesis showed that this element is bifunctional: this led to de-repression (i.e. weak constitutive splicing) as well as rendering the exon unable to respond to Tra2 β .

How does this arrangement of *cis*-acting sequences in the HIPK3 'T' exon respond to changes in the cellular splicing code, and how does this provide cell type-specific regulation in the testis? A physiological shift in the balance of positive and negative splicing regulators within the testis most probably act together to induce splicing of HIPK3 'T' (Fig. 3c). In somatic cells, splicing is inhibited by hnRNPA1 which acts to exclude Tra2 β binding, possibly in cooperation with hnRNPH. In germ cell development, hnRNPA1 is physiologically down-regulated. This is coupled with a Tra2 β protein expression level several fold higher in the testis than in somatic cells, and an increase in the hypophosphorylated form of Tra2 β , which preferentially binds to the HIPK3 'T' exon. This enables Tra2 β to activate the closest 5' splice site even though this is weaker, rather than use the stronger downstream 5' splice site, and cause splicing of the HIPK3 'T' exon.

Conclusions and future directions

A good analogy for understanding splicing regulation might be found in the Chinese philosophical notion of Yin and Yang which proposes complementary and polar forces, whose tension and conflict resolve in balance, order and change. This global analogy seems

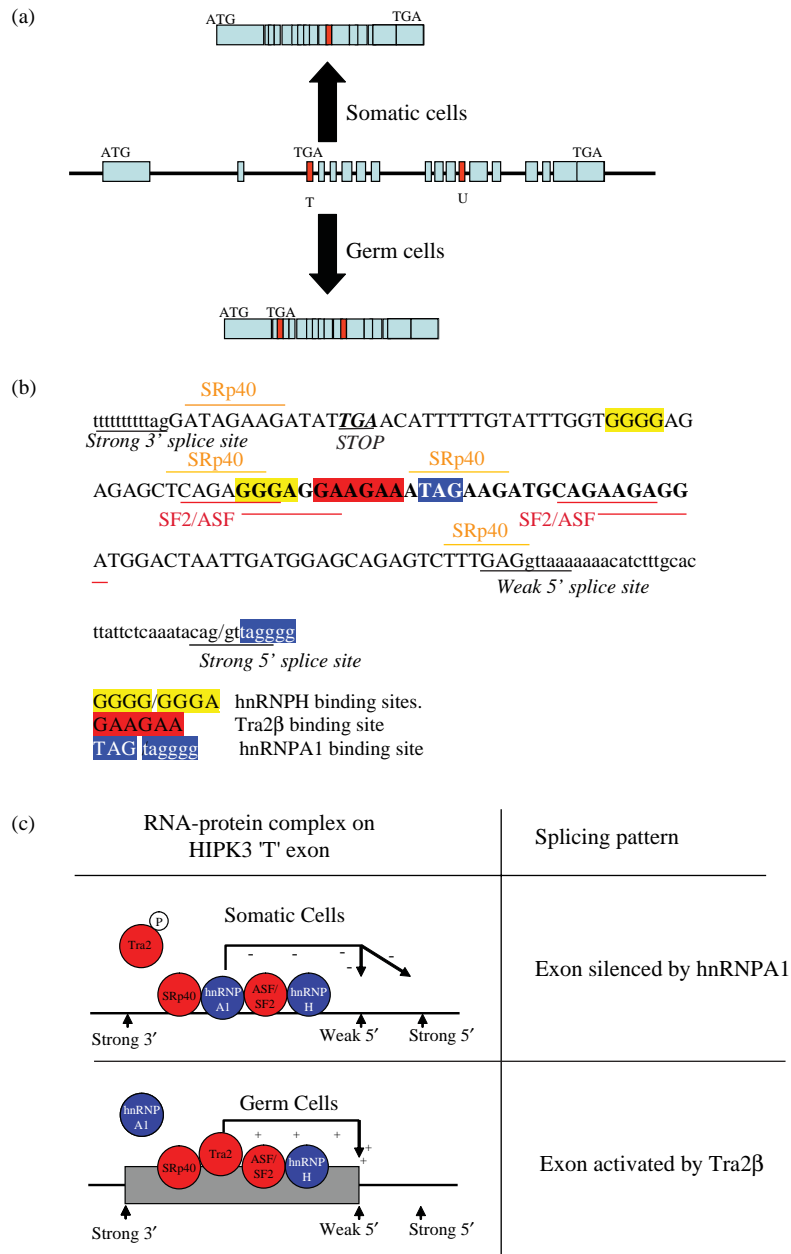


Figure 3 Mechanism of splicing of a testis-specific exon. (a) The 'T' exon in the HIPK3 transcript is only spliced in the testis and not in somatic cells. This results in the inclusion of a premature stop codon (TGA), while the somatic-spliced transcript uses a downstream TGA stop codon. A second alternatively spliced 'U' transcript is variably included in different tissues (both somatic and germ cells). (b) Sequence of the HIPK3T exon, showing the position of high affinity sites for *trans*-acting splicing regulators identified as binding in nuclear extracts and to regulate splicing of HIPK3 'T' exon in *in vitro* splicing extracts and/or in transfected cells. Note that while the hnRNPA1 site identified downstream of the regulated exon at the unused strong 5' splice site is a good match to the matrix derived from high affinity SELEX winners, the exonic UGA hnRNPA1 site is only a partial match but has been shown to be sufficient for mediating hnRNPA1 function within an alternative splice in the hnRNPA1 pre-mRNA itself. It is also likely to be functional since splicing activity mediated by the purine-rich element responds to hnRNPA1. Note most of the control elements are downstream of the stop codon, the sequence of which is not required for coding. (c) Cartoon showing splicing regulation of the HIPK3 'T' exon. Splicing activators are shown in red and splicing repressors in blue. Full details are given in the text.

appropriate given the power of alternative splicing in development and evolution.

What are the major challenges at the moment for understanding regulation of alternative splicing regulation in the testis? One important priority for

deciphering the germ cell splicing code will be to identify the RNA target sequences of testis-specific splicing regulators, such as RBM, T-STAR and hnRNPG-T. This can be achieved by SELEX (systematic evolution of ligands through exponential enrichment)

approaches which involve the reiterative selection of high affinity binding sites such as the approach discussed above to identify high affinity hnRNPA1 sites (Burd & Dreyfuss 1994); by identifying mis-regulated transcripts in mutant mice (Ule *et al.* 2005b); by examining the response of minigenes encoding known testis-specific alternative splicing substrates to *trans*-acting factors (Venables *et al.* 2005) and by crosslinking immunoprecipitation (CLIP) to directly link RNA to proteins, followed by immunoprecipitation to identify RNA targets of specific proteins in germ cells (Ule *et al.* 2005a). The functional role of these proteins in germ cell development can be tested in knockout mice, similar to the experiments in which the gene encoding the cytoplasmic RNA-binding protein Dazl was shown to be required for germ cell development (Ruggiu *et al.* 1997). Germ cells are particularly suited for a genetic approach to examine splicing changes in the context of a developmental pathway, since male germ cell development continues in the adult. Global patterns of exon incorporation over germ cell development in the mouse can be mapped using exon or splice junction microarrays using RNA from both wild-type and mutant mice (Johnson *et al.* 2003, Pan *et al.* 2004). In addition to the behaviour of whole populations of exons, detailed mapping of the response elements within single transcripts and establishing how these mediate responses to changes in the germ cell splicing code will be important to understand the molecular basis of splicing regulation in the testis.

Acknowledgements

The work in our group has been funded by the Wellcome Trust, BBSRC, Cancer research UK, Royal Society and Newcastle Healthcare Charities. We thank our colleagues for their contribution to the work described in this review, particularly Caroline Dalgliesh and Drs Julian Venables, Cyril Bourgeois and James Stevenin. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References

- Azubel M, Habib N, Sperling R & Sperling J 2006 Native spliceosomes assemble with pre-mRNA to form supraspliceosomes. *Journal of Molecular Biology* **356** 955–966.
- Batsche E, Yaniv M & Muchardt C 2006 The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nature Structural & Molecular Biology* **13** 22–29.
- Black DL 2003 Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* **72** 291–336.
- Burd CG & Dreyfuss G 1994 RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO Journal* **13** 1197–1204.
- Dellaire G & Bazett-Jones DP 2004 PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* **26** 963–977.
- Ding JH, Xu X, Yang D, Chu PH, Dalton ND, Ye Z, Yeakley JM, Cheng H, Xiao RP, Ross J *et al.* 2004 Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. *EMBO Journal* **23** 885–896.
- Elliott DJ 2000 RBMY genes and AZFb deletions. *Journal of Endocrinological Investigation* **23** 652–658.
- Elliott DJ 2004 The role of potential splicing factors including RBMY, RBMX, hnRNP-G-T and STAR proteins in spermatogenesis. *International Journal of Andrology* **27** 328–334.
- Elliott DJ, Venables JP, Newton CS, Lawson D, Boyle S, Eperon IC & Cooke HJ 2000 An evolutionarily conserved germ cell-specific hnRNP is encoded by a retrotransposed gene. *Human Molecular Genetics* **9** 2117–2124.
- Foulkes NS, Mellstrom B, Benusiglio E & Sassone-Corsi P 1992 Developmental switch of CREM function during spermatogenesis: from antagonist to activator. *Nature* **355** 80–84.
- Fu XD 2004 Towards a splicing code. *Cell* **119** 736–738.
- Gilbert SC 2000 In *Developmental Biology*, 6th edition. Sinauer Associates Inc: Sunderland, MA, USA.
- Gornemann J, Kotovic KM, Hujer K & Neugebauer KM 2005 Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular Cell* **19** 53–63.
- Grellscheid SN & Smith CW 2006 An apparent pseudo-exon acts both as an alternative exon that leads to nonsense-mediated decay and as a zero-length exon. *Molecular and Cellular Biology* **26** 2237–2246.
- Hackstein JH, Hochstenbach R & Pearson PL 2000 Towards an understanding of the genetics of human male infertility: lessons from flies. *Trends in Genetics* **16** 565–572.
- Han K, Yeo G, An P, Burge CB & Grabowski PJ 2005 A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS Biology* **3** e158.
- Iguchi N, Tobias JW & Hecht NB 2006 Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. *PNAS* **103** 7712–7717.
- Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loecher PM, Armour CD, Santos R, Schadt EE, Stoughton R & Shoemaker DD 2003 Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302** 2141–2144.
- Jurica MS & Moore MJ 2003 Pre-mRNA splicing: awash in a sea of proteins. *Molecular Cell* **12** 5–14.
- Kamma H, Portman DS & Dreyfuss G 1995 Cell type-specific expression of hnRNP proteins. *Experimental Cell Research* **221** 187–196.
- Kan Z, Garrett-Engele PW, Johnson JM & Castle JC 2005 Evolutionarily conserved and diverged alternative splicing events show different expression and functional profiles. *Nucleic Acids Research* **33** 5659–5666.
- Lacadie SA & Rosbash M 2005 Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5' splice base pairing in yeast. *Molecular Cell* **19** 65–75.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al.* 2001 Initial sequencing and analysis of the human genome. *Nature* **409** 860–921.
- Lynch KW 2004 Consequences of regulated pre-mRNA splicing in the immune system. *Nature Reviews. Immunology* **4** 931–940.
- Matlin AJ, Clark F & Smith CW 2005 Understanding alternative splicing: towards a cellular code. *Nature Reviews. Molecular Cell Biology* **6** 386–398.
- Mazeyrat S, Saut N, Mattei MG & Mitchell MJ 1999 RBMY evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. *Nature Genetics* **22** 224–226.
- Pan Q, Shai O, Misquitta C, Zhang W, Saltzman AL, Mohammad N, Babak T, Siu H, Hughes TR, Morris QD *et al.* 2004 Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. *Molecular Cell* **16** 929–941.
- Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, Frey BJ & Blencowe BJ 2006 Quantitative microarray profiling provides

- evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes and Development* **20** 153–158.
- Park JW, Parisky K, Celotto AM, Reenan RA & Graveley BR** 2004 Identification of alternative splicing regulators by RNA interference in *Drosophila*. *PNAS* **101** 15974–15979.
- Ruggiu M, Speed R, Taggart M, McKay SJ, Kilanowski F, Saunders P, Dorin J & Cooke HJ** 1997 The mouse Dazl gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* **389** 73–77.
- Schmidt EE** 1996 Transcriptional promiscuity in testes. *Current Biology* **6** 768–769.
- Schmidt EE & Schibler U** 1995 High accumulation of components of the RNA polymerase II transcription machinery in rodent spermatids. *Development* **121** 2373–2383.
- Stamm S** 2002 Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Human Molecular Genetics* **11** 2409–2416.
- Swinburne IA, Meyer CA, Liu XS, Silver PA & Brodsky AS** 2006 Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome Research* **16** 912–921.
- Ule J, Jensen K, Mele A & Darnell RB** 2005a CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37** 376–386.
- Ule J, Ule A, Spencer J, Williams A, Hu JS, Cline M, Wang H, Clark T, Fraser C, Ruggiu M et al.** 2005b Nova regulates brain-specific splicing to shape the synapse. *Nature Genetics* **37** 844–852.
- Venables JP, Vernet C, Chew SL, Elliott DJ, Cowmeadow RB, Wu J, Cooke HJ, Artzt K & Eperon IC** 1999 T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis. *Human Molecular Genetics* **8** 959–969.
- Venables JP, Dalgliesh C, Paronetto MP, Skitt L, Thornton JK, Saunders PT, Sette C, Jones KT & Elliott DJ** 2004 SIAH1 targets the alternative splicing factor T-STAR for degradation by the proteasome. *Human Molecular Genetics* **13** 1525–1534.
- Venables JP, Bourgeois CF, Dalgliesh C, Kister L, Stevenin J & Elliott DJ** 2005 Up-regulation of the ubiquitous alternative splicing factor Tra2beta causes inclusion of a germ cell-specific exon. *Human Molecular Genetics* **14** 2289–2303.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al.** 2001 The sequence of the human genome. *Science* **291** 1304–1351.
- Wallace AM, Dass B, Ravnik SE, Tonk V, Jenkins NA, Gilbert DJ, Copeland NG & MacDonald CC** 1999 Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells. *PNAS* **96** 6763–6768.
- Wang HY, Xu X, Ding JH, Bermingham JR Jr & Fu XD** 2001 SC35 plays a role in T cell development and alternative splicing of CD45. *Molecular Cell* **7** 331–342.
- Wyckoff GJ, Wang W & Wu CI** 2000 Rapid evolution of male reproductive genes in the descent of man. *Nature* **403** 304–309.
- Xia H** 2005 Regulation of gamma-fibrinogen chain expression by heterogeneous nuclear ribonucleoprotein A1. *Journal of Biological Chemistry* **280** 13171–13178.
- Xu X, Yang D, Ding JH, Wang W, Chu PH, Dalton ND, Wang HY, Bermingham JR Jr, Ye Z, Liu F et al.** 2005 ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell* **120** 59–72.
- Yeo G, Holste D, Kreiman G & Burge CB** 2004 Variation in alternative splicing across human tissues. *Genome Biology* **5** R74.
- Zhang QS, Manche L, Xu RM & Krainer AR** 2006 hnRNP A1 associates with telomere ends and stimulates telomerase activity. *RNA* **12** 1116–1128.
- Zhou Z, Licklider LJ, Gygi SP & Reed R** 2002 Comprehensive proteomic analysis of the human spliceosome. *Nature* **419** 182–185.

Received 14 August 2006

First decision 8 September 2006

Accepted 25 September 2006