Regulation of plant gene expression by alternative splicing

Craig G. Simpson*, Sujatha Manthri†, Katarzyna Dorota Raczynska‡, Maria Kalyna§, Dominika Lewandowska*, Branislav Kusenda§, Monika Maronova§, Zofia Szweykowska-Kulinska‡, Artur Jarmolowski‡, Andrea Barta§ and John W.S. Brown*†[1]

*Genetics Programme, SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., †Plant Sciences Division, University of Dundee at SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., ‡Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Miedzychodzka 5, 60-371 Poznan, Poland, and §Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

Abstract

AS (alternative splicing) is a post-transcriptional process which regulates gene expression through increasing protein complexity and modulating mRNA transcript levels. Regulation of AS depends on interactions between *trans*-acting protein factors and *cis*-acting signals in the pre-mRNA (precursor mRNA) transcripts, termed 'combinatorial' control. Dynamic changes in AS patterns reflect changes in abundance, composition and activity of splicing factors in different cell types and in response to cellular or environmental cues. Whereas the SR protein family of splicing factors is well-studied in plants, relatively little is known about other factors influencing the regulation of AS or the consequences of AS on mRNA levels and protein function. To address fundamental questions on AS in plants, we are exploiting a high-resolution RT (reverse transcription)–PCR system to analyse multiple AS events simultaneously. In the present paper, we describe the current applications and development of the AS RT–PCR panel in investigating the roles of splicing factors, cap-binding proteins and nonsense-mediated decay proteins on AS, and examining the extent of AS in genes involved in the same developmental pathway or process.

Introduction

AS (alternative splicing) generates more than one spliced mRNA isoform from the same gene through the selection of alternative splice sites. Different types of AS event include alternative 5′ and 3′ splice site selection, intron retention, exon skipping and mutually exclusive exon splicing, resulting in the inclusion or exclusion of intronic or exonic sequences [1–3]. The consequences of such changes in transcript sequence are altered protein sequence and functionality (activity, protein–protein or protein–substrate interactions, localization, protein modification, etc.) or the introduction of premature termination codons giving truncated proteins or leading to degradation of AS isoforms by NMD (nonsense-mediated decay) [1–5]. In humans, 95 % of multi-exon genes undergo AS [6], whereas, in plants, current estimates are that approx. 35 % of genes (in *Arabidopsis* and rice) undergo AS [7–11]. This level of AS is likely to be underestimated owing to the relatively lower level of ESTs (expressed sequence tags) available in plants when compared with, for example, humans or mice, and because many AS events are not represented or are under-represented in EST collections as they occur only in specific cells and tissues, growth conditions or stages of development [10–13]. AS is found in genes involved in a wide range of cellular processes, including transcription, splicing, development, signal transduction, and responses to biotic and abiotic stress contributing to, for example, seed development and quality, germination, disease resistance, flowering time and the circadian clock [10,11,14–20].

Although more and more examples of AS in plant genes are being discovered, very little is known about (i) the dynamic changes in AS of the majority of annotated alternatively spliced genes, (ii) how such changes are regulated, or (iii) the functions of protein isoforms produced from the AS transcripts. In order to begin to address the aspects of dynamic modulation and regulation of AS, we developed a high-resolution RT (reverse transcription)–PCR system to monitor changes in AS isoform abundance of approx. 300 *Arabidopsis* AS events [21]. The AS events represent different types of AS and are mainly in transcription factor, RNA-binding protein or stress-related genes. By using overexpression lines and knockout mutants of factors involved in splicing regulation [e.g. SR (serine/arginine-rich) proteins [21] and PTB (polypyrimidine tract-binding protein)-like proteins] and mRNA biogenesis [e.g. CBPs (cap-binding proteins)] [22], the AS RT–PCR system detects whether the factors influence AS and which events and types of AS event are affected. Similarly, by analysing mutants impaired in NMD, it is possible to identify AS isoforms which are subject to degradation by NMD. Although the AS RT–PCR system does not report on

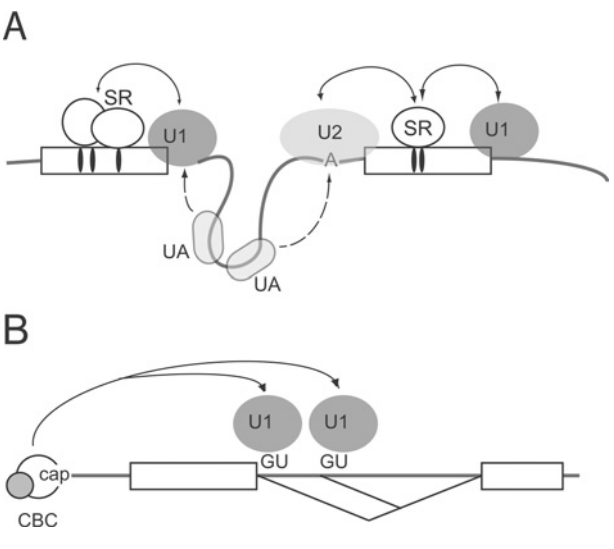
Key words: alternative splicing, *Arabidopsis*, precursor mRNA, splicing factor.

Abbreviations used: AS, alternative splicing; AtGRP, *Arabidopsis thaliana* glycine-rich protein; CBC, cap-binding complex; CBP, cap-binding protein; EST, expressed sequence tag; NMD, nonsense-mediated decay; PTC, premature termination codon; RT, reverse transcription; snRNP, small nuclear ribonucleoprotein; SR, serine/arginine-rich; UPF, up-frameshift; UTR, untranslated region.

[1]To whom correspondence should be addressed (email j.w.s.brown@dundee.ac.uk or john.brown@scri.ac.uk).

Figure 1 | Regulation by AS

(**A**) Splice site selection depends on multiple interactions between splicing factors and components and splicing signals in the transcript. SR proteins interact with other splicing factors and U1 and U2 snRNPs early in spliceosome assembly to select splice sites. In plants, this process may further involve UA-binding proteins interacting with the UA-rich introns. Differential interactions of different SR proteins with exon splicing enhancers promote alternative splice site selection. (**B**) In plants, the CBC can influence AS and does this preferentially at the 5′ splice site of the first intron, but can promote splicing at either the proximal or distal 5′ splice site [22]. Boxes represent exons; lines represent introns; vertical lines in exons represent exon splicing enhancer sequences; U1, U1 snRNP at the 5′ splice site; U2, U2 snRNP at the branchpoint (A); UA, UA-binding protein.



AS across the whole genome, the number of different events (approx. 300) is large enough to indicate the extent to which particular proteins affect AS. Finally, in addition to exploiting the genetic resources of *Arabidopsis* (in terms of mutant lines and transgenic overexpression lines), the AS RT–PCR system is being used to assess splicing efficiency and AS in different plant organs, developmental stages and under different growth and stress conditions (e.g. temperature). For example, a subset of intron-retention events have been used to examine splicing efficiency in plants under anoxic conditions which cause a core exon junction complex protein to relocalize to the nucleolus [23].

Regulation of AS

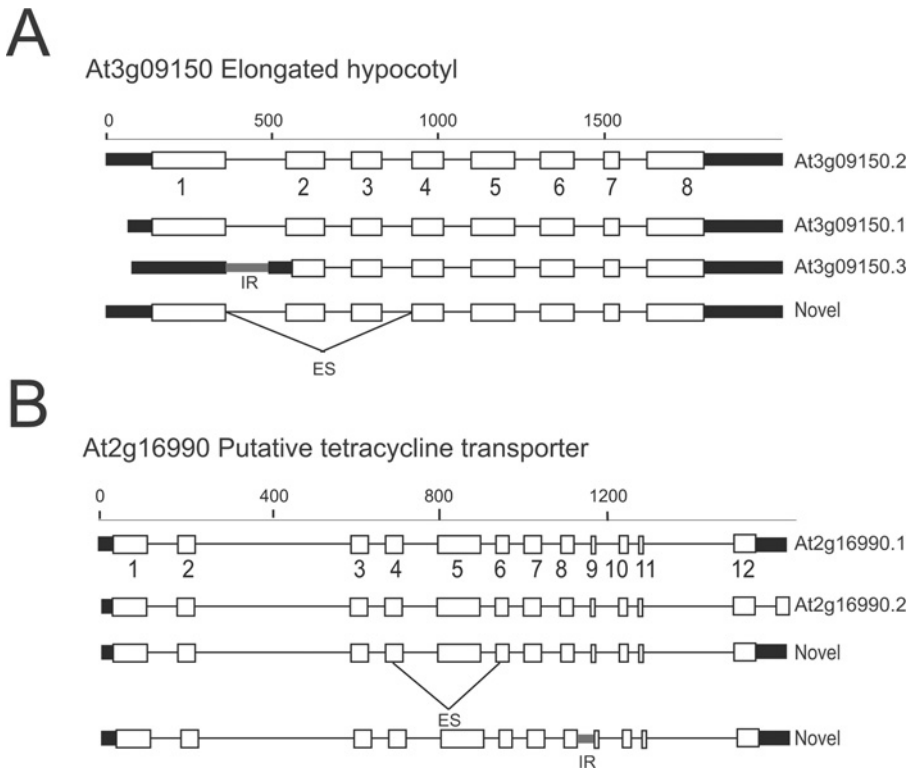
Regulation of AS is determined by the interactions of multiple *trans*-acting factors with intronic and exonic signals in the transcript giving rise to the concept of combinatorial control and a splicing code [24,25] (Figure 1A). Most progress in terms of *trans*-acting factors involved in plant splicing regulation has been made in the characterization of *Arabidopsis* SR proteins [26,27]. SR proteins are involved in both constitutive splicing and AS and generally promote splice site selection by binding to splicing enhancer sequences in the pre-mRNA and through protein–protein interactions (Figure 1A). In *Arabidopsis*, the essential role of these proteins is reflected in the range of developmental changes in plants where SR proteins are overexpressed or mutated [27–30]. SR protein genes are also of particular interest because the majority are themselves alternatively spliced and some of the AS events are evolutionarily conserved, suggesting that their regulation by AS is important for plant development [31]. Moreover, protein isoforms, produced from different AS transcripts of the *Arabidopsis* SR protein gene, *SR45*, are able to complement different developmental phenotypes of an *sr45* mutant demonstrating functional differences for the AS products [32]. Changes in expression of SR proteins are, not surprisingly, associated with changes in AS of many genes, including other SR protein genes themselves [20,28,29], and the AS RT–PCR panel is currently being used to analyse the subsets of AS events which are affected by particular SR proteins.

In addition to splicing factors, other proteins influence splicing. For example, the CBC (cap-binding complex) has a number of functions in mRNA biogenesis, including splicing, mRNA export, protection of the transcripts from nuclease degradation and 3′-end formation by stabilizing the interaction of the 3′-end processing machinery. In mammals and yeast, the CBC at the 5′ end of the pre-mRNA transcript promotes the initial interaction between U1 snRNP (small nuclear ribonucleoprotein) and the 5′ splice site of the first intron in the transcript and enhances the formation of spliced mRNAs [33]. However, the CBC has not been shown to affect AS to date. The single-knockout mutants of CBP20 and CBP80 and the double mutant are viable, but have slow growth, have serrated leaf margins, are late-flowering, are hypersensitive to abscisic acid and have enhanced drought tolerance [34–37]. Effects on splicing by knocking out CBC proteins have been observed in flowering time genes and in the increased occurrence of unspliced introns detected on microarrays [36–38]. To investigate effects on AS, the AS RT–PCR panel has been used in conjunction with single and double mutants of CBP20 and CBP80 [22]. Of ~250 AS events, 101 showed significant changes in AS in at least one mutant, and 41 of these showed changes in all three mutant lines. In the latter group, the number of events involving the first intron in a transcript and alternative 5′ splice site selection was significantly increased, suggesting that the *Arabidopsis* CBC preferentially (but not exclusively) affected AS of the first intron and particularly at the 5′ splice site [22] (Figure 1B). Moreover, we observed that CBP80 exerts a stronger influence on splicing than CBP20.

The NMD pathway targets mRNAs containing PTCs (premature termination codons) for degradation. In both mammals and plants, mRNAs with an increased distance between the PTC and the 3′ end of the mRNA [long 3′ UTRs (untranslated regions)] and/or with an exon junction complex downstream of a PTC are targeted for degradation [4,39,40]. NMD is thought to reduce the consequences of mutations or mistakes in expression, to reduce genomic noise

Figure 2 | Identification of novel AS events for process-specific RT–PCR panels

Cloning and sequencing of RT–PCRs covering the gene sequences has identified novel AS events of flowering time genes (an example is shown in **A**) and cold-stress induced genes (an example is shown in **B**). AS events in TAIR (The *Arabidopsis* Information Resource) are shown below the exon–intron structure of the gene, and novel events are labelled. White boxes represent exons; black boxes represent 5′ and 3′ UTRs; horizontal lines represent introns; lines below gene structures represent splicing events. IR, intron retention; ES, exon skip.



and prevent production of potentially detrimental truncated proteins. In addition, NMD modulates the levels of AS isoforms which contain PTCs. Approx. 10 % of the human and yeast transcriptome is turned over by NMD [41,42] and 30 % of human AS transcripts contain PTCs and are degraded by NMD [5]. In *Arabidopsis*, examples of regulation by AS and NMD include the genes encoding the circadian clock proteins, AtGRP (*Arabidopsis thaliana* glycine-rich protein) 7 and AtGRP8 [43,44]. By analysing mutants of the NMD proteins, UPF (up-frameshift) 1 and UPF3 (*upf3-1* and *upf1-5*), the AS RT–PCR system readily detects increases in the abundance of specific AS isoforms in a significant number of AS events. These potentially represent transcripts which are turned over by the NMD degradation pathway and are currently being characterized.

Process-specific AS

Ultimately, we wish to understand how transcriptional and post-transcriptional control systems are integrated in the regulation of gene expression. Current models suggest that post-transcriptional control networks such as AS are super-imposed on transcriptional networks and that regulation of AS must be co-ordinated with regulation of transcription [45]. For example, the fine-tuning of developmental or metabolic pathways will reflect the integrated outputs of both transcription and AS networks such that changes in expression of one gene at the transcriptional or post-transcriptional levels could feed back on its own expression and the expression of other genes in the pathway. An excellent example of this type of control is the auto- and cross-regulation of the clock proteins AtGRP7 and AtGRP8 mentioned above.

Our current knowledge of AS events in plants is incomplete, making global approaches to studying AS (such as the use of exon arrays or splice junction arrays) presently unfeasible. Indeed, one of the outcomes of the use of the AS RT–PCR panel has been the unexpected number of novel RT–PCR products, some of which have been characterized and shown to be novel unannotated AS events. The number of novel RT–PCR products which we detect in our amplicons suggests that the level of AS is much higher in plants than currently estimated. Therefore, to be able to address co-ordination of post-transcriptional control by AS in genes in the same pathway, it is necessary to identify all the possible AS events in these genes. To this end, we are systematically cloning and sequencing AS events for all genes in particular pathways (e.g. the flowering time pathway) (Figure 2). AS

events are assembled from existing transcript information available in databases and from cloning and sequencing of RT–PCR products covering the coding sequence and UTRs of each gene. Novel AS events have been observed and characterized (Figure 2) and, combining this information, the feasibility of generating process- or pathway-specific AS RT–PCR panels is currently being tested. In the future, next-generation sequencing is expected to provide a platform for the global identification and quantification of AS isoforms in different plant species, but robust methods of data analysis for this particular function are still under development. The detailed characterization of AS events by the AS RT–PCR system will be invaluable both upstream in developing the next-generation sequencing data-analysis systems and downstream in validating changes in AS.

Funding

We acknowledge funding for this research from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant numbers BB/G000212/1 and BB/G024979/1]; the Scottish Government Rural and Environment Research and Analysis Directorate (RERAD) [grant number WP114]; the European Union Framework Programme 6 Network of Excellence on Alternative Splicing (EURASNET) [grant number LSHG-CT-2005-518238]; European Research Area Network (ERA-NET) Plant Genomics; the Polish Government [grant numbers PBZ-KBN-110/P04/2004; PBZ/2/3/2006; PZ-19/2006/4/ZSK; 312/N-COST/2008/0]; the Austrian Science Foundation [grant number SFB-F017/10/11] and the Genome Research in Austria (GEN-AU) programme.

References

1 Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. **72**, 291–336
2 Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science **302**, 2141–2144
3 Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. and Soreq, H. (2005) Function of alternative splicing. Gene **344**, 1–20
4 Maquat, L.E. (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev. Mol. Cell Biol. **5**, 89–99
5 Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. U.S.A. **7**, 189–192
6 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
7 Xiao, Y.-L., Smith, S.R., Ishmael, N., Redman, J.C., Kumar, N., Monaghan, E.L., Ayele, M., Haas, B.J., Wu, H.C. and Town, C.D. (2005) Analysis of cDNAs of hypothetical genes on *Arabidopsis* chromosome 2 reveals numerous transcript variants. Plant Physiol. **139**, 1323–1337
8 Wang, B.B. and Brendel, V. (2006) Genome-wide comparative analysis of alternative splicing in plants. Proc. Natl. Acad. Sci. U.S.A. **103**, 7175–7180
9 Campbell, M.A., Haas, B.J., Hamilton, J.H., Mount, S.M. and Buell, C.R. (2006) Comprehensive analysis of alternative splicing in rice and comparative analysis with *Arabidopsis*. BMC Genomics **7**, 327–343
10 Barbazuk, W.B., Fu, Y. and McGinnis, K.M. (2008) Genome-wide analyses of alternative splicing in plants: opportunities and challenges. Genome Res. **18**, 1381–1392
11 Reddy, A.S.N. (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annu. Rev. Plant Biol. **58**, 267–294
12 Hirose, T., Sugita, M. and Sugiura, M. (1993) cDNA structure, expression and nucleic acid-binding properties of three RNA-binding proteins in tobacco: occurrence of tissue-specific alternative splicing. Nucleic Acids Res. **21**, 3981–3987
13 Yoshimura, K., Yabuta, Y., Ishikawa, T. and Shigeoka, S. (2002) Identification of a *cis* element for tissue-specific alternative splicing of chloroplast ascorbate peroxidase pre-mRNA in higher plants. J. Biol. Chem. **277**, 40623–40632
14 Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y. and Staiger, D. (2008) Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in *Arabidopsis*. Nucleic Acids Res. **36**, 6977–6987
15 Larkin, P.D. and Park, W.D. (1999) Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. Plant Mol. Biol. **40**, 719–727
16 Egawa, C., Kobayashi, F., Ishibashi, M., Nakamura, T., Nakamura, C. and Takumi, S (2006) Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. Genes Genet. Syst. **81**, 77–91
17 Balasubramanian, S., Sureshkumar, S., Lempe, J. and Wiegel, D. (2006) Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. PLoS Genet. **2**, e106
18 Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A. and and Shinozaki, K. (2004) Genome-wide analysis of alternative splicing in *Arabidopsis thaliana* based on full-length cDNA sequences. Nucleic Acids Res. **32**, 5096–5103
19 Bove, J., Kim, C.Y., Gibson, C.A. and Assmann, S.M. (2008) Characterization of wound-responsive RNA-binding proteins and their splice variants in *Arabidopsis*. Plant Mol. Biol. **67**, 71–88
20 Dinesh-Kumar, S.P and Baker, B.J. (2000) Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. Proc. Natl. Acad. Sci. U.S.A. **97**, 1908–1913
21 Simpson, C.G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Barta, A. and Brown, J.W.S. (2008) Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts. Plant J. **53**, 1035–1048
22 Raczynska, K.D., Simpson, C.G., Ciesiolka, A., Szewc, L., Lewandowska, D., McNicol, J., Szweykowska-Kulinska, Z., Brown, J.W.S. and Jarmolowski, A. (2009) Involvement of the nuclear cap-binding protein complex in alternative splicing in *Arabidopsis thaliana*. Nucleic Acids Res. **38**, 265–278
23 Koroleva, O.A., Calder, G., Pendle, A.F., Kim, S.H., Lewandowska, D., Simpson, C.G., Jones, I.M., Brown, J.W.S. and Shaw, P.J. (2009) Dynamic behaviour of the eIF4A-III putative core protein of the exon junction complex: fast relocation to nucleolus and speckles under hypoxia. Plant Cell **21**, 1592–1606
24 Smith, C.W. and Valcárcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem. Sci. **25**, 381–388
25 Matlin, A.J., Clark, F. and Smith, C.W.J. (2005) Understanding alternative splicing: towards a cellular code. Nat. Rev. Mol. Cell Biol. **5**, 89–99
26 Kalyna, M. and Barta, A. (2004) A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions. Biochem. Soc. Trans. **32**, 561–564
27 Barta, A., Kalyna, M. and Lorković, Z.J. (2009) Plant SR proteins and their functions. Curr. Top. Microbiol. Immunol. **326**, 83–102
28 Kalyna, M., Lopato, S. and Barta, A. (2003) Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. Mol. Biol. Cell **14**, 3565–3577
29 Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R. and Barta, A. (1999) atSRp30, one of two SF2/ASF-like proteins from *Arabidopsis thaliana*, regulates splicing of specific plant genes. Genes Dev. **13**, 987–1001
30 Ali, G.S., Palusa, S.G., Golovkin, M., Prasad, J., Manley, J.L. and Reddy, A.S. (2007) Regulation of plant developmental processes by a novel splicing factor, PLoS One, e471
31 Kalyna, M., Lopato, S., Voronin, V. and Barta, A. (2006) Evolutionary conservation and regulation of particular alternative splicing events in plant SR proteins. Nucleic Acids Res. **34**, 4395–4405
32 Zhang, X.-N. and Mount, S.M. (2009) Two alternatively spliced isoforms of the *Arabidopsis* SR45 protein have distinct roles during normal plant development. Plant Physiol. **150**, 1450–1458
33 Lewis, J.D. and Izaurralde, E. (1997) The role of the cap structure in RNA processing and nuclear export. Eur. J. Biochem. **247**, 461–469

34 Hugouvieux, V., Kwak, J.M. and Schroeder, J.I. (2001) An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in *Arabidopsis*. Cell **106**, 477–487
35 Papp, I., Mur, L.A., Dalmadi, A., Dulai, S. and Koncz, C. (2004) A mutation in the cap binding protein 20 gene confers drought tolerance to *Arabidopsis*. Plant Mol. Biol. **55**, 679–686
36 Bezerra, I.C., Michaels, S.D., Schomburg, F.M. and Amasino, R.M. (2004) Lesions in the mRNA cap-binding gene *ABA HYPERSENSITIVE 1* suppress FRIGIDA-mediated delayed flowering in *Arabidopsis*. Plant J. **40**, 112–119
37 Kuhn, J.M., Hugouvieux, V. and Schroeder, J.I. (2008) mRNA cap binding proteins: effects on abscisic acid signal transduction, mRNA processing, and microarray analyses. Curr. Top. Microbiol. Immunol. **326**, 139–150
38 Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rätsch, G. and Weigel, D. (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. **105**, 8795–8800
39 Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A., Gyula, P., Lacomme, C., Barta, E., Nagy, F. and Silhavy, D. (2008) Inter-kingdom conservation of mechanism of nonsense-mediated decay. EMBO J. **27**, 1585–1595
40 Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Palfy, T., Barta, E. and Silhavy, D. (2006) Both introns and long 3′-UTRs operate as *cis*-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res. **34**, 6147–6157
41 Mendall, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F. and Dietz, H.C. (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. **36**, 1073–1078
42 He, F., Li, X., Spatrick, P., Casillo, R., Dong, S. and Jacobson, A. (2003) Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. Mol. Cell **12**, 1439–1452
43 Schöning, J.C., Streitner, C., Page, D.R., Hennig, S., Uchida, K., Wolf, E., Furuya, M. and Staiger, D. (2007) Auto-regulation of the circadian slave oscillator component AtGRP7 and regulation of its targets is impaired by a single RNA recognition motif point mutation. Plant J. **52**, 1119–1130
44 Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y. and Staiger, D. (2008) Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in *Arabidopsis*. Nucleic Acids Res. **36**, 6977–6987
45 Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. Cell **126**, 37–47

Received 29 September 2009
doi:10.1042/BST0380667