sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815
14 Yamasaki, H. and Sakihama, Y. (2000) Simultaneous production of nitric oxide and peroxynitrite by plant nitrate reductase: *in vitro* evidence for the NR-dependent formation of active nitrogen species. *FEBS Lett.* 468, 89–92
15 Chandok, M.R. *et al.* (2003) The pathogen-inducible nitric oxide synthase (iNOS) in plants is a variant of the P protein of the glycine decarboxylase complex. *Cell* 113, 469–482
16 Douce, R. *et al.* (2001) The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci.* 6, 167–176
17 Foissner, I. *et al.* (2000) *In vivo* imaging of an elicitor-induced nitric oxide burst in tobacco. *Plant J.* 23, 817–824

1360-1385/$ - see front matter
doi:10.1016/j.tplants.2003.09.003

Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged

Kemal Kazan

Commonwealth Scientific and Industrial Research Organisation (CSIRO) Plant Industry, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, Brisbane, Qld 4067, Australia

Alternative splicing has recently emerged as one of the most significant generators of functional complexity in several relatively well-studied animal genomes, but little is known about the extent of this phenomenon in higher plants. However, recent computational and experimental studies discussed here suggest that alternative splicing probably plays a far more significant role in the generation of proteome diversity in plants than was previously thought.

Many genes in higher eukaryotes contain introns that are removed post-transcriptionally from the precursor mRNA (pre-mRNA) during the production of functional mRNAs. However, in some instances, the splicing machinery can process the same pre-mRNA differently by selectively joining different protein coding elements (exons). This potentially leads to the generation of structurally and/or functionally distinct proteins. The extent to which alternative splicing can contribute to the proteome (the full protein complement of a genome) diversity in humans has become evident with the sequencing of the human genome. The unexpectedly small number of genes, which do not fully account for the presence of many protein isoforms, suggested that alternative splicing contributes significantly to human proteome diversity. According to recent estimates, ~40% of human genes are alternatively spliced [1]. Although alternative splicing is thought to be less prevalent in plants than it is in higher eukaryotes, recent computational analyses of alternative splicing in *Arabidopsis thaliana* are beginning to question this assertion.

Multiple complementary DNAs (cDNAs) or expressed sequence tags (ESTs) that map to the same genomic locus but deviate from the predicted mRNA sequence for that locus suggest that of the estimated 27 170 protein-coding genes in *Arabidopsis*, 1267 have alternate splice isoforms generating 2678 proteins in total (http://www.tigr.org/tdb/e2k1/ath1/altsplicing/splicing_variations.shtml). However, the estimation that 5% of all predicted genes are alternatively spliced according to computational analyses is probably an underestimate of alternative splicing in this species. First, the number of publicly available ESTs from *Arabidopsis* (~179 000) is nowhere near that from relatively well-studied genomes (e.g. ~5 million ESTs in human). Second, the full-length cDNAs are available for only approximately half of all predicted genes [2]. More importantly, ESTs from genes with low expression levels or with expression at a particular developmental stage or physiological condition, which are likely to be alternatively spliced, might be scarcely represented in the databases.

Functional groups of alternatively spliced genes

In parallel with the large-scale detection of alternatively spliced genes, various commonalities are beginning to emerge between animals and plants. Approximately 75% of the alternatively spliced genes in humans encode proteins involved in signalling and gene regulation, such as receptors, and signal transduction and transcription factors [1]. Similarly, the great majority of alternatively spliced genes in *Arabidopsis* encode proteins with regulatory functions (http://www.tigr.org/tdb/e2k1/ath1/altsplicing/splicing_variations.shtml). In addition, the genes associated with various stress responses seem to be particularly prone to alternative splicing in both plants and animals (recent examples are given in [3] and Table 1). For instance, alternative splicing of Toll-like plant disease resistance genes is well known [4]. The list for the alternatively spliced genes associated with abiotic stress responses is also rapidly expanding. A couple of recent examples are: the relative abundance of the two splicing forms of the *Arabidopsis SOS4* (*salt overly sensitive 4*) gene is regulated by salt stress [5]; and *OsIM* from rice (*Oryza sativa*), which encodes an alternative oxidase functioning in the removal of stress-induced reactive oxygen species, is alternatively spliced under salt stress [6]. The ratio of the two splicing forms, *OsIM1* and *OsIM2*,

Corresponding author: Kemal Kazan (kemal.kazan@csiro.au).

Table 1. Recent examples of alternatively spliced stress-associated plant genes

Species	Gene	Protein product	Refs
	Biotic stress		
Flax	*L6*	Resistance protein	[17]
Flax	*M*	Resistance protein	[18]
Arabidopsis	*RPS4*	Resistance protein	[19]
Tobacco	*N*	Resistance protein	[20]
Arabidopsis	*RPP5*	Resistance protein	[21]
Tomato	*Bs4*	Resistance protein	[22]
Barley	*Mla*	Resistance protein	[23]
Rice	*OsMAPK5*	MAP kinase	[7]
	Water stress		
Rice	*OsMAPK5*	MAP kinase	[7]
	Salt stress		
Arabidopsis	*SOS4*	Pyridoxal kinase	[5]
Rice	*OsIM1*	Alternative oxidase	[6]
	Wounding		
Tomato	*Prosys*	Systemin precursor	[24]
	Heavy metal stress		
Maize	*Bronze 2*	Glutathione-*S*-transferase	[25]
	Heat stress		
Maize	*HSP22*	Heat shock protein	[26]
	Oxidative stress		
Spinach	*chlAPX*	Ascorbate peroxidase	[27]
Rice	*photolyase*	DNA repair protein	[28]
	Light		
Tobacco	*ZGT*	Circadian-clock protein	[29]
Arabidopsis	*AtGRP7*	RNA-binding protein	[30]
Pumpkin	*HPR*	Hydroxypryuvate reductase	[31]
	Other stresses		
Arabidopsis	*AtMMH*	DNA repair protein	[32]
Tomato	*LeCBDGK*	Diacylglycerol kinase	[33]
Apple	*MdSPDS*	Spermidine synthase	[34]

differs between salt-tolerant and salt-sensitive varieties. In the salt-tolerant variety, *OsIM1* is maintained at a high level, whereas in the salt-sensitive variety, its transcript level decreases rapidly after a short exposure to salt stress [6]. Similarly, a water stress and plant defence-associated MAP-kinase gene of rice is alternatively spliced to generate at least two splicing forms. One of these forms has auto-phosphorylation and kinase activities associated with stress tolerance; the exact role of the second one remains unknown [7].

Comparative studies of large numbers of human, mouse and rat genes suggest that evolutionary change measured as exon creation or loss occurs at a faster rate in alternatively spliced genes than those not subjected to alternative splicing [8]. Assuming that alternative splicing in plants is also associated with an increased evolutionary change, alternative splicing in stress-associated regulatory genes might be useful for acquiring certain adaptive benefits that are important for survival under stress conditions.

Future prospects

The key issues in the future are: (i) the large-scale identification of alternative splicing events, (ii) the determination of the developmental stage and/or physiological conditions under which splicing patterns for a given gene are altered, (iii) the mechanisms of how splice-site selection is achieved, and (iv) the identification of the differential function of splicing isoforms produced by alternative splicing (Figure 1).

The lack of extensive sequence information in many plant species currently prohibits large-scale identification and comparative analysis of alternative splicing events in plants. Thus, additional insights into the extent of alternative splicing in plants would benefit from the availability of genomic and large collections of EST sequences that preferably originate from diverse cDNA libraries. However, in the short term, an approach recently developed for high-throughput detection of alternative splicing in yeast [9,10] can be adapted for *Arabidopsis* and rice, which already have extensive sequence information available. This method takes advantage of the situation that alternative splicing of a gene creates different exon–exon junctions. DNA chips carrying short micro-array probes spanning these junctions are then used to screen the RNA from specific tissues, developmental stages or physiological functions for genome-wide detection of alternative splicing events.

So far, almost none of the splicing forms identified based on computational analyses in *Arabidopsis* have been verified and characterized experimentally. Determining the potential functions of splicing forms and how different splice forms are regulated are other areas of future research. Determining the potential functions of splicing forms could be studied by overexpression [11] or silencing of the alternative splicing forms in transgenic plants and testing the potential effects of these modifications on plant phenotype. Currently, little is known about the mechanisms regulating alternative splicing in plant genes. Specific recognition of splice sites involves interaction between *cis*-acting elements and *trans*-acting factors. Although few *trans*-acting factors have been isolated from plants based on similarities to the animal *trans*-acting factors [12,13], virtually nothing is known about the nature of *cis*-acting elements. A highly conserved putative *cis*-acting element recently identified in the genes encoding chloroplast-specific ascorbate peroxidase (chlAPX) isoenzymes of higher plants probably represents one of the first examples of a plant *cis*-acting element that modulates the tissue-specific alternative splicing patterns of the *chlAPX* gene [14]. Mutational analysis of this element revealed the functional significance of AU- or U-rich motifs [14]. This is in consistent with the suggestion that the abundance of AU- or U-rich motifs in plant introns possibly contributes to intron recognition and splice-site selection.

So far, several proteins [e.g. serine-rich (SR) and heterogeneous nuclear ribonucleoprotein (hnRNP)-like proteins] associated with either regulation or execution of splicing in plants have also been identified [12,13]. It is likely that the relative levels of these factors in each tissue or physiological condition determine the splice-site selection in plant cells. To date, the potential roles of few of these factors have been investigated in plants. For instance, the *Arabidopsis* splicing factor AtSRp30 causes several developmental changes by altering the splicing patterns of various pre-mRNAs when overexpressed in transgenic plants [15]. Similarly, overexpressing a kinase containing a LAMMER motif, conserved in proteins associated with the splicing machinery, results in differential expression of genes associated with stress and

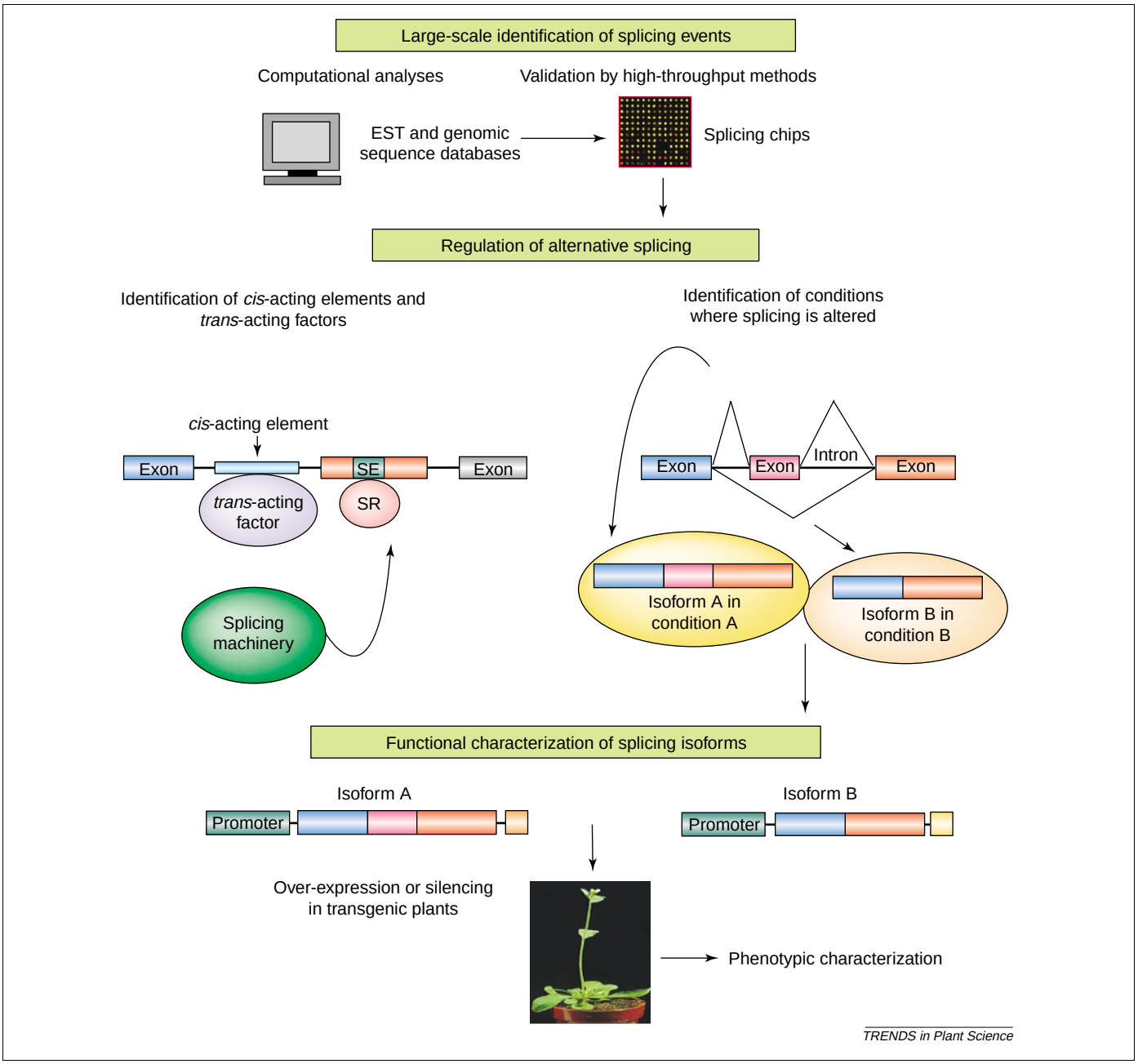

Figure 1. Potential areas of future research in plant alternative splicing research. A combinatorial approach involving computational analyses of available sequence information and experimental analyses using splicing-specific DNA chips could be used for large-scale detection and cataloguing of splicing events in plants. Identification of conserved *cis*-acting elements and *trans*-acting factors could help better understand how alternative splicing is regulated. Putative functions of distinct splicing isoforms that resulted from alternative splicing under a particular developmental stage and/or physiological condition are evaluated by overexpression or silencing of the expression of these isoform in transgenic plants. Abbreviations: SE, splicing enhancer; SR, serine-rich protein.

defence responses as well as up-regulation of genes with roles in protein turnover as revealed by DNA microarray analysis [16]. However, it is likely that some of the effects observed after constitutive overproduction of these regulatory proteins might be because of non-specific activation of various signal transduction pathways with regulatory effects on several other genes.

Conclusion

Although much research is needed to uncover the extent and regulation of alternative splicing in plants, it is already clear that this phenomenon occurs in plants on a much wider scale than was previously thought. The recently completed sequence of the rice genome and increased availability of sequence information from several other plant species should provide additional impetus to research alternative splicing in plants.

Acknowledgements

I thank Sukumar Chakraborty and an anonymous referee for useful comments on the manuscript.

References

1 Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nat. Genet.* 30, 13–19
2 Seki, M. *et al.* (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* 296, 141–145

3 Xie, J. and McCobb, D.B. (1998) Control of alternative splicing of potassium channels by stress hormones. *Science* 280, 443–446
4 Jordan, T. *et al.* (2002) Alternative splicing of transcripts encoding Toll-like plant resistance proteins – what is the functional relevance to innate immunity? *Trends Plant Sci.* 7, 392–398
5 Shi, H. *et al.* (2002) The *Arabidopsis salt overly sensitive 4* mutants uncover a critical role for vitamin B in plant salt tolerance. *Plant Cell* 14, 575–588
6 Kong, J. *et al.* (2003) A new AOX homologous gene *OsIM1* from rice (*Oryza sativa* L.) with an alternative splicing mechanism under salt stress. *Theor. Appl. Genet.* 107, 326–331
7 Xiong, L. and Yang, Y. (2003) Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. *Plant Cell* 15, 745–759
8 Modrek, B. and Lee, C.J. (2003) Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation/or loss. *Nat. Genet.* 34, 177–180
9 Hu, G.K. *et al.* (2001) Predicting splice variant from DNA chip expression data. *Genome Res.* 11, 1237–1245
10 Barrass, D.J. and Beggs, J.D. (2003) Splicing goes global. *Trends Genet.* 19, 295–298
11 Macknight, R. *et al.* (2002) Functional significance of the alternative transcript processing of the *Arabidopsis* floral promoter FCA. *Plant Cell* 14, 877–888
12 Lorkovic, Z.J. *et al.* (2000) Pre-mRNA splicing in higher plants. *Trends Plant Sci.* 5, 160–167
13 Lambermon, M.H.L. *et al.* (2000) UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* 19, 1638–1649
14 Yoshimura, K. *et al.* (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
15 Lopato, S. *et al.* (1999) atSRp30, one of two SF2/ASF-like proteins from *Arabidopsis thaliana*, regulates splicing of specific plant genes. *Genes Dev.* 13, 987–1001
16 Savaldi-Goldstein, S. *et al.* (2003) Alternative splicing modulation by a LAMMER kinase impinges on developmental and transcriptome expression. *Plant Cell* 15, 926–938
17 Ayliffe, M.A. *et al.* (1999) Analysis of the alternative transcripts of the flax *L6* rust resistance gene. *Plant J.* 17, 287–292
18 Anderson, P.A. *et al.* (1997) Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* 9, 641–651
19 Gassmann, W. *et al.* (1999) The *Arabidopsis RPS4* bacterial-resistance gene is a member of the TIR–NBS–LRR family of disease resistance genes. *Plant J.* 20, 265–277
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 Parker, J.E. *et al.* (1997) The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. *Plant Cell* 9, 879–894
22 Ballvora, A. *et al.* (2001) Genetic mapping and functional analysis of the tomato *Bs4* locus, governing recognition of the *Xanthomonas campestris* pv. *vesicatoria* AvrBs4 protein. *Mol. Plant–Microbe Interact.* 14, 629–638
23 Halterman, D.A. *et al.* (2003) Powdery mildew-induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. *Plant Physiol.* 131, 558–567
24 Li, L. and Howe, G.A. (2001) Alternative splicing of prosystemin pre-mRNA produces two isoforms that are active as signals in the wound response pathway. *Plant Mol. Biol.* 46, 409–419
25 Marrs, K.A. and Walbot, V. (1997) Expression and RNA splicing of the maize glutathione *S*-transferase *Bronze 2* gene is regulated by cadmium and other stresses. *Plant Physiol.* 113, 93–102
26 Lund, A.A. *et al.* (2001) *In vivo* modifications of the maize mitochondrial small heat shock protein, HSP22. *J. Biol. Chem.* 276, 29924–29949
27 Yoshimura, K. *et al.* (1999) Alternatively spliced mRNA variants of chloroplast ascorbate peroxidase isoenzymes in spinach leaves. *Biochem. J.* 338, 41–48
28 Hirouchi, T. *et al.* (2003) A gene for Class II DNA photolyase from *Oryza sativa*: cloning of the cDNA by dilution amplification. *Mol. Genet. Genomics* 269, 508–516
29 Xu, J. and Johnson, C.H. (2001) A clock- and light-regulated gene that links the circadian oscillator to LHCB gene expression. *Plant Cell* 13, 1411–1425
30 Staiger, D. *et al.* (2003) The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. *Plant J.* 33, 361–371
31 Mano, S. *et al.* (2000) A leaf peroxisomal protein, hydroxypyruvate reductase, is produced by light-regulated alternative splicing. *Cell Biochem. Biophys.* 32, 147–154
32 Murphy, T.M. and Gao, M.J. (2001) Multiple forms of formamido-pyrimidine-DNA glycosylase produced by alternative splicing in *Arabidopsis thaliana*. *J. Photochem. Photobiol. B.* 61, 87–93
33 Snedden, W.A. and Blumwald, E. (2000) Alternative splicing of a novel diacylglycerol kinase in tomato leads to a calmodulin-binding isoforms. *Plant J.* 24, 317–326
34 Zhang, Z. *et al.* (2003) Structure and expression of spermidine synthase genes in apple: two cDNAs are spatially and developmentally regulated through alternative splicing. *Mol. Genet. Genomics* 268, 799–807

doi:10.1016/j.tplants.2003.09.001

CAF-1 and MSI1-related proteins: linking nucleosome assembly with epigenetics

Steven van Nocker

Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA

A key question in developmental biology is how epigenetic states of gene activity are maintained faithfully through successive rounds of cell division. Recent observations in *Arabidopsis thaliana* point to an important and conserved role for chromatin assembly factor-1 and the MSI1 class of chromatin-related proteins in perpetuating epigenetic marks on chromatin.

The maintenance of correct gene activity patterns in specified cell lineages is crucial for proper development of multicellular organisms. Much of the information needed to maintain repressed or active states of transcription is

Corresponding author: Steven van Nocker (vannocke@msu.edu).