

## REVIEW

## Open Access

# Upstream regulatory architecture of rice genes: summarizing the baseline towards genus-wide comparative analysis of regulatory networks and allele mining

Benildo G de los Reyes<sup>1\*</sup>, Bijayalaxmi Mohanty<sup>2</sup>, Song Joong Yun<sup>3</sup>, Myoung-Ryoul Park<sup>1</sup> and Dong-Yup Lee<sup>2</sup>**Abstract**

Dissecting the upstream regulatory architecture of rice genes and their cognate regulator proteins is at the core of network biology and its applications to comparative functional genomics. With the rapidly advancing comparative genomics resources in the genus *Oryza*, a reference genome annotation that defines the various cis-elements and trans-acting factors that interface each gene locus with various intrinsic and extrinsic signals for growth, development, reproduction and adaptation must be established to facilitate the understanding of phenotypic variation in the context of regulatory networks. Such information is also important to establish the foundation for mining non-coding sequence variation that defines novel alleles and epialleles across the enormous phenotypic diversity represented in rice germplasm. This review presents a synthesis of the state of knowledge and consensus trends regarding the various cis-acting and trans-acting components that define spatio-temporal regulation of rice genes based on representative examples from both foundational studies in other model and non-model plants, and more recent studies in rice. The goal is to summarize the baseline for systematic upstream sequence annotation of the rapidly advancing genome sequence resources in *Oryza* in preparation for genus-wide functional genomics. Perspectives on the potential applications of such information for gene discovery, network engineering and genomics-enabled rice breeding are also discussed.

**Keywords:** Cis-elements; Trans-acting factors; Spatio-temporal regulation; Regulatory network; Comparative genomics

**Introduction**

The ‘gene-centric’ and subsequently the ‘genome-enabled’ paradigms of functional biology have led to concerted efforts to understand the biochemical synergies in the plant cell, by examining how genes function in a spatio-temporal scale. Understanding the multi-layered complexity of gene regulation is central to this goal, from the process of chromatin remodeling to transcription initiation, post-transcriptional processing and modulation, and protein modification and interaction. The process of transcription initiation is a critical crossroad of such multi-layered complexity of gene regulation, because it is an early step in the process where coordination could be interfaced with various intrinsic and extrinsic signals for

growth, development, reproduction and adaptation. Thus, the study of the architecture of promoters and the nature of the interacting general and specialized regulatory transcription factors has always been an important aspect of understanding the intricate wirings of gene function (Kaufmann *et al.*, 2010; Mejia-Guerra *et al.*, 2012).

A large number of rice genes of potential agronomic value are regulated largely at the transcriptional level. Looking back to as early as the mid-1980’s, rice geneticists and molecular biologists have always been interested in studying the architecture of promoters, initially for the primary objective of discovering regulatory sequences for spatio-temporal engineering of transgenes. Examples of seminal studies include the dissection of the regulatory elements for precise endosperm-specific expression of seed storage proteins, which also elucidated the mechanisms of how seed development and maturation are regulated, and

\* Correspondence: [benildo.de@maine.edu](mailto:benildo.de@maine.edu)

<sup>1</sup>School of Biology and Ecology, University of Maine, Orono, ME 04469, USA  
Full list of author information is available at the end of the article

led to applications in seed bioengineering (Takaiwa et al., 1987; Okita et al., 1989; Takaiwa et al., 1991).

Subsequently, the elucidation of the entire genome of a reference genotype of cultivated rice (*O. sativa* ssp. japonica, cv. Nipponbare) made the establishment of direct links between agronomically important physiological processes and gene regulation at the whole transcriptome level possible (Ohyanagi et al., 2006; Yun et al., 2010; Sato et al., 2011; Hamada et al., 2011). The enormous amount of transcriptome data generated to date has routinely spark the same old question that rice geneticists has long been trying to address even prior to the advent of the genome-enabled era of biology, that is the question of: *What unifying DNA signals determine the spatio-temporal transcription patterns of rice genes and how do these elements work?* In today's context, we often ask a more elaborate question of: *What cis-elements and trans-acting factors facilitate the coordinate expression of functionally related rice genes and what can we learn from such synergy to understand the network relationships among genes in a grander scale that can explain phenotypic variation?* As we get more advanced in the era of integrative, predictive and comparative systems biology in rice research, such a question has become quite relevant in the context of subsequent translation of genomics to various applications in molecular breeding, allele mining and network engineering for the improvement of yield and nutritional value of rice.

This review addresses a fundamental question that is central to the understanding of how transcription of rice genes is programmed. Its primary motivation is the need to present a synthesis of the current state of knowledge and consensus trends regarding the cis-acting and trans-acting components of transcriptional regulatory modules using representative examples from both seminal and more recent studies. It is hoped that this review will catalyze concerted efforts to continuously compile and update information in the same dynamic pace as in the dicot model *Arabidopsis*. While the information contained in this review is by no means inclusive of the entire body of information available about this topic, it is to our knowledge the first attempt to integrate the results of seminal studies with more recent updates about the structure, organization and evolutionary conservation of cis-elements in rice. Integration of this body of knowledge should nucleate new concepts on the application of regulatory sequences to comparative genomics, allele mining, and genetic manipulation by network engineering.

## Review

### Core promoter architecture and basal transcription machinery

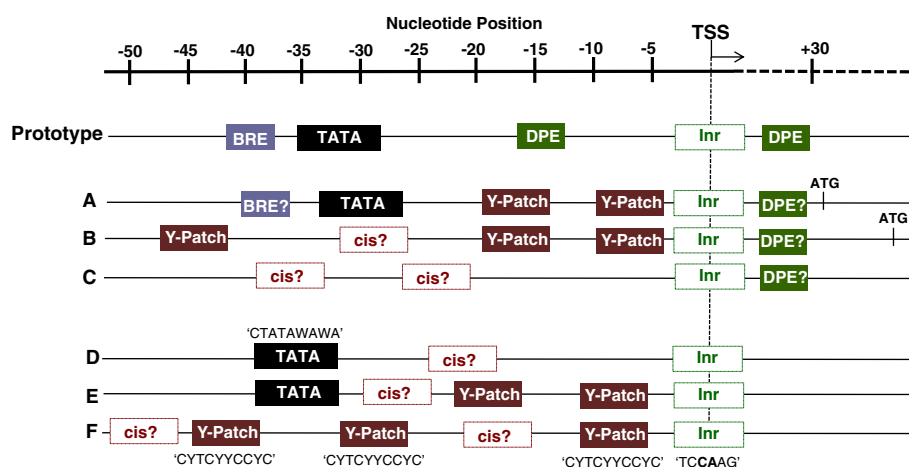
With the annotated reference genome sequence of *O. sativa* ssp. japonica (cv. Nipponbare) and a rapidly increasing

volume of whole-genome resequencing data on representative germplasm diversity panel across the genus, one important feature that needs to be defined for a biologically meaningful interrogation of protein-coding genes is the proximal regulatory region or '*core promoter*'. The core promoter is at the very heart of the process of transcription, being the site for the assembly of various protein components of RNA polymerase II pre-initiation complex, and its cooperation with other cis-elements and trans-acting factors that facilitate precise spatio-temporal regulation.

So what constitutes a core promoter? The consensus core promoter of eukaryotic genes is a contiguous stretch of DNA comprised of various cis-acting elements (Figure 1). The most well characterized is the *TATA-box*, defined by the consensus sequence 'TATAWAWAR', where the upstream 'T' is often located around -30 relative to A +1 (or G +1) position of the initiator (*Inr*) sequence for the transcription start site or TSS (Smale and Baltimore, 1989). The *TATA-box* is recognized by a sub-unit of the general transcription factor TFIID called the TATA-binding protein (TBP) to facilitate the formation of pre-initiation complex (Butler and Kadonaga, 2001; Smale, 2001). Occasionally found upstream to the *TATA-box* is a less frequently occurring motif called *BRE* (TFIIB recognition element), which facilitates efficient binding of transcription factor TFIIB to the core promoter. About 30-nt downstream of *Inr* is the Downstream Promoter Element (*DPE*), which is recognized by the TBP-associated factor (TAF) to enhance TBP binding to the core promoter (Burke and Kadonaga, 1997). In some cases *TATA-box*, *Inr* and *DPE* all occur in the same promoter (Juven-Gershon and Kadonaga, 2010).

Structural conservation of TFIID in metazoans and plants suggests that the cis-acting elements involved in RNA polymerase II-dependent transcription are similar among eukaryotes (Vogel et al., 1993; Smale, 2001). Pioneering studies on the rice phenylalanine ammonia-lyase (*PAL*) promoter *ZB8* showed that 5'-truncation up to -35 could maintain an accurate initiation of basal transcription (Zhu et al., 1995). The 'TATTTAA' motif within -35 to -28 that resembles the consensus sequence for eukaryotic *TATA-box* is critical for minimal promoter activity. Replacement of this sequence either with G-rich motif ('GCGGGTT') or with 2-nt substituted versions (TCGTTAA or TATGGAA) could lead to complete loss of promoter function. It has been shown that a recombinant TBP of rice (*OsTBP2*) interacts with *OsTFIIB* for efficient binding to the *TATA-box* in *ZB8* promoter, causing highly efficient TATA-dependent basal transcription (Zhu et al., 2002). Moreover, the functionality of -35 to -28 region of *ZB8* is in context with the consensus 'TCCAAG' *Inr* sequence. Substitution of the conserved core 'CA' motif at -1 and +1 positions of *Inr* leads to erroneous initiation at multiple locations.

Contrary to earlier notions that *TATA-box* is a universal feature of most eukaryotic core promoters, subsequent



**Figure 1 Comparison of the core promoter architectures of rice genes with the prototype core promoters of metazoan and Arabidopsis genes.**

In addition to the *TATA-box*, the composite prototype model of metazoan core promoter shows all previously identified cis-elements (*BRE*= TFIIIB recognition element; *DPE*= Downstream promoter element; *Inr*= Initiator sequence) occurring in various combinations with the *TATA-box*. In Arabidopsis genes (**A, B, C**), *TATA-box*-containing core promoters (**A**) represent about 30% of all the protein-coding genes encoded by the genome. Non-*TATA-box*-containing promoters (**B, C**) represent a larger proportion of Arabidopsis genes. Small groups of genes contain novel motifs (*cis?*) that appear to be specific to higher plants. Evidence of the importance of *DPE*-like and/or *BRE*-like sequences in the functionality of Arabidopsis core promoters have not been established so far. In rice (**D, E, F**), *TATA-box*-containing core promoters (**D**) represent only about 18% of all protein-coding genes in the rice genome. *Y-Patch* is found in as much as 50% of the total protein-coding genes either in combination with or independent of *TATA-box* (**E, F**). *DPE*-like and *BRE*-like sequences are insignificantly represented in the core promoters of rice genes.

surveys of Arabidopsis, human and Drosophila genomes have shown that this was not the case (Molina and Grotewold, 2005; Lenhard et al., 2012). For instance, although the *TATA-boxes* among Arabidopsis genes were contextually positioned at the correct upstream location, their occurrence in such context was established only in less than 30% of the total validated genes. While the occurrence of *TATA*-related motifs was established reliably among highly expressed genes, other more prominent sequence motifs distinct from *BRE* and *DPE* have been identified in smaller subsets of *TATA*-less Arabidopsis genes (Figure 1).

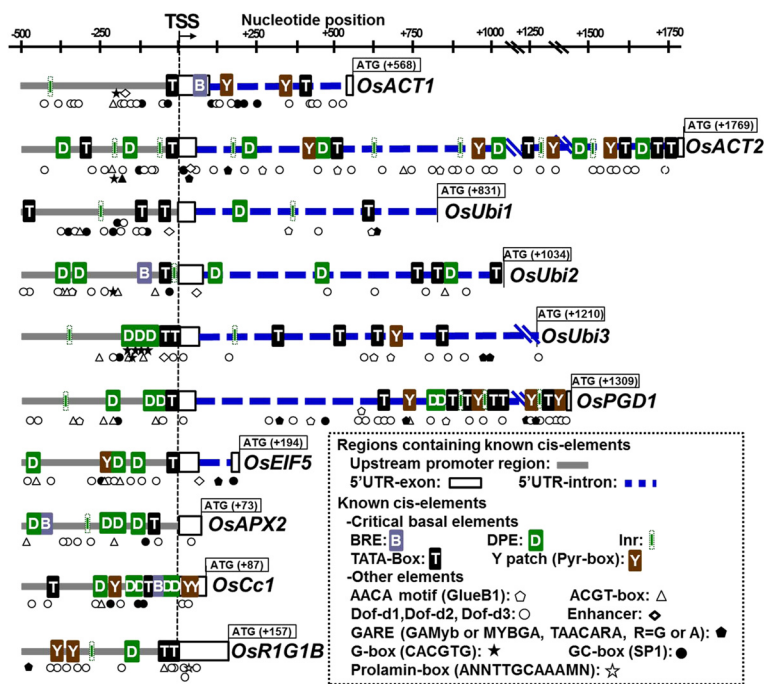
Indeed, more recent surveys of *TATA-box* distribution in the rice genome established a trend similar to what has been established in Arabidopsis (Figure 1). Independent analysis with Gibbs Motif Sampler and Dragon Motif Builder showed that less than 20% of annotated rice genes contain *TATA box*-like motifs in the right context. Furthermore, *DPE*-like and/or *BRE*-like sequences are not significantly represented in rice core promoters (Clivan and Svec, 2009). The consensus sequence for rice *TATA-box* is defined as 'CTATAWAWA' located within a C-rich region around -49 to -20. However, much higher occurrence and broader distribution of a T/C-rich motif called *Y-Patch* (pyrimidine patch) was established in a global survey of rice core promoters. Although it has lesser degree of sequence conservation than *TATA-box*, *Y-Patch* elements occur in correct upstream context in more than 50% of annotated rice genes, either in combination with or independent of *TATA-box*. The consensus *Y-Patch*

sequence has been defined as 'CYTCYCCYC', occurring in one or more copies and with strict directional sensitivity around -49 to +1 region (Yamamoto et al., 2007; Clivan and Svec, 2009).

#### Key regulatory modules for constitutive expression of rice genes

A number of constitutive promoters have been particularly well investigated in rice including actin (*OsAct*), ubiquitin (*OsUbi*), cytochrome-c (*OsCc1*), ascorbate peroxidase (*OsAPX*), translation initiation factor-5 (*OseIF5*), phosphoglucate dehydrogenase (*OsPGD*), and R1G1-domain containing protein B (*OsRIG1B*) genes. The shared structural feature of these promoters is a short exon and large intron in the untranslated region (UTR) near the TSS, in addition to upstream regulatory modules (Figure 2, Table 1). The 5'-UTR introns are A/T-rich but share little sequence identity and with variable lengths (Jeong et al., 2006).

The strong inducing effects of *OsAct* and *OsUbi* promoters are to a large degree due to the UTR introns. Intron-mediated enhancement (*IME*) of transcription requires the transcribed sequences near the TSS in the right orientation. In general, *IME*-type introns in rice are longer than non-enhancing introns and contain more regulatory elements (Bradnam and Korf, 2008; Parra et al., 2011). Dispersed within the *IME*-type intron are potential cis-elements including the predominant pentamer motif 'CGATT'. It has been suggested that *IME*-type introns act through gene loops that facilitate steady reinitiation of



**Figure 2** Models of the regulatory sequence architectures of well characterized promoters of constitutively expressed rice genes such as actin (*OsACT1*, *OsACT2*), ubiquitin (*OsUbi1*, *OsUbi2*, *OsUbi3*), phosphogluconate dehydrogenase (*OsPGD1*), Initiation factor (*OsEIF5*), ascorbic peroxidase (*OsAPX2*), cytochrome-c (*OsCc1*), and R1G1-domain-containing protein (*OsR1G1B*). Locations of critical cis-elements are indicated with colored symbols in both the upstream regions (gray lines) and downstream regions (5'UTR-exon, 5'UTR-intron; blue lines) relative to the location of the transcription start site (TSS).

transcription by RNA Pol II (Moabbi et al., 2011; Rose et al., 2011). The coding sequences of ubiquitin (*OsUbi3*) monomer often involving the first nine nucleotides enhance promoter activity over 4-fold (Clancy et al., 1994). In addition, the most commonly found elements in the 5' upstream and UTR intron are Dof, GC-box, and ACGT-box (Figure 2, Table 1).

Typical cis-elements in canonical core promoters are also important elements required for strong constitutive expression of *OsAct*. For example, *OsAct1* promoter contains a *TATA-box* at -35, a *BRE-like* motif at position +88 and two copies of *Y-Patch* element at positions +138 and +384. In addition, the region from position -50 are highly enriched with other potential regulatory elements

**Table 1** List of experimentally or computationally defined upstream regulatory sequences, and their cognate transcriptional regulators that determine basal and constitutive expression of rice genes

Cis-element	Core motif	Transcriptional regulator	Biological function
<b>TATA-box</b>	CTATAWAWA	TATA-binding protein ( <i>OsTBP2</i> )	Core promoter; Pre-initiation complex
<b>Y-Patch</b> (Pyrimidine patch)	CYTCYCCYC	Unknown	Core promoter; Pre-initiation complex
<b>Actin (Act) elements</b>	CCCAA (tandem octa-repeats)	Unknown	Fine-tuner of constitutive <i>OsAct</i> genes
	CAAT (CCAT-box)	Unknown	Positive regulator of constitutive <i>OsAct</i> genes
	GTGAC (intronic)	Unknown	Positive regulator of constitutive <i>OsAct</i> genes
	CCGCGTTGGC (ABA-responsive)	Unknown	Positive regulator of constitutive <i>OsAct</i> genes
	ATTAAT, CACGTA (light-responsive)	Unknown	Positive regulator of constitutive <i>OsAct</i> genes
	(A/T)GCC (dideca-repeats)	Unknown	Positive regulator of constitutive <i>OsAct</i> genes
<b>Ubiquitin (Ubi) elements</b>	GTTGTGTTTG	Unknown	Positive regulator of constitutive <i>OsUbi</i> genes

List was based on seminal studies published in the literature and annotated in public databases.



including *Sp1*, light-responsive *SORLIP*, mesophyll-specific *CACTFTPPCA1*, *Alfin1*, a novel zinc-finger binding element found in the salt stress inducible genes, a signature element of nitrate reductase gene *AMMORESIIUDCRNIA1* and a drought-responsive *SBOXATRBCS* (Figure 2, Table 1).

The critical regulatory elements of rice actin genes (*OsAct1*, *OsAct2*) include a 5'-upstream sequence, non-coding exon, 5'-UTR region intron, and a short non-coding region of the second exon (Figure 2). The first non-coding exon and intron are 79-bp and 58-bp, and 447-bp and 1,736bp for *OsAct1* and *OsAct2*, respectively (McElroy et al., 1991; Zhang et al., 1991; He et al., 2009). Many cis-elements required for constitutive expression are located around the core promoter region, including a 38-bp poly(dA-dT) between positions -146 and -186 which act as a positive regulator, and eight tandemly repeated copies of imperfect pentamer motif with consensus 'CCCAA' at -294, acting as a negative regulator. In addition, the branch-point site ('GTGAC') for mRNA splicing located in the intron between the positions 359 and 363 is essential for promoter activity (McElroy et al., 1990).

A negative regulator between positions +96 and +274 within the intron is responsible for the repression of *OsAct2* expression. On the other hand, the two *CAAT-boxes* between positions -448 and -445 and positions -635 and -632 act as possible positive regulators. Other cis-elements in the 5'-upstream region include ABA-responsive element ('CCGCGTTGGC'), two different light responsive elements ('ATTAAT', 'CACGTA'), and 20 repeated 'A/TGCC' triplets. Numerous cis-elements are also present between positions -50 and TSS, including *CAATBOX1*, *SORLIPIAT*, *Alfin1*, *GATABOX*, *UPRMOTIFIIAT*, *C-rich Q*, *BOXCP-SAS1*, and 'CGTGG' motif.

Ubiquitin (*OsUbi*) promoter includes the 5'-upstream sequence and 5'-UTR containing a short first exon and intron (Wang and Oard, 2003; Bhattacharyya et al., 2012). Critical elements include a signature enhancer motif ('GTTGTGGTTTG') found in both *OsUbi1* and *OsUbi2* (Christensen and Quail, 1996; Wang and Oard, 2003). In addition to a *TATA-box*-like element, core promoters of *OsUbi* contain other signature cis-elements associated with inducible expression such as *Alfin1*, *Dof*, *CARGCW8GAT*, *TAAAGSTKST1*, *PRECONSCRHSP70A*, *GT-box*, *POLLENILELAT52*, *WUSATAg*, and *ARRIAT* (Figure 2, Table 1).

The promoter of cytochrome-c (*OsCcl1*) includes the 5'-upstream region and 86-bp untranslated sequence without UTR intron (Figure 2). In addition to the signature core promoter elements such as *TATA-box* and *CAAT box*, it contains the cAMP response element (CRE), nuclear respiratory factor (NRF) binding site, and light responsive G-box (CACGTV) (Jang et al., 2002). Promoters of ascorbate peroxidase (*OsAPX2*), phosphogluconate dehydrogenase (*OsPGD1*), and *RIG1*-domain-containing protein

(*OsRIG1B*) have similar constitutive strengths as the promoter of *OsAct1*, with various enhancers such as light responsive G-box and TGACG-element associated with high-level gene expression. Endosperm-specific regulatory elements such as 'AACA' and 'ACGT' are also important characteristics of *OsPGD1* and *OsRIG1B* (Park et al., 2010a; Figure 2, Table 1). Based on the trends established from the analysis of housekeeping genes, it is apparent that constitutive expression requires many of the elements of the core promoter in various combinatorial complexities with various UTR structural features and a plethora of cis-elements involved in developmental and environmental signal perception.

### Integration of hormonal, developmental and stress response signaling

Plant growth, development and adaptation are consequences of cooperative and antagonistic actions of different regulatory hormones, starting from germination to the development and maturation of seeds. Integration of hormonal and other intrinsic developmental and environmental signals leads to changes in gene expression.

### Root development and auxin signaling

Auxin is recognized as the universal plant growth hormone because of its principal role in the regulation of cell expansion, division and differentiation (Kieffer et al., 2010). The principal receptor of auxin is a class of nuclear F-box protein that belongs to the TIR1/AFB family (Darmasiri et al., 2005). Auxin perception leads to rapid induction of three classes of auxin early responsive genes, which include the *Aux/IAA*, *GH3* and the small auxin-up RNA or *SAUR* (Hagen and Guilfoyle, 2002). *Aux/IAA* regulates its own transcription, and functions as transcriptional regulator of other auxin responsive genes in cooperation with Auxin Response Factors (ARF) (Chapman and Estelle, 2009). ARF binds specifically to the promoters of *GH3*, *SAUR* and other auxin-responsive genes leading to rapid induction (Ulmasov et al., 1995; Ulmasov et al., 1997). In the absence of auxin, *Aux/IAA* inhibits ARF function and represses other auxin early responsive genes. Auxin binding activates TIR1/AFB, a component of ubiquitin E3 ligase SCF1, mediating ubiquitination and degradation of *Aux/IAA*. In effect ARF is relieved from repression, hence expression of auxin early responsive genes (Tiwari et al., 2003; Quint and Gray, 2006). In rice, 25 *OsARF* homologs have been identified, indicating a high degree of conservation across the dicot and monocot lineages (Wang et al., 2007).

Analysis of the japonica rice reference genome identified 31, 12 and 58 members of the *OsAux/IAA*, *OsGH3* and *OsSAUR* gene families, respectively, many of which exhibit the characteristic rapid induction in response to auxin (Thakur et al., 2001; Jain et al., 2006a; Jain et al., 2006b; Jain et al., 2006c; Song et al., 2009). Moreover, it

has been estimated that between 2% to 5% of the total genes in Arabidopsis and rice are either upregulated or downregulated by auxin (Nemhauser et al., 2004). The signature of the promoter of auxin early responsive genes is a 6-bp motif with the consensus sequence 'TGTCTC', often referred to as auxin-response element (*AuxRE*), occurring in either simple or composite configurations (Ulmasov et al., 1997). Simple *AuxRE* contains either direct or palindromic repeats of 'TGTCTC' motif separated by 3 to 7-nt spacer with no apparent consensus sequence. In contrast, composite *AuxRE* is functional only in modular form with CG-rich 'coupling' or 'constitutive' element (Table 2).

The analysis of the *crown rootless* (*cr11*) mutant, defective in auxin-mediated formation of crown roots facilitated the elucidation of cis-elements and trans-acting factors involved in auxin-regulated gene expression in rice (Inukai et al., 2005; Liu et al., 2005a). The *cr11* mutation exhibits various auxin-related abnormalities including reduced lateral roots, insensitivity to auxin during lateral root initiation, and impaired root gravitropism. Wild-type *Cr11* encodes an Asymmetric leaves-2/Lateral Organ Boundaries (AS2/LOB) domain protein containing the N-terminus C-motif and GAS-motif at the C-terminus. The mutant protein contains a single amino acid substitution (A to T) within the AS2/LOB domain. The wild-type *Cr11* is strongly induced by auxin, dependent on the degradation of Aux/IAA proteins, as shown by the effects of mutation in the degradation domain (domain II) of *OsIAA3*. The fact that Aux/IAA functions as a repressor of the transcriptional activator ARF has implied that *Cr11* is a direct target of *OsARF* transcription factors. The promoter of wild-type *Cr11* has two copies of *AuxRE* signature sequence. The first copy (*AuxRE1*) located upstream and distal from the *TATA-box* has the 'TGTCTC' sequence, while the second copy (*AuxRE2*) located upstream but more proximal to the *TATA-box* has the inverted sequence 'GACACA' (Table 2). The primary auxin-response transcriptional activator in rice (*OsARF1*) binds with high affinity with *AuxRE2* but not with *AuxRE1* (Waller et al., 2002).

Another *crown rootless* mutant in rice (*cr15*) has impaired initiation of crown root primordia (Kitomi et al., 2011). Mutants have abnormal root gravitropic response, delayed flowering, various types of floral anatomical abnormalities and small panicles. In wild-type plants, *Cr15* is rapidly induced by exogenous auxin in the nodes or in regions of the stem where root growth is initiated. Like *Cr11*, rapid expression of *Cr15* is also dependent upon the degradation of Aux/IAA proteins, thus *Cr15* is downstream to *OsAux/IAA* and *OsARF*. The 642-aa residue CRL5 protein contains two copies of AP2 DNA-binding domain, related to AP2/ERF transcription factor *AINTEGUMENTA* (*ANT*). The promoter of *Cr15* contains three copies of the 'TGTC' core motif of

*AuxRE*. The first copy (*RE1*), which is upstream and most distal to the *TATA-box* consists of a tandem array of three inverted 'TGTC' motifs ('GACAGACAGACA'). The second copy (*RE2*) is a single base substituted version ('TGTCGC') of the consensus for *AuxRE*. The third copy (*RE3*), which is upstream but most proximal to the *TATA-box* has two tandem 'TGTC' motifs with a dinucleotide spacer in between, i.e., 'GACActGACA' (Table 2). Direct binding of recombinant *OsARF1* was demonstrated for the most proximal wild-type *RE3* but not for its mutated version 'GATAActGATAct', indicating that *Cr15* is directly activated by *OsARF1* like *Cr11*. The consensus from the analysis of *Cr11* and *Cr15* is that integration of auxin-responsive gene expression during root development requires a proximal *AuxRE*.

### **Seed germination and gibberellic acid signaling**

First discovered in rice, the signature effect of gibberellic acid (GA) is the ability to promote stem elongation. Bioactive GAs affect virtually all aspects of plant growth and development from germination to hypocotyl elongation, stem growth, circadian rhythm, and reproductive organ and seed development (Lovegrove and Hooley, 2000; Hartweck and Olszewski, 2006).

Discovered first in rice, *GID1/GID2* genes encode the nuclear localized primary receptor of GA (Gomi et al., 2004; Ueguchi-Tanaka et al., 2005; Griffiths et al., 2006; Ueguchi-Tanaka et al., 2007). Another important component of GA signaling are the DELLA proteins encoded by *SLR1* in rice and *GAI/RGA-like* genes in Arabidopsis, which function as negative regulators of GA response (Ikeda et al., 2001; Fleet and Sun, 2005; Itoh et al., 2005). The rice DELLA protein is stable in the absence of GA ensuring that transcription of GA-early responsive genes are repressed by blocking the activity of a transcriptional activator bound to the promoters of target genes. GA binds to the nuclear localized GID1 receptor, triggering its interaction with the DELLA repressor. The GID1-DELLA complex recruits the GID2 F-box protein, a component of the SCF ubiquitin ligase complex, resulting in the degradation of the DELLA repressor and subsequent transcription of GA-early responsive genes (Hartweck, 2008). Analysis of the GA-response transcriptome estimated that about 250 and 300 genes are regulated by GA in Arabidopsis and rice, respectively (Ogawa et al., 2003; Yazaki et al., 2004).

In rice and other cereals, available sugars in the embryo are rapidly depleted at the onset of germination, triggering synthesis of GA which is then released to the aleurone layer. Subsequent expression of hydrolases ensures sufficient supply of sugar from the breakdown of starch to support seedling growth (Kaneko et al., 2002). Expression of  $\alpha$ -amylases during rice germination is repressed by high sugar and abscisic acid (ABA) in the embryo, while they are induced by GA in the

**Table 2 List of experimentally and computationally defined upstream regulatory sequences, and their cognate transcriptional regulators that determine spatio-temporal control of rice genes**

Cis-element	Consensus core motif	Transcriptional regulator	Biological function
<b>Auxin responseelement (AuxRE1)</b>	TGTCTC	Auxin responsefactor ( <i>ARF1</i> , <i>ARF2</i> )	Primary regulator of auxin-regulated expression
<b>Auxin responseelement (AuxRE2)</b>	GACACA	Auxin responsefactor ( <i>ARF1</i> , <i>ARF2</i> )	Primary regulator of auxin-regulated expression
<b>Auxin responseelement (AuxRE3)</b>	GACActGACA	Auxin responsefactor ( <i>ARF1</i> )	Primary regulator of auxin-regulated expression
<b>Gibberellin responseelement (GARE)</b>	TAACCACC	<i>GAMYb</i> factors	Primary regulator of GA-regulated expression; primary regulator of sugar sensitivity of GA-responsive genes
<b>Glut element</b>	AACA/ACGT	Unknown	Regulator of endosperm-specific expression of rice seed storage protein glutelin ( <i>OsGlut</i> )
<b>Gt1/Gt2 box-I</b>	ATATCATGAGTCACTTCA	<i>GT1</i> , <i>GT2</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>Gt1/Gt2 box-II</b>	CTTCGTGTACCACA	<i>GT1</i> , <i>GT2</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>Gt1/Gt2 box-III</b>	ACAATGCTGCTCAATTA	<i>GT1</i> , <i>GT2</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>Gt1/Gt2 box-IV</b>	ATTATCCAATGTCATATTG	<i>GT1/GT2</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>Gt1/Gt2 box-V</b>	TAAGTACGTTTGATGA	<i>GT1</i> , <i>GT2</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>Gt3 box</b>	ATATCATGAGTCACTTCA	<i>GT1</i> , <i>GT2</i> , <i>GT3</i>	Positive regulator of endosperm-specific expression of rice glutelin ( <i>OsGlut</i> ) genes
<b>PB-1 element</b>	AAGCAACACACAAC	Dof-type P-boxBinding factor ( <i>PBF</i> )	Positive regulator of endosperm-specific expression
<b>Prolamin box</b>	TGTAGAA, TGTTTTAATATGACGTGG	Dof-type P-boxbinding factor ( <i>PBF</i> )	Positive regulator of endosperm-specific expression
<b>SP8/SURE</b>	TACTATT, TCACTATT,	<i>Opaque-2</i> transcriptionfactor	Positive regulator of endosperm-specific expression
<b>LPSE1</b>	ATTTGAGCTGCC	Unknown	Positive spatial regulator of expression in green leaves, stem and young panicle
<b>LPSE2</b>	TTGATATATTTGT	Unknown	Positive regulator of expression on green leaves and negative regulator of expression in young panicle and roots
<b>LPSRE1</b>	CGGCGCGCCAC	Unknown	Positive regulator of expression in green leaves and negative regulator of expression in panicle and stem
<b>LPSRE2</b>	TTAGATAATGGA	Unknown	Positive spatial regulator of expression in green leaves, stem root, and young panicles
<b>PSE1</b>	TTTATCTATTCC	Unknown	Negative regulator of expression in young panicle and stem
<b>PSE2</b>	TCTTTGGCAGAG	Unknown	Positive regulator of expression in green leaves and negative regulator of expression in stem and young panicles
<b>Ethylene and jasmonic acid responsive element</b>	GCC-CORE	Ethylene response factor	Regulator of senescence-induced gene expression ( <i>OsERF</i> )
<b>GS2</b>	TTCGAAATTGAACGTGCT-TAACCAAGAGAACAC	<i>WRKY</i> transcription factors	Senescence box
<b>ABRE (ACGT-type)</b>	CGMCACGTB	<i>OsABIS</i> , <i>OREB1</i>	ABA-response element for ABA regulated gene expression during stress response and seed maturation
<b>CE1</b>	TGCCACCGG	<i>OsABIS</i> , <i>OREB1</i>	ABRE coupling factor
<b>CE3</b>	ACGCGTGTCTCTC	<i>OsABIS</i> , <i>OREB1</i>	ABRE coupling factor
<b>DRT/CRE</b>	(a/g)CCGAC	<i>OsDREB1</i> , <i>OsDREB2</i>	Dehydration response element for abiotic stress regulated gene expression independent of ABA

**Table 2 List of experimentally and computationally defined upstream regulatory sequences, and their cognate transcriptional regulators that determine spatio-temporal control of rice genes (Continued)**

<b>Abiotic/oxidative stress Element</b>	CACG	<i>OsNAC6, OsNAC5, OsNAC2</i>	Regulator of abiotic stress-responsive gene expression independent of the DREB regulon
<b>as1/ocs/TGA-like element</b>	AATTTGAT, TAATTTGA	<i>OsTGA10</i>	Regulator of chilling mediated oxidative stress signaling
<b>Pyrimidine-box/Myb2-box element</b>	AAAGAAAA, TAGTTTT	<i>OsMyb4</i>	Regulator of chilling mediated oxidative stress signaling
<b>JARE</b>	CGTCA, TGACG, VCGCGB	Ethylene response factors ( <i>ERF</i> )	Jasmonic acid response element; regulator of responses to pathogens and herbivores
<b>GCC-box</b>	TAAGAGCCGCC	Ethylene response factors ( <i>ERF</i> )	Negative regulator of responses to <i>Magnaporthe oryzae</i>
<b>P-box</b>	CCGCCCTCCA	Ethylene response factors ( <i>ERF</i> )	Disease responsive element
<b>W-box</b>	TTGAC element	<i>OsWRKY45, OsWRKY71, OsWRKY03, OsWRKY53, OsWRKY31, OsWRKY30, OsWRKY76</i>	Disease responsive
<b>E-box</b>	CAGTCG, CACCTG	<i>OsbHLH65</i>	Pathogen and herbivore response element; regulator of responses to <i>Magnaporthe Oryzae</i> and <i>Nilaparvata lugens</i>

List was based on seminal studies published in the literature and annotated in public databases.

endosperm. Dissection of the promoter architecture of  $\alpha$ -amylases established a clear picture of the interacting cis-elements that mediate differential regulation (Lu et al., 1998; Chen et al., 2002; Sutoh and Yamauchi, 2003; Chen et al., 2006). Among the members of the germination-associated  $\alpha$ -amylase gene family,  $\alpha$ Amy3 (*RAmy3D*) and  $\alpha$ Amy8 (*RAmy3E*) provide an interesting contrast because both genes are synchronously induced by sugar depletion in the embryo, while in the endosperm  $\alpha$ Amy3 is induced only by sugar depletion and  $\alpha$ Amy8 only by GA. Induction of  $\alpha$ Amy8 by GA is also correlated with the expression of *GAMyb* transcription factors.

The sugar response complex (SRC) was identified within the -186 to -82 region of  $\alpha$ Amy3, comprised of a *GC-box* (~30-nt GC-rich motif) distal to the *TATA-box*, two tandemly repeated copies of *TA-box* ('TATCCA') upstream but more proximal to *TATA-box*, and *G-box* ('CTACGTGG') in between the two elements. The sugar response complex/GA-response complex (SRC/GARC) of  $\alpha$ Amy8 is located within -318 to -89, comprised of a *GC-box* upstream and distal to *TATA-box*, single copy of proximal *TA-box*, and a gibberellic acid response element *GARE* ('TAACCACC') in between the two elements (Table 2). Mutation of *GARE* in the promoter of  $\alpha$ Amy8 confirmed its critical role for GA-regulation as well as sugar-insensitivity (Table 2). Therefore, differential regulation of  $\alpha$ -amylases by sugar depletion and GA in embryo and endosperm is facilitated by the interaction of *GARE* with *GC-box* and *TA-box*.

Expression of *OsGAMyb* transcription factors and  $\alpha$ Amy8 are highly coordinated in the endosperm. In rice *gamyb-2* mutant with impaired GA-dependent expression of  $\alpha$ -amylases in the endosperm,  $\alpha$ Amy8 becomes

non-responsive to GA and repressible by glucose. Similarly, overexpression of *OsGAMyb* transcription factor has no effect on glucose repression of  $\alpha$ Amy3, while an increase in expression of *OsGAMyb* renders  $\alpha$ Amy8 insensitive to glucose repression. All these results established that *OsGAMyb* is a key regulator allowing differential responses of  $\alpha$ -amylase genes to sugar depletion and GA in the endosperm and embryo by virtue of its interaction with *GARE* in *GARC*-promoters (Table 2).

#### **Seed storage protein regulation during seed development**

Making up as much as 80% of the total endosperm proteins of rice, glutelins serve as major nitrogen and carbon reservoir for the initial energy requirements during germination (Okita et al., 1989). Rice glutelins are classified into three groups: *GluA*, *GluB*, and *GluC*. The spatio-temporal regulation of three representative members of *GluA* sub-subfamily (*Gt1*, *Gt2*, *Gt3*) have been extensively investigated in an effort to understand their roles in the regulation of seed development. *Gt1* and *Gt2* are transcribed about 5 days after anthesis and continue virtually throughout the entire duration of seed development. In contrast, *Gt3* transcripts accumulate maximally between 5 to 10 days post-anthesis followed by a steady decline (Okita et al., 1989; Zhao et al., 1994).

Promoter studies established that both 5' distal and proximal cis-elements in *Gt1* promoter are critical for precise spatio-temporal regulation (Zheng et al., 1993). The distal region (-5.1 to -1.8 kb) contains positive elements for temporal regulation and confer as much as 20-fold increase in expression. Leaky expression in leaves, stems, sheaths and roots was observed in promoter-GUS reporter driven by -507 or -154 truncated fragments of *Gt1* promoter. This indicates that



while the distal positive element within the -5.1 kb promoter contributes to temporal control of *Gt1*, proximal elements are critical for spatial specification in the endosperm.

DNase-I footprinting identified five protein-binding motifs (box-I to box-V) in *Gt1* promoter ordered sequentially downstream from -434 position: 'TAAGTCACGTTTGA TGA' (box-V), 'ATTATCCAATGTCATATTG' (box-IV), 'ACAATGCTGCTCAATTA' (box-III), 'CTTTCGTGTAC CACA' (box-II), and 'ATATCATGAGTCACTTCA' (box-I) (Table 2). Promoter activity is completely abolished when boxes I to IV are mutated simultaneously. Box-II is the most critical, causing about 35-fold reduction in expression when mutated. Consistent with similar spatio-temporal patterns of *Gt1* and *Gt2*, boxes-I to V are highly conserved between the two promoters, except that box-V has varied locations (Kim and Wu, 1990).

Similar to *Gt1*, the functional elements responsible for spatial, temporal and quantitative effects of *Gt3* promoter are located at distal (-945 to -726) and proximal (-346 to -263) regions (Zhao et al., 1994; Croissant-Sych and Okita 1996). The critical elements responsible for precise temporal control post-anthesis are located within +7 to -181 and -278 to -320 regions. Removal of either of these proximal elements causes as much as 50% reduction in expression. The -106 to -88 and -125 to -111 regions of *Gt3* promoter contain sequences similar to box-I and box-II in *Gt1* and *Gt2* promoters, with the same relative positions. Box-I contains the signature binding sequence ('ATGAc/gTCAT') for the yeast *GCN4*, *AP-1*, *Jun* and *Fos* transcription factors, while box-II contains another motif found in seed-specific  $\alpha$ -amylase genes. *Gt1*, *Gt2* and *Gt3* promoters also share the signature motif 'AACA' around -75 to -64, which is essential for endosperm-specific expression (Table 2).

*Gt3* promoter differed from *Gt1* and *Gt2* promoters by the absence of boxes-III, IV and V. However, its -482 to -278 region contains several functional elements that may be critical for its unique expression including the 'AAGCAACACACAAC' motif (-279 to -259) known as the binding site for nuclear factor PB-1, two prolamin-box-like (*P-box*-like) signals 'TGTAGAA' (-302 to -295) and 'TGTTTTAA' (-322 to -314) known to interact with Dof-type *P-box* binding factor (*PBF*), and 'TATGACGTGG' motif, which resembles the consensus binding sequence for the seed-development transcription factor *Opaque-2* (Table 2). *Gt3* promoter also contains additional sequences around -482 region that suppresses its expression in the leaves and other vegetative organs. Studies on other glutelin gene sub-families also indicated the critical roles of *GCN4*, AACA-motif and P-box for endosperm-specific expression of *GluB-3* and *GluB-5* (Washida et al., 1999; Qu et al., 2008; Table 2).

### Tissue/organ-specific gene expression

In rice, the nature of cis-elements and trans-acting factors that confer tissue or organ-specific programs is not yet well understood. However, results of the analysis of few candidate rice genes are starting to reveal interesting trends. One good example is *Oshox1*, encoding a homeo-domain leucine zipper (HD-Zip) transcription factor that plays an important role in vascular differentiation. *Oshox1* expression initiates when procambial cells have already acquired distinctive anatomical properties but before achieving terminal differentiation (Scarpella et al., 2000).

Upstream region (-1,600) of *Oshox1* directs vascular-specific expression through a process that involves auxin and sucrose (Scarpella et al., 2005). *Oshox1* promoter drives expression in a tissue-specific fashion, showing particularly high expression in the apical region of primary and secondary procambial strands of cotyledon, but not in the axis of mature embryo. It also directs expression in the vascular cylinder of the basal mature region of the root, and in the apical hydathode of developing first leaves. Various proximal and distal sections of the -1,600 upstream sequence of *Oshox1* confer different patterns of vascular cell, tissue and organ-specific expression and these sections contain modular combinations of 'CCA/TTG' repeats with stretches of 'CCCC' or 'AACA' motifs (Table 2). The auxin-responsive complex of *Oshox1* promoter is located within -1,621 to -898 region defined by 11 copies of 'AAGG' *Dof* element and 'GGTCCAT' signature element typical of many known auxin-responsive genes. The same region (-1,621 to -898) is also critical for sucrose-regulated expression, consistent with the occurrence of several known sucrose responsive elements including *SP8 element* ('TACTATT'), three copies of *SURE* ('TACTATT', 'TCACTATT', 'TACTAT') and one copy of *B-box* ('CTAAC') (Grierson et al., 1994; Ishiguro and Nakamura, 1994; Baumann et al., 1999; Zourelidou et al., 2002).

Another example of organ/tissue-specific expression is illustrated by a photosystem-II 10-kDa polypeptide-encoding gene (Os08g10020), which is expressed strongly in leaves and sheaths but very weakly in non-green vegetative tissues of the root, stem, mature panicle and seed (Cai et al., 2007). The promoter of this gene spans the -2,134 to +41 region, punctuated by proximal core promoter signals such as '*TATA-box*' and '*CAAT-box*'. The full-length promoter confers 14-fold, 90-fold, 14-fold and 129-fold higher expression in the leaves than in the stem, endosperm/embryo, panicle and root, respectively. As positive regulator, the '*leaf and panicle/stem-specific element*' *LPSE1* ('ATTTGAGCTGCC') within -106 to -96 defines the expression in leaves and young panicles. As negative regulators, the '*leaf, panicle/stem and root element*' *LPSRE2* ('TTAGATAATGGA') within -545 to -534 suppresses expression in the leaf, root, young panicle and stem, while the '*panicle/stem-specific element*' *PSE1* ('TTTATCTATTTCC') within -708 to -696

suppresses expression in young panicle and stem. Elements with dual roles as positive and negative regulators include the *LPSRE1* ("CGGCGCGCCAC") within -222 to -214, *LPSE2* ("TTGATATATTTGT") within -383 to -371, and *PSE2* ("TCTTTGGCAGAG") within -670 to -659. All three elements function as enhancers in the leaf while *LPSRE1*, *LPSE2* and *PSE2* are silencers in the stem, young panicle/root, and young panicle/stem, respectively (Table 2).

#### **Aging and senescence**

Leaf senescence is a consequence of either developmentally programmed or stress-induced cell deterioration and declining rate of photosynthesis (Buchanan-Wollaston, 1997; Gan and Amasino, 1997; Chandlee, 2001). Cytokinin and ethylene are two major hormones that regulate the process of leaf senescence by virtue of their promoting and inhibiting effects, respectively. Other hormones and signaling molecules such as jasmonic acid, salicylic acid, abscisic acid (ABA) and reactive oxygen species also play either major or minor role singly or in combination (Nam, 1997). In rice, age-dependent leaf senescence is an important process for grain filling and seed maturation while developmentally uncoupled programmed cell death (premature senescence) is often induced by environmental stresses (Buchanan-Wollaston et al., 2003; Lim et al., 2007).

The ability to delay senescence has always been a major goal of plant biology for yield improvements by enhancing photosynthesis, i.e., 'stay green' phenotype (Thomas and Howart, 2000). Discovery of several classes of senescence-associated genes (*SAG*) provided a window to the mechanisms of senescence-regulated gene expression (Robatzek and Somssick, 2002). It was estimated that about 40 different types of transcription factors are involved in the regulation of *SAGs* including several *Myb*, *zinc finger*, *MADS-box*, *WRKY* and *bZIP* transcription factors (Chen et al., 2002). Knock-out mutation of senescence-associated *WRKY6* has been shown to alter the spatio-temporal expression of certain *SAGs* delaying onset of senescence.

The cysteine protease-encoding *AtSAG12* of Arabidopsis, which has a functional homolog in rice (*OsSAG39*) is highly expressed in senescing leaves, downregulated by auxin, cytokinin, and sucrose but not affected by desiccation, dark incubation, wounding, ABA or ethylene (Noh and Amasino, 1999). The senescence box or 'GS2' in the promoter of *AtSAG2* is defined by the motif "TTCGAAATTGAACGTGCTTAACCAAGAGAACAC" within -603 to -571 region that binds nuclear protein extracts from senescing leaves (Table 2). Recently, it was found that the upstream region of *OsSAG39* contains positive regulatory elements within -2,021 to +24 for maximal expression when chlorophyll content is down to about 40% (Liu et al., 2010). Within its proximal and distal regions are arrays of known cis-elements such as 'GCCCORE' for ethylene and jasmonic

acid regulated expression, ABA-boxes *RYREPEATBNNA PA*, *MYCATRD22* and *DPBFCOREDCDC3*, drought elements *MYBCORE* and *MYBST1*, *WRKY7IOS* site for *OsWRKY71*, and elicitor responsive element *HBOXCONSENSUSPVCHS* (Urao et al., 1993; Baranowskij et al., 1994; Abe et al., 1997). *OsSAG39* promoter is inducible by the ethylene precursor ACC, ABA, GA, and low temperature. Protein extracts from senescing leaves bind to the *HBOXCONSENSUSPVCHS* and *WRKY7IOS* elements (Table 2).

#### **Abiotic stress and abscisic acid signaling**

Drought, low temperature and salinity are essentially overlapping forms of stresses characterized by similar types of physiological perturbations and responses. At the cellular level, they trigger 'physiological dehydration', which affects almost every aspect of cellular physiology. As a major stress hormone in plants, the various components of abscisic acid (ABA) signal transduction pathway has been elucidated using Arabidopsis as a model (Razem et al., 2006). Seminal studies in Arabidopsis also led to the elucidation of major cis and trans components of stress response mechanisms that are evolutionarily conserved in flowering plants. Exposure to low temperature, dehydration and salinity stresses induces a common subset of genes referred to as *cor*, *cold* regulated or *rd*, responsive to desiccation genes through an ABA-dependent or ABA-independent pathway (Liu et al., 1998; Thomashow, 1999; Shinozaki and Yamaguchi-Shinozaki, 2000; Yamaguchi-Shinozaki and Shinozaki, 2005).

Subsequent transduction of ABA towards gene expression involves phospholipase-D-dependent signaling and expression of various *MYB*, *MYC* and *bZIP* transcription factors that control the expression of different ABA responsive genes through the various combinations of cis-elements including *MybR*, *MycR* and the most predominant *ABRE* (ABA Response Element) (Finkelstein, 2006). *ABRE* interacts with bZIP-type ABRE-binding factors or ABF that function as regulators of stress response and seed maturation (Giraudat et al., 1994; Guedes Correa et al. 2008). The *ABF*-target genes include many *cor/rd* genes with the *ABRE* core motif 'ACGT', also referred to as *G-box* or *C-box* (Guiltinan et al., 1990). Additionally, *ABRE* functionality is combinatorial with other types of 'non-ACGT' coupling elements such as *CE1*, *CE3* and *DRE* (drought response element), comprising the minimal ABA-responsive complex or *ABRC* (Shen and Ho, 1995; Shen et al., 1996; Hobo et al., 1999; Narusaka et al., 2003).

In rice, the dominant ABRE-binding transcription factors have been identified as *OsABIS* and *OREB1* (Nakashima et al., 2009; Todaka et al., 2012). A recent genome-wide survey of co-expressed rice genes estimated that more than 10% (>400) of total genes are regulated through the ABA-signaling pathway, based on over-representation of the

consensus sequence 'CGMCACGTB' within -1,000 upstream regions (Cheng et al., 2007; Lenka et al., 2009; Yun et al., 2010; Table 2). Included among the ABRE-enriched genes are various components of stress signaling and defense mechanisms, further reiterating the important role of ABA as a central signal that integrates various types of stress defenses in rice (Cheng et al., 2007; Yun et al., 2010). Furthermore, genome-wide comparative analysis of ABRC composition between in rice and Arabidopsis revealed striking differences between the two species (Gomez-Porrás et al., 2007; Xu et al., 2012a). While ABRCs defined by ABRE-ABRE duplex occurred at similarly high frequencies among ABA-regulated genes in both species, the occurrence of CE3 coupling element with the consensus sequence '(c/a)CG(c/g)(c/g)g(c/a)gC(t/g)' is nearly specific to the ABA-responsive genes of rice (Table 2). These trends indicate that ABA-regulated genes in rice and Arabidopsis have evolved distinct combinatorial cis-element modules.

The drought, low temperature and salt-inducible expression of *cor/rd* genes in Arabidopsis requires a class of element defined by the core motif '(a/g)CCGAC' named as *Dehydration Responsive Element/C-repeat* (*DRE/CRT*) (Yamaguchi-Shinozaki and Shinozaki, 1994; Table 2). This element does not function as *ABRE* but could act as a coupling element for *ABRE* to form a functional *ABRC* (Hobo et al., 1999; Yazaki et al., 2004). The trans-activators *DREB1/CBF* and *DREB2* that bind to the *DRE/CRT* (i.e., *DREB* for **DRE**-Binding factor or *CBF* for **C**-repeat Binding Factor) are members of the AP2/ERF family of transcription factors that are highly conserved between the dicot and monocot groups of plants (Stockinger et al., 1997; Liu et al., 1998; Stockinger et al., 2001). Constitutive expression of *DREB1/CBF* and *DREB2* have led to the coordinated activation of target *cor/rd* genes, and subsequently increased stress tolerance (Jaglo-Ottosen et al., 1998; Kasuga et al., 1999; Gilmour et al., 2000). Overexpression studies estimated that more >300 genes encoded by the Arabidopsis genome are regulated through the *DREB/CBF* pathway with >100 genes confirmed as direct targets (Seki et al., 2001; Fowler and Thomashow, 2002).

The *DREB/CBF* regulon is highly conserved across the temperate and tropical groups of flowering plants as evidenced by the occurrence of functional *DREB1/CBF* and *DREB2* orthologs in both monocots (including rice) and dicots (Jaglo et al., 2001; Choi et al., 2002; Dubouzet et al., 2003; Nakashima et al., 2009; Yun et al., 2010; Todaka et al., 2012). Genome-wide survey of the enrichment of '(a/g)CCGAC' core motifs with correct spatial distribution revealed that *DRE/CRT* are as ubiquitous among the stress-regulated genes encoded by the rice genome as *ABRE*, i.e., significantly enriched in >10% of total rice genes, similar to what has been observed in Arabidopsis (Rabbani et al., 2003; Fowler and Thomashow, 2002; Yazaki et al., 2004; Vogel et al., 2005; Cheng et al., 2007;

Lenka et al., 2009). For example, *DRE/CRT* is the dominant class of cis-elements that regulates a late embryogenesis abundant protein (*OsLEA*) network involved in dehydration, low temperature and salinity stress mechanisms in japonica rice (Meier et al., 2008).

The promoter of Arabidopsis *CBF/DREB* (e.g., *CBF3*) is characterized by a highly conserved 125-bp module as binding sites (ICE Boxes I, II, III, IV, V) for a Myc-like bHLH transcription factor *ICE1* (Shinwari et al., 1998; Chinnusamy et al., 2003; Zarka et al., 2003). *ICE1* is regulated by phosphorylation and its activated form binds to the *c-Myc* target element ('CANNTG') located in the highly conserved 125-bp region of *CBF* promoters. Box IV within the critical 125-bp region of *CBF2* promoter contains the motif 'CACATG'. Similar survey of upstream regions of all members of the orthologous *DREB/CBF* gene families in rice and sorghum revealed common signatures that include cis-elements like Ca/Calmodulin element ('VCGCGB'), *ABRE* ('CGTGG', 'MACGYGB'), jasmonic acid signaling ('CGTCA', 'TGACG') and MYC ('CANNTG') (Srivastav et al., 2010).

While the *DREB/CBF* network appears to play the dominant role low temperature (4°C and below) and drought response mechanisms in rice, several recent studies revealed other regulons with likely overlap with biotic stress response networks through oxidative signaling (Table 2). An example is the NAC regulon controlled by the NAC-type (*NAM*, *ATAF*, *CUC*) transcription factors *OsNAC6/SNAC2* and *OsNAC5*, with target cis-elements having the core motif 'CACG' (Nakashima et al., 2007; Hu et al., 2008; Nakashima et al., 2009). Another example is the network of genes regulated by bZIP transcription factor *OsTGA10* (Os06g41100) and R2R3-type Myb transcription factor *OsMyb4* (Os04g43680), which is activated primarily in by oxidative signal stress during exposure to chilling and configures various components of oxidative defenses (Jakoby et al., 2002; Cheng et al., 2007; Park et al., 2010b; Yun et al., 2010; Xu et al., 2012). This network is comprised of a large number of target genes defined by several combinations of *as1/ocs/TGA*-like elements ('AATTTGAT', 'TAATTTGA') and *GARE/Pyr-box/Myb2-box*-like elements ('AAAGAAAA', 'TAGTTTTT').

#### **Biotic stress and ethylene, salicylic acid and jasmonic acid signaling**

Ethylene, salicylic acid and jasmonic acid are key components of host plant signaling and defenses against pathogens and insect herbivores (Abeles et al. 1992; Pieterse et al., 1998; Johansson et al., 2006; Dempsey and Klessig, 2012). Ethylene signaling during host-pathogen or host-herbivore interactions lead to changes in gene expression through various types of ERF (Ethylene Response Factor) transcription factors, while the salicylic acid and jasmonic



acid mediated gene expression involves the action of various types of WRKY and bHLH transcription factors, respectively (Helliwell and Yang 2013). For instance, the early responses of rice against the striped stemborer *Chilo suppressalis* is known to involve ethylene, salicylic acid, jasmonic acid and H<sub>2</sub>O<sub>2</sub>, mediating a MAPK-phosphorylation cascade that regulates the expression of several *OsWRKY* and *OsERF3* transcription factors (Lu *et al.*, 2011). Similarly, the early responses of rice to both virulent and avirulent pathovars of *Magnaporthe oryzae* (blast) is mediated by ethylene leading to the upregulation of *OsERF922* transcription factor and its suite of target defense-related genes (Singh *et al.*, 2004; Liu *et al.*, 2012).

Examples of known targets of defense-related ERF transcription factors are the pathogenesis-related proteins *PR3*, *PR4* and *PDF1.2*. A common feature shared by many of these PR-protein genes and other types of ethylene-regulated defense genes is the canonical *GCC-box* cis-element defined by the core motif 'AGCCGCC' (Hart *et al.*, 1993; Ohme-Takagi and Shinshi, 1995; Manners *et al.*, 1998; Suzuki *et al.*, 1998; Table 2). A more recent example of ERF-type transcription factors that directly target defense-related genes in rice is *OsERF922*, which acts on *GCC-box* containing genes to negatively regulate resistance to *Magnaporthe oryzae* (Liu *et al.*, 2012). In Arabidopsis, other types of ERF transcription factors such as *Pti4/5/6* have also been shown to bind to the canonical *GCC-box* as well as other *GCC-box* related core-motifs such as 'TAAGAGCCGCC' (Table 2). These ERF transcription factors and their cognate cis-elements are highly conserved in both dicot and monocot plants including rice (Hao *et al.*, 1998; Gu *et al.*, 2000). Additionally, the conserved '*GCC-box*' has also been shown to function as JA-responsive cis-element involved in both pathogen and herbivore defense mechanisms in Arabidopsis and rice (Brown *et al.*, 2003). Many such ERF-type transcription factors that act on '*GCC-box*' cis-element have been shown to modulate the expression of their target ethylene-responsive and jasmonic acid-responsive genes either positively or negatively, by virtue of their dual activation and repression domains (Rojo *et al.*, 1999; Fujimoto *et al.*, 2000; Ohta *et al.*, 2001; Ohme-Takagi *et al.*, 2000; Bodenhausen and Reymond, 2007).

Recent studies revealed other defense-related ERF transcription factors that regulate ethylene-inducible gene expression through the *P-box* element, defined by the core motif 'CCGCCCTCCA' (Table 2). Initially identified in the promoter of ethylene and pathogen-inducible putrescine-N-methyltransferase-2 (*PMT2*) gene of tobacco, *P-box* element binds various ERF- transcription factors that are not involved in *GCC-box*-mediated gene expression (e.g., *ERF189*, *ERF115*, *ERF179*, *ERF163*, *ORCA3*, *AtERF13*). *P-box*-like motifs have also been shown to be highly enriched among putative ERF-target defense-related genes in rice (Shoji *et al.*, 2010; Shoji and Hashimoto 2012).

Salicylic acid plays a major role in defense mechanisms particularly against biotrophic pathogens that parasitize a living host (Alvarez, 2000; Desveaux *et al.*, 2004; Garcion and Métraux, 2006; Ciolkowski *et al.*, 2008; Vlot *et al.*, 2009). In rice, salicylic acid-mediated gene expression has been shown to involve similar types of cis-elements and transcriptional regulators that have been identified in other dicot plants such as Arabidopsis and tobacco (Silverman *et al.*, 1995; Yang *et al.*, 2004). Primarily, salicylic acid-mediated defense involves WRKY transcription factors acting on the *SA-responsive* element defined by the core motif 'TGACG' and *W-box* element defined by the core motif 'TTGAC(c/t)' (Chen *et al.*, 2002; Li *et al.*, 2004). In rice, several WRKY transcription factors (*OsWRKY45*, *OsWRKY71*, *OsWRKY03*) have been shown to bind to *W-box* of several PR-protein genes such as *OsPR10a* (Liu *et al.* 2005b; Shimono *et al.* 2007; Hwang *et al.* 2008; Table 2).

Many members of WRKY family of transcription factors in rice are transcriptionally activated in response to *M. oryzae*, the causal agent of blast and *Xanthomonas oryzae pv. oryzae*, the causal agent of bacterial leaf blight (Ryu *et al.*, 2006; Bagnaresi *et al.*, 2012).

For example, *OsWRKY45* has been shown to play a dual role in mechanisms of resistance to both blast and bacterial leaf blight, while many other WRKY transcription factors such as *OsWRKY53*, *OsWRKY31*, and *OsWRKY30* have been shown to function specifically as positive regulators of defenses against blast by virtue of their ability to activate a suite of target genes with common features of having either *W-box* and/or *SA-response* elements (Zhang *et al.*, 2008; Peng *et al.*, 2012; Shimono *et al.*, 2012). Some WRKY transcription factors such as *OsWRKY76* function as negative regulators of disease resistance also by virtue of their interaction with the *W-box*. This indicates that the *W-box* functions both as activator and repressor of SA-mediated gene expression in rice (Yokotani *et al.* 2013).

Jasmonic acid is an important regulator of responses to pathogens, insect herbivores and other related factors such as wounding and mechanical injury (Creelman and Mullet, 1995; Farmer *et al.*, 2003; Wasternack, 2007; Koo and Howe, 2009). The primary activators of jasmonic acid-mediated signaling are bHLH-type transcription factors that are conserved between Arabidopsis and rice. The precise identity of the associated activator and repressor cis-elements for jasmonic acid-mediated gene expression are just starting to be revealed, but a number of candidates have recently been suggested. For example, three copies of an *E-box* element defined by the core motif 'CANNTG' has been suggested to function as JA-responsive enhancer within -515 to -265 region of rice chitinase gene *OsChia4a* (Table 2). The *E-box* element has been suggested as the binding site for the bHLH-type *RERJ1* transcription factor (Nishizawa *et al.*, 1999; Miyamoto *et al.*,



2012; Wager and Browse 2012; Santino et al., 2013). Recently, another bHLH-type transcription factor *OsbHLH65* that responds to *Magnaporthe grisea*, *Nilaparvata lugens*, salicylic acid and jasmonic acid directly binds to several potential *E-boxes* with the signature motifs ‘CAGCTG’ and ‘CACCTG’ (Santino et al., 2013).

#### **Cis-element and transcription factor databases**

Systematic curation of cis-elements and their cognate trans-acting factors that define regulated transcription in eukaryotes is central to computational modelling of transcriptional networks (Buchler et al., 2003; Xie et al., 2005; Mohanty et al., 2012). Our recent survey of regulatory sequences of rice genes across the available public databases indicate that rice information has been steadily increasing given the level of global activity in rice functional genomics (Tables 2 and 3). During the last two decades, several

data repositories have been developed to provide an annotated reference of cis-elements and their cognate trans-acting factors established from the results of wet-lab experiments. The first generation information repositories which include *TRANSFAC*, *PLACE*, *PlantCARE* and *AGRIS* served as initial foundation for large-scale computational prediction of the unifying regulatory features among co-expressed genes revealed by genome-wide transcriptome studies (Wingender et al., 1996; Higo et al., 1999; Lescot et al., 2002; Davuluri et al., 2003; Matys et al., 2003). However, while *TRANSFAC* has been utilized for the analysis of rice data, its predictive power for novel plant-specific elements is quite limited. On the other hand, *PLACE*, *PlantCARE* and *AGRIS* have offered relatively higher resolutions for the analysis of rice data, given their stronger emphasis on monocot and dicot plants.

More recent generations of cis-element databases like the Database of Arabidopsis Transcription Factors (*DATF*),

**Table 3 Partial list of recently annotated regulatory sequences from rice based on recent updates in public cis-element and transcription factor databases**

Cis-element	Consensus core motif	Transcriptional regulator	Biological function
<b>AACA element</b>	AACAAAC	<i>OsMYB5</i>	Regulator of endosperm-specific expression
<b>ABRE</b>	GTACGTGGCGC	Unknown	Positive regulator of GA response; Negative regulator of ABA response
	ACGTGKC	Unknown	ABA-response element for ABA regulated gene expression
	GCCGCGTGCC	Unknown	ABA-response element for ABA regulated gene expression
	AGTACGTGGC	Unknown	ABA-response element for ABA regulated gene expression
<b>ABRE/hex-3</b>	GGACGCGTGCC	<i>OsABF1</i>	ABA-response element for ABA regulated gene expression
<b>ABRE-like</b>	TACGTGTC	<i>OsTRAB1</i>	ABA-response element for ABA regulated gene expression
<b>G-box</b>	RTACGTGGCR	<i>OsTAF-1</i>	Positive regulator for response to ABA and desiccation in vegetative tissues
	CTTCCACGTGGCA	<i>OsBZ8</i>	ABA-response element for ABA regulated gene expression
<b>RY repeat</b>	CATGCATG	Unknown	ABA-response element for ABA regulated gene expression
<b>AH1 element</b>	CAATTATTG	<i>OsHox1</i>	Enhancer of provascular-specific expression
<b>ARR1 element</b>	AGATT	<i>OsARR1</i>	Non-symbiotic hemoglobin-2 promoter
<b>as1-like</b>	GCATCTTTACTTTAGCATC	as1-like box binding factor	Phloem-specific expression
<b>Ilb element</b>	TGTGGGACCATG	<i>OsPCF1</i> , <i>OsPCF2</i>	Enhancer of meristem-specific expression
<b>BIHD10S element</b>	TGTCA	<i>OsBIHD1</i>	Pathogen-regulated expression
<b>BP-73 element</b>	AACGT	<i>OsBP-73</i>	Regulation of cell proliferation
<b>E2F element</b>	GCGGGAAA	<i>OsE2F</i>	Enhancer in actively dividing cells
<b>GATA box</b>	GATA	<i>OsASF-2</i>	Light regulated expression
<b>GCC-box</b>	AGCCGCC	<i>OsEREBP1</i>	Pathogen-regulated expression
	CATAAGAGCCGCCACT	<i>OsBIERF3</i>	Response to biotic and abiotic stress
<b>IDE1 element</b>	ATCAAGCATGCTTCTTGC	<i>OsIDE1</i> , <i>OsIDE2</i>	Iron deficiency regulated expression
<b>OsIRO2 element</b>	CACGTGG	<i>OsIRO2</i>	Iron deficiency regulated expression
<b>PR2 element</b>	ACGCTGCCG	<i>OsWOX9</i>	Root meristem-specific expression
<b>TGTCA element</b>	TGTCA	<i>OsBIHD1</i>	Pathogen-regulated expression
<b>W-box</b>	TTGAC	<i>OsWRKY13</i> , <i>OsWRKY51</i> , <i>OsWRKY71</i>	Pathogen-regulated expression

Database of Rice Transcription Factor (*DRTF*), Plant Transcription Factor Database (*Plant TFDB*), and Arabidopsis Stress Responsive Transcription Factor Database (*STIFDB*) have significantly improved the resolution of regulatory sequence prediction in rice genomic sequences given the broader emphasis on plant physiological processes (Guo et al., 2005; Gao et al., 2006; Riaño-Pachón et al., 2007; Shameer et al., 2009; Perez-Rodriguez et al., 2010; Wang et al., 2010; Zhang et al., 2011). *DRTF* in particular holds information on more than 2,000 putative transcription factors and their computationally predicted target cis-elements anchored to the reference genome sequence of rice. Other more recent additions to available databases include the Plant Transcription Factor database (*PlantTFDB*), *GRASSIUS* and *Osiris* (Guo et al., 2008; Morris et al., 2008; Yilmaz et al., 2009; Zhang et al., 2011).

In summary, the number of public databases that curate information about regulatory sequences and transcription factors in plants has increased substantially over the last decade. The information contained in these resources are evolving constantly and their maximum utility for high-resolution computational prediction of transcriptional networks in rice relies on systematic annotation based on experimentally validated data. What is currently necessary for more biologically meaningful analysis of regulatory sequence variation in the genus *Oryza* is a systematic means to integrate and synthesize the data from all resources in order to resolve inconsistencies in annotation, redundancy, and lack of uniform nomenclature.

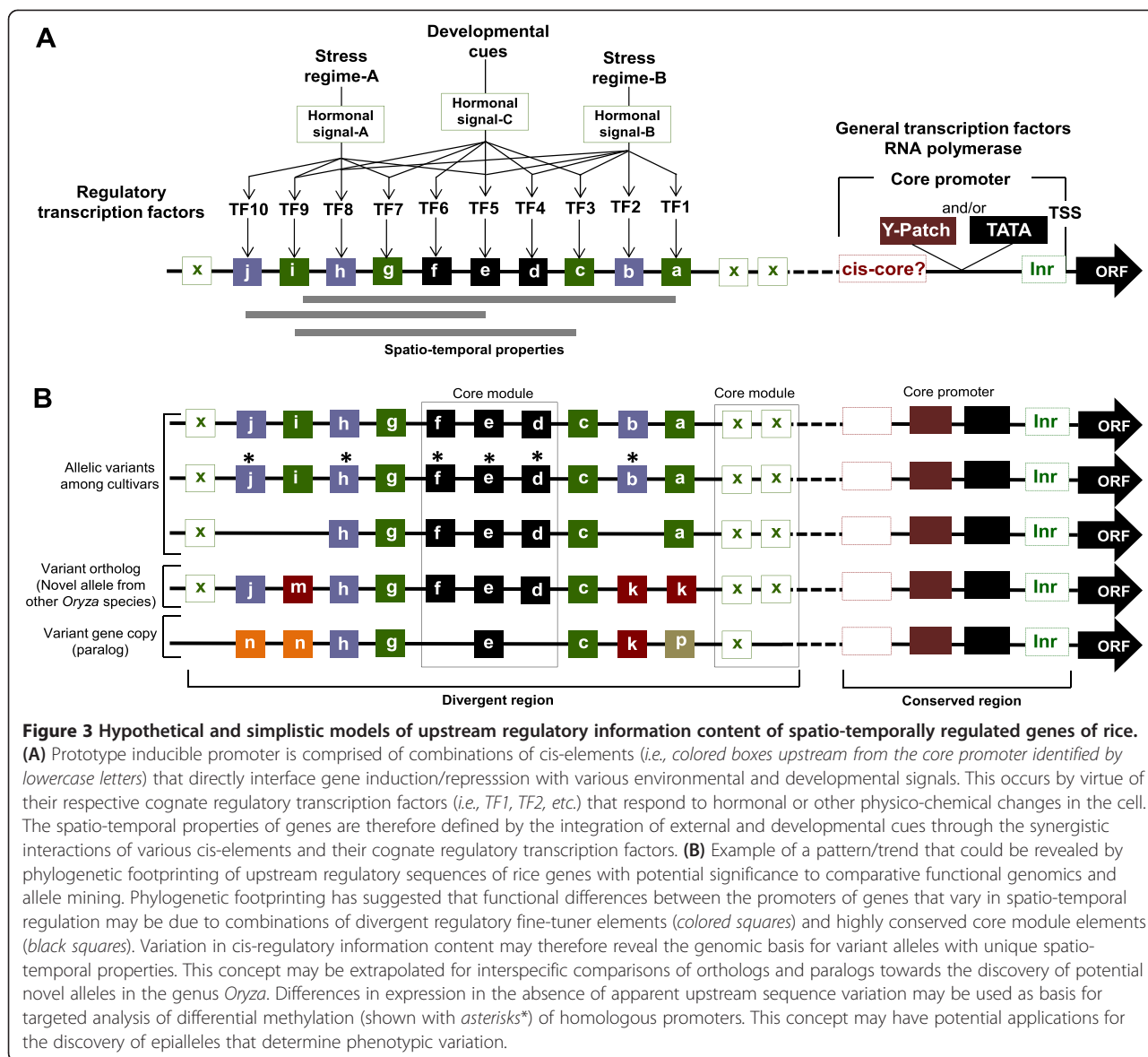
## Conclusions and perspectives

### Applications of upstream sequence variation to genus-wide gene discovery, allele mining and network engineering

The fundamental question about the architecture of rice promoters and the nature of their cognate regulator proteins remains very relevant in the era of comparative functional genomics. Our recent survey of the major public eukaryotic cis-element and transcription factor databases have indicated that sequences and annotation of cis-elements and their cognate transcriptional regulators in the reference rice genome are still far from being comprehensive to guide a systematic scrutiny of regulatory networks at a genome-wide and comparative scales. The 'finished' reference genome sequence of rice has been available to the community now for a decade, serving as nucleator for hypothesis-driven interrogation of gene function and evolution. As we enter the next decade of reference genome-enabled experimentation in rice genetics, which has directed its focus on genotypic comparisons, allele and novel gene mining, the importance of a comprehensive annotation of cis-acting elements and trans-acting factors that define regulatory networks cannot be over-emphasized.

Creating novel combinations of superior alleles by integrative use of conventional and biotechnological methods is the overarching goal of 21<sup>st</sup> century rice breeding. The genome-enabled research paradigm in rice biology has contributed to this goal through the understanding of gene function and regulation. Moreover, at the very core of the reference genome-guided research paradigm is the quest to pinpoint useful allelic variation in the germplasm that explains simple or quantitative traits to guide targeted introgression and selection. Broadly, gene function can be described in the context of a gene's specific molecular, biochemical and/or biological roles in the cell and in context of its role in a broader molecular or biochemical network. The upstream regulatory sequence signature of a gene is a window to the combinatorial complexity of its regulation and reflects how it is interfaced with the various intrinsic and extrinsic signals for growth, development, reproduction and adaptation (Figure 3A). In these contexts, a truly biologically meaningful annotation of the reference genomes of rice that could facilitate a holistic understanding of the nature of allelic and epiallelic variation for phenotypes of interest must include information on the critical cis-elements and trans-acting factors that define the spatio-temporal regulatory properties of each gene. Establishing a modular map of regulatory elements within the upstream regions of each gene locus as part of reference genome annotations in a similar manner that we understand protein domain architectures of coding sequences will be an important tool for various applications in comparative functional genomics both at the intraspecific and interspecific levels.

The specific applications of upstream sequence annotation are numerous and some of the more important concepts are discussed in this review. First, in-depth understanding of the modular architectures of upstream regulatory regions for each annotated gene locus (i.e., open-reading frame) in the reference genome could nucleate direct comparison of orthologs and paralogs by phylogenetic footprinting across a meaningful germplasm diversity panel. Phylogenetic footprinting is based on the assumption that conserved sequence motifs or modular combinations of motifs within the upstream sequences of homologous genes represent functional cis-elements that define the spatio-temporal regulation of that gene (Hardison, 2000; Hong et al., 2003; Zhang and Gerstein, 2003; De Bodt et al., 2006; Freeling and Subramaniam, 2009; Xu et al., 2012a). In other words, patterns of cis-element conservation or divergence revealed by phylogenetic footprints define the integrated effects of various developmental and external signals on the expression potential of a given gene (Figure 3A). An offshoot of that concept is that upstream sequences containing critical regulatory elements can be the subject of targeted scrutiny to reveal meaningful sequence variation both at the intraspecific and interspecific levels for the purpose of uncovering



conserved or divergent patterns associated with expression variability and elite phenotypes (Figure 3B). This is an area in rice comparative functional genomics that has not been actively pursued given the rapidly growing resources for comparative functional genomics for mining novel alleles. As more germplasm resequencing data become available, meaningful comparative analysis of upstream regulatory sequence variation by phylogenetic footprinting must be pursued towards the discovery of novel alleles and associated networks with potential applications in rice breeding.

An example of a seminal study that highlights the potential of phylogenetic footprinting for novel gene discovery and allele mining was a recent comparison of upstream regulatory sequence signatures across orthologous and paralogous groups of stress-associated bZIP transcription factors of rice, Arabidopsis and sorghum, three species

representing more than 140 million years of evolutionary history and unique geo-climatic distribution (Xu et al., 2012a). It was revealed that while orthologs shared very similar basal developmental programming by virtue of highly conserved 'core modules' of cis-elements, they also exhibit unique cis-element footprints hypothesized to act as 'regulatory fine tuners' in conjunction with the 'core modules'. It was inferred that differences in upstream sequence footprints contributed by the 'regulatory fine tuners' may serve to highlight the unique spatio-temporal characteristics of orthologs and paralogs, underscoring the potential uniqueness of each ortholog and paralog in a regulatory context hence expression variability. Indeed, the trends established from such comparison has been supported by similar patterns established based on phylogenetic footprints of upstream non-coding sequences between rice and other

syntenic cereal genomes including maize, sorghum and barley, suggesting that phylogenetic footprints could be a powerful tool for revealing some of the major determinants of novel allelic variation (Guo and Moose, 2003; Wang et al., 2009).

The genus *Oryza* represents about 15 million years of evolutionary divergence that has been established as a 'one-of-its-kind' tractable comparative genomics system. Capitalizing on the 'finished' reference sequence of *O. sativa* ssp. *japonica*, an international consortium is currently developing a genus-wide platform for the discovery of novel genes and for mining novel alleles with potential use in genomics-enabled introgression breeding (Wing et al., 2005; Jacquemin et al., 2012). To this date, a comprehensive set of sequence scaffolds representing all diploid genomes (i.e., AA, BB, CC, EE, FF, and GG) has been developed and in various stage of sequence assembly and completion, with few already released including the African cultivated species *O. glaberrima* (AA) and the wild species *O. barthii* (AA), *O. longistaminata* (AA), *O. nivara* (AA), *O. punctata* (AA), *O. rufipogon* (AA), and *O. brachyantha* (FF) (Rounsley et al., 2009; Xu et al., 2012b; Chen et al., 2013; Jacquemin et al., 2014). Given this rapidly unfolding genomics resources, it would soon be possible to conduct a meaningful global and cross-genomic analysis of orthologous and paralogous upstream sequence and expression variation in *Oryza*. The resulting knowledge may provide an important window to the nature of allelic variation for rice genes that are regulated at the transcriptional level, and could set the stage for targeted gene introgression and allele replacement in the future.

Second, it is important to articulate that the true essence of functional genomics in the context of rice breeding must be deeply rooted to the ability to identify meaningful sequence variation across an allelic series, understanding how such sequence variation contributes to differences in gene function, and whether differences in gene function are due simply to sequence variation (alleles) or differential DNA methylation in the absence of drastic sequence variation (epialleles). More and more examples of homologous genes that vary in expression in the absence of meaningful sequence variation are being reported in the literature (Peng and Zhang, 2009; Hai et al., 2013; Zheng et al., 2013). Understanding the organization and distribution of critical regulatory element modules within the upstream sequences of intraspecific or interspecific homologs has the potential to provide another layer of information for understanding the contribution of 'epiallelic variation' to phenotypic variation. In this context, homologous genes that differ in expression characteristics despite the absence of any meaningful differences in upstream regulatory sequence architecture should justify examination of the contribution of differential upstream DNA methylation in the expression of the desired phenotype (Figure 3B). While the concept of

epiallelic introgression to phenotypic gain is still vague and yet to be explored in rice genetics and breeding, the establishment of reference genome sequences with well annotated upstream regulatory sequences should set the stage for future large-scale association between DNA methylation patterns and upstream sequence variation. Certainly, the rapidly unfolding reference genome sequence resources across the genus *Oryza* will serve as foundation for the discovery of potential epiallelic series that are due to differential upstream sequence methylation.

Third, it is well established that the major transcriptional regulatory networks involved in stress response mechanisms are highly conserved or universal in flowering plants. For instance, the *DREB/CBF* regulon has functioning equivalents in plant species with contrasting sensitivities to low temperature such as rice (intolerant) and *Arabidopsis* (tolerant), or to drought such as rice (intolerant) and sorghum (tolerant). *DREB/CBF* regulons are also equally functional even between any two rice genotypes that differ significantly in regard to relative stress tolerance. What remains unclear is how variation in the activity of the *DREB/CBF* regulatory network contributes to phenotypic gradient for stress tolerance at the intraspecific and interspecific levels. One theory is that regulatory networks like the *DREB/CBF* regulon may contribute to phenotypic gradient by virtue of variation in size and compositional complexity of downstream genes that are controlled by the master regulator. It has been proposed that this may be dependent on the relative distribution of the target cis-elements across individual genomes (Yun et al., 2010; Zhang et al., 2004). Testing this hypothesis in rice will require a systematic survey of genome-wide distribution of *bona-fide* *DREB/CBF* target cis-element modules across a meaningful diversity panel that represents the gradient of stress sensitivity in the germplasm.

The outcome of such type of analysis could also contribute to the understanding of how regulatory networks can be reconfigured by altering the distribution of critical cis-regulatory modules by homologous recombination and genome shuffling during the process of parental intercrossing. This concept could also lead to future exploration of the potential of transcription factor and target gene cluster complementation during introgression. For instance, a novel transcription factor homolog from a donor may be combined with a much more elaborate array of downstream target genes that contain the critical cis-element in the recipient genome. The possible effect could be the further elaboration of the regulon hence novel configuration of the transcriptional network due to such novel combination of master switch and downstream target cis-elements. This may lead to physiological reconfiguration and may be explored as foundation for the creation of novel and/or transgressive phenotypes by breeding.



Finally, one major application of understanding cis-element function that has been explored since the early days of rice biotechnology is the concept of 'designer promoters' (Mehrotra et al., 2011). Given the growing comparative genomics resources in rice, this concept may be revisited with a slightly different set of questions. The classical concept asked the question regarding the feasibility of finely optimizing the spatio-temporal expression of a regulon that defines the expression of a complex phenotype such as drought tolerance by driving its major regulatory gene (i.e., regulatory hub) under the control of a novel and ideal chimeric promoter. Conversely in today's context, given the emerging technology for gene editing and replacement, it may be relevant to ask the question on the feasibility of using our knowledge of cis-element architecture to alter the distribution of critical cis-elements in the genome to create a more elaborate transcriptional regulatory network and novel phenotypes. These are some of the forward-looking questions in rice regulon engineering that would require in-depth understanding of the nature and distribution cis-elements and their cognate regulators in a genome-wide scale. The rapid progress in rice genetics facilitated by reference sequence-enabled comparative genomics in *Oryza* should nucleate future studies addressing some of these concepts.

#### Abbreviations

TSS: Transcription start site; TBP: TATA-binding protein; BRE: TFIIB recognition element; DPE: Downstream promoter element; IME: Intron-mediated enhancement; IAA: Indole 3-acetic acid or auxin; ABA: Abscisic acid; GA: Gibberellic acid; SRC: Sugar response complex; JA: Jasmonic acid; SA: Salicylic acid.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

BGDR was responsible for the overall concept, data integration and synthesis of data, and wrote the manuscript. BM and DYL contributed information and to the writing of the sections on biotic stress and databases. SJY and MRP contributed information and to the writing of the sections on constitutive promoters. All authors read and approved the final manuscript.

#### Acknowledgement

This study was supported by a grant to BGDR from the USDA-National Research Initiative (2006-35604-1669) and base funding from the Maine Agricultural and Forest Experiment Station (Publication 3392). BM and DYL were supported by the National Research Foundation of Singapore (NRF-CRP5-2009-03) and Next-Generation BioGreen 21 Program (PJ009520), Rural Development Administration, Republic of Korea. MRP was supported by Korea Research Foundation Postdoctoral Fellowship (KRF-2006-352-F00002). SJY was supported by a grant from the Regional Subgenebank Support Program of Rural Development Administration, Republic of Korea.

#### Author details

<sup>1</sup>School of Biology and Ecology, University of Maine, Orono, ME 04469, USA. <sup>2</sup>Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117576, Singapore. <sup>3</sup>Department of Crop Science and Institute of Agricultural Science and Technology, Chonbuk National University, Chonju 561-756, Korea.

Received: 2 November 2014 Accepted: 12 January 2015

Published online: 28 February 2015

#### References

- Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of Arabidopsis MYC and MYB homologs in drought and abscisic acid regulated gene expression. *Plant Cell* 9:1859–1868
- Abeles FB, Morgan PW, Saltvet ME (1992) Ethylene in Plant Biology, 2nd edn. CA, Academic Press Inc., San Diego, 414 pp
- Alvarez ME (2000) Salicylic acid in the machinery of hypersensitive cell death and disease resistance. *Plant Mol Biol* 44:429–442
- Bagnaresi P, Biselli C, Orrù L, Urso S, Crispino L, Abbruscato P, Piffanelli P, Lupotto E, Cattivelli L, Valè G (2012) Comparative transcriptome profiling of the early response to *Magnaporthe oryzae* in durable resistant vs susceptible rice (*Oryza sativa* L.) genotypes. *PLoS ONE* 7(12):e51609
- Baumann K, De Paolis A, Constantino P, Gualberti G (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the *rolB* oncogene in plants. *Plant Cell* 11:323–334
- Baranowskij N, Frohberg C, Prat S, Willmitzer L (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J* 13:5383–5392
- Bhattacharyya J, Chowdhury AH, Ray S, Jha JK, Das S, Gayen S, Chakraborty A, Mitra J, Maiti MK, Basu A, Sen SK (2012) Native polyubiquitin promoter of rice provides increased constitutive expression in stable transgenic rice plants. *Plant Cell Rep* 31:271–279
- Bodenhausen N, Reymond P (2007) Signaling pathways controlling induced resistance to insect herbivores in Arabidopsis. *Mol Plant-Micro Interact* 20:1406–1420
- De Bodt S, Thiessen G, Van de Peer Y (2006) Promoter analysis of MADS-box genes in eudicots through phylogenetic footprinting. *Mol Biol Evol* 23:1293–1303
- Bradnam KR, Korf I (2008) Longer first introns are a general property of eukaryotic gene structure. *PLoSOne* 3(8):e3093
- Brown RL, Kazan K, McGrath KC, Maclean DJ, Manners JM (2003) A role for the GCC-box in jasmonate-mediated activation of the *PDF1.2* gene of Arabidopsis. *Plant Physiol* 132:1020–1032
- Buchanan-Wolaston V, Earl S, Harrison E, Mathas E, Navabpour S, Page T, Pink D (2003) The molecular analysis of leaf senescence- A genomics approach. *Plant Biotech J* 1:3–22
- Buchanan-Wollaston V (1997) The molecular biology of leaf senescence. *J Exptl Bot* 48:181–199
- Buchler NE, Gerland U, Hwa T (2003) On schemes of combinatorial transcription logic. *Proc Natl Acad Sci U S A* 100:5136–5141
- Burke TW, Kadonaga JT (1997) The downstream core promoter element DPE is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes Dev* 11:3020–3031
- Butler JEF, Kadonaga JT (2001) Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev* 15:2515–2519
- Cai M, Wei J, Li X, Caiguo X, Wang S (2007) A rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. *Plant Biotech J* 5:664–674
- Chandlee JM (2001) Current molecular understanding of the genetically programmed process of leaf senescence. *Physiol Plant* 113:1–8
- Chapman EJ, Estelle M (2009) Mechanism of auxin-regulated gene expression in plants. *Annu Rev Genet* 43:265–285
- Chen J, Huang Q, Gao D, Wang J, Lang Y, Liu T, Li B, Bai Z, Goicoechea JL, Liang C, Chen C, Zhang W, Sun S, Liao Y, Zhang X, Yang L, Song C, Wang M, Shi J, Liu G, Liu J, Zhou H, Zhou W, Yu Q, An N, Chen Y, Cai Q, Wang B, Liu B, Min J, et al (2013) Whole-genome sequencing of *Oryza brachyantha* reveals mechanisms underlying *Oryza* genome evolution. *Nature Comm* 4:1595
- Chen PW, Lu CA, Yu TS, Tseng TH, Wang CS, Yu SM (2002) Rice  $\alpha$ -amylase transcriptional enhancers direct multiple mode regulation of promoters in transgenic rice. *J Biol Chem* 277:13641–13649
- Chen PW, Chiang CM, Tseng TH, Yu SM (2006) Interaction between rice *MYBGA* and the gibberellin response element controls tissue-specific sugar sensitivity of  $\alpha$ -amylase genes. *Plant Cell* 18:2326–2340
- Cheng C, Yun KY, Ransom H, Mohanty B, Bajic VB, Jia Y, Yun SJ, De los Reyes BG (2007) An early response regulatory cluster induced by low temperature and hydrogen peroxide in seedlings of chilling-tolerant japonica rice. *BMC Genomics* 8:e175
- Chinnusamy V, Ohta M, Kanrar S, Lee B, Hong X, Agarwal M, Zhu JK (2003) *ICE1*: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes Dev* 17:1043–1054
- Choi DW, Rodriguez EM, Close TM (2002) Barley *Cbf3* gene identification, expression pattern and map location. *Plant Physiol* 129:1781–1787

- Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high levels of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5:213–218
- Ciolkowski I, Wanke D, Birkenbihl RP, Somssich IE (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol Biol* 68:81–92
- Clancy M, Vasil V, Hannah LC, Vasil IK (1994) Maize *shrunken-1* intron and exon regions increase gene expression in maize protoplasts. *Plant Sci* 98:151–161
- Clivan P, Svec M (2009) Genome-wide analysis of rice (*O. sativa* ssp. japonica) TATA box and Y-patch promoter elements. *Genome* 52:294–297
- Creelman RA, Mullet JE (1995) Jasmonic acid distribution and action in plants: Regulation during development and response to biotic and abiotic stress. *Proc Natl Acad Sci U S A* 92:4114–4119
- Croissant-Sych Y, Okita T (1996) Identification of positive and negative regulatory cis-elements of the rice glutelin *Gt3* promoter. *Plant Sci* 116:27–35
- Darmasiri N, Darmasiri S, Estelle M (2005) The F-box protein TIR1 is an auxin receptor. *Nature* 435:441–445
- Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E (2003) AGRIS: Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinform* 4:e25
- Dempsey DA, Klessig DF (2012) SOS - too many signals for systemic acquired resistance? *Trends Plant Sci* 17:538–545
- Desveaux D, Subramaniam R, Despres C (2004) A 'Whirly' transcription factor is required for salicylic acid-dependent disease resistance in *Arabidopsis*. *Dev Cell* 6:229–240
- Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki Y (2003) *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high salt-, and cold-responsive gene expression. *Plant J* 33:751–763
- Farmer EE, Alm eras E, Krishnamurthy V (2003) Jasmonates and related oxylipins in plant responses to pathogenesis and herbivory. *Curr Opin Plant Biol* 6:372–378
- Finkelstein R (2006) Studies of abscisic acid perception finally flower. *Plant Cell* 18:786–791
- Fleet CM, Sun T (2005) A DELLAcate balance: The role of gibberellins in plant morphogenesis. *Curr Opin Plant Biol* 8:77–85
- Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the *CBF* cold response pathway. *Plant Cell* 14:1675–1690
- Freeling M, Subramaniam S (2009) Conserved non-coding sequences (CNSs) in higher plants. *Curr Opin Plant Biol* 12:126–132
- Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M (2000) Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* 12:393–404
- Gan S, Amasino RM (1997) Making sense of senescence. *Plant Physiol* 133:313–319
- Gao G, Zhong Y, Guo A, Zhu Q, Tang W, Zheng W, Gu X, Wei L, Luo J (2006) DRTF: A database of rice transcription factors. *Bioinformatics* 22(10):1286–1287
- Garcion C, M traux JP (2006) Salicylic acid. *Plant Hormone Signal* 24:229–255
- Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the *Arabidopsis CBF3* transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol* 124:1854–1865
- Giraudat J, Parcy F, Bertandre N, Gosti F, Leung J, Morris PC, Bouvier-Durand M, Vartanian N (1994) Current advances in abscisic acid action and signaling. *Plant Mol Biol* 26:1557–1577
- Gomez-Porras JL, Riano-Pachon DM, Dreyer J, Mayer JE, Mueller-Roeber B (2007) Genome-wide analysis of ABA-responsive elements ABRE and CE3 reveals divergent patterns in Arabidopsis and rice. *BMC Genomics* 8:260
- Gomi K, Sasaki A, Itoh H, Ueguchi-Tanaka M, Ashikari M, Kitano H, Matsuoka M (2004) GID2, an F-box subunit of the SCF E3 complex, specifically interacts with phosphorylated SLR1 protein and regulates the gibberellin-dependent degradation of SLR1 in rice. *Plant J* 37:626–634
- Grierson C, Du JS, De Torres ZM, Beggs K, Smith C, Holdsworth M, Bevan M (1994) Separate cis sequences and trans factors direct metabolic and developmental regulation of a potato tuber storage protein gene. *Plant J* 5:815–826
- Griffiths J, Murase K, Rieu I, Zentella R, Zhang ZL, Powers SJ, Gong F, Philips AL, Hedden P, Sun TP, Thomas SG (2006) Genetic characterization and functional analysis of the GID1 gibberellin receptors in Arabidopsis. *Plant Cell* 18:3399–3414
- Gu Y, Yang C, Thara YK, Zhou J, Martin GB (2000) *Pti4* is induced by ethylene and salicylic acid, and its product is phosphorylated by the *Pto* kinase. *Plant Cell* 12:771–786
- Guedes Correa LG, Riano-Pachon DM, Schrago CG, Vincenzini dos Santos R, Mueller-Roeber B, Vincents M (2008) The role of bZIP transcription factors in green plants evolution: Adaptive features emerging from four founder genes. *PLoSOne* 8:e2994
- Guiltinan MJ, Marcotte WR Jr, Quatrano RS (1990) A plant leucine zipper protein that recognizes and abscisic acid response element. *Science* 250:267–271
- Guo AY, Chen X, Gao G, Zhang H, Zhu QH, Liu XC, Zhong YF, Gu X, He K, Luo J (2008) PlantTFDB: a comprehensive plant transcription factor database. *Nucleic Acids Res* 36:D966–D969
- Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: A Database of Arabidopsis Transcription Factors. *Bioinformatics* 21:2568–2569
- Guo H, Moose SP (2003) Conserved noncoding sequences among cultivated cereal genomes identify candidate regulatory sequence elements and patterns of promoter evolution. *Plant Cell* 15:1143–1158
- Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: Genes, promoters and regulatory factors. *Plant Mol Biol* 49:373–385
- Hai P, Guanghai J, Jing Z, Weixiong Z, WenXue Z (2013) DNA methylation polymorphism and stability in Chinese indica hybrid rice. *Sci China Life Sci* 56:1097–1106
- Hamada K, Hongo K, Suwabe K, Shimizu A, Nagayama T, Abe R, Kikuchi S, Yamamoto N, Fujii T, Yokoyama K, Tsuchida H, Sano K, Mochizuki T, Oki N, Horiuchi Y, Fujita M, Watanabe M, Matsuoka M, Kurata N, Yano K (2011) OryzaExpress: An integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol* 52:220–229
- Hao D, Ohme-Takagi M, Sarai A (1998) Unique mode of GCC-box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *J Biol Chem* 273:26857–26861
- Hardison RC (2000) Conserved noncoding sequences are reliable guides to regulatory elements. *Trends Genet* 16:369–372
- Hart CM, Nagy F, Meins F Jr (1993) A 61-bp enhancer element of the tobacco beta-1,3-glucanase B gene interacts with one or more regulated nuclear proteins. *Plant Mol Biol* 21:121–131
- Hartweck LM (2008) Gibberellin signaling. *Planta* 229:1–13
- Hartweck LM, Olszewski NE (2006) Rice gibberellin insensitive dwarf1 is a gibberellin receptor that illuminates and raises questions about GA signaling. *Plant Cell* 18:278–282
- He C, Lin Z, McElroy D, Wu R (2009) Identification of a rice *Actin2* gene regulatory region for high-level expression of transgenes in monocots. *Plant Biotech J* 7:227–239
- Helliwell EE, Yang Y (2013) Molecular strategies to improve rice disease resistance. *Methods Molec Biol* 956:285–309
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nucl Acids Res* 27:297–300
- Hobo T, Koyama Y, Hatori Y (1999) A bZIP factor TRAB1 interacts with VP1 and mediates abscisic acid induced transcription. *Proc Natl Acad Sci USA* 96:15348–15353
- Hong RL, Hamaguchi L, Busch MA, Weigel D (2003) Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *Plant Cell* 15:1296–1309
- Hu H, You J, Fang Y, Zhu X, Xiong L (2008) Characterization of transcription factor gene *SNAC2* conferring cold and salt tolerance in rice. *Plant Mol Biol* 67:169–181
- Hwang SH, Lee IA, Yie SW, Hwang DJ (2008) Identification of an *OsPR10a* promoter region responsive to salicylic acid. *Planta* 227:1141–1150
- Inukai Y, Sakamoto T, Ueguchi-Tanaka M, Shibata Y, Gomi K, Umemura I, Hasegawa Y, Ashikari M, Kitano H, Matsuoka M (2005) Crown rootless1, which is essential for crown root formation in rice, is a target of an auxin response factor in auxin signaling. *Plant Cell* 17:1387–1396
- Ikeda A, Ueguchi-Tanaka M, Sonoda Y, Kitano H, Koshioka M, Futshuhara Y, Matsuoka M, Yamaguchi J (2001) Slender rice, a constitutive gibberellins response mutant is caused by a null mutation in the *SLR1* gene, an orthologue of the height-regulating gene *GAI/RGA/RHT/D8*. *Plant Cell* 13:999–1010
- Ishiguro S, Nakamura K (1994) Characterization of a cDNA encoding a novel DNA-binding protein SPF1 that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol Gen Genet* 244:563–571
- Itoh H, Sasaki A, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Hasegawa Y, Minami E, Ashikari M, Matsuoka M (2005) Dissection of the phosphorylation of rice DELLA protein SLENDER RICE1. *Plant Cell Physiol* 46:1392–1399
- Jacquemin J, Bhatia D, Singh K, Wing RA (2012) The International *Oryza* Map Alignment Project: Development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. *Curr Opin Plant Biol* 16:147–156

- Jacquemin J, Ammiraju SS, Haberer G, Billheimer DD, Yu Y, Liu LC, Rivera LF, Mayer K, Chen M, Wing RA (2014) Fifteen million years of evolution in the *Oryza* genus shows extensive gene family expansion. *Molec Plant* 7(4):642–656
- Jaglo KR, Kleff S, Amundsen KL, Zhang X, Haake V, Zhang JZ, Deits T, Thomashow MF (2001) Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiol* 127:910–917
- Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF (1998) *Arabidopsis* CBF1 overexpression induces *COR* genes and enhances freezing tolerance. *Science* 280:104–106
- Johansson A, Staal J, Dixelius C (2006) Early responses in the *Arabidopsis-Verticillium longisporum* pathosystem are dependent on NDR1, JA- and ET-associated signaling via cytosolic NPR1 and RFO1. *Mol. Plant Microbe Interact* 19:958–969
- Jain M, Kaur N, Tyagi AK, Khurana JP (2006a) The auxin-responsive *GH3* gene family in rice (*Oryza sativa*). *Funct Integr Genomics* 6:36–46
- Jain M, Tyagi AK, Khurana JP (2006b) Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive *SAUR* gene family in rice (*Oryza sativa*). *Genomics* 88:360–371
- Jain M, Kaur N, Garg R, Thakur JK, Tyagi AK, Khurana JP (2006c) Structure and expression analysis of early auxin-responsive *Aux/IAA* gene family in rice (*Oryza sativa*). *Funct Integr Genomics* 6:47–59
- Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F (2002) bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci* 7:106–111
- Jang IC, Choi WB, Lee KH, Song SJ, Nahm BH, Kim JK (2002) High-level and ubiquitous expression of the rice cytochrome c gene *OsCc1* and its promoter activity in transgenic plants provides a useful promoter for transgenesis of monocots. *Plant Physiol* 129:1473–1481
- Jeong YM, Mun JH, Lee I, Woo JC, Hong CB, Kim SG (2006) Distinct roles of the first introns on the expression of *Arabidopsis* profilin gene family members. *Plant Physiol* 140:196–209
- Juven-Gershon T, Kadonaga JT (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol* 339:225–229
- Kaneko M, Itoh H, Euguchi-Tanaka M, Ashikari M, Matsuoka M (2002) The  $\alpha$ -amylase induction in endosperm during rice seed germination is caused by gibberellin synthesized in epithelium. *Plant Physiol* 128:1264–1270
- Kasuga M, Kiu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotech* 17:287–291
- Kaufmann K, Pajoro A, Angenent GC (2010) Regulation of transcription in plants: Mechanisms controlling developmental switches. *Nature Rev Genet* 11:830–842
- Kieffer M, Neve J, Kepinski S (2010) Defining auxin response contexts in plant development. *Cur Opin Plant Biol* 13:12020
- Kim SY, Wu R (1990) Multiple protein factors bind to a rice glutelin promoter region. *Nucl Acids Res* 18:6845–6852
- Kitomi Y, Ito H, Hobo T, Aya K, Kitano H, Inikai Y (2011) The auxin responsive AP2/ERF transcription factor CROWN ROOTLESS5 is involved in crown root initiation in rice through the induction of *OsRR1*, a type-A response regulator of cytokinin signaling. *Plant J* 67:472–484
- Koo AJK, Howe GA (2009) The wound hormone jasmonate. *Phytochem* 70:1571–1580
- Lenhard B, Sandelin A, Carninci P (2012) Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nature Rev Genet* 13:233–245
- Lenka SK, Lohia B, Kumar A, Chinnusamy V, Bansal KC (2009) Genome-wide targeted prediction of ABA responsive genes in rice based on over-represented cis-motif in co-expressed genes. *Plant Molec Biol* 69:261–271
- Lescot M, Déhais P, Thijs G, Marchal M, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* 30:325–327
- Lim PO, Kim HJ, Nam HG (2007) Leaf senescence. *Ann Rev Plant Biol* 58:115–136
- Li J, Brader G, Palva ET (2004) The *WRKY70* transcription factor: a node of convergence for jasmonate mediated and salicylate-mediated signals in plant defense. *Plant Cell* 16:319–331
- Liu H, Wang S, Yu X, Yu J, He X, Zhang S, Wu P (2005a) ARL1, a LOB domain protein required for adventitious root formation in rice. *Plant J* 43:47–56
- Liu XQ, Bai XQ, Qian Q, Wang XJ, Chen MS, Chu CC (2005b) *OsWRKY03*, a rice transcriptional activator that functions in defense signaling pathway upstream of *OsNPR1*. *Cell Res* 15:593–603
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, *DREB1* and *DREB2*, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low temperature-responsive gene expression, respectively in *Arabidopsis*. *Plant Cell* 10:1391–1406
- Liu L, Zhou Y, Szczesna MW, Li S, Lin Y (2010) Identification and application of a rice senescence-associated promoter. *Plant Physiol* 153:1239–1249
- Liu D, Chen X, Liu J, Ye J, Guo Z (2012) The rice ERF transcription factor *OsERF922* negatively regulates resistance to *Magnaporthe oryzae* and salt tolerance. *J Exp Bot* 63:3899–3911
- Lovegrove A, Hooley R (2000) Gibberellin and abscisic acid signaling in aleurone. *Trends Plant Sci* 5:102–110
- Lu J, Ju HP, Zhou GX, Zhu CS, Erb M, Wang XP, Wang P, Lou YG (2011) An EAR-motif-containing ERF transcription factor affects herbivore-induced signaling, defense and resistance in rice. *Plant J* 68:583–596
- Lu CA, Lim EK, Yu SM (1998) Sugar response sequence in the promoter of a rice  $\alpha$ -amylase gene serves as a transcriptional enhancer. *J Biol Chem* 273:10120–10131
- Manners JM, Penninckx IAMA, Vermaere K, Kazan K, Brown RL, Morgan A, Maclean DJ, Curtis MD, Cammue BPA, Broekaert WF (1998) The promoter of the plant defensin gene *PDF1.2* from *Arabidopsis* is systemically activated by fungal pathogens and responds to methyl jasmonate but not to salicylic acid. *Plant Mol Biol* 38:1071–1080
- Matys V, Fricke E, Geffers R, Göbling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucl Acids Res* 31:374–378
- McElroy D, Blowers AD, Jeness B, Wu R (1991) Construction of expression vectors based on the rice actin1 (*Act1*) 5' region for use in monocot transformation. *Mol Gen Genet* 231:150–160
- McElroy D, Zhang W, Wu R (1990) Isolation of an efficient actin promoter for in rice transformation. *Plant Cell* 2:163–171
- Mehrotra R, Gupta G, Sethi R, Bhalothia P, Kumar N, Mehrotra S (2011) Designer promoter: An artwork of cis engineering. *Plant Mol Biol* 75:527–536
- Meier S, Gehring C, MacPherson CR, Kaur M, Maqungo M, Reuben S, Mayunga S, Shih MD, Wei FJ, Wanchana S, Mauleon R, Radovanovic A, Bruskiwicz R, Tanaka T, Mohanty B, Itoh T, Wing R, Gojobori T, Sasaki T, Swarup S, Hsing Y, Bajic VB (2008) The promoter signatures in rice LEA genes can be used to build a co-expressing LEA gene network. *Rice* 1:177–187
- Mejia-Guerra MK, Pomeranz M, Morohashi K, Grotewold E (2012) From plant gene regulatory grids to network dynamics. *Biochim Biophys Acta* 1819:454–465
- Miyamoto K, Shimizu T, Lin F, Sainsbury F, Thuenemann E, Lomonosoff G, Nojiri H, Yamane H, Okada K (2012) Identification of an E-box motif responsible for the expression of jasmonic acid-induced chitinase gene *OsChia4a* in rice. *J Plant Physiol* 169:621–627
- Moabbi AM, Agarwal N, Kaderi BE, Ansari A (2011) Role for gene looping in intron-mediated enhancement of transcription. *Proc Natl Acad Sci U S A* 109:8505–8510
- Mohanty B, Heath V, Wijaya E, Yeo HC, De los Reyes BG, Lee DY (2012) Patterns of cis-element enrichment revealed potential regulatory modules involved in the transcriptional regulation of anoxia response in japonica rice. *Gene* 511:235–242
- Molina C, Grotewold E (2005) Genome-wide analysis of *Arabidopsis* core promoters. *BMC Genomics* 6:25e
- Morris RT, O'Connor TR, Wyrick JJ (2008) Osiris: an integrated promoter database for *Oryza sativa* L. *Bioinformatics* 24:2915–2917
- Nakashima K, Tran LSP, Van Nguyen D, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K (2007) Functional analysis of a NAC-type transcription factor *OsNAC6* involved in abiotic and biotic stress-responsive gene expression in rice. *Plant J* 51:617–630
- Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol* 149:88–95
- Nam HG (1997) The molecular genetic analysis of leaf senescence. *Cur Opin Biotech* 8:200–207
- Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis rd29A* gene in response to dehydration and high-salinity stresses. *Plant J* 34:137–144
- Nemhauser JL, Mockler TC, Chory J (2004) Interdependency of brassinosteroid and auxin signaling in *Arabidopsis*. *PLoS Biol* 2:e258
- Nishizawa Y, Kawakami A, Hibi T, He DY, Shibuya N, Minami E (1999) Regulation of the chitinase gene expression in suspension cultured rice cells by N-acetylchitooligosaccharides: differences in the signal transduction pathways leading to the activation of elicitor-responsive genes. *Plant Mol Biol* 39:907–914
- Noh YS, Amasino RM (1999) Identification of a promoter region responsible for the senescence-specific expression of *SAG12*. *Plant Molec Biol* 41:181–194

- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. *Plant Cell* 15:1591–1604
- Ohme-Takagi M, Shinshi H (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* 7:173–182
- Ohme-Takagi M, Suzuki K, Shinshi H (2000) Regulation of ethylene-induced transcription of defence genes. *Plant and Cell Physiol* 41:1187–1192
- Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M (2001) Repression domains of class II ERF transcriptional repressor share an essential motif for active repression. *Plant Cell* 13:1959–1968
- Ohyanaagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, Habara T, Fujii Y, Antonio BA, Nagamura Y, Imanishi T, Ikeo K, Itoh T, Gajobori T, Sasaki T (2006) The rice annotation project (RAP-DB) Hub for *Oryza sativa ssp. japonica* genome information. *Nucl Acids Res* 34:D741–D744
- Okita TW, Whang YS, Huilo J, Kim WT, Aryan AP, Larsen R, Krishnan HB (1989) Structure and expression of the rice glutelin multigene family. *J Biol Chem* 264:12573–12581
- Parra G, Bradnam K, Rose AB, Korf I (2011) Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucl Acids Res* 39:5328–5337
- Park SH, Yi N, Kim YS, Jeong MH, Bang SW, Choi YD, Kim JK (2010a) Analysis of five novel putative constitutive gene promoters in transgenic rice plants. *J Exptl Bot* 61:2459–2467
- Park MR, Yun KY, Herath V, Mohanty B, Xu F, Bajic VB, Yun SJ, De los Reyes BG (2010b) Supra-optimal expression of the cold-regulated *OsMyb4* transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. *Plant Cell Environ* 33:2209–2230
- Peng H, Zhang J (2009) Plant genomic DNA methylation in response to stresses: Potential applications and challenges in plant breeding. *Prog Nat Sci* 19:1037–1045
- Peng X, Hu Y, Tang X, Zhou P, Deng X, Wang H, Guo Z (2012) Constitutive expression of rice *WRKY30* gene increases the endogenous jasmonic acid accumulation, PR gene expression and resistance to fungal pathogens in rice. *Planta* 236:1485–1498
- Perez-Rodriguez P, Riano-Pachon DM, Correa LG, Rensing SA, Kersten B, Mueller-Roeber B (2010) PlnTFDB: Updated content and new features of the plant transcription factor database. *Nucl Acids Res* 38:D822–D827
- Pieterse CMJ, Van Wees SCM, Van Pelt JA, Knoester M, Laan R, Gerrits H, Weisbeek PJ, Van Loon LC (1998) A novel signalling pathway controlling induced systemic resistance in *Arabidopsis*. *Plant Cell* 10:1571–1580
- Qu LQ, Xing YP, Liu W, Xu XP, Song YR (2008) Expression pattern and activity of six glutelin gene promoters in transgenic rice. *J Exptl Bot* 59:2417–2424
- Quint M, Gray WM (2006) Auxin signaling. *Cur Opin Plant Biol* 9:448–453
- Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol* 133:1755–1767
- Razem FA, El Kereamy A, Abrams SR, Hill RD (2006) The RNA-binding protein FCA is an abscisic acid receptor. *Nature* 439:290–294
- Riño-Pachón DM, Ruzicic S, Dreyer J, Mueller-Roeber B (2007) PlnTFDB: An integrative plant transcription factor database. *BMC Bioinformatics* 8:e42
- Robatzek S, Somssick IE (2002) Targets of *AtWRKY6* regulation during plant senescence and pathogen defense. *Genes Dev* 16:1139–1149
- Rojo E, Leon J, Sanchez-Serrano JJ (1999) Cross-talk between wound signaling pathways determines local versus systemic gene expression in *Arabidopsis thaliana*. *Plant J* 20:135–142
- Rose AB, Emami S, Bradnam K, Korf I (2011) Evidence for a DNA-based mechanism of intron-mediated enhancement. *Front Plant Sci* 2:1–9
- Rounsley S, Reddy Marri P, Yu Y, He R, Sisneros N, Goicoechea JL, Lee SJ, Angelova A, Kudrna D, Luo M, Affourtit J, Desany B, Knight J, Niazi F, Egholm M, Wing RA (2009) *De novo* next generation sequencing of plant genomes. *Rice* 2:35–43
- Ryu HS, Han M, Lee SK, Cho JI, Ryoo N, Heu S, Lee YH, Bhoo SH, Wang GL, Hahn TR, Jeon JS (2006) A comprehensive expression analysis of the *WRKY* gene superfamily in rice plants during defense response. *Plant Cell Rep* 25:836–847
- Santino A, Taurino M, DeDomenico S, Bonsegna S, Poltronieri P, Pastor V, Flors V (2013) Jasmonate signaling in plant development and defense response to multiple (ab)iotic stresses. *Plant Cell Rep* 32:1085–1098
- Sato K, Shimizu T, Kondoh H, Hirguri A, Sasaya T, Choi IR, Omura T, Kikuchi S (2011) Relationship between symptoms and gene expression induced by the infection of three strains of rice dwarf virus. *PLoS One* 6:e18094
- Scarpella E, Rueb S, Boot KJM, Hoge JHC, Meijer AH (2000) A role of the rice homeobox gene *OsHox1* in provascular fate commitment. *Development* 127:3655–3669
- Scarpella E, Simons EJ, Meijer AH (2005) Multiple regulatory elements contribute to the vascular-specific expression of the rice HD-ZIP gene *OsHox1* in Arabidopsis. *Plant Cell Physiol* 46:1400–1410
- Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Caminci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell* 13:61–72
- Shameer K, Ambika S, Varghese SM, Karaba N, Udayakumar M, Sowdhamini R. (2009) STIFDB—Arabidopsis Stress-responsive Transcription Factor DataBase. *Int J Plant Genomics* ID583429: doi:10.1155/2009/583429.
- Shen QJ, Zhang P, Ho TH (1996) Modular nature of abscisic acid (ABA) response complexes: Composite promoter units that are necessary and sufficient for ABA induction of gene expression in barley. *Plant Cell* 8:1107–1119
- Shen QJ, Ho TH (1995) Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing G-box and a novel cis-acting element. *Plant Cell* 7:295–307
- Shimono M, Sugano S, Nakayama A, Jiang CJ, Ono K, Toki S, Takatsuji H (2007) Rice *WRKY45* plays a crucial role in benzothiadiazole-inducible blast resistance. *Plant Cell* 19:2064–2076
- Shimono M, Koga H, Akagi A, Hayashi N, Goto S, Sawada M, Kurihara T, Matsushita A, Sugano S, Jiang CJ, Kaku H, Inoue H, Takatsuji H (2012) Rice *WRKY45* plays important roles in fungal and bacterial disease resistance. *Mol Plant Pathol* 13:83–94
- Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: Differences and cross-talk between two stress signaling pathways. *Curr Opin Plant Biol* 3:217–223
- Shinwari ZK, Nakashima K, Miura S, Masuga M, Seki M, Yamaguchi-Shinozaki K, Shinozaki K (1998) An *Arabidopsis* gene family encoding DRE/CRT binding proteins involved in low temperature responsive gene expression. *Biochem Biophys Res Comm* 250:161–170
- Shoji T, Kajikawa M, Hashimoto T (2010) Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *Plant Cell* 22:3390–3409
- Shoji T, Hashimoto T (2012) DNA-binding and transcriptional activation properties of tobacco NIC2-locus ERF189 and related transcription factors. *Plant Biotechnol* 29:35–42
- Singh MP, Lee FN, Counce PA, Gibbons JH (2004) Mediation of partial resistance to rice blast through anaerobic induction of ethylene. *Phytopathology* 94:819–825
- Silverman P, Seskar M, Kanter D, Schweizer P, Metraux JP, Raskin I (1995) Salicylic acid in rice. Biosynthesis, conjugation, and possible role. *Plant Physiol* 108:633–639
- Smale ST (2001) Core promoters: Active contributors to combinatorial gene regulation. *Genes Dev* 15:2503–2508
- Smale ST, Baltimore D (1989) The 'initiator' as a transcription control element. *Cell* 57:103–113
- Song Y, You J, Xiong L (2009) Characterization of *OsiAA* gene, a member of rice *Aux/IAA* family involved in auxin and brassinosteroid hormone responses and plant morphogenesis. *Plant Mol Biol* 70:297–309
- Srivastav AQ, Mehta S, Lindlof A, Bhargava S (2010) Over-represented promoter motifs in abiotic stress-induced DREB genes of rice and sorghum and their probable role in regulation of gene expression. *Plant Signal Behav* 5:775–784
- Stockinger EJ, Yaopan M, Regier MK, Triezenberg SJ, Thomashow MF (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with *CBF1*, a transcriptional activator involved in cold-regulated gene expression. *Nucl Acids Res* 29:1524–1533
- Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana* *CBF1* encodes an AP2 domain-containing transcription activator element that binds to the C-repeat/DRE, a cis-acting regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci U S A* 94:1035–1040
- Sutoh K, Yamauchi D (2003) Two cis-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds. *Plant J* 34:635–645
- Suzuki K, Suzuki N, Ohme-Takagi M, Shinshi H (1998) Immediate early induction of mRNAs for ethylene-responsive transcription factors in tobacco leaf strips after cutting. *Plant J* 15:657–665
- Takaiwa F, Kikuchi S, Oono K (1987) Nucleotide sequence of a rice glutelin gene. *FEBS Lett* 206:33–35
- Takaiwa F, Oono K, Wing D, Kato A (1991) Sequence of three members and expression of a new major subfamily of glutelin genes from rice. *Plant Molec Biol* 17:875–885
- Thakur JK, Tyagi AK, Khurana JP (2001) *OsiAA1*, an *Aux/IAA* cDNA from rice and changes in its expression as influenced by auxin and light. *DNA res* 8:193–203
- Thomas H, Howart CJ (2000) Five ways to stay green. *J Exptl Bot* 5:329–337



- Thomashow MF (1999) Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Ann Rev Plant Physiol Plant Molec Biol* 50:571–599
- Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. *Plant Cell* 43:265–285
- Todaka D, Nakashima K, Shinozaki K, Yamaguchi-Shinozaki K (2012) Toward understanding transcriptional regulatory networks in abiotic stress responses and tolerance in rice. *Rice* 5:e6
- Ueguchi-Tanaka M, Nakajima M, Katoh E, Ohmiya H, Asano K, Saji S, Hongyu X, Ashikari M, Kitano H, Yamaguchi I, Matsuoka M (2007) Molecular interactions of a soluble gibberellins receptor, GID1 with a rice DELLA protein SLR1 and gibberellins. *Plant Cell* 19:2140–2155
- Ueguchi-Tanaka M, Ashikari M, Nakajima M, Itoh H, Katoh E, Kobayashi M, Chow TY, Hsing YI, Kitano H, Yamaguchi I, Matsuoka M (2005) GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* 437:693–698
- Ulmasov T, Hagen G, Guilfoyle TJ (1997) *ARF1*, a transcription factor that binds to auxin response elements. *Science* 276:1865–1868
- Ulmasov T, Liu ZB, Hagen G, Guilfoyle TJ (1995) Composite structure of auxin response elements. *Plant Cell* 7:1611–1623
- Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993) An *Arabidopsis* myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell* 5:1529–1539
- Vlot AC, Dempsey DMA, Klessig DF (2009) Salicylic acid, a multifaceted hormone to combat disease. *Annu Rev Phytopathol* 47:177–206
- Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF (2005) Roles of *CBF2* and *ZAT12* transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *Plant J* 41:195–211
- Vogel JM, Roth B, Cigan M, Freeling M (1993) Expression of the two maize TATA binding protein genes and functions of the encoded TBP proteins by complementation in yeast. *Plant Cell* 5:1627–1638
- Wager A, Browse J (2012) Social network: JAZ protein interactions expand our knowledge of jasmonate signaling. *Front Plant Sci* 3:1–11
- Waller F, Furuya M, Nick P (2002) *OsARF1*, an auxin response factor from rice is an auxin-regulated and classifies as a primary auxin responsive gene. *Plant Mol Biol* 50:415–425
- Wang Z, Libault M, Joshi T, Valliyodan B, Nguyen HT, Xu D, Stacey G, Cheng J (2010) SoyDB: A knowledge database of soybean transcription factors. *BMC Plant Biol* 10:e14
- Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Han B, Tao Y (2007) Genome-wide analysis of the auxin response factors (*ARF*) gene family in rice (*Oryza sativa*). *Gene* 394:13–24
- Wang X, Haberer G, Mayer KFX (2009) Discovery of cis-elements between sorghum and rice using co-expression and evolutionary conservation. *BMC Genom* 10:284e
- Wang J, Oard JH (2003) Rice ubiquitin promoters: Deletion analysis and potential usefulness in plant transformation systems. *Plant Cell Rep* 22:129–134
- Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F (1999) Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene *GluB-1*. *Plant Molec Biol* 40:1–12
- Wasternack C (2007) Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann Bot* 100:681–697
- Wing RA, Ammiraju SS, Luo M, Kim HR, Yu Y, Kudrna D, Goicoechea JL, Wang W, Nelson W, Rao K, Brar DS, Mackill DJ, Han B, Soderlund C, Stein L, SanMiguel P, Jackson S (2005) The *Oryza* Map Alignment Project: the golden path to unlocking the genetic potential of wild rice species. *Plant Molec Biol* 59:53–62
- Wingender E, Dietze P, Karas H, Knüppel R (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucl Acids Res* 24:238–241
- Xie X, Lu J, Kullbøkas EJ, Golub T, Mootha V, Lindblad-Toh K, Lander ES, Kellis M (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 434:338–345
- Xu F, Park MR, Mohanty B, Herath V, Kitazumi A, Yun SJ, De los Reyes BG (2012a) Cis-regulatory signatures among orthologous groups of stress-associated bZIP transcription factors from rice, sorghum and *Arabidopsis* based on phylogenetic footprints. *BMC Genomics* 13:e497
- Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, Li J, He W, Zhang G, Zheng X, Zhang F, Li Y, Yu C, Kristiansen K, Zhang X, Wang J, Wright M, McCouch S, Nielsen R, Wang J, Wang W (2012b) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotech* 30:105–111
- Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic and cold stress responsive promoters. *Trends Plant Sci* 10:88–94
- Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low temperature or high salt stress. *Plant Cell* 6:251–264
- Yamamoto YY, Ichida H, Matusi M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T (2007) Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* 8:e67
- Yang Y, Qi M, Mei C (2004) Endogenous salicylic acid protects rice plants from oxidative damage caused by aging as well as biotic and abiotic stress. *Plant J* 40:909–919
- Yazaki J, Shimatani Z, Hashimoto A, Nagata Y, Fujii F, Kojima K, Suzuki K, Taya T, Tonouchi M, Nelson C, Nakagawa A, Otomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S (2004) Transcriptional profiling of genes responsive to abscisic acid and gibberellin in rice: phenotyping and comparative analysis between rice and *Arabidopsis*. *Physiol Genomics* 17:87–100
- Yilmaz A, Nishiyama MY Jr, Fuentes BG, Souza GM, Janies D, Gray J, Grotewold E (2009) GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiol* 149:171–180
- Yokotani N, Sato Y, Tanabe S, Chujo T, Shimizu T, Okada K, Yamane H, Shimono M, Sugano S, Takatsuji H, Kaku H, Minami E, Nishizawa Y (2013) *WRKY76* is a rice transcriptional repressor playing opposite roles in blast disease resistance and cold stress tolerance. *J Expt Bot* 64:5085–5097
- Yun KY, Park MR, Mohanty B, Herath V, Xu F, Mauleon R, Wijaya E, Bajic VB, Bruskiewich R, De los Reyes BG (2010) Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. *BMC Plant Biol* 10:e16
- Zarka DG, Vogel JT, Cook D, Thomashow MF (2003) Cold induction of *Arabidopsis* *CBF* genes involves multiple *ICE* (Inducer of *CBF* Expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiol* 133:910–918
- Zhang Z, Gerstein M (2003) Of mice and men: Phylogenetic footprinting aids the discovery of regulatory elements. *BMC J Biol* 2:e11
- Zhang X, Fowler SG, Cheng H, Lou Y, Rhee SY, Stockinger EJ, Thomashow MF (2004) Freezing-sensitive tomato has a functional *CBF* cold response pathway, but a *CBF* regulator that differs from that of freezing-tolerant *Arabidopsis*. *Plant J* 39:905–919
- Zhang H, Jin JP, Tang L, Zhao Y, Gu XC, Gao G, Luo JC (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Res* 39:D1114–D1117
- Zhang J, Peng Y, Guo Z (2008) Constitutive expression of pathogen-inducible *OsWRKY31* enhances disease resistance and affects root growth and auxin response in transgenic rice plants. *Cell Res* 18:508–521
- Zhang W, McElroy D, Wu R (1991) Analysis of rice *Act1 5'* region activity in transgenic rice plants. *Plant Cell* 3:1155–1165
- Zhao Y, Leisy DJ, Okita TW (1994) Tissue-specific expression and temporal regulation of the rice glutelin *Gt3* gene are conferred by at least two spatially separated cis-regulatory elements. *Plant Molec Biol* 25:429–436
- Zheng X, Chen L, Li M, Lou Q, Xia H, Wang P, Li T, Liu H, Luo L (2013) Transgenerational variations in DNA methylation induced by drought stress in two rice varieties with distinguished difference to drought resistance. *PLoSOne* 8:e80253
- Zheng Z, Kawagoe Y, Xiao S, Li Z, Okita T, Hau TL, Lin A, Murai N (1993) 5' distal and proximal cis-acting regulator elements are required for developmental control of a rice seed storage protein glutelin gene. *Plant J* 4:357–366
- Zhu Q, Dabi T, Lamb C (1995) TATA-box and initiator functions in the accurate transcription of a plant minimal promoter in vitro. *Plant Cell* 7:1681–1689
- Zhu Q, Ordiz MI, Dabi T, Beachy RN, Lamb C (2002) Rice TATA binding protein interacts functionally with transcription factor IID and the RF2a bZIP transcriptional activator in an enhanced plant in vitro transcription system. *Plant Cell* 14:795–803
- Zourelidou M, De-Torres Zabala M, Smith C, Bevan MW (2002) Store-keeper defines a new class of plant-specific DNA binding proteins and is a putative regulator of patatin expression. *Plant J* 30:489–497