Genome-wide analysis of bromodomain gene family in Arabidopsis and Rice

Abiraami T.V.[1], Ravi Prakash Sanyal[1], Hari Sharan Misra[1,2] and Ajay Saini[1,2,*]

[1]Molecular Biology Division, Bhabha Atomic Research Centre, Trombay,

Mumbai, Maharashtra, India

[2]Homi Bhabha National Institute, Anushaktinagar, Trombay,

Mumbai, Maharashtra, India

Emails:	Abiraami T.V.	: abiraami@barc.gov.in
	Ravi Prakash Sanyal	: ravis@barc.gov.in
	Hari Sharan Misra	: hsmisra@barc.gov.in
	Ajay Saini *	: ajays@barc.gov.in

***Corresponding Author:**

Dr. Ajay Saini

Tel: +912225595078

Email: ajays@barc.gov.in

Abstract

The bromodomain containing proteins (Brd-proteins) comprise a diverse family of 'epigenetic mark readers', which are an integral to epigenetic gene regulation, important for diverse cellular processes and environmental stress responses. The Brd-proteins harbour a conserved 'bromodomain (BRD)' that interacts with the acetylated histones in the chromatin. Presence of several additional domains enhances the structural-functional diversity and the associated functions of the Brd-proteins. Plants contains multiple Brd-proteins, however the extent of diversity and the molecular events responsible are relatively less understood. The present genome-wide analysis of Brd-members in *A. thaliana* and *O. sativa* showed extensive diversity in structure of gene and regulatory regions, expression dynamics, and domains/motifs, including bromodomain. More than 40% Brd-genes in both the species were affected by genomic duplication events, while, ~60% *A. thaliana* and ~40% *O. sativa* genes showed alternative splicing-mediated changes. These molecular events affected one or more features of the Brd-members (regulatory elements, untranslated regions, and coding regions), which could potentially alter their expression, and associated functions. The phylogenetic analysis divided the *A. thaliana* and *O. sativa* Brd-homologs into six major clusters, and displayed consistency with certain domain/motif combination architectures. Homology modelling of most *A. thaliana* and *O. sativa* Brd-homologs displayed typical BRD-fold. However, the duplicate/divergent sequences displayed important structural variations, which can potentially alter its interaction with the chromatin, and affect associated function. The results also suggested substantial contribution of various genome duplication events (predominantly block events) in the evolution and expansion of Brd-family among diverse plants, including both monocot and dicot species.

Keywords: Alternative splicing, *Arabidopsis thaliana*, bromodomain, epigenetic regulation, genome duplication events, homology modeling, *Oryza sativa*

1. Introduction

Gene expression in eukaryotes including plants is a complex process, regulated at multiples levels and dynamically modulated in response to physiological requirements and stress conditions (Floris et al., 2009; Haak et al., 2017; Merchante et al., 2017; Withers and Dong, 2017). The transcription status of genes is influenced by both genetic and epigenetic components (Singh, 1998). Unlike the regulation mediated by sequence-elements (e.g., core promoter, *cis*-elements, enhancers/silencers), the epigenetic controls involve non-sequence based modifications to alter the expression of genes with short-term as well as long-term implications (Gibney and Nolan, 2010; Lämke and Bäurle, 2017; Rendina González et al., 2018). Epigenetic modifications of DNA and/or histones affects the state of chromatin and transcription activity (Iwasaki and Paszkowski, 2014), and enhances the response potential of the genetic material (Strahl and Allis 2000; Loidl 2004). Various post-translational modifications (e.g., PTMs, acetylation, methylation, phosphorylation, ubiquitylation) affect the characteristics of several proteins including histones where the PTMs modulate nucleosome dynamics (Bowman and Poirier, 2015). Lysine acetylation are crucial for protein-protein interactions, nuclear transport, and chromatin state during epigenetic regulation by affecting the positive charge and steric bulk of nucleosomes (Bowman and Poirier, 2015). The cellular epigenetic system is based on coordinated action of 'writer', 'reader' and 'eraser' proteins for management of PTM marks (Musselman et al., 2012; Zhao et al., 2018). For lysine acetylation lysine acetyltransferases and lysine deacetylases serves as 'writers' and 'erasers' (Musselman et al., 2012), while a conserved bromodomain (BRD) present in several chromatin-associated proteins serve as 'reader' of histone lysine acetylation (Drazic et al., 2016),

The bromodomain containing genes (Brd-genes) were first identified in the *D. melanogaster* (Tamkun, et al., 1992), and subsequently reported in diverse eukaryotes (Rao et al., 2014). The ~110 amino acid bromodomain region folds into four α-helices (αZ, αA, αB, αC) and three loops (ZA, AB, BC) to form a conserved bromodomain/BRD-fold with a hydrophobic pocket that recognizes acetylated lysine residues in the C-terminal tail of the histones (Ferri et al., 2016; Mujtaba et al. 2007; Marmorstein and Berger 2001; Bottomley et al. 2004). A single bromodomain is capable to recognize acetylated lysine residues on different histones (Josling et al. 2012). Brd-containing proteins alone or as a part of multi-protein complexes regulates gene expression by different mechanisms viz. chromatin remodeling, histone modifications, transcriptional machinery regulation (Fujisawa et al., 2017). The Brd-gene family is a large and diverse family in different organisms, and in humans a total of forty-six members are divided into eight structurally and functionally diverse groups due to presence of several additional domains (Filippakopoulos et al., 2012). Human Brd-genes have been well studies for roles in chromatin dynamics and diverse cellular functions, and have gained attention as promising druggable targets for important pathological conditions (Zeng and Zhou, 2002; Sanchez and Zhou, 2009; Taniguchi, 2016; Cochran et al., 2019; Uppal et al., 2019; Boyson et al., 2021).

Epigenetics-mediated changes on chromatin structure and its impact on regulation of gene expression is equally important in physiological processes as well as stress responses in plants (Rosa et al. 2013, Bhadouriya et al. 2021, Vergara et al. 2017, Fransz and De Jong 2002, Pei et al. 2021). Plants also harbour multiple Brd-gene members, however, unlike animals their functional significance is relatively less explored. Studies on few Brd-homologs from Arabidopsis and other plants have shown their involvement in seed germination (Duque and Chua 2003), leaf development (Chua et al. 2005), mitotic cell cycle (Airoldi et al. 2010), sugar and ABA responses (Misra et al. 2017), development (Martel et al. 2017; Rao et al. 2014; Kalantidis et al. 2007), and pathogen perception (Sukarta et al. 2020). The Brd-gene family members show variable numbers in different organisms (Rao et al. 2014). In most plants genomes, various duplication events have contributed towards generation of additional gene copies, which often diverge towards regulatory, structural and functional difference (Flagel and Wendel 2009; Barker et al. 2012; Qiao et al, 2019). Also, alternative splicing (AS) also play roles in regulation of genes in diverse conditions, and altering important features of the protein isoforms (Reddy et al., 2013). Information of involvement of both these mechanisms on Brd-gene diversity among plants is not known, and is worth investigating.

In the present study, we carried out genome-wide analysis of Brd-family of *A. thaliana* (dicot) and *O. sativa* (monocot), to understand the extent of diversity of genes, proteins, regulatory regions, expression dynamics, domains-motifs combinations, and variations in the BRD-fold itself. The genome duplications and AS-events affected substantial number of Brd-homologs, and affected their regulatory and protein characteristics, in both the plant species. Analysis of several photosynthetic organisms further showed the role of duplications in Brd-gene copy number differences among lower and higher plants. To our knowledge, this is the first study on analysis of diversity of Brd-homologs of model plants, *A. thaliana* and *O. sativa*, will be useful for further studies on deciphering their functional significance in different conditions including stress responses.

2. Materials and Methods

2.1. Identification of bromodomain genes in *A. thaliana* and *O. sativa* genome databases

Multiple databases were used for retrieval of sequence data of Brd-genes from *A. thaliana* (referred to as 'AtBrd') and *O. sativa* (referred to as 'OsBrd'). The Arabidopsis Information Resource (TAIR, https://www.arabidopsis.org/index.jsp) and PLAZA dicots (https://bioinformatics.psb.ugent.be/plaza/versions/plaza _v4_5_dicots/, Van Bel et al., 2018) databases were used for Arabidopsis, and whereas Rice Genome Annotation Project (RGAP, http://rice.uga.edu/) and PLAZA monocots (version 4.5, https://bioinformatics.psb.ugent.be/ plaza/versions/plaza_v4_5_monocost/, Van Bel et al., 2018) databases were used for rice. Keyword search and Brd-family ID (SCOP database) based approaches were used for identification of Brd-family members in respective databases.

2.2. *In silico* analysis of characteristics of genes, proteins and transcripts

The structure and organization of AtBrd and OsBrd genes in terms of untranslated regions (UTRs), exons, and introns was analyzed using the Gene Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/, Hu et al. 2014). The important characteristics (molecular weight, MW; isoelectric point, pI etc.) of the AtBRD and OsBRD proteins were estimated using the PROTPARAM tool on the ExPASy website (http://web.expasy.org/protparam/). The alternative transcript/protein isoforms of the AtBrd and OsBrd genes were retrieved from the respective databases, and compared with the corresponding constitutive isoforms by pair-wise alignment using ClustalX (Thompson et al. 1997).

2.3. *In silico* analysis of promoter structure and *cis*-regulatory elements

Upstream (up to 2000 bp) regulatory regions of AtBrd and OsBrd genes were retrieved from the TAIR, RGAP and PLAZA databases. Presence and organization of CpG islands, transcription factor binding sites (TFBS), and tandem repeats motifs was analyzed at Plant Promoter Analysis Navigator online resource (PlantPAN3.0, http://plantpan3.itps.ncku.edu.tw/), whereas *cis*-elements (types, location, copy number) were analyzed at Plant *Cis*-Acting Regulatory Elements databases (PlantCARE, http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

2.4. *In silico* analysis of conserved domains, functional sites and motifs

Presence of conserved domains, functional sites and amino acid patterns in the BRD proteins was analysed using Conserved Domain Database (CDD, NCBI) (https://www.ncbi.nlm.nih.gov/cdd/), and PROSITE (https://prosite.expasy.org/, Sigrist et al., 2012). Domain analysis was carried out using default search parameters and only significant hits were considered for analysis. Conserved motifs were analyzed at MEME online Suite (version 5.4.1, http://meme-suite.org/tools/meme, Timothy et al., 2015) using following parameters, minimum and maximum motif width: 6-50, number of motifs: 15.

2.5. *In silico* analysis of gene expression using RNA-Seq data

The RNA-Seq expression data (as FPKM values, Fragments Per Kilobase of transcript, per Million mapped reads) of respective Brd-genes was retrieved from the Arabidopsis RNA-seq Database (V2, http://ipf.sustech.edu.cn/pub/athrna/) and Rice Expression Database (http://expression.ic4r.org/). The gene/locus names were used for search and retrieval of FPKM data. For tissue-specific analysis in Arabidopsis, data was retrieved for shoot, root, stem, meristem, seedling, embryo, leaf, silique, silique, endosperm, flower and pollen. In rice, data was retrieved for root, shoot, panicle, anther, pistil, aleurone and seed. In case of multiple libraries, the average FPKM values were used. FPKM values were log2-transformed and used for generation of heat map-based transcript profiles by Heatmap Illustrator software (HemI, version 1.0, Deng et al., 2014).

2.6. Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment of the bromodomain region of AtBRDs and OsBRDs was done by ClustalX software (Thompson et al. 1997) using default parameters. Multiple alignment was used for estimating sequence divergence, and analysis of genetic relationships using neighbour-joining approach (Saitou and Nei, 1987) in Molecular Evolutionary Genetic Analysis X software (MEGAX, version 10.0.5, Kumar et al. 2018). Statistical analysis was carried out by bootstrap method (Felsenstein 1985). Bromodomain region of few human homologs containing single (UniProtKB accession numbers: Q9NR48, ASH1L; Q9NPI1, BRD7; Q9H0E9-2, BRD8B; P55201-1, BRPF1A; Q92830, GCN5L2; Q03164, MLL; Q13342, SP140; Q13263, TRIM28; Q9UPN9, TRIM33A; O15016, TRIM66; P51531, SMCA2; P51532, SMCA4) or two bromodomains (P25440, BRD2; Q15059, BRD3; O60885, BRD4; Q58F21, BRDT; Q6RI45, BRWD3; P21675, TAF1; Q9NS16, WDR9) were also included for analysis.

2.7. Homology modelling and comparison

The homology models of bromodomain regions of several *A. thaliana* and *O. sativa* BRD proteins were generated using automated mode and compared for structural differences at SWISS-MODEL workspace (http://swissmodel.expasy.org). For identification of structural differences due to variations among BRD-folds, the bromodomain homology models of duplicate Brd-pairs or divergent Brds were superposed using structure comparison tools at SWISS-MODEL workspace.

2.8. Analysis of duplication events among plant genomes

The Brd-gene family members from *A. thaliana* and *O. sativa* were analyzed for duplication events at PLAZA dicots and monocots database (https://bioinformatics.psb.ugent.be/plaza/, Van Bel et al., 2018). Synteny analysis was carried out to identify the chromosomal segment-pairs originated due to block and tandem duplication events. InterPro identifier IPR001487 (bromodomain) was used to identify the locations of all the Brd-genes including the duplicates, and the chromosomal representations of the Brd-genes (including duplicates), segments, and chromosomes involved duplications were made using the Circle Plot tool available at PLAZA server. Also, more than 80 plant genomes (seven lower photosynthetic organisms, 28 monocots and 46 dicots) available at PLAZA monocots and dicots databases were analyzed for prevalence of duplication events leading to multiple copies of Brd-genes.

3. Results

3.1. Genome-wide analysis of Brd-gene members *A. thaliana* and *O. sativa*

The keyword search and bromodomain superfamily-based search (SCOP ID: 3001843, SCOP database) identified twenty-eight Brd-gene family members in Arabidopsis, and twenty-two in rice. The resultant hits were further confirmed for the presence of 'bromodomain' using Conserved Domain (CD)-Search at NCBI. The corresponding sequences of genes, coding regions, promoter regions, and encoded proteins were designated using first two letter for species (At: *Arabidopsis thaliana* and Os: *Oryza*

sativa), followed by 'Brd' for bromodomain gene/transcript or 'BRD' for bromodomain protein, and finally the numerals indicating particular gene member (Arabidopsis: 1-28 and Rice: 1-22). Important characteristics of the AtBrd and OsBrd-genes, transcripts (constitutive and alternative), and encoded proteins are summarized in Table 1 and Table 2. The Brd-genes showed unequal chromosomal distribution in both the species. Out of 28 AtBrd-genes, nine were localized on Chr1, seven each on Chr3 and Chr5, and five on Chr2, and none on Chr4 (Table 1). The 22 OsBrd-genes were distributed on nine of total twelve chromosomes, ranging from one (Chr10), to four genes (Chr2), whereas Chr5, Chr11, Chr12 did not harbour any OsBrd-gene (Table 2).

3.2. Block and tandem duplicated Brd-genes in *A. thaliana* and *O. sativa*

Syntenic analysis on PLAZA databases identified that six AtBrd-gene pairs have originated due to six block duplication events (BD1 to BD6, Table 1) in *A. thaliana*, whereas in *O. sativa*, of the five OsBrd-gene pairs, one-pair originated due to a tandem duplication (TD1 in Chr 5) and remaining due to four block duplication events (BD1 to BD4, Table 2). Three duplicate AtBrd-gene pairs were outcome of intra-chromosomal block duplication events BD1 (AtBrd3-AtBrd8) and BD2 (AtBrd4-AtBrd9) within Chr1, and BD6 (AtBrd22-AtBrd28) within Chr5. The remaining duplicate AtBrd-gene pairs originated due to inter-chromosomal block duplication events viz. BD3 (AtBrd11-AtBrd20, Chr2 and Chr3), BD4 (AtBrd12-AtBrd21, Chr2 and Chr3) and BD5 (AtBrd15-AtBrd23, Chr3 and Chr5) (Figure 1A). In rice, one OsBrd-gene pair originated due to an intra-chromosomal block duplication event, BD1 (OsBrd1-OsBrd2, within Chr1), while three OsBrd-gene pairs due to inter-chromosomal block duplication events, BD2 (OsBrd11-OsBrd18, Chr4 and Chr8), BD3 (OsBrd13-OsBrd17, Chr6 and Chr8), BD4 (OsBrd19-OsBrd20, Chr8 and Chr9) (Figure 1B). Interestingly, OsBrd11 was also involved in a tandem duplication event (TD1, in Chr4) leading to OsBrd10 (Figure 1B).

3.3. Heterogeneity of OsBrd-gene family and role of duplication and splicing events

The 28 AtBrd-genes showed extensive heterogeneity in the length of gene (1,919 bp to 10,519 bp), coding region (1,110 bp to 6,579 bp), and encoded protein (369 AA to 2,192 AA) (Table 1). However, the OsBrd-gene family showed relatively higher variability in length of gene (2,119 bp to 16,548 bp), coding region (717 bp to 6,603 bp), and protein size (238 AA to 2,200 AA) (Table 2). In general, length and number of exons and introns and UTRs (5´and 3´- ends) contributed towards the observed length heterogeneity of Brd-genes in both the plants. Among AtBrd-genes the number of exon ranged from two to a maximum of twenty-two (Figure 2A), and among OsBrd-genes it ranged from one to twenty-one (Figure 2B). The 5´ and 3´UTRs also contributed toward the length variation of genes. Two AtBrd and five OsBrd-genes lacked 5´-UTRs, and in general OsBrd-genes (OsBrd1, OsBrd2, OsBrd5, OsBrd13, OsBrd15, OsBrd17) harbored longer UTRs than AtBrd-members (Figure 2).

The duplicated Brd-gene pairs showed variability in gene length and organization including exons, introns and UTRs at 5´- and 3´- ends. All the six block duplicated AtBrd-gene pairs showed variation

in length of gene and coding regions (Table 1, Figure 2A). In general, the long gene size correlated with the long coding sequence, except AtBrd-pairs of BD1 and BD4 events, while the duplicate gene pairs AtBrd11-AtBrd20 and AtBRD15-AtBrd23 showed differed in the number of exons (Table 1, Figure 2A). The duplicate OsBrd-gene pairs displayed relatively higher divergence in all characteristics like gene length, coding region and organization of UTRs, exons and introns (Figure 2B). Among the tandem and block-duplicated OsBrd-gene pairs, the length variation in gene ranged from 31 bp to 3,360 bp, while coding region variation ranged from 75 bp to 1,614 bp (Table 2). The longer genes in the OsBrd-pairs of BD1, BD3, BD4 events also showed long coding region, while OsBrd-pairs generated by BD2 and TD1 events showed inverse trend (Table 2, Figure 2B).

Apart from the role of gene duplication events, involvement of AS-mediated impact on diversity of transcript and proteins was further investigated. It was observed that around 60% AtBrd-genes displayed AS-events, generating a maximum of six transcripts in AtBrd6, AtBrd13 and AtBrd14 genes (Table 1, Supplementary Figure 1). AS affected UTRs in three genes (AtBrd2, AtBrd25, and AtBrd26), coding region in eight genes (AtBrd7, AtBrd8, AtBrd10, AtBrd15-18, AtBrd20), and both UTR and coding regions in remaining genes (Table 1, Supplementary Figure 1). Interestingly, only one AtBrd copy from block duplications BD1, BD2, BD3 showed AS, while both copies from BD5 showed different AS-transcripts (Table 1). On the contrary, ~41% OsBrd-genes showed AS-events, that affected only UTRs (OsBrd13, OsBrd17), coding regions (OsBrd1, OsBrd6, OsBrd12, OsBRD16, OsBrd18), and both (OsBrd5, OsBrd15) (Table 2, Supplementary Figure 2). In rice also, only one of the copy of BD1 and BD2 duplicated gene-pairs showed AS, while both the copies of BD3 gene pair showed AS-transcripts with differences in the UTRs (Table 2, Supplementary Figure 2).

3.4. BRD-proteins showed heterogeneity in domain and motif organization

Analysis of 28 AtBRD and 22 OsBRD proteins identified differences in domains in addition to the conserved bromodomain (BRD). The AtBRDs were divided into four major groups containing members with, (i) only BRD domain, (ii) BRD and additional domains/functional sites (MYB-DBD, SANT, GNAC, ATPase, Ub, WD40, Helicase CTD etc.), (iii) BRD and ET (extra-terminal) domain, and (iv) BRD, ET and other domains (MDN1, Ser/Pro/Ala repeat regions etc.) (Figure 3A, top panel). Likewise, 22 OsBRDs formed similar groups with members containing, (i) only BRD domain, (ii) BRD with additional domains/functional sites, (iii) BRD-ET combination, and (iv) BRD, ET and additional domains (Figure 3A, bottom panel). In addition to the variability in position of BRD (and other domains), domains specific to AtBRDs (MYB-DBD, MDN1) and OsBRDs (PHD, WHIM1, Spo-VK, Med15, Znf) were also identified. Further analysis of BRDs of the two species identified 15 conserved motifs, M1 to M15 (Supplementary Table 1 and 2), which divided the BRDs of both species into six groups (Figure 3B). While 12 AtBRDs (groups a, b) showed 6-10 motifs, remaining members (including the largest BRDs) harbored only 1-4 motifs (Figure 3B, top panel). Among OsBRDs, groups

‘a’ and ‘b’ specific BRDs harbored 7-9 motifs, while other members showed 1-6 motifs (Figure 3B, bottom panel), and motifs M1 and M2 were most prevalent in both species.

3.5. Duplication and AS-events affected the domain heterogeneity of BRDs

Block and tandem duplications affected the domains/motifs profiles of both AtBRD and OsBRD pairs. Among the six block duplicated (BD1-BD6) AtBRD-pairs, divergence of one of the copy resulted in variability of domains/motifs, with least heterogeneity in the AtBRD22-AtBRD28 pair (Figure 3A, top panel). Similarly, divergence was also evident among the domain composition of five duplicated OsBRD-pairs, with maximum heterogeneity in OsBRD19-OsBRD20 and OsBRD11-OsBRD18 pairs (Figure 3A, bottom panel). The variation in the motifs were more pronounced among the OsBRDs generated by duplication events TD1, BD1, BD2, and BD4 (Figure 3B, bottom panel).

AS-events in the coding region also affected the domains heterogeneity of several BRDs of both species. Fourteen AtBRDs affected by AS-events showed loss of domains or functional sites (Myb-DBD in AtBRD7.2; Ser-RR in AtBRD8.2; MDN1 in AtBRD27.2, 27.3), while loss at N- and C-terminal region (AtBRD10.2; AtBRD14.3, 14.4, 14.6; AtBRD17.2, 17.4; AtBRD18.1) formed small sized isoforms (Figure 4A). Among OsBRDs, the AS-events resulted in the loss of Ser-RR (OsBRD5.3; OsBRD6.6; OsBRD18.2), BRD domain (completely in OsBRD1.2; partially in OsBRD12.3; OsBRD16.2), and C-terminal truncation with no impact on domains (OsBRD6.3, 6.4 and 6.5; OsBRD15.3; OsBRD16.2) (Figure 4B). Interestingly, the AS-mediated loss of BRD domain was specific to three OsBRDs, and not in AtBRDs.

3.6. *Cis*-elements indicates involvement of Brd-genes in important functions

Analysis of upstream regions of Brd-genes at PlantCARE database identified *cis*-elements associated with diverse functions. In both the species, the major groups of *cis*-elements are associated with response to phytohormones (abscisic acid, auxin, salicylic acid, jasmonic acid, ethylene, gibberellin), light, abiotic (low temperature, anaerobic condition) and biotic stress (wound, defense, elicitor-mediated activation), along with some physiological functions and transcription factor binding (Figure 5A and 5B), with differences in copy number and locations in the upstream region (Supplementary Figure 3 and 4). In both species, certain genes contained higher number of *cis*-elements related to biotic stress (AtBrd12, AtBrd21, AtBrd26, OsBrd19, OsBrd21) and physiological functions (AtBrd12, AtBrd20 AtBrd23, OsBrd7, OsBrd21, OsBrd22), while some lacked elements specific to phytohormone response (OsBrd7), abiotic stress (OsBrd2) and light (OsBrd9). Moreover, specificity of some *cis*-elements to certain genes was also observed viz. NON box (OsBrd5), motif 1 (OsBrd14), TATC-box (OsBrd15, OsBrd17), MBSI (OsBrd10) (Supplementary Figure 3 and 4).

3.7. Duplication events affected the promoter structure and *cis*-element diversity of Brd-genes

The tandem/block duplications affected the *cis*-elements diversity of the AtBrd and OsBrd duplicate gene-pairs (Figure 5). For example, AtBrd3-AtBrd8 pair differed in elements associated with light, defense, low temperature, phytohormones (ethylene, gibberellin), and physiological functions (zein metabolism, endosperm expression), whereas AtBrd22-AtBrd28 pair differed in elements for phytohormones (abscisic acid, ethylene, gibberellin), anaerobic induction, defense, light, phytohormones (ethylene, gibberellin), and physiological mechanisms (cell cycle, endosperm expression, flavonoid biosynthesis, zein metabolism). Similar differences were evident in other AtBrd duplicate-pairs (Figure 5A, Supplementary Figure 3). The duplicate OsBrd-gene pairs also showed *cis*-element differences. The OsBrd1-OsBrd2 pair differed in elements responsive to abiotic and biotic stresses, and physiological functions (tissue specific expression, circadian rhythm), whereas OsBrd11-OsBrd18 pair differed in light, abiotic and biotic stresses, phytohormones (gibberellin, salicylic acid) responsive elements (Figure 5B, Supplementary Figure 4).

In general, the length of promoter region of duplicate AtBrd-gene was not affected by duplication events, however substantial difference in the promoter length of OsBrd-duplicates was evident, except OsBrd1-OsBrd2 (Figure 6). Differences in CpG islands, transcription factor binding sites (TFBS), and repeat motifs were also evident. The six AtBrd-duplicate pairs displayed differences in TFBS arrangement, repeat motifs (AtBrd4-AtBrd9; AtBrd11, AtBrd20), and CpG islands (AtBrd11-AtBrd20; AtBrd15-AtBrd23) (Figure 6A). All duplicated OsBrd-pairs showed differences in the CpG islands and TFBS locations, and repeat motif heterogeneity was among OsBrd1-OsBrd2, OsBrd13-OsBrd17, and OsBrd19-OsBrd20 pairs (Figure 6B). Collectively these results indicate that the regulatory elements of the duplicated AtBrd and OsBrd genes have diverged for differential responses towards internal and external factors.

3.8. AtBrd and OsBrd-genes showed tissue-specific expression differences

RNA-Seq data based analysis showed differential abundance of AtBrd-genes among *A. thaliana* tissues (Figure 7A). The AtBrd10 and AtBrd15 showed highest levels in seeds, whereas in pollen AtBrd17, AtBrd22 and AtBrd27 expressed at higher levels. Also, in most tissues AtBrd22 showed higher levels, whereas AtBrd16 showed low transcript levels (Figure 7A). The six duplicated AtBrd-pairs (BD1-BD6) also showed expression level differences, with maximum variation within Brd-members of AtBrd11-AtBrd20 and AtBrd12-AtBrd21 pairs (Figure 7A). The OsBrd-genes also showed differences in tissue-specific transcript levels, with most of them expressing at lower levels in anther I stage, and at higher levels at panicle II and anther II stages (Figure 7B). In most tissues, OsBrd3, OsBrd6, OsBrd15, OsBrd18 showed higher basal levels, while OsBrd2, OsBrd9 and OsBrd16 showed lower transcript levels. As in AtBrd-duplicates, the duplicate OsBrd-pairs originated due to block (BD1-BD4) and

tandem (TD1) events also displayed tissue-specific differences, with maximum variation in OsBrd1-OsBrd2, followed by OsBrd19-OsBrd20 pair (Figure 7B).

3.9. Sequence divergence of bromodomain (BRD) region of Brd-genes

The sequence divergence in the conserved bromodomain (BRD) region of *A. thaliana* and *O. sativa* BRD-homologs was analyzed by multiple sequence alignment of 50 sequences, including tandem and block duplicates. The BRD-regions of OsBRD-homologs showed more length variation (57-133 AA) than AtBRDs (94-133 AA) (Figure 8). The BRD-regions of three BRD-duplicates of *A. thaliana* (AtBRD11-AtBRD20; AtBRD12-AtBRD21; AtBRD15-AtBRD23) and *O. sativa* (OsBRD1-OsBRD2, OsBRD11-OsBRD18; OsBRD19-OsBRD20) differed in length (Figure 8). Analysis identified conserved and variable regions, and also showed that the indels at the ends contributed towards the length variation of the BRD-regions. Also, exceptionally small BRD-regions in two rice-specific BRDs (OsBRD7, OsBRD4) were due to long deletions towards N-terminal region (Figure 8).

The phylogenetic analysis based on conserved BRD region placed the AtBRDs and OsBRDs into six clusters (I to VI) supported by high bootstrap values, with representation of both the species (Figure 9A). The number of variable sites within groups ranged from 21% (II) to 67.3% (VI), whereas between groups it ranged from 60.4% (I vs IV) to 77.8% (III vs VI). The largest cluster I (13 AtBRDs and 11 OsBRDs) was subdivided into five subgroups (A-E), with a single divergent OsBRD2 in subgroup C. The BRD-regions of the six block duplicate AtBRD-pairs (BD1 to BD6) and three OsBRD-pairs from tandem (TD1) and block duplications (BD2, BD3) displayed less divergence and grouped together in respective clusters (Figure 9A). On the contrary, OsBRD1-OsBDR2 and OsBRD19-OsBRD20 block-duplicated pairs were highly divergent and placed in different subgroups of cluster I, or different clusters (III, V) (Figure 9A). Certain groups based on small BRD-region were consistent with the domain signatures of full-length AtBRD and OsBRD-proteins viz. cluster I (BRD and ET, alone/with other motifs), cluster II (BRD and GNAC), and cluster VI (BRD with one or more additional domains like WD, Helicase-CTD, TLD, ATPase, QLQ, Ub, PHD) (Figure 3 and Figure 9A).

Analysis including BRD-regions of representative human homologs (including single and dual BRDs) identified plant BRD-sequences that grouped together (Figure 9B). Also, groups/subgroups specific to plants (GII, subgroups IA, IB, ID and IF, Figure 9B) and humans (GVII, subgroups IC, IE, IG, Figure 9B) were evident. Although, AtBRDs/OsBRDs single BRD-domains, sequences clustered close to the two domains of dual-BRD human homologs were evident. Subgroup IH containing BRD(1) domains of human BRD2-4 and BRDT was close to subgroup IG (AtBRD6, OsBRD13 and OsBRD 17) (Figure 9B), whereas the BRD(2) domain subgroup (IE) was close to subgroup ID (OsBRD12, AtBRD10 and AtBRD 18) (Figure 9B). Likewise, two BRD-domains of human WDR9 were clustered with plant BRD sequences in two different clusters (subgroup IF and GIII. Figure 9B).

3.10. Heterogeneity mediated structural variations in the bromodomain fold (BRD-fold)

The BRD-region folds into four α-helices (αZ, αA, αB, αC) and three loops (ZA, AB, BC loops) to form a conserved BRD-fold (Figure 10A). Effect of length and sequence heterogeneity in the BRD-region of AtBRDs and OsBRDs (Figure 8) on important structural elements of the BRD-fold was evaluated by homology modeling. A typical BRD-fold was observed for most AtBRDs and OsBRDs (Figure 10B). The divergent BRD-regions, however displayed structural variations in important regions such as, a) truncated αZ helix due to N-terminal deletion in the BRD-region of OsBRD13 and OsBRD17 (Figure 10C), b) an extended region prior to αZ-helix (AtBRD14, OsBRD8) and variation in ZA-loop (AtBRD14) (Figure 10D), c) an extra α-helical region after a sharp bend subsequent to αC-helix due to long C-terminal region (ATBRD16, OsBRD14) (Figure 10E), and d) complete loss of αZ-helix and ZA-loop in two small BRD-regions (OsBRD4 and OsBRD7) (Figure 10F). Superposition of homology models revealed structural changes in the BRD-folds of duplicated BRDs. Less divergent duplicate AtBRD-pairs (AtBRD11-AtBRD20, AtBRD15-AtBRD23) showed minor structural variations (Figure 10G and 10H), compared to more divergent OsBRD-pairs, which showed complete loss of αZ-helix in OsBRD1 (OsBRD1-OsBDR2) (Figure 10I), and structural variations in αZ, αC, and BC loop regions in the OsBRD19-OsBRD20 pair (Figure 10J). Such structural changes in the BRD-fold can affect the interaction and epigenetic function of the BRD-homologs.

3.11. Duplication events affected the Brd-gene copy number among higher plants

A. thaliana and *O. sativa* harbored six and five pairs of the Brd-genes that originated due to duplications that corresponds to > 40% of the Brd-family members (Figure 1). Analysis of additional 79 genomes including lower and higher photosynthetic organisms revealed prevalence of different duplication events, and their contribution in Brd-gene copy number (Figure 11). Four lower photosynthetic organisms showed no duplication events and harbored only 09-16 Brd-genes (Figure 11A). Further, *A. trichopoda* (early-diverging flowering plant) showed one tandem Brd-duplicate and *P. patens* (moss) showed four block and eight tandemly duplicated Brd-genes (Figure 11A). Interestingly, *P. abies* harboured 23 Brd-genes without any evidence of block/tandem duplications. The number of Brd-genes among 27 monocots ranged from 14 (*A. shenzhenica*) to 79 (*T. aestivum*), with three species without duplication generated Brd-gene copies (Figure 11B). Genomes of other monocots showed different combinations of duplication events, a) only block event (BD), b) only tandem event (TD), c) both tandem and block events (TD, BD), d) tandem and combined events (TD, TD+BD), e) block and combined events (BD, TD+BD), and f) all duplication events. Most prevalent block duplications (BD) along with combined (TD+BD) and tandem (TD) events, lead to higher Brd-gene copies in many monocots viz. *Z. mays*, *M. acuminata*, *E. guineensis*, *M. sinesis*, *T. turgidum*, *S. spontaneum* and *T. aestivum* (Figure 11B). On the contrary, all dicots analyzed showed one or more duplication events, with Brd-genes ranging from 20 (*B. vulgaris*) to 62 (*G. max*). Dicots also showed prevalence of different duplication events, a) only TD events, b) only BD events, c) BD and TD events, d) BD and combined

(TD+BD) events, e) and all types of events (Figure 11C). The block duplications (BD) along with tandem (TD) and combined (TD+BD) were contributing factor towards in higher Brd-genes among dicots like *D. carota, P. trichocarpa, M. esculanta, C. arietinum, B. rapa, B. oleracea, C. quinoa, P. bretscheneideri, A. chinensis* and *G. max* (Figure 11C). These results show that genome duplications have contributed substantially towards Brd-gene copies among plants.

4. Discussion

The epigenetic changes in the chromatin state modulates the gene expression, where the coordinated action of epigenetic mark 'readers', 'writers' and 'erasers' (Lauria and Rossi, 2011; Bourbousse et al. 2020; Samo et al., 2021) are crucial for the responses towards metabolic, developmental and environmental cues (Strahl and Allis, 2000; Loidl, 2004; Ojolo et al., 2018; Samo et al., 2021). The Brd-genes function as epigenetic mark readers among diverse organisms, and are involved in chromatin dynamics, diverse cellular functions, and conditions (Tamkun et al., 1992; Rao et al., 2014; Zeng and Zhou, 2002; Sanchez and Zhou, 2009; Taniguchi, 2016; Uppal et al., 2019; Boyson et al., 2021). However, the Brd-genes among plants are relatively less studied.

The present study reveals the extent of diversity of Brd-gene families of *A. thaliana* and *O. sativa*, along with role of duplications and alternative splicing events. Presence of multiple, heterogeneous (w.r.t. domain combination/organization) Brd-homologs indicates their involvement in diverse functions, as seen in human homologs (Filippakopoulos et al., 2012; Fujisawa and Filippakopoulos, 2016). In both plants, the Brd-homologs displayed single BRD-architecture (alone/with other domains) and lacked dual/poly BRDs (Sanchez and Zhou, 2009; Filippakopoulos et al., 2012). Exceptions were observed in Brd-homologs of lower photosynthetic organisms, *M. commoda* (MCO15G409, dual-BRD) and *C. reinhardtii* (Cre05.g247000, triple-BRD). The conserved BRD-fold that interacts with acetylated lysine on histones (Marmorstein and Berger 2001; Bottomley et al. 2004; Mujtaba et al. 2007) is a characteristic feature of several chromatin-associated proteins involved in chromatin remodeling in response to intrinsic and extrinsic factors (Filippakopoulos et al., 2012; Rendina González et al., 2018; Zhang et al., 2018; Bäurle and Trindade, 2020; Ojolo et al, 2018; Samo et al., 2021). Interestingly, few *A. thaliana* and *O. sativa* Brd-homologs lacked typical BRD-fold elements or showed structural variations, which might alter their interaction capability/affinity with the chromatin.

The diversity in promoter structure, CpG islands, and *cis*-elements of the *A. thaliana* and *O. sativa* Brd-genes can alter their expression levels, as seen in the tissue and developmental stage-specific expression differences. In fact, the CpG islands often contain transcription start sites and can support dispersed transcription initiation events (Deaton and Bird 2011). As the reprogramming of cellular machinery in response to metabolic/environmental cues also involve epigenetic changes (Lämke and Bäurle, 2017; Ojolo et al., 2018; Zhang et al., 2018; Chang et al., 2019), alternations in the relative levels of 'Brd-epigenetic-mark readers' might also be crucial for the chromatin dynamics.

Present analysis revealed that the genome duplication events contributed towards the diversity of *A. thaliana* and *O. sativa* Brd-homologs, as also seen in case of other multi-copy genes in plants (Flagel and Wendel, 2009; Barker et al., 2012; Qiao et al, 2019). The duplications-mediated divergence (domains, motifs and promoters) seems more pronounced in the Brd-gene copies of *O. sativa*. Furthermore, the BRD-region heterogeneity caused minor structural variations in the BRD-folds of most duplicate-pairs, and lead to the loss of αZ-helix in OsBrd1 (block duplicate of OsBrd2). In addition, the N-terminal truncation resulted in loss of αZ-helix and ZA-loop in OsBRD4 and OsBRD7 (Figure 11). Such loss of these two elements was also evident in an uncharacterized human protein (Supplementary Figure 6). These structural alterations in the BRD-folds are likely to affect the epigenetic functions of the proteins (Josling et al. 2012). Interestingly, tandem duplicate of OsBrd4 (BRD-PHD type, LOC_Os02g09920) retained only PHD-domain (LOC_Os02g09910) and lost regions coding for BRD-domain and WHIM1 and Znf motifs (Supplementary Figure 7). Since, several homologs lacking specific BRD-features are transcriptionally active (Figure 7), it will be interesting to decipher their characteristics at protein level.

The AS that enhances transcriptome/proteome diversity and is important for physiological processes and stress responses (Ali et al., 2007; Syed et al., 2012; Reddy et al., 2013; Laloum et al., 2018) affected several *A. thaliana* and *O. sativa* Brd-homologs. The AS-events affected UTRs (important for transcript stability, translation and localization, Mignone et al. 2002) and coding regions (partial/complete loss of domains/motifs, and in certain cases (OsBrd1.2, OsBrd5.2, OsBrd13.2, Figure 7) showed predominance over CS-events. Among the different AS-modes (independent, functionally shared, accelerated), which the paralogs can undergo (Iñiguez, 2017), the *A. thaliana* and *O. sativa*, Brd-duplicates displayed non-shared isoforms, indicating divergence towards sub-functionalization (Iñiguez, 2017). In both plants, the duplication and AS-events enhanced the heterogeneity of the Brd-homologs. AS-events were also shown to responsible for the differential fate and interaction of two GCN5 isoforms with similar localization (Martel et al., 2017). Therefore, understanding the significance of AS-events of At/OsBrds-homologs is important, as in certain cases AS-isoforms (e.g., OsBrd1.2, lacks BRD) are more abundant than the CS-isoforms, in both tissue- and stress-specific environment (Supplementary Figure 8).

While human Brd-homologs are important in several molecular mechanisms (Sanchez and Zhou, 2009; Sanchez et al., 2014; Filippakopoulos et al., 2012; Taniguchi, 2016; Cochran et al., 2019; Morgado-Pascual et al., 2019; Uppal et al., 2019; Boyson et al., 2021), only few plant Brd-homologs have been analyzed, including BRD-ET proteins of *A. thaliana* (GTE1/IMB1, Duque and Chua, 2003; GTE6, Chua et al. 2005; GTE4, Airoldi et al., 2010; GTE9; GTE 11, Misra et al. 2017), *B. distachyon* GCN5 (Martel, et al., 2017), and *N. benthamiana* NbDBCP (Sukarta et al. 2020). The plant Brd-homologs with multi-domain architecture, are likely to be involved in diverse cellular mechanisms. The diversity of the bromodomain-region and the corresponding *A. thaliana* and *O. sativa* Brd-homologs seems consistent, suggesting that it can be used to analyze Brd-family of a species, to overcome difficulties

associated with full-length multi-domain proteins (Nakano et al., 2006). The *A. thaliana* and *O. sativa* Brd-homologs showed representation in all clusters, and although they lacked dual-BRD-homologs (human homologs Brd2, Brd3, Brd4, TAF1, WDR9; Filippakopoulos et al., 2012), similar domains were present in separate proteins. It would be interesting to find out if they differ in interaction capabilities as shown for the dual-BRD domains (Miller et al., 2016).

Contribution of genome duplication events, known in enhanced copy number/diversity of plant genes (Qiao et al, 2019), was evident in Brd-gene copy number among most plants analyzed, which also seem associated with genome complexity. The lower photosynthetic organisms (*M. commoda*, *S. moellendorffii*, *C. reinhardtii*, *M. polymorpha*) harbored few Brd-genes and no duplications, whereas the tandem or/and block events enhanced the Brd-gene copies in *A. trichopoda* (single duplication event, Amborella Genome Project, 2013) and *P. patens* (two whole genome duplications, WGD, Lang et al., 2017). Interestingly, without any duplication, *P. abies* showed higher Brd-gene copies, which is likely due to the transposons, also attributed for its large genome size (Nysted et al., 2020). Among higher plants, more rounds of genome duplications have contributed towards higher gene copies (Qiao et al., 2019). Monocots affected by genome duplications events like ζ (ancestral), ε (paleohexaploidization) σ and ρ (predating Poaceae divergence), lineage-specific events (e.g., *M. acuminata*) and polyploidy (e.g., *T. turgidum*, tetraploid; *T. aestivum*, hexaploid), show higher Brd-gene copies. Similarly, dicots with initial duplication events (ζ and ε), triplication (WGT, γ), and lineage-specific WGD/ploidy events (α and β, crucifer lineage; *Gossypium* specific ploidy, WGDs specific to poplar, legumes, and *Glycine*) also showed high Brd-gene copies among many species. Also, loss/deletion of copies immediately after duplication is also a known phenomenon (Qiao et al., 2019), and was evident in *A. shenzhenica*, *P. equestris*, *Z. marina* which lack duplicate Brd-gene copies despite an ancient WGD event (Cai et al., 2014; Olsen et al., 2016; Zhang et al., 2017).

Present genome-wide analysis revealed extensive diversity among the *A. thaliana* and *O. sativa* OsBrd-family members, regulatory regions, domains/motifs, expression, and sequence/structural differences in the BRD-fold itself. Duplication and AS-events enhanced the Brd-gene numbers and diversity, and various duplication events and Brd-gene copies seems related to the complexity of photosynthetic organisms. The plant Brd-gene family is less investigated, however its diversity (including duplication/AS-mediated), domain signatures, suggests involvement in diverse cellular mechanisms, and needs a thorough analysis for understanding their functional significance.

Supplementary data files

Supplementary Table 1: List of conserved motifs identified among twenty-eight *A. thaliana* bromodomain (BRD) containing proteins.

Supplementary Table 2: List of conserved motifs identified among twenty-two *O. sativa* bromodomain (BRD) containing proteins.

Supplementary Figure 1: Schematic representation of alternative splicing events affecting bromodomain containing genes of *A. thaliana*.

Supplementary Figure 2: Schematic representation of alternative splicing events affecting bromodomain containing genes of *O. sativa*.

Supplementary Figure 3: Diversity of *cis*-elements in the upstream regions of *A. thaliana* Brd-genes as per analysis as PlantCARE database.

Supplementary Figure 4: Diversity of *cis*-elements in the upstream regions of *Oryza sativa* Brd-genes as per analysis as PlantCARE database.

Supplementary Figure 5: Comparison of homology model of a normal bromodomain (BRD) region containing all typical features of the BRD-fold and homology model of BRD-region of a human protein (K2026_Human, UniProt ID: Q5HYC2) with a deletion at N-terminal region leading to loss of αZ and ZA-loop elements.

Supplementary Figure 6: CDD-NCBI based conserved domain analysis of OsBrd4 (LOC_Os02g09920, BRD-homolog with BRD-PHD combination and its tandem duplicate (LOC_Os02g09910) encoding protein with only PHD.

Supplementary Figure 7: Heatmap-based analysis of RNA-Seq data of constitutive and alternative transcripts of OsBrd-genes in different tissues and stress conditions.

Acknowledgements and Funding Source

This work was supported by the institutional funding of Bhabha Atomic Research Centre, Mumbai, India. No separate funding was obtained from any other National/International funding body for this study. We thank Dr Sheetal Uppal, MBD, BARC for suggestions and comments.

Authors contribution

ATV: analysis of gene and protein sequences, RNA-Seq data, phylogeny, RPS: analysis of splicing events and domain heterogeneity; HSM: data analysis and review, manuscript writing; AS: planning and execution, data analysis and review, manuscript compilation and communication.

Data availability

The sequences used in the study are available from the public databases, GenBank-NCBI (https://www.ncbi.nlm.nih.gov/genbank), Rice Genome Annotation Project (RGAP, http://rice.uga.edu/), The Arabidopsis Information Resource (TAIR, https://www.arabidopsis.org/index.jsp) and PLAZA monocot and dicots online resources (https://bioinformatics.psb.ugent.be/plaza/). All the locus numbers and accession numbers are listed in manuscript, Table 1 and Table 2.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Airoldi CA, Rovere FD, Falasca G, Marino G, Kooiker M, Altamura MM, Citterio S, Kater MM (2010) The Arabidopsis BET bromodomain factor GTE4 is involved in maintenance of the mitotic cell cycle during plant development. Plant Physiol. 152(3):1320-34. doi: 10.1104/pp.109.150631.

Ali GS, Palusa SG, Golovkin M, Prasad J, Manley JL, Reddy AS (2007) Regulation of plant developmental processes by a novel splicing factor. PLoS One 2(5): e471. doi: 10.1371/journal.pone.0000471.

Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants. Science 342(6165):1241089. doi: 10.1126/science.1241089.

Barker MS, Baute GJ, Liu SL (2012). Duplications and Turnover in Plant Genomes. In: Wendel J, Greilhuber J, Dolezel J, Leitch I, eds. Plant Genome Diversity Volume 1. Springer, Vienna, 155-169, doi: 10.1007/978-3-7091-1130-7_11.

Bäurle I, Trindade I (2020) Chromatin regulation of somatic abiotic stress memory. J. Exp. Bot. 71(17):5269-5279. doi: 10.1093/jxb/eraa098

Bhadouriya SL, Mehrotra S, Basantani MK, Loake GJ, Mehrotra R (2021) Role of Chromatin Architecture in Plant Stress Responses: An Update. Front. Plant Sci. 11:603380. doi: 10.3389/fpls.2020.603380.

Bottomley, MJ (2004). Structures of protein domains that create or recognize histone modifications. EMBO Rep. 5(5), 464-469.

Bourbousse C, Barneche F, Laloi C (2020). Plant chromatin catches the sun. Front. Plant Sci. 10, 1728.

Bowman GD, Poirier MG (2015). Post-translational modifications of histones that influence nucleosome dynamics. Chem. Rev. 115(6), 2274–2295. https://doi.org/10.1021/cr500350x

Boyson SP, Gao C, Quinn K, Boyd J, Paculova H, Frietze S, Glass KC (2021). Functional Roles of Bromodomain Proteins in Cancer. Cancers 13(14), 3606. https://doi.org/10.3390/cancers13143606.

Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, Chen LJ, He Y, Xu Q, Bian C, Zheng Z, Sun F, Liu W, Hsiao YY, Pan ZJ, Hsu CC, Yang YP, Hsu YC, Chuang YC, Dievart A, Dufayard JF, Xu X, Wang JY, Wang J, Xiao XJ, Zhao XM, Du R, Zhang GQ, Wang M, Su YY, Xie GC, Liu GH,

Li LQ, Huang LQ, Luo YB, Chen HH, Van de Peer Y, Liu ZJ (2015) The genome sequence of the orchid Phalaenopsis equestris. Nat. Genet. 47(1):65-72. doi: 10.1038/ng.3149.

Chang Y, Zhu C, Jiang J, Zhang H, Zhu JK Duan CG (2019). Epigenetic regulation in plant abiotic stress responses: Epigenetic codes of plant abiotic stress. J. Integr. Plant Biol. 62. 10.1111/jipb.12901.

Chua YL, Channelière S, Mott E, Gray JC (2005). The bromodomain protein GTE6 controls leaf development in Arabidopsis by histone acetylation at ASYMMETRIC LEAVES1. Genes Dev. 19(18), 2245-2254.

Cochran AG, Conery AR, Sims RJ 3rd. (2019) Bromodomains: a new target class for drug development. Nat. Rev. Drug Discov. 18(8):609-628. doi: 10.1038/s41573-019-0030-7.

Deaton AM, Bird A. (2011). CpG islands and the regulation of transcription. Genes Dev. 25:1010–1022.

Deng W, Wang Y, Liu Z, Cheng H, Xue Y. (2014) HemI: a toolkit for illustrating heatmaps. PLoS One, 9(11): e111988. doi: 10.1371/journal.pone.0111988.

Drazic A, Myklebust LM, Ree R, Arnesen T. (2016) The world of protein acetylation. Biochim. Biophys. Acta. 1864(10):1372-401. doi: 10.1016/j.bbapap.2016.06.007.

Duque P, Chua NH (2003). IMB1, a bromodomain protein induced during seed imbibition, regulates ABA- and phyA-mediated responses of germination in Arabidopsis. Plant J. 35: 787-799. https://doi.org/10.1046/j.1365-313X.2003.01848.x

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39(4):783-791. doi: 10.1111/j.1558-5646.1985.tb00420.x.

Ferri E, Petosa C, McKenna CE. (2016) Bromodomains: Structure, function and pharmacology of inhibition. Biochem. Pharmacol. 106:1-18. doi: 10.1016/j.bcp.2015.12.005.

Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, Felletar I, Volkmer R, Müller S, Pawson T, Gingras AC, Arrowsmith CH, Knapp S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 149(1):214-31. doi: 10.1016/j.cell.2012.02.013.

Flagel LE, Wendel JF (2009). Gene duplication and evolutionary novelty in plants. New Phytol. 183, 557-564.

Florence B, Faller DV (2001). You bet-cha: a novel family of transcriptional regulators. Front. Biosci. 6(1), D1008-18.

Floris M, Mahgoub H, Lanet E, Robaglia C, Menand B. (2009) Post-transcriptional regulation of gene expression in plants during abiotic stress. Int. J. Mol. Sci. 10(7):3168-85. doi: 10.3390/ijms10073168.

Fransz P F, De Jong JH (2002). Chromatin dynamics in plants. Curr. Opin. Plant Biol. 5(6), 560-567.

Fujisawa T, Filippakopoulos P (2017) Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. Nat. Rev. Mol. Cell. Biol. 18, 246–262. https://doi.org/10.1038/nrm.2016.143.

Gibney E, Nolan C (2010) Epigenetics and gene expression. Heredity 105, 4-13. https://doi.org/10.1038/hdy.2010.54.

Haak DC, Fukao T, Grene R, Hua Z, Ivanov R, Perrella G, Li S (2017) Multilevel regulation of abiotic stress responses in plants. Front. Plant Sci. 8, 1564.

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015). GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics, 31(8), 1296-1297. doi: 10.1093/bioinformatics/btu817.

Iñiguez LP, Hernández G (2017). The Evolutionary Relationship between Alternative Splicing and Gene Duplication. Front. Genet. 8, 14. https://doi.org/10.3389/fgene.2017.00014.

Iwasaki M, Paszkowski J. (2014) Epigenetic memory in plants. EMBO J. 33(18):1987-98. doi: 10.15252/embj.201488883.

Josling GA, Selvarajah SA, Petter M, Duffy MF (2012). The role of bromodomain proteins in regulating gene expression. Genes 3(2), 320–343. https://doi.org/10.3390/genes3020320

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35(6), 1547.

Laloum T, Martín G, Duque P (2018) Alternative splicing control of abiotic stress responses. Trends Plant Sci. 23, 140–150.

Lämke J, Bäurle I (2017) Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. Genome Biol. 18, 124. https://doi.org/10.1186/s13059-017-1263-6.

Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, Piednoel M, Gundlach H, Van Bel M, Meyberg R, Vives C, Morata J, Symeonidi A, Hiss M, Muchero W, Kamisugi Y, Saleh O, Blanc G, Decker EL, van Gessel N, Grimwood J, Hayes RD, Graham SW, Gunter LE, McDaniel SF, Hoernstein SNW, Larsson A, Li FW, Perroud PF, Phillips J, Ranjan P, Rokshar DS, Rothfels CJ, Schneider L, Shu S, Stevenson DW, Thümmler F, Tillich M, Villarreal Aguilar JC, Widiez T, Wong GK, Wymore A, Zhang Y, Zimmer AD, Quatrano RS, Mayer KFX, Goodstein D, Casacuberta JM, Vandepoele K, Reski R, Cuming AC, Tuskan GA, Maumus F, Salse J, Schmutz

J, Rensing SA (2018). The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. Plant J. 93(3):515-533. doi: 10.1111/tpj.13801.

Lauria M, Rossi V (2011). Epigenetic control of gene regulation in plants. Biochim. Biophys. Acta. 1809(8):369-78. doi: 10.1016/j.bbagrm.2011.03.002.

Loidl P (2004). A plant dialect of the histone language. Trends Plant Sci. 9(2), 84-90.

Marmorstein R, Berger SL (2001). Structure and function of bromodomains in chromatin-regulating complexes. Gene 272(1-2), 1-9.

Martel A, Brar H, Mayer BF, Charron JB. (2017) Diversification of the histone acetyltransferase GCN5 through alternative splicing in Brachypodium distachyon. Front. Plant Sci. 8, 2176.

Mignone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. Genome Biol. 3(3): REVIEWS0004. doi: 10.1186/gb-2002-3-3-reviews0004.

Miller TC, Simon B, Rybin V, Grötsch H, Curtet S, Khochbin S, Carlomagno T, Müller CW (2016). A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT. Nat Commun. 7:13855. doi: 10.1038/ncomms13855.

Misra A, McKnight TD, Mandadi KK (2018) Bromodomain proteins GTE9 and GTE11 are essential for specific BT2-mediated sugar and ABA responses in *Arabidopsis thaliana*. Plant Mol. Biol. 96, 393–402. Doi: 10.1007/s11103-018-0704-2

Morgado-Pascual JL, Rayego-Mateos S, Tejedor L, Suarez-Alvarez B, Ruiz-Ortega M (2019). Bromodomain and Extraterminal Proteins as Novel Epigenetic Targets for Renal Diseases. Front. Pharmacol. 10, 1315. https://doi.org/10.3389/fphar.2019.01315

Mujtaba S, Zeng L, Zhou MM. (2007). Structure and acetyl-lysine recognition of the bromodomain. Oncogene 26(37), 5521-5527.

Muller S, Filippakopoulos P, Knapp S (2011). Bromodomains as therapeutic targets. Expert Rev. Mol. Med. 13, e29.

Musselman CA, Lalonde ME, Côté J, Kutateladze TG. (2012). Perceiving the epigenetic landscape through histone readers. Nat. Struct. Mol. Biol. 19(12):1218-27. doi: 10.1038/nsmb.2436.

Nakano T, Suzuki K, Fujimura T, Shinshi H. (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 140(2):411-32. doi: 10.1104/pp.105.073783.

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad

L, Bhalerao R, Bohlmann J, Bousquet J, Garcia Gil R, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, Van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497(7451):579-84. doi: 10.1038/nature12211.

Ojolo SP, Cao S, Priyadarshani SVGN, Li W, Yan M, Aslam M, Zhao H, Qin Y. (2018) Regulation of Plant Growth and Development: A Review From a Chromatin Remodeling Perspective. Front. Plant Sci. 9:1232. doi: 10.3389/fpls.2018.01232.

Olsen JL, Rouzé P, Verhelst B, Lin YC, Bayer T, Collen J, Dattolo E, De Paoli E, Dittami S, Maumus F, Michel G, Kersting A, Lauritano C, Lohaus R, Töpel M, Tonon T, Vanneste K, Amirebrahimi M, Brakel J, Boström C, Chovatia M, Grimwood J, Jenkins JW, Jueterbock A, Mraz A, Stam WT, Tice H, Bornberg-Bauer E, Green PJ, Pearson GA, Procaccini G, Duarte CM, Schmutz J, Reusch TB, Van de Peer Y. (2016) The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. Nature 530(7590):331-5. doi: 10.1038/nature16548.

Pandey R, Müller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA. (2002) Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic Acids Res. 30(23):5036-55. doi: 10.1093/nar/gkf660.

Pei L, Li G, Lindsey K, Zhang X Wang M (2021) Plant 3D genomics: the exploration and application of chromatin organization. New Phytol. 230: 1772-1786. https://doi.org/10.1111/nph.17262.

Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH (2019). Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol. 20(1), 38. doi: 10.1186/s13059-019-1650-2.

Rao RSP, Thelen JJ Miernyk JA (2014) In silico analysis of protein Lys-N"-acetylation in plants. Front. Plant Sci. 5, 381.

Reddy ASN, Marquez Y, Kalyna M, Barta A (2013). Complexity of the alternative splicing landscape in plants. Plant Cell 25, 3657–3683.

Rendina González AP, Preite V, Verhoeven KJF, Latzel V. (2018) Transgenerational Effects and Epigenetic Memory in the Clonal Plant *Trifolium repens*. Front Plant Sci. 9:1677. doi: 10.3389/fpls.2018.01677.

Rosa S, Shaw P. (2013) Insights into chromatin structure and dynamics in plants. Biology (Basel), 2(4):1378-1410. Published 2013 Nov 28. doi:10.3390/biology2041378,

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4(4):406-25. doi: 10.1093/oxfordjournals.molbev.a040454.

Samo N, Ebert A, Kopka J, Mozgová I (2021). Plant chromatin, metabolism and development - an intricate crosstalk. Curr. Opin. Plant Biol. 61:102002. doi: 10.1016/j.pbi.2021.102002.

Sanchez R, Zhou MM (2009). The role of human bromodomains in chromatin biology and gene transcription. Curr. Opin. Drug Discov. Dev. 12(5), 659-665.

Sanchez R, Meslamani J, Zhou, M M (2014). The bromodomain: from epigenome reader to druggable target. Biochim. Biophy. Acta Gene Regul. Mech. 1839(8), 676-685.

Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. (2012) New and continuing developments at PROSITE. Nucleic Acids Res. doi: 10.1093/nar/gks1067

Singh KB (1998) Transcriptional regulation in plants: the importance of combinatorial control. Plant Physiol. 118(4):1111-20. doi: 10.1104/pp.118.4.1111.

Strahl BD, Allis CD (2000). The language of covalent histone modifications. Nature 403(6765): 41-45. doi: 10.1038/47412.

Sukarta OCA, Townsend PD, Llewelyn A, Dixon CH, Slootweg EJ, Pålsson LO, Takken FLW, Goverse A, Cann MJ. (2020) A DNA-Binding Bromodomain-Containing Protein Interacts with and Reduces Rx1-Mediated Immune Response to Potato Virus X. Plant Commun. 1(4):100086. doi: 10.1016/j.xplc.2020.100086.

Syed NH, Kalyna M, Marquez Y, Barta A, Brown JWS (2012) Alternative splicing in plants—Coming of age. Trends Plant Sci. 17, 616–623

Tamkun JW, Deuring R, Scott M.P, Kissinger M, Pattatucci AM, Kaufman TC, Kennison JA (1992) brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell 68(3): 561-572. doi: 10.1016/0092-8674(92)90191-e.

Taniguchi Y (2016). The Bromodomain and Extra-Terminal Domain (BET) Family: Functional Anatomy of BET Paralogous Proteins. Int. J. Mol. Sci. 17(11), 1849. https://doi.org/10.3390/ijms17111849.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25(24), 4876-4882.

Timothy LB, James J, Charles EG, William SN (2015) "The MEME Suite", Nucleic Acids Res. 43(W1): W39-W49.

Uppal S, Gegonne A, Chen Q, Thompson PS, Cheng D, Mu J, Meerzaman D, Misra HS, Singer DS. (2019). The Bromodomain Protein 4 Contributes to the Regulation of Alternative Splicing. Cell Rep. 29(8):2450-2460.e5. doi: 10.1016/j.celrep.2019.10.066.

Van Bel M, Diels T, Vancaester E, Kreft L, Botzki A, Van de Peer Y, Coppens F, Vandepoele K. (2018) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. Nucleic Acids Res. 46(D1): D1190-D1196. doi: 10.1093/nar/gkx1002.

Vergara Z, Gutierrez C (2017) Emerging roles of chromatin in the maintenance of genome organization and function in plants. Genome Biol. 18, 96. https://doi.org/10.1186/s13059-017-1236-9.

Wang W, Zhang X, Deng F, Yuan R, Shen F. (2017) Genome-wide characterization and expression analyses of superoxide dismutase (SOD) genes in Gossypium hirsutum. BMC Genom. 18(1):376. doi: 10.1186/s12864-017-3768-5.

Withers J, Dong X. (2017) Post-translational regulation of plant immunity. Curr. Opin. Plant Biol. 38:124-132. doi: 10.1016/j.pbi.2017.05.004

Zeng L, Zhou MM (2002). Bromodomain: an acetyl-lysine binding domain. FEBS Lett. 513(1), 124-128.

Zhang H, Lang Z, Zhu JK. (2018) Dynamics and function of DNA methylation in plants. Nat. Rev. Mol. Cell Biol. 19(8):489-506. doi: 10.1038/s41580-018-0016-z.

Zhang GQ, Liu KW, Li Z, Lohaus R, Hsiao YY, Niu SC, Wang JY, Lin YC, Xu Q, Chen LJ, Yoshida K, Fujiwara S, Wang ZW, Zhang YQ, Mitsuda N, Wang M, Liu GH, Pecoraro L, Huang HX, Xiao XJ, Lin M, Wu XY, Wu WL, Chen YY, Chang SB, Sakamoto S, Ohme-Takagi M, Yagi M, Zeng SJ, Shen CY, Yeh CM, Luo YB, Tsai WC, Van de Peer Y, Liu ZJ. The Apostasia genome and the evolution of orchids. Nature 549(7672):379-383. doi: 10.1038/nature23897.

Zhao S, Zhang B, Yang M, Zhu J, Li H (2018) Systematic profiling of histone readers in Arabidopsis thaliana. Cell Rep. 22, 1090-1102.

Table 1. An overview of characteristics of bromodomain containing genes (Brd-genes) in *Arabidopsis thaliana* genome.

S. No.	Locus No[1] and Designation[2]	Chr No	Gene Length (bp)	Number of Transcripts, IDs and (Designation)		CDS Length (bp)	Protein Length (aa)	Molecular Weight (Da)	Isoelectric point (pI)
1	AT1G05910 (AtBrd1)	1	6500	1	AT1G05910.1 (AtBrd1.1)	3633	1210	133782.5	5.64
2	AT1G06230 (AtBrd2)		4094	4	AT1G06230.1 (AtBrd2.1)	2301	766	84093.9	5.01
					AT1G06230.2 (AtBrd2.2)*	2301	766	84093.9	5.01
					AT1G06230.3 (AtBrd2.3)*	2301	766	84093.9	5.01
					AT1G06230.4 (AtBrd2.4)*	2301	766	84093.9	5.01
3	AT1G17790 (AtBrd3)[BD1]		2290	1	AT1G17790.1 (AtBrd3.1)	1464	487	53454.3	6.66
4	AT1G20670 (AtBrd4)[BD2]		3960	1	AT1G20670.1 (AtBrd4.1)	1959	652	72955.5	6.86
5	AT1G32750 (AtBrd5)		10519	1	AT1G32750.1 (AtBrd5.1)	5760	1919	217191.7	5.55
6	AT1G58025 (AtBrd6)		4577	6	AT1G58025.1 (AtBrd6.1)	1719	572	64849.3	6.80
					AT1G58025.2 (AtBrd6.2)**	1749	582	66001.7	6.63
					AT1G58025.3 (AtBrd6.3)*	1722	573	64920.4	6.80
					AT1G58025.4 (AtBrd6.4)*	1722	573	64920.4	6.80
					AT1G58025.5 (AtBrd6.5)*	1722	573	64920.4	6.80
					AT1G58025.6 (AtBrd6.6)**	1719	572	64849.3	6.80
7	AT1G61215 (AtBrd7)		2975	2	AT1G61215.1 (AtBrd7.1)	1428	475	52706.5	10.29
					AT1G61215.2 (AtBrd7.2)**	1371	456	50768.3	10.44
8	AT1G73150 (AtBrd8)[BD1]		2434	2	AT1G73150.1 (AtBrd8.1)	1386	461	50811.8	6.29
					AT1G73150.2 (AtBrd8.2)**	1299	432	48308.8	6.66
9	AT1G76380 (AtBrd9)[BD2]		3592	3	AT1G76380.1 (AtBrd9.1)	1740	579	64850.0	7.43
					AT1G76380.2 (AtBrd9.2)**	1743	580	64907.1	7.43
					AT1G76380.3 (AtBrd9.3)	1740	579	64792.0	7.80
10	AT2G34900 (AtBrd10)	2	2672	2	AT2G34900.1 (AtBrd10.1)	1161	386	43441.9	6.30
					AT2G34900.2 (AtBrd10.2)**	831	276	31625.9	8.37
11	AT2G42150 (AtBrd11)[BD3]		2375	1	AT2G42150.1 (AtBrd11.1)	1896	631	70445.1	8.61
12	AT2G44430 (AtBrd12)[BD4]		2995	1	AT2G44430.1 (AtBrd12.1)	1941	646	72404.9	8.86
13	AT2G46020 (AtBrd13)		9523	6	AT2G46020.1 (AtBrd13.1)	6579	2192	245437.4	9.30
					AT2G46020.2 (AtBrd13.2)**	6582	2193	245467.4	9.23
					AT2G46020.3 (AtBrd13.3)**	6582	2193	245467.4	9.23
					AT2G46020.4 (AtBrd13.4)**	6582	2193	245467.4	9.23
					AT2G46020.5 (AtBrd13.5)*	6579	2192	245437.4	9.30
					AT2G46020.6 (AtBrd13.6)*	6579	2192	245437.4	9.30
14	AT2G47410 (AtBrd14)		9156	6	AT2G47410.1 (AtBrd14.1)	4563	1520	171534.9	7.07
					AT2G47410.2 (AtBrd14.2)**	4560	1519	171447.9	7.07
					AT2G47410.3 (AtBrd14.3)**	4047	1348	151704.5	7.05
					AT2G47410.4 (AtBrd14.4)**	4017	1338	150692.5	7.11
					AT2G47410.5 (AtBrd14.5)**	4593	1530	172546.9	7.02
					AT2G47410.6 (AtBrd14.6)**	4047	1348	151704.5	7.05
15	AT3G01770 (AtBrd15)[BD5]	3	3509	2	AT3G01770.1 (AtBrd15.1)	1863	620	69880.7	5.11
					AT3G01770.2 (AtBrd15.2)**	1470	489	54637.2	6.74
16	AT3G19040 (AtBrd16)		8593	2	AT3G19040.1 (AtBrd16.1)	5361	1786	202250.4	7.66
					AT3G19040.2 (AtBrd16.2)**	5379	1792	202813.7	7.37
17	AT3G27260 (AtBrd17)		5232	4	AT3G27260.1 (AtBrd17.1)	2442	813	90232.5	4.53
					AT3G27260.2 (AtBrd17.2)**	2295	764	85090.1	4.58
					AT3G27260.3 (AtBrd17.3)**	2469	822	91181.7	4.52
					AT3G27260.4 (AtBrd17.4)**	2157	718	79762.2	4.75
18	AT3G52280 (AtBrd18)		2455	2	AT3G52280.1 (AtBrd18.1)	1110	369	42392.4	8.01
					AT3G52280.2 (AtBrd18.2)**	1161	386	44382.7	8.56
19	AT3G54610 (AtBrd19)		4248	1	AT3G54610.1 (AtBrd19.1)	1707	568	63123.0	6.42
20	AT3G57980 (AtBrd20) [BD3]		2504	2	AT3G57980.1 (AtBrd20.1)	1953	650	72310.1	9.86

					AT3G57980.2 (AtBrd20.2)**	1959	652	72551.4	9.86
21	AT3G60110 (AtBrd21)[BD4]		3923	1	AT3G60110.1 (AtBrd21.1)	1926	641	72045.1	9.74
22	AT5G10550 (AtBrd22)[BD6]	5	1919	1	AT5G10550.1 (AtBrd22.1)	1557	518	64102.7	6.55
23	AT5G14270 (AtBrd23)[BD5]		4314	3	AT5G14270.1 (AtBrd23.1)	2067	688	75894.3	4.61
					AT5G14270.2 (AtBrd23.2)**	2070	689	75991.4	4.61
					AT5G14270.3 (AtBrd23.3)**	2067	688	75894.3	4.61
24	AT5G46550 (AtBrd24)		2743	1	AT5G46550.1 (AtBrd24.1)	1485	494	55618.7	9.82
25	AT5G49430 (AtBrd25)		9715	2	AT5G49430.1 (AtBrd25.1)	5034	1677	186918.3	7.08
					AT5G49430.2 (AtBrd25.2)*	5034	1677	186918.3	7.08
26	AT5G55040 (AtBrd26)		5139	2	AT5G55040.1 (AtBrd26.1)	2751	916	103414.1	6.00
					AT5G55040.2 (AtBrd26.2)*	2751	916	103414.1	6.00
27	AT5G63320 (AtBrd27)		5343	3	AT5G63320.1 (AtBrd27.1)	3186	1061	118972.0	4.43
					AT5G63320.2 (AtBrd27.2)**	1434	477	53000.4	7.20
					AT5G63320.3 (AtBrd27.3)**	1434	477	53000.4	7.20
28	AT5G65630 (AtBrd28)[BD6]		3743	1	AT5G65630.1 (AtBrd28.1)	1770	589	65076.7	6.70

Note: [1]: The locus numbers are as per TAIR database (The Arabidopsis Infromation Resource at https://www.arabidopsis.org/); [2]: Simplified designation of the genes in a sequential manner w.r.t. chromosome number; BD1-6: Block duplication events 1-6; Chr No: Chromosome number; CDS: Coding DNA sequence; *: Alternatively spliced transcripts with differences in UTR; **: Alternatively spliced transcripts with difference in exons/coding region.

Table 2. An overview of characteristics of bromodomain containing genes (Brd-genes) in *Oryza sativa* genome.

S. No.	Locus No[1] and Designation[2]	Chr No	Gene Length (bp)	Number of Transcripts, Transcript IDs and (Designation)		CDS Length (bp)	Protein Length (aa)	Molecular Weight (Da)	Isoelectric Point (pI)
1	LOC_Os01g11580 (OsBrd1) BD1	1	5988	2	LOC_Os01g11580.1 (OsBrd1.1)	1068	355	39749.8	4.96
					LOC_Os01g11580.2 (OsBrd1.2)**	663	220	24655.7	4.42
2	LOC_Os01g46040 (OsBrd2) BD1		4046	1	LOC_Os01g46040.1 (OsBrd2.1)	717	238	26206.1	6.95
3	LOC_Os02g02290 (OsBrd3)	2	10570	1	LOC_Os02g02290.1 (OsBrd3.1)	6603	2200	246212.0	9.07
4	LOC_Os02g09920 (OsBrd4)		5584	1	LOC_Os02g09920.1 (OsBrd4.1)	2940	979	110342.0	4.79
5	LOC_Os02g15220 (OsBrd5)		7999	3	LOC_Os02g15220.1 (OsBrd5.1)	1971	656	71476.3	10.00
					LOC_Os02g15220.2 (OsBrd5.2)*	1971	656	71476.3	10.00
					LOC_Os02g15220.4 (OsBrd3.3)**	1875	624	68946.6	10.20
6	LOC_Os02g38980 (OsBrd6)		5209	5	LOC_Os02g38980.1 (OsBrd6.1)	2145	714	78630.4	4.84
					LOC_Os02g38980.3 (OsBrd6.3)**	1728	575	63133.5	5.62
					LOC_Os02g38980.4 (OsBrd6.4)**	1701	566	62093.3	5.61
					LOC_Os02g38980.5 (OsBrd6.5)**	1443	480	52786.3	6.67
					LOC_Os02g38980.6 (OsBrd6.6)**	1302	433	48256.5	8.01
7	LOC_Os03g03870 (OsBrd7)	3	6059	1	LOC_Os03g03870.1 (OsBrd7.1)	1452	483	52387.9	8.43
8	LOC_Os03g19340 (OsBrd8)		16548	1	LOC_Os03g19340.1 (OsBrd8.1)	4881	1626	183000	6.64
9	LOC_Os03g21450 (OsBrd9)		2605	1	LOC_Os03g21450.1 (OsBrd9.1)	1677	558	59620.1	10.63
10	LOC_Os04g53130 (OsBrd10) TD1	4	3427	1	LOC_Os04g53130.1 (OsBrd10.1)	1068	355	39412.4	4.81
11	LOC_Os04g53170 (OsBrd11) TD1-BD2		2150	1	LOC_Os04g53170.1 (OsBrd11.1)	1371	456	50290.7	6.94
12	LOC_Os06g04640 (OsBrd12)	6	5300	3	LOC_Os06g04640.1 (OsBrd12.1)	1083	360	40658.1	7.05
					LOC_Os06g04640.2 (OsBrd12.2)**	819	272	31690.2	6.19
					LOC_Os06g04640.3 (OsBrd12.3)**	684	227	26418.0	6.32
13	LOC_Os06g24870 (OsBrd13) BD3		5765	3	LOC_Os06g24870.1 (OsBrd13.1)	1137	378	40896.8	8.15
					LOC_Os06g24870.2 (OsBrd13.2)*	1137	378	40896.8	8.15
					LOC_Os06g24870.3 (OsBrd13.3)*	1137	378	40896.8	8.15
14	LOC_Os06g43790 (OsBrd14)		14210	1	LOC_Os06g43790.1 (OsBrd14.1)	5475	1824	206046	5.46
15	LOC_Os07g32420 (OsBrd15)	7	8036	3	LOC_Os07g32420.1 (OsBrd15.1)*	1455	484	53992.3	8.33
					LOC_Os07g32420.2 (OsBrd15.2)*	1455	484	53992.3	8.33
					LOC_Os07g32420.1 (OsBrd15.3)*	1455	484	53992.3	8.33
16	LOC_Os07g37800 (OsBrd16)		4117	3	LOC_Os07g37800.1 (OsBrd16.1)	1485	494	53351.1	10.54
					LOC_Os07g37800.2 (OsBrd16.2)**	1407	464	50177.5	10.50
					LOC_Os07g37800.3 (OsBrd16.3)**	1035	344	37209.5	8.15
17	LOC_Os08g01794 (OsBrd17) BD3	8	7581	2	LOC_Os08g01794.1 (OsBrd17.1)	1773	590	65099.3	5.49
					LOC_Os08g01794.2 (OsBrd17.2)*	1773	590	65099.3	5.49
18	LOC_Os08g09340 (OsBrd18) BD2		2119	2	LOC_Os08g09340.1 (OsBrd18.1)	1446	481	53638.1	6.63
					LOC_Os08g09340.2 (OsBrd18.2)**	1260	419	47311.1	9.65
19	LOC_Os08g39980 (OsBrd19) BD4		3367	1	LOC_Os08g39980.1 (OsBrd19.1)	1983	660	68668.1	9.38
20	LOC_Os09g33980 (OsBrd20) BD4	9	7027	1	LOC_Os09g33980.1 (OsBrd20.1)	3597	1198	133903	6.49
21	LOC_Os09g37760 (OsBrd21)		5783	1	LOC_Os09g37760.1 (OsBrd21.1)	1245	414	45957.8	9.84
22	LOC_Os10g28040 (OsBrd22)	10	6439	1	LOC_Os10g28040.1 (OsBrd22.1)	1536	511	56685.3	6.79

Note: [1]: The locus numbers are as per RGAP database (Rice Genome Annotation Project at http://rice.uga.edu/); [2]: Simplified designation of the genes in a sequential manner w.r.t to chromosome number; BD1-4: Block duplication events 1-6; TD1: Tandem duplication event; Chr No: Chromosome number; CDS: Coding DNA sequence; *: Alternatively spliced transcripts with differences in UTR; **: Alternatively spliced transcripts with difference in exons/coding region

Figure legends

Figure 1: Circle-plot representation of distribution of Brd-genes, and block/tandem duplicated gene-pairs in *A. thaliana* (A) and *O. sativa* (B) genomes. Designations Chr1-5 and Chr1-12 indicate the chromosome numbers, while AtBrd1-28 and OsBrd1-22 indicate the Brd-genes in the two species, as per details in Table 1 and Table 2. Curved lines with double-sided arrowheads indicate the block duplicated Brd-gene pairs, and the type of event is indicated by designations 'BD' (block duplication) and 'TD' (tandem duplication).

Figure 2: Length and heterogeneity of Brd-genes of *A. thaliana* (A) and *O. sativa* (B), as per details in Table 1 and Table 2. Scale on the top indicates the length (kilobase, kb), orange colored region indicates upstream/downstream regions including UTRs, green boxes show the exons, while introns is indicated by dashed line. Duplicate Brd-gene pairs are indicated with the designations 'BD' (block duplication) and 'TD' (tandem duplication) in the names, and arranged one below the other for ease of comparison.

Figure 3: (A) Domain heterogeneity among BRD-proteins of *A. thaliana* (top panel) and *O. sativa* (bottom panel), as per details in Table 1 and Table 2. Major groups based on domain combinations are indicated by numerals 'i-iv', while the domains/functional sites are shown in different colors. Scale on the top indicates the protein length (number of amino acids) and the relative positions of the domains/functional sites. Duplicate BRD-pairs are indicated with the designations 'BD' (block duplication) and 'TD' (tandem duplication) in the names, and arranged one below the other for ease of comparison.

(B) Motif heterogeneity among Brd-proteins of *A. thaliana* (top panel) and *O. sativa* (bottom panel), as per details listed in Table 1 and Table 2. Major groups based on different combination of motifs M1-M15 are indicated by designations 'a-f', while the motifs (listed in Supplementary Table 1 and 2) are shown in different colors. Scale on the top indicates the protein length (number of amino acids) and the relative positions of the motifs. Duplicate BRD-pairs are indicated with the designations 'BD' (block duplication) and 'TD' (tandem duplication) in the names, and arranged one below the other for ease of comparison.

Figure 4: Domain heterogeneity of AS-transcripts encoded BRD-isoforms of *A. thaliana* (A) and *O. sativa* (B), as per details in the Table 1 and Table 2. Different designations indicate constitutive (.1) and alternative (0.2 to 0.6) isoforms, while domains/functional sites are shown in different color codes. Scale on the top indicates the protein length (number of amino acids) and the relative positions of the domains/functional sites. Designations 'BD' (block duplication), and 'TD' (tandem duplication) in the names of BRD-homolog indicate the type of duplication event involved.

Figure 5: Diversity of *cis*-regulatory elements in the upstream region (up to -2000 bp) of *A. thaliana* (A) and *O. sativa* (B) Brd-genes, as per details in Table 1 and Table 2. The analysis was carried out at PlantCARE database, and the major categories of *cis*-elements are represented by different colors. The duplicated Brd-gene pairs are shown by double sided arrows and designations 'BD' (block duplication), and 'TD' (tandem duplication) indicate the type of event involved, and are arranged side-by-side for easy comparison of *cis*-element differences.

Figure 6: Comparative organization of CpG islands, transcription factor binding sites (TFBS, 1-6), and repetitive motifs (R) in the upstream region (up to -2000bp) of block/tandem duplicated Brd-genes of *A. thaliana* (A) and *O. sativa* (B), as per analysis on PlantPAN3.0 database. The Brd-gene designations and corresponding locus numbers are indicated. Scale on the top indicates the length of the promoter region and relative positions of CpG islands, TFBS and repeat motifs, while the designations 'BD' (block duplication), and 'TD' (tandem duplication) indicate the type of duplication event involved.

Figure 7: Heatmap based analysis of RNA-Seq data for tissue-specific expression pattern of Brd-genes of *A. thaliana* (A) and *O. sativa* (B), as per details in the Table 1 and Table 2. The Brd-genes are listed on the left side, while the tissue types are indicated on the top of the heatmap. Duplicate Brd-gene pairs are indicated with names containing designations 'BD' (block duplication) and 'TD' (tandem duplication), and are arranged one below the other for ease of comparison. A color gradient scale indicates the expression level from blue (low) to red (high).

Figure 8: Multiple sequence alignment of the amino acid sequences of the bromodomain-region of 28 AtBRD and 22 OsBRD-homologs, as per details in the Table 1 and Table 2. Scale on the top indicates the sites in the MSA file. The duplicate-pairs are indicated with names with designations 'BD' (block duplication) and 'TD' (tandem duplication), while numerals (I-VI) indicates the six major clusters as per phylogenetic analysis (Figure 9).

Figure 9: Bootstrap consensus phylogenetic tree based on the bromodomain regions of BRD-homologs of *A. thaliana* and *O. sativa* including the block/tandem duplicates, generated by neighbor-joining method in MEGA-X software (A). Numerals (I-VI) indicate six major clusters, while sub-clusters are indicated letters (A-E). The numbers at the nodes represents bootstrap values (in %) for a 500-replicate analysis. Both the designations and locus numbers (in parenthesis) of the BRD-homologs of two species are shown, with AtBRDs in bold font and OsBRDs in regular font. Designations 'BD' (block duplication) and 'TD' (tandem duplication) in the names indicate the type of duplication event involved.

(B) Radiation tree of the bromodomain regions of *A. thaliana*, *O. sativa*, some representative human BRD-homologs. Seven major groups are indicated by designation GI-GVII, while subgroups are indicated letters (A-H). The AtBRDs are written in bold, OsBRDs in regular font, and human BRDs in grey font. Designations 'BD' (block duplication) and 'TD' (tandem duplication) indicate the type of duplication event involved, and numerals (1,2) indicate the two different bromodomain of the dual-Brd domain containing homologs.

Figure 10: Homology modeling of bromodomain regions of Brd-homologs of *A. thaliana* and *O. sativa* and structural superposition at SWISS-MODEL workspace. (A) AtBRD17, (B) AtBRD4, ArBRD7, OsBRD21 and OsBRD16, (C) OsBRD13 and OsBRD17, (D) AtBRD14 and OsBRD8, (E) AtBRD16 and OsBRD14, (F) OsBRD4 and OsBRD7. Homology models of duplicate Brd-pairs, and superposition analysis: (G) AtBRD11-AtBRD20, (H) AtBRD15-AtBRD23, (I) OsBRD1-OsBRD2, and (J) OsBRD19-OsBRD20. Important regions of the bromodomain (α-helices: αZ, αA, αB, αC and connecting loops: ZA, AB and BC) are indicated in (A), whereas the structural variations due to sequence/length heterogeneity or duplication events are indicated by arrows.

Figure 11: Comparative assessment of different types of duplication events affecting Brd-gene copy numbers among genomes of lower photosynthetic organisms (A), monocots (B) and dicots (C), as per analysis at PLAZA monocots and dicots databases. Types of duplication events are indicated by different shades and designations ND (no evidence of duplication), TD (tandem duplication), BD (block duplication), and TD+BD (combine tandem and block duplication).

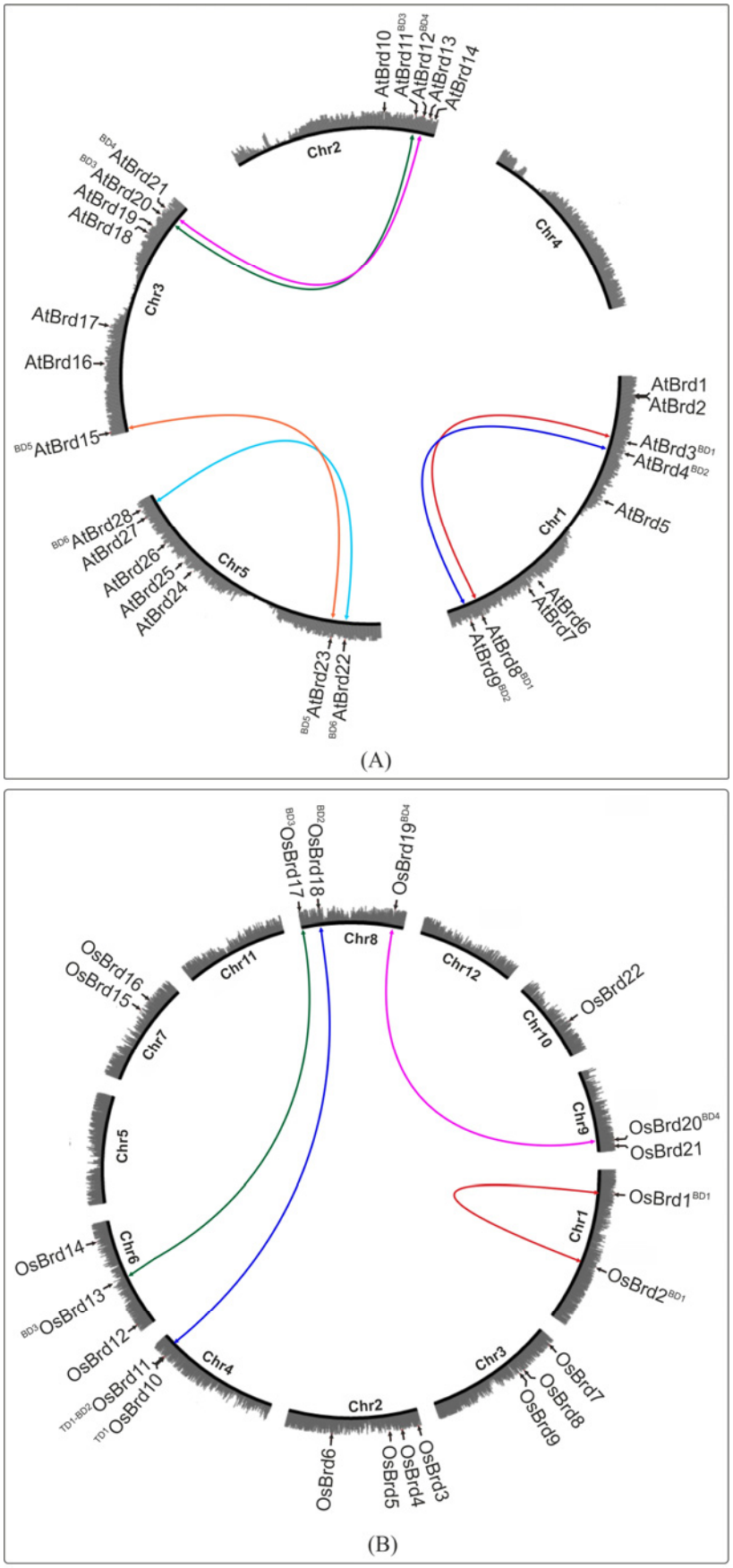
Chr1
Chr2
Chr3
Chr4
Chr5
AtBrd1
AtBrd2
AtBrd3BD1
AtBrd4BD2
AtBrd5
AtBrd6
AtBrd7
AtBrd8BD1
AtBrd9BD2
AtBrd10
AtBrd11BD3
AtBrd12BD4
AtBrd13
AtBrd14
BD4AtBrd21
BD3AtBrd20
AtBrd19
AtBrd18
AtBrd17
AtBrd16
BD5AtBrd15
BD6AtBrd28
AtBrd27
AtBrd26
AtBrd25
AtBrd24
BD5AtBrd23
BD6AtBrd22
(A)
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
OsBrd1BD1
OsBrd2BD1
OsBrd7
OsBrd8
OsBrd9
OsBrd3
OsBrd4
OsBrd5
OsBrd6
TD1-BD2OsBrd11
TD1OsBrd10
OsBrd12
BD3OsBrd13
OsBrd14
OsBrd15
OsBrd16
OsBrd20BD4
OsBrd21
OsBrd22
OsBrd19BD4
OsBrd18BD2
BD1OsBrd17
(B)

Figure 1

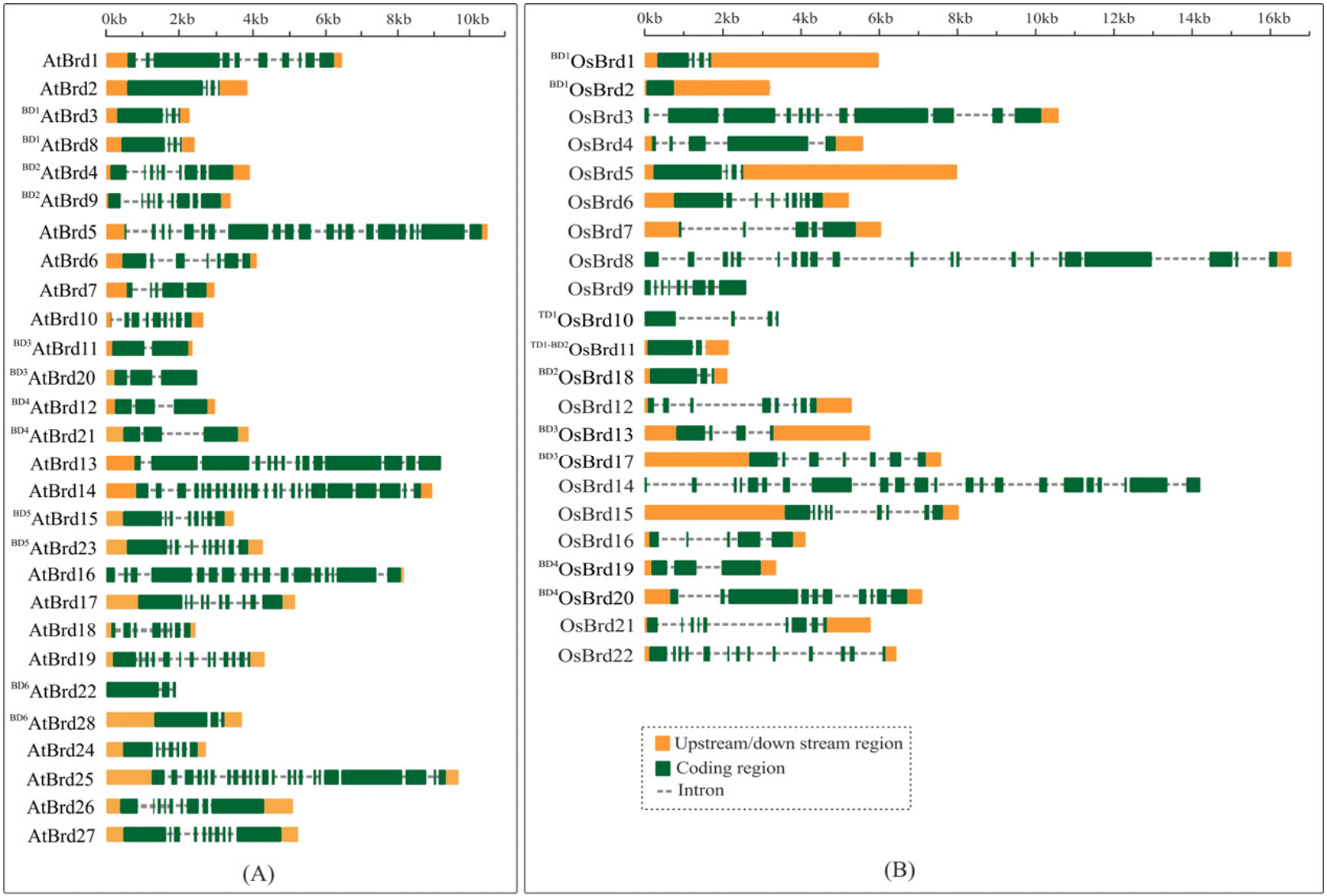
0kb
2kb
4kb
6kb
8kb
10kb
AtBrd1
AtBrd2
BD1 AtBrd3
BD1 AtBrd8
BD2 AtBrd4
BD2 AtBrd9
AtBrd5
AtBrd6
AtBrd7
AtBrd10
BD3 AtBrd11
BD3 AtBrd20
BD4 AtBrd12
BD4 AtBrd21
AtBrd13
AtBrd14
BD5 AtBrd15
BD5 AtBrd23
AtBrd16
AtBrd17
AtBrd18
AtBrd19
BD6 AtBrd22
BD6 AtBrd28
AtBrd24
AtBrd25
AtBrd26
AtBrd27
(A)
0kb
2kb
4kb
6kb
8kb
10kb
12kb
14kb
16kb
BD1 OsBrd1
BD1 OsBrd2
OsBrd3
OsBrd4
OsBrd5
OsBrd6
OsBrd7
OsBrd8
OsBrd9
TD1 OsBrd10
TD1-BD2 OsBrd11
BD2 OsBrd18
OsBrd12
BD3 OsBrd13
BD3 OsBrd17
OsBrd14
OsBrd15
OsBrd16
BD4 OsBrd19
BD4 OsBrd20
OsBrd21
OsBrd22
Upstream/down stream region
Coding region
Intron
(B)

Figure 2

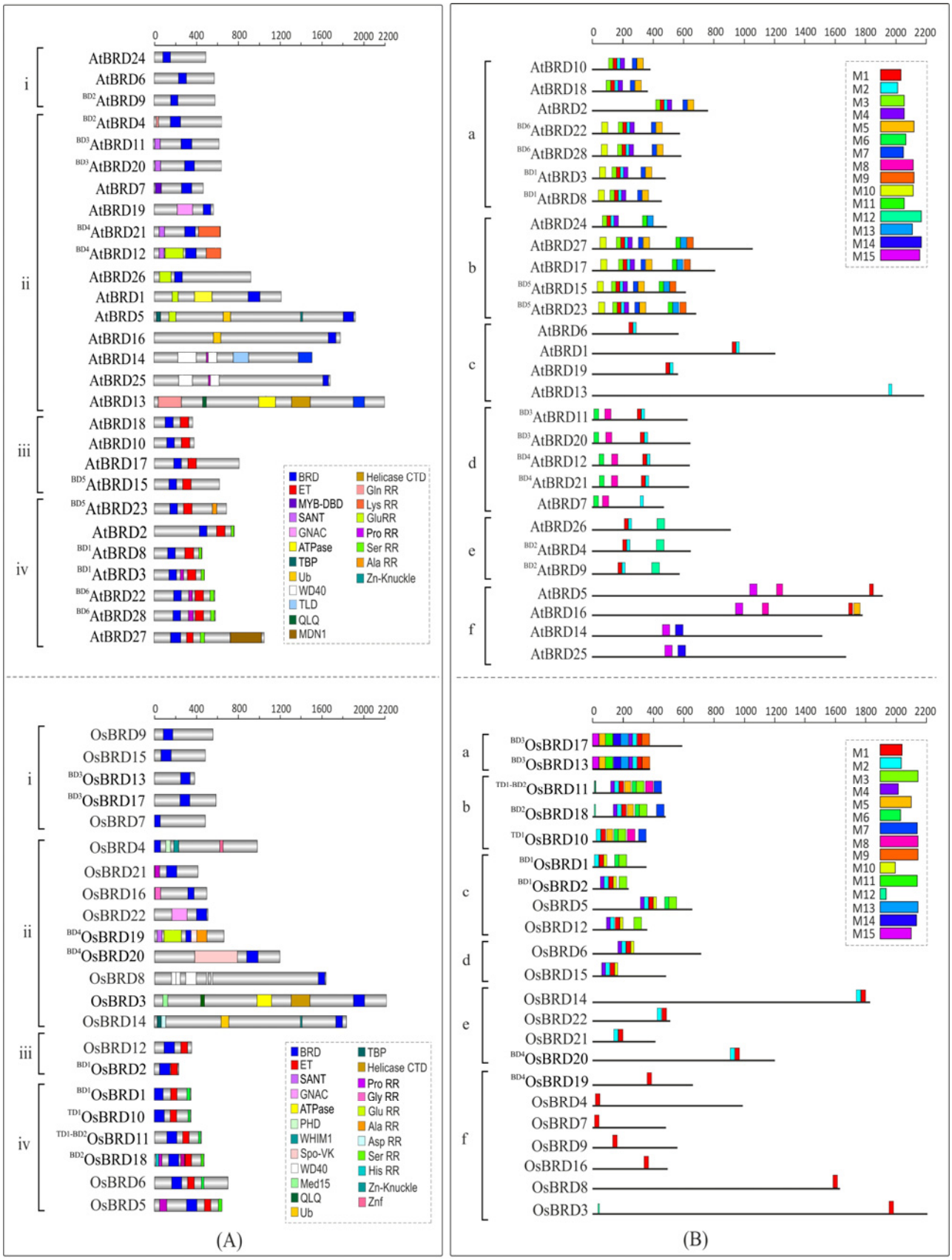
0 400 800 1200 1600 2000 2200
i
AtBRD24
AtBRD6
BD2AtBRD9
ii
BD2AtBRD4
BD3AtBRD11
BD3AtBRD20
AtBRD7
AtBRD19
BD4AtBRD21
BD4AtBRD12
AtBRD26
AtBRD1
AtBRD5
AtBRD16
AtBRD14
AtBRD25
AtBRD13
iii
AtBRD18
AtBRD10
AtBRD17
BD5AtBRD15
iv
BD5AtBRD23
AtBRD2
BD1AtBRD8
BD1AtBRD3
BD6AtBRD22
BD6AtBRD28
AtBRD27
BRD
ET
MYB-DBD
SANT
GNAC
ATPase
TBP
Ub
WD40
TLD
QLQ
MDN1
Helicase CTD
Gln RR
Lys RR
GluRR
Pro RR
Ser RR
Ala RR
Zn-Knuckle
0 200 400 600 800 1000 1200 1400 1600 1800 2000 2200
a
AtBRD10
AtBRD18
AtBRD2
BD6AtBRD22
BD6AtBRD28
BD1AtBRD3
BD1AtBRD8
b
AtBRD24
AtBRD27
AtBRD17
BD5AtBRD15
BD5AtBRD23
c
AtBRD6
AtBRD1
AtBRD19
AtBRD13
d
BD3AtBRD11
BD3AtBRD20
BD4AtBRD12
BD4AtBRD21
AtBRD7
e
AtBRD26
BD2AtBRD4
BD2AtBRD9
f
AtBRD5
AtBRD16
AtBRD14
AtBRD25
M1
M2
M3
M4
M5
M6
M7
M8
M9
M10
M11
M12
M13
M14
M15
(A)
(B)

Figure 3

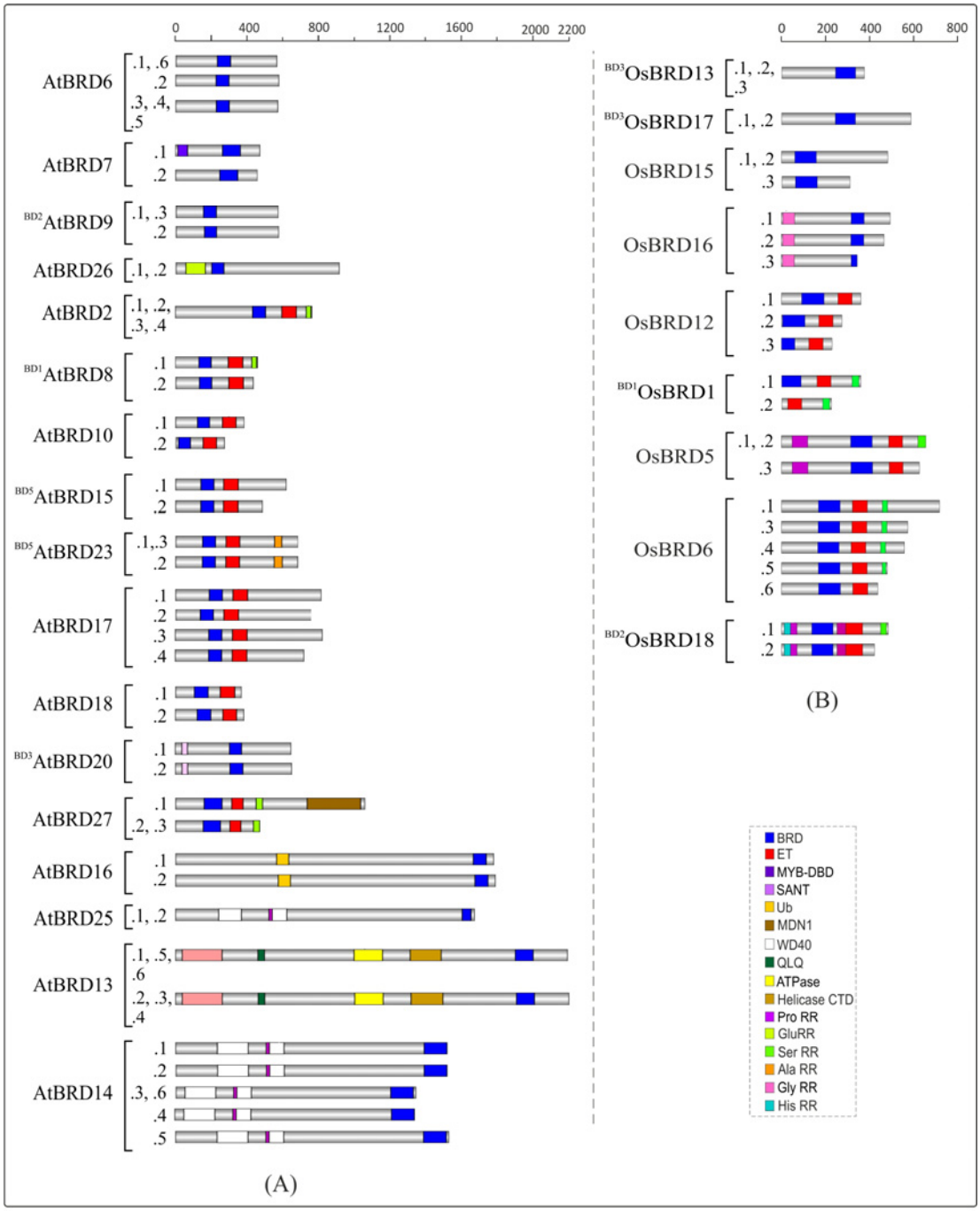
0
400
800
1200
1600
2000
2200
AtBRD6
.1, .6
.2
.3, .4,
.5
AtBRD7
.1
.2
BD2AtBRD9
.1, .3
.2
AtBRD26
.1, .2
AtBRD2
.1, .2,
.3, .4
BD1AtBRD8
.1
.2
AtBRD10
.1
.2
BD5AtBRD15
.1
.2
BD5AtBRD23
.1,.3
.2
AtBRD17
.1
.2
.3
.4
AtBRD18
.1
.2
BD3AtBRD20
.1
.2
AtBRD27
.1
.2, .3
AtBRD16
.1
.2
AtBRD25
.1, .2
AtBRD13
.1, .5,
.6
.2, .3,
.4
AtBRD14
.1
.2
.3, .6
.4
.5
(A)
0
200
400
600
800
BD3OsBRD13
.1, .2,
.3
BD3OsBRD17
.1, .2
OsBRD15
.1, .2
.3
OsBRD16
.1
.2
.3
OsBRD12
.1
.2
.3
BD1OsBRD1
.1
.2
OsBRD5
.1, .2
.3
OsBRD6
.1
.3
.4
.5
.6
BD2OsBRD18
.1
.2
(B)
BRD
ET
MYB-DBD
SANT
Ub
MDN1
WD40
QLQ
ATPase
Helicase CTD
Pro RR
GluRR
Ser RR
Ala RR
Gly RR
His RR

Figure 4

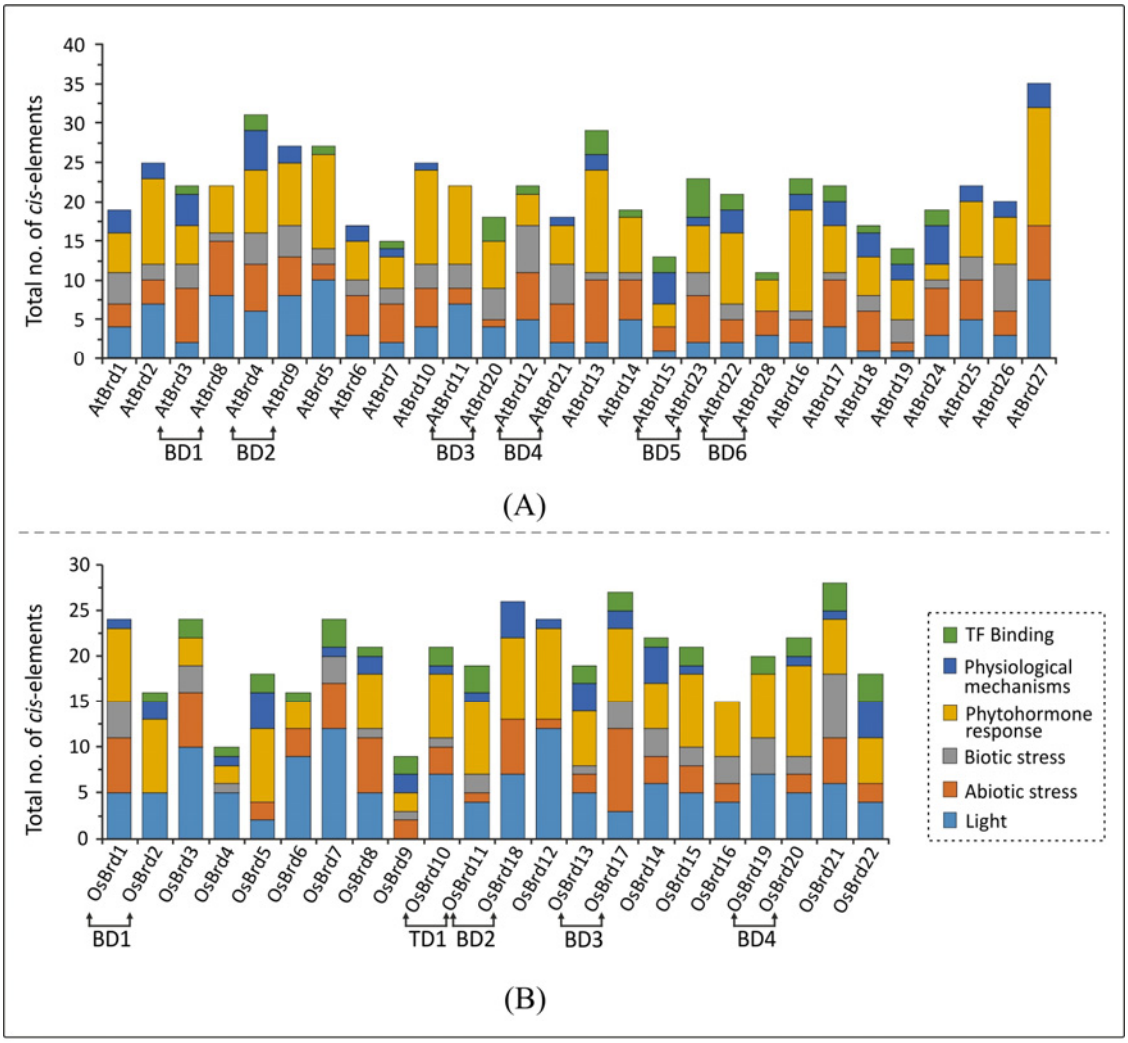
Total no. of cis-elements
40
35
30
25
20
15
10
5
0
AtBrd1
AtBrd2
AtBrd3
AtBrd8
AtBrd4
AtBrd9
AtBrd5
AtBrd6
AtBrd7
AtBrd10
AtBrd11
AtBrd20
AtBrd12
AtBrd21
AtBrd13
AtBrd14
AtBrd15
AtBrd23
AtBrd22
AtBrd28
AtBrd16
AtBrd17
AtBrd18
AtBrd19
AtBrd24
AtBrd25
AtBrd26
AtBrd27
BD1
BD2
BD3
BD4
BD5
BD6
(A)
Total no. of cis-elements
30
25
20
15
10
5
0
OsBrd1
OsBrd2
OsBrd3
OsBrd4
OsBrd5
OsBrd6
OsBrd7
OsBrd8
OsBrd9
OsBrd10
OsBrd11
OsBrd18
OsBrd12
OsBrd13
OsBrd17
OsBrd14
OsBrd15
OsBrd16
OsBrd19
OsBrd20
OsBrd21
OsBrd22
BD1
TD1
BD2
BD3
BD4
TF Binding
Physiological mechanisms
Phytohormone response
Biotic stress
Abiotic stress
Light
(B)

Figure 5

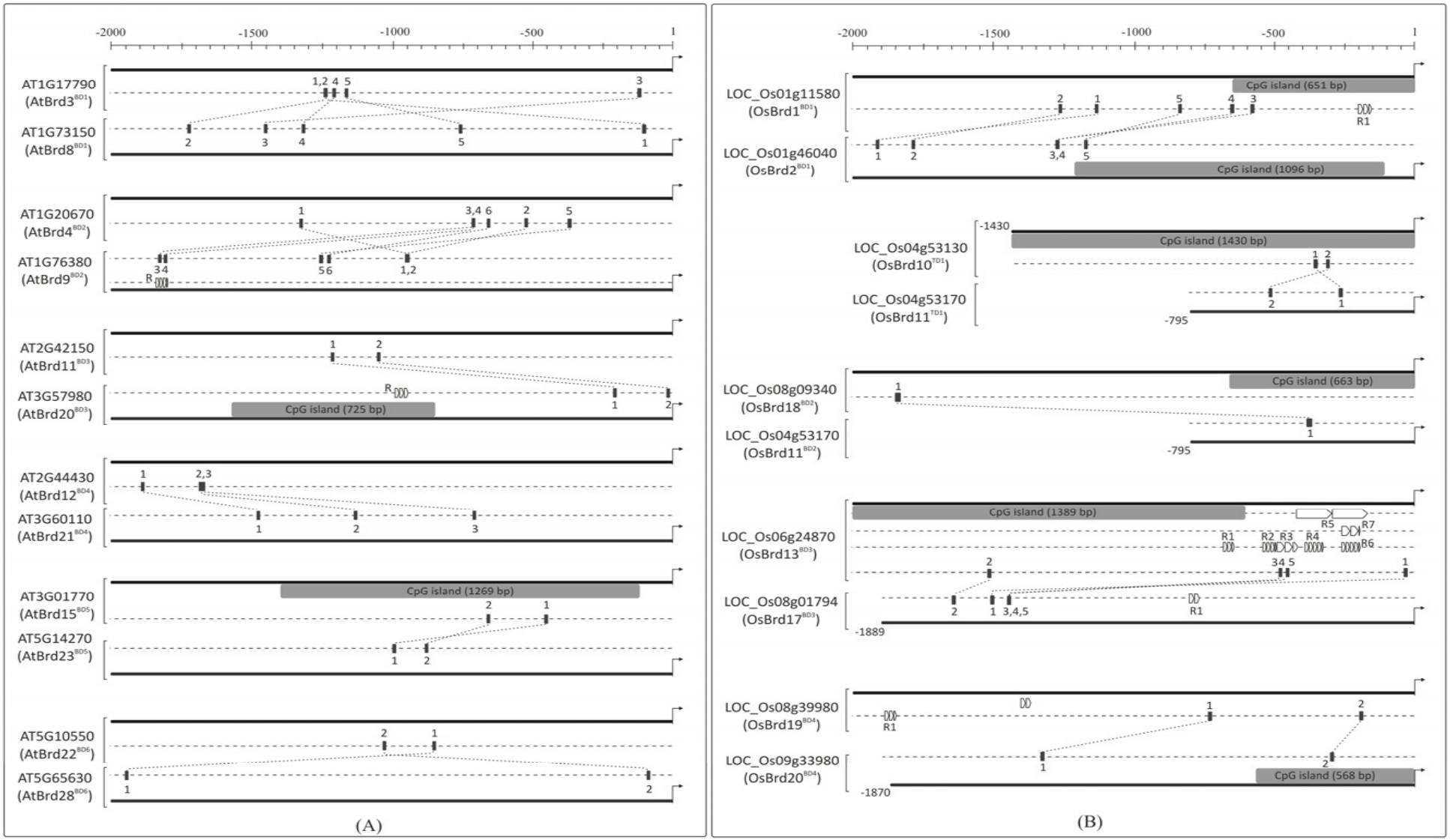
-2000
-1500
-1000
-500
1
AT1G17790 (AtBrd3BD1)
AT1G73150 (AtBrd8BD1)
AT1G20670 (AtBrd4BD2)
AT1G76380 (AtBrd9BD2)
AT2G42150 (AtBrd11BD3)
AT3G57980 (AtBrd20BD3)
CpG island (725 bp)
AT2G44430 (AtBrd12BD4)
AT3G60110 (AtBrd21BD4)
AT3G01770 (AtBrd15BD5)
CpG island (1269 bp)
AT5G14270 (AtBrd23BD5)
AT5G10550 (AtBrd22BD6)
AT5G65630 (AtBrd28BD6)
(A)
LOC_Os01g11580 (OsBrd1BD1)
CpG island (651 bp)
LOC_Os01g46040 (OsBrd2BD1)
CpG island (1096 bp)
-1430
LOC_Os04g53130 (OsBrd10TD1)
CpG island (1430 bp)
LOC_Os04g53170 (OsBrd11TD1)
-795
LOC_Os08g09340 (OsBrd18BD2)
CpG island (663 bp)
LOC_Os04g53170 (OsBrd11BD2)
-795
LOC_Os06g24870 (OsBrd13BD3)
CpG island (1389 bp)
LOC_Os08g01794 (OsBrd17BD3)
-1889
LOC_Os08g39980 (OsBrd19BD4)
LOC_Os09g33980 (OsBrd20BD4)
CpG island (568 bp)
-1870
(B)

Figure 6

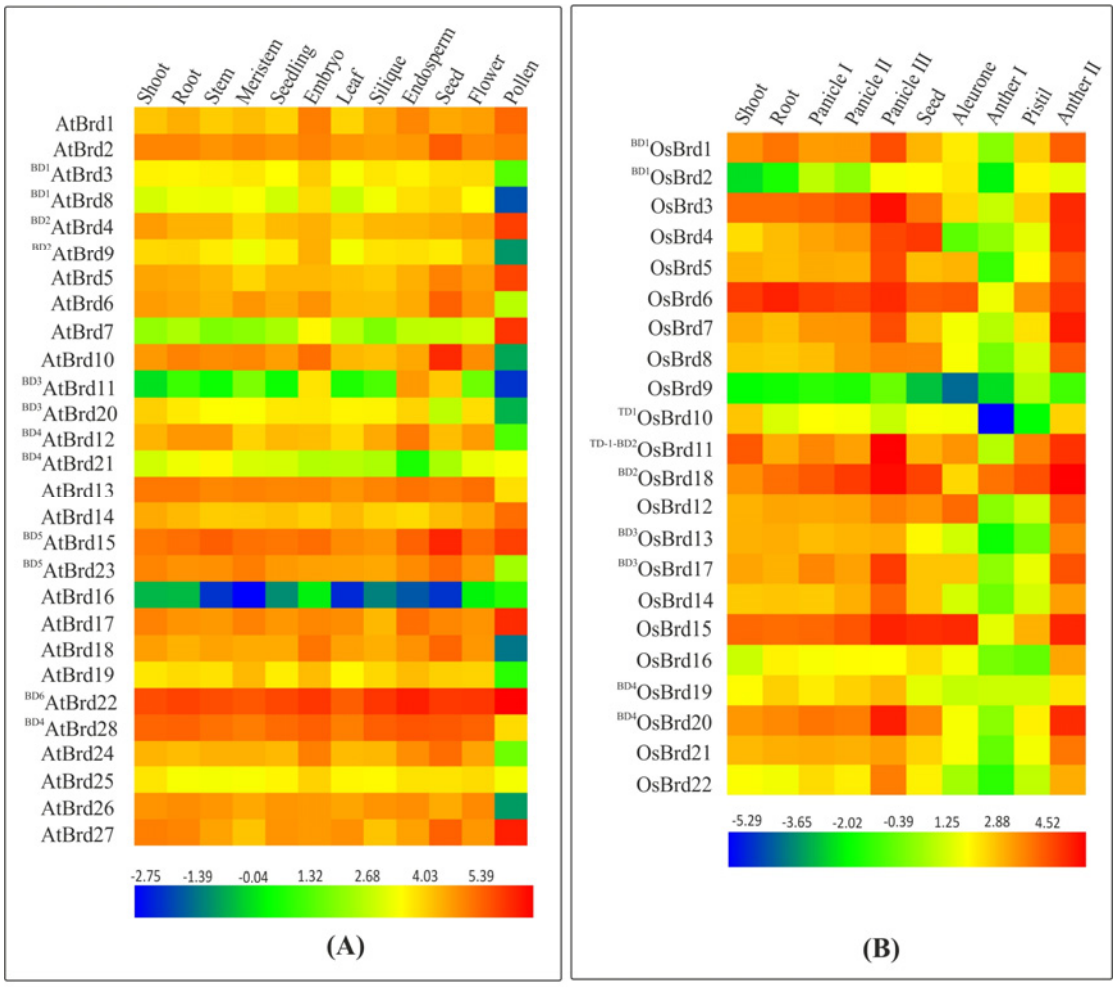
Shoot
Root
Stem
Meristem
Seedling
Embryo
Leaf
Silique
Endosperm
Seed
Flower
Pollen
AtBrd1
AtBrd2
BD1 AtBrd3
BD1 AtBrd8
BD2 AtBrd4
BD2 AtBrd9
AtBrd5
AtBrd6
AtBrd7
AtBrd10
BD3 AtBrd11
BD3 AtBrd20
BD4 AtBrd12
BD4 AtBrd21
AtBrd13
AtBrd14
BD5 AtBrd15
BD5 AtBrd23
AtBrd16
AtBrd17
AtBrd18
AtBrd19
BD6 AtBrd22
BD4 AtBrd28
AtBrd24
AtBrd25
AtBrd26
AtBrd27
-2.75 -1.39 -0.04 1.32 2.68 4.03 5.39
(A)
Shoot
Root
Panicle I
Panicle II
Panicle III
Seed
Aleurone
Anther I
Pistil
Anther II
BD1 OsBrd1
BD1 OsBrd2
OsBrd3
OsBrd4
OsBrd5
OsBrd6
OsBrd7
OsBrd8
OsBrd9
TD1 OsBrd10
TD-1-BD2 OsBrd11
BD2 OsBrd18
OsBrd12
BD3 OsBrd13
BD3 OsBrd17
OsBrd14
OsBrd15
OsBrd16
BD4 OsBrd19
BD4 OsBrd20
OsBrd21
OsBrd22
-5.29 -3.65 -2.02 -0.39 1.25 2.88 4.52
(B)

Figure 7

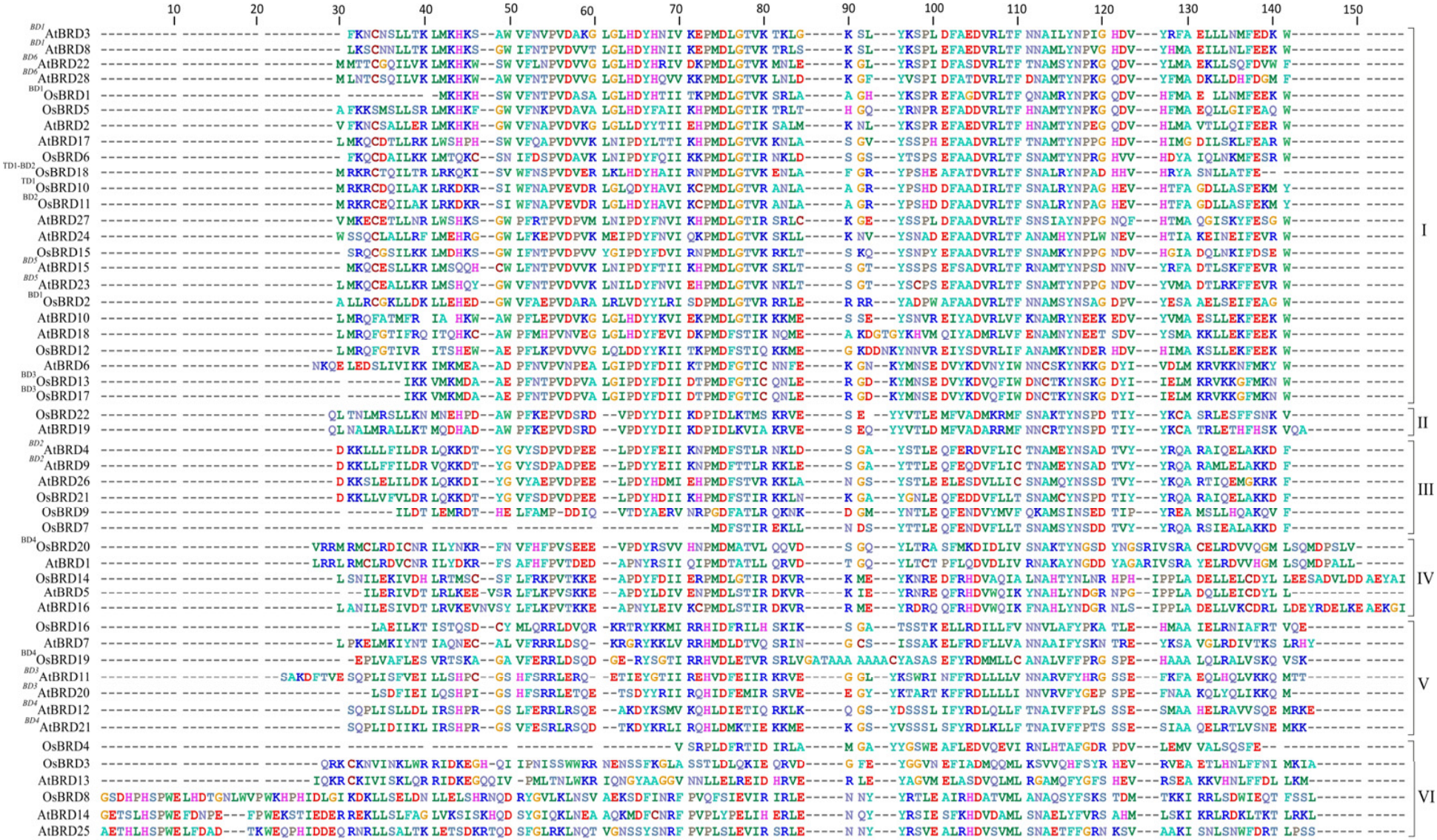
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150
BD1 AtBRD3
BD1 AtBRD8
BD6 AtBRD22
BD6 AtBRD28
BD1 OsBRD1
OsBRD5
AtBRD2
AtBRD17
OsBRD6
TD1-BD2 OsBRD18
TD1 OsBRD10
BD2 OsBRD11
AtBRD27
AtBRD24
OsBRD15
BD5 AtBRD15
BD5 AtBRD23
BD1 OsBRD2
AtBRD10
AtBRD18
OsBRD12
AtBRD6
BD3 OsBRD13
BD3 OsBRD17
OsBRD22
AtBRD19
BD2 AtBRD4
BD2 AtBRD9
AtBRD26
OsBRD21
OsBRD9
OsBRD7
BD4 OsBRD20
AtBRD1
OsBRD14
AtBRD5
AtBRD16
OsBRD16
AtBRD7
BD4 OsBRD19
BD3 AtBRD11
BD3 AtBRD20
BD4 AtBRD12
BD4 AtBRD21
OsBRD4
OsBRD3
AtBRD13
OsBRD8
AtBRD14
AtBRD25
I
II
III
IV
V
VI

Figure 8.

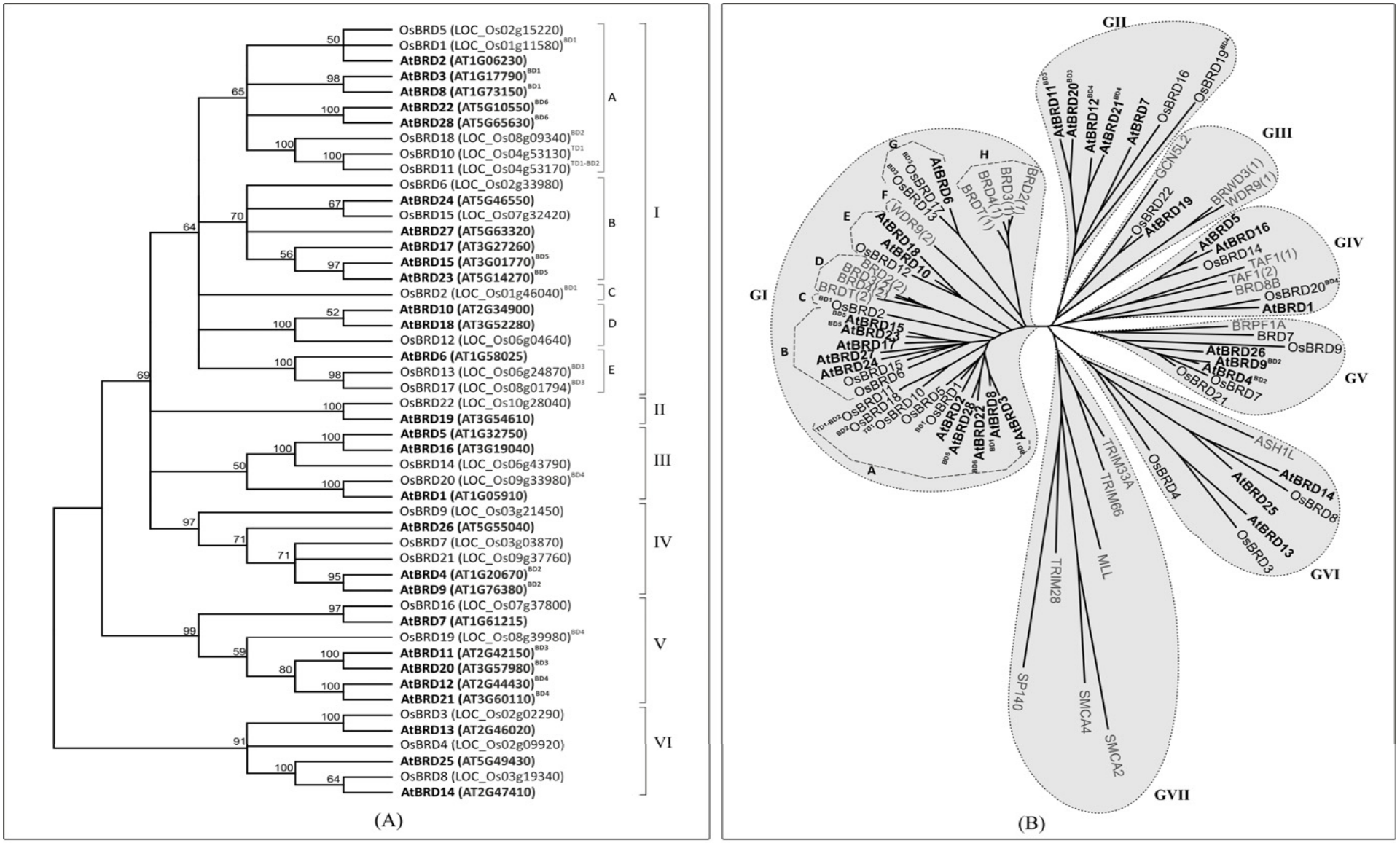
OsBRD5 (LOC_Os02g15220)
OsBRD1 (LOC_Os01g11580)
AtBRD2 (AT1G06230)
AtBRD3 (AT1G17790)
AtBRD8 (AT1G73150)
AtBRD22 (AT5G10550)
AtBRD28 (AT5G65630)
OsBRD18 (LOC_Os08g09340)
OsBRD10 (LOC_Os04g53130)
OsBRD11 (LOC_Os04g53170)
OsBRD6 (LOC_Os02g33980)
AtBRD24 (AT5G46550)
OsBRD15 (LOC_Os07g32420)
AtBRD27 (AT5G63320)
AtBRD17 (AT3G27260)
AtBRD15 (AT3G01770)
AtBRD23 (AT5G14270)
OsBRD2 (LOC_Os01g46040)
AtBRD10 (AT2G34900)
AtBRD18 (AT3G52280)
OsBRD12 (LOC_Os06g04640)
AtBRD6 (AT1G58025)
OsBRD13 (LOC_Os06g24870)
OsBRD17 (LOC_Os08g01794)
OsBRD22 (LOC_Os10g28040)
AtBRD19 (AT3G54610)
AtBRD5 (AT1G32750)
AtBRD16 (AT3G19040)
OsBRD14 (LOC_Os06g43790)
OsBRD20 (LOC_Os09g33980)
AtBRD1 (AT1G05910)
OsBRD9 (LOC_Os03g21450)
AtBRD26 (AT5G55040)
OsBRD7 (LOC_Os03g03870)
OsBRD21 (LOC_Os09g37760)
AtBRD4 (AT1G20670)
AtBRD9 (AT1G76380)
OsBRD16 (LOC_Os07g37800)
AtBRD7 (AT1G61215)
OsBRD19 (LOC_Os08g39980)
AtBRD11 (AT2G42150)
AtBRD20 (AT3G57980)
AtBRD12 (AT2G44430)
AtBRD21 (AT3G60110)
OsBRD3 (LOC_Os02g02290)
AtBRD13 (AT2G46020)
OsBRD4 (LOC_Os02g09920)
AtBRD25 (AT5G49430)
OsBRD8 (LOC_Os03g19340)
AtBRD14 (AT2G47410)
(A)
GI
GII
GIII
GIV
GV
GVI
GVII
(B)

Figure 9

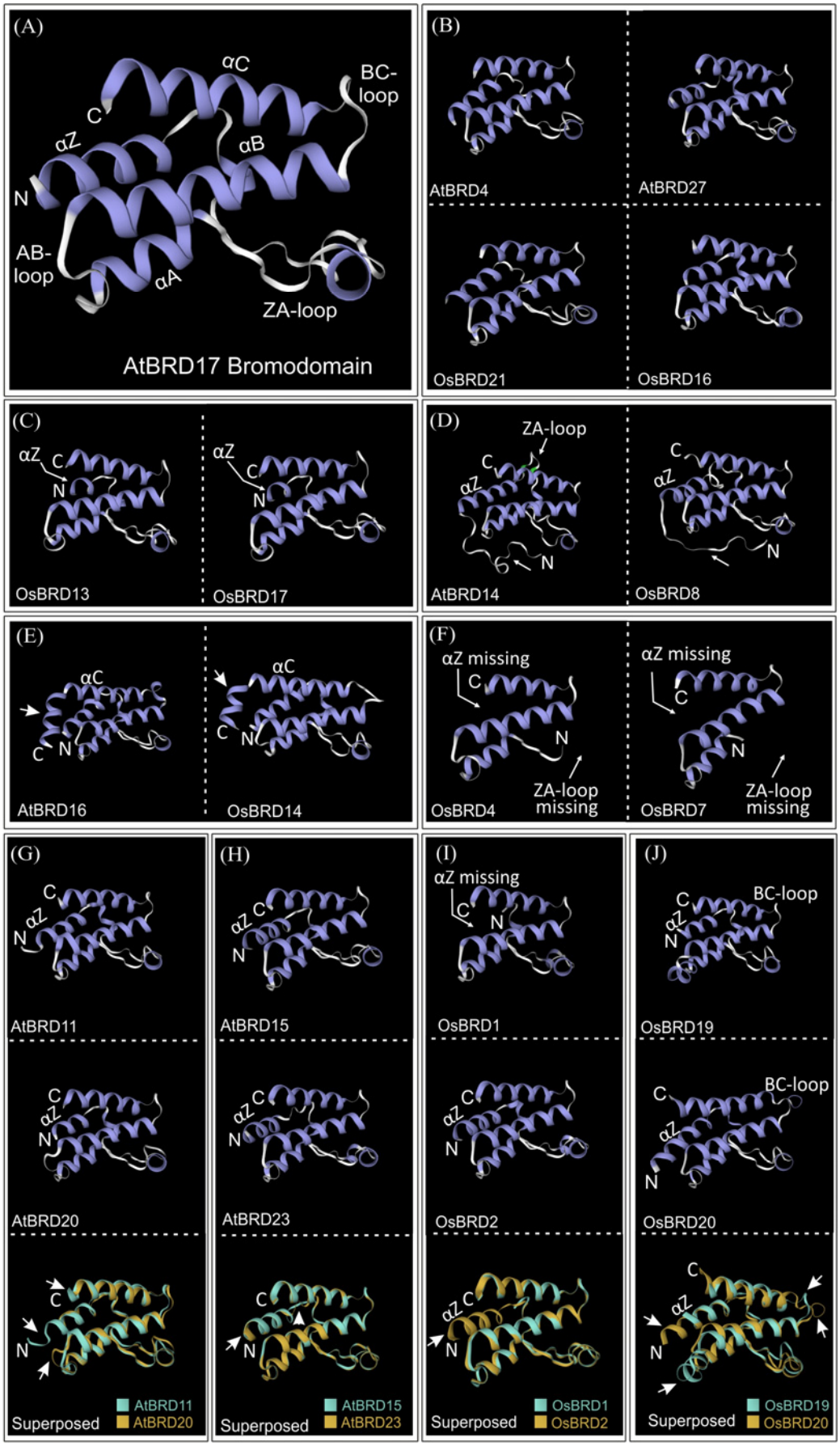
(A)
αC
BC-loop
C
αZ
αB
N
AB-loop
αA
ZA-loop
AtBRD17 Bromodomain
(B)
AtBRD4
AtBRD27
OsBRD21
OsBRD16
(C)
αZ
C
N
OsBRD13
αZ
C
N
OsBRD17
(D)
ZA-loop
C
αZ
N
AtBRD14
C
αZ
N
OsBRD8
(E)
αC
C
N
AtBRD16
αC
C
N
OsBRD14
(F)
αZ missing
C
N
ZA-loop missing
OsBRD4
αZ missing
C
N
ZA-loop missing
OsBRD7
(G)
C
αZ
N
AtBRD11
C
αZ
N
AtBRD20
C
N
Superposed
AtBRD11
AtBRD20
(H)
αZ
C
N
AtBRD15
αZ
C
N
AtBRD23
C
N
Superposed
AtBRD15
AtBRD23
(I)
αZ missing
C
N
OsBRD1
C
αZ
N
OsBRD2
αZ
C
N
Superposed
OsBRD1
OsBRD2
(J)
C
BC-loop
αZ
N
OsBRD19
C
BC-loop
αZ
N
OsBRD20
C
αZ
N
Superposed
OsBRD19
OsBRD20

Figure 10

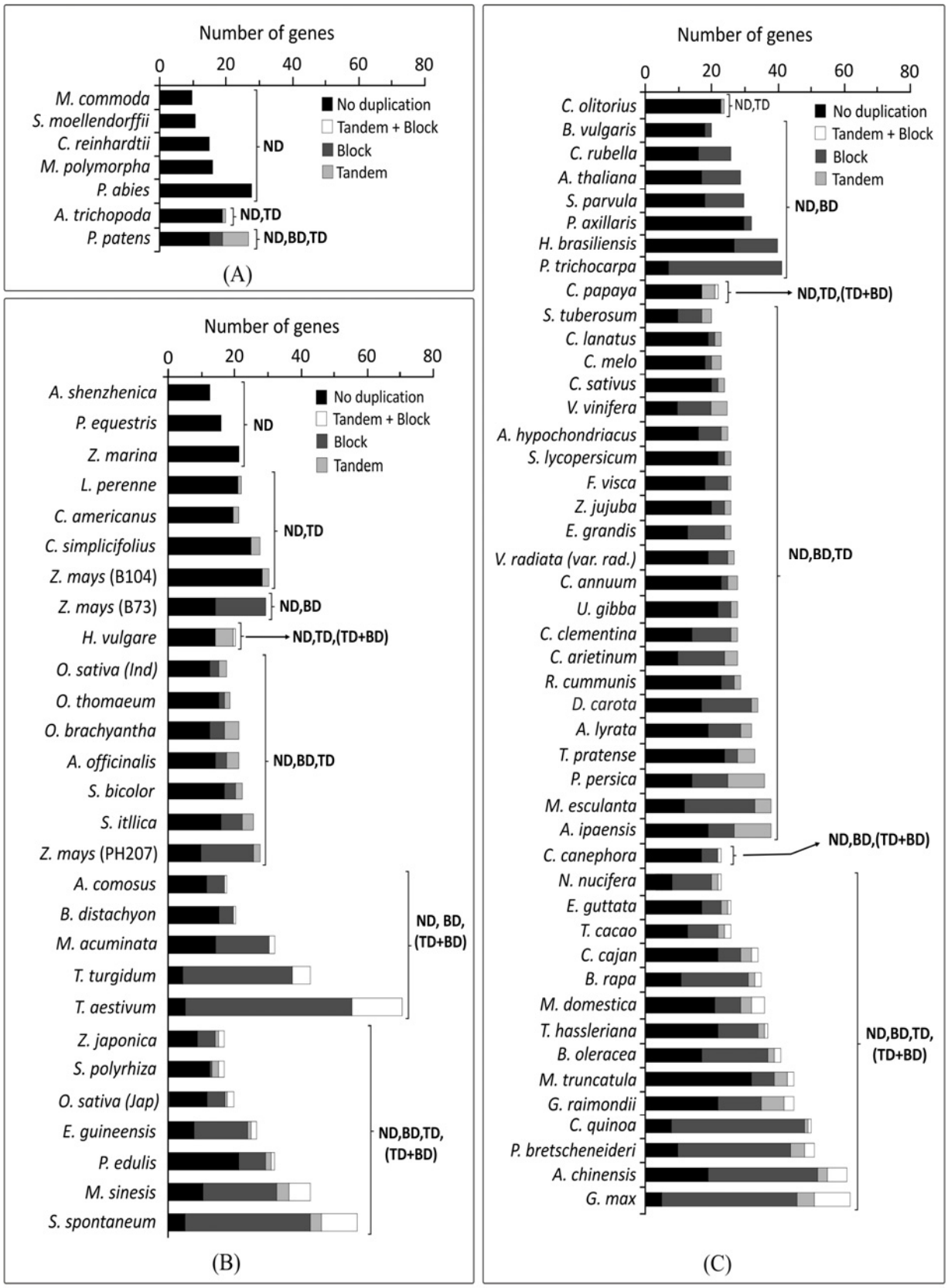
Number of genes
0
20
40
60
80
No duplication
Tandem + Block
Block
Tandem
M. commoda
S. moellendorffii
C. reinhardtii
M. polymorpha
P. abies
A. trichopoda
P. patens
ND
ND,TD
ND,BD,TD
(A)
A. shenzhenica
P. equestris
Z. marina
L. perenne
C. americanus
C. simplicifolius
Z. mays (B104)
Z. mays (B73)
H. vulgare
O. sativa (Ind)
O. thomaeum
O. brachyantha
A. officinalis
S. bicolor
S. itllica
Z. mays (PH207)
A. comosus
B. distachyon
M. acuminata
T. turgidum
T. aestivum
Z. japonica
S. polyrhiza
O. sativa (Jap)
E. guineensis
P. edulis
M. sinesis
S. spontaneum
ND,TD
ND,BD
ND,TD,(TD+BD)
ND,BD,TD
ND, BD, (TD+BD)
ND,BD,TD, (TD+BD)
(B)
C. olitorius
B. vulgaris
C. rubella
A. thaliana
S. parvula
P. axillaris
H. brasiliensis
P. trichocarpa
C. papaya
S. tuberosum
C. lanatus
C. melo
C. sativus
V. vinifera
A. hypochondriacus
S. lycopersicum
F. visca
Z. jujuba
E. grandis
V. radiata (var. rad.)
C. annuum
U. gibba
C. clementina
C. arietinum
R. cummunis
D. carota
A. lyrata
T. pratense
P. persica
M. esculanta
A. ipaensis
C. canephora
N. nucifera
E. guttata
T. cacao
C. cajan
B. rapa
M. domestica
T. hassleriana
B. oleracea
M. truncatula
G. raimondii
C. quinoa
P. bretscheneideri
A. chinensis
G. max
ND,TD
ND,BD
ND,TD,(TD+BD)
ND,BD,TD
ND,BD,(TD+BD)
ND,BD,TD, (TD+BD)
(C)

Figure 11