Planta (2014) 239:299–312
DOI 10.1007/s00425-013-1979-9

ORIGINAL ARTICLE

Genomic surveys and expression analysis of *bZIP* gene family in castor bean (*Ricinus communis* L.)

Zhengwei Jin · Wei Xu · Aizhong Liu

Received: 3 September 2013 / Accepted: 11 October 2013 / Published online: 29 October 2013

Abstract The basic leucine zipper (bZIP) transcription factors comprise a family of transcriptional regulators present extensively in plants, involved in regulating diverse biological processes such as flower and vascular development, seed maturation, stress signaling and pathogen defense. Castor bean (*Ricinus communis* L. Euphorbiaceae) is one of the most important non-edible oilseed crops and its seed oil is broadly used for industrial applications. We performed a comprehensive genome-wide identification and analysis of the bZIP transcription factors that exist in the castor bean genome in this study. In total, 49 RcbZIP transcription factors were identified, characterized and categorized into 11 groups (I–XI) based on their gene structure, DNA-binding sites, conserved motifs, and phylogenetic relationships. The dimerization properties of 49 RcbZIP proteins were predicted on the basis of the characteristic features in the leucine zipper. Global expression profiles of 49 *RcbZIP* genes among different tissues were examined using high-throughput sequencing of digital gene expression profiles, and resulted in diverse expression patterns that may provide basic information to further reveal the function of the 49 *RcbZIP* genes in castor bean. The results obtained from this study would provide valuable information in understanding the molecular basis of the RcbZIP transcription factor family and their potential function in regulating the growth and development, particularly in seed filling of castor bean.

Keywords Digital gene expression · High-throughput sequencing · Phylogenetic analysis · *Ricinus communis* · bZIP transcription factor

Abbreviations

Leu	Leucine
Asn	Asparagines
bZIP	Basic leucine zipper
DGE	Digital gene expression profiling

Introduction

The basic leucine zipper (bZIP) transcription factor family is one of the largest and most diverse transcriptional regulators among eukaryotic organisms. The bZIP proteins are characterized by a 40–80 amino acids conserved bZIP domain, which possess a basic region that binds DNA and an adjacent Leu (leucine) zipper region that mediates protein dimerization (Talanian et al. 1990; Hurst 1995). The basic region is highly conserved and consists of around 16 amino acid residues with an invariant N-x7-R/K-x9 motif, while the Leu zipper region is less conserved, consisting of several heptad repeats of Leu or other bulky hydrophobic amino acids. The basic region preferentially binds to DNA sequences with an ACGT core *cis*-element,

Electronic supplementary material The online version of this article (doi:10.1007/s00425-013-1979-9) contains supplementary material, which is available to authorized users.

Z. Jin · A. Liu (✉)
Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, 88 Xuefu Road, Kunming 650223, China
e-mail: liuaizhong@mail.kib.ac.cn

Z. Jin
School of Life Sciences, University of Science and Technology of China, 443 Huangshan Road, Hefei 230027, China

W. Xu · A. Liu
Kunming Institute of Botany, Chinese Academy of Sciences, 132 Lanhei Road, Kunming 650204, China

especially like G-box (CACGTG), C-box (GACGTC) and A-box (TACGTA) (Izawa et al. 1993; Foster et al. 1994). The N-terminal of the basic region can bind to the major groove of double-stranded DNA, whereas the C-terminal of the Leu zipper mediates dimerization to form a superimposed coiled-coil structure and mediates the homo- and/or hetero-dimerization of bZIP proteins (Landschulz et al. 1988; Ellenberger et al. 1992).

Studies have demonstrated that bZIP transcription factors were involved in diverse physiological processes in plant growth and development. As known from previous studies, bZIP genes participated in transcriptional regulation of abiotic stress responses, including drought, osmotic, cold, oxidative stress, and phytohormone like ABA. In *Arabidopsis*, for instance, *AtbZIP36* (*ABF2/AREB1*), *AtbZIP38* (*ABF4/AREB2*), and *AtbZIP37* (*ABF3*) genes were up-regulated in response to the ABA signal, dehydration, and salinity (Uno et al. 2000). The AtbZIP11 (ATB2) might play an important role in the sugar signaling pathway during metabolism (Rook et al. 1998). In rice, *OsbZIP12* (*OsABF1*) and *OsbZIP46* (*OsABF2*) were strongly induced in response to drought, salinity and cold stresses, as well as the ABA signal (Hossain et al. 2010a, b). Over-expression of *ABF3* improved the drought tolerance in both *Arabidopsis* and rice (Kang et al. 2002; Oh et al. 2005). Over-expression of *OsbZIP72* displayed a hypersensitivity to ABA, elevated levels of expression of ABA response gene such as *LEA*s, and increased the ability of drought tolerance in rice (Lu et al. 2009). The function of OsbZIP10 (OsABI5) was involved in the metabolic pathway which could influence the fertility and enhance stress tolerance in rice (Zou et al. 2008). In tobacco, co-expression of *LEA* and *ThbZIP1* genes resulted in an additive enhancement of stress tolerance (Qu et al. 2012). In grape, expression of *VvbZIP23* was induced by various abiotic stresses, including drought, salt and coldness (Tak and Mhatre 2013). However, the global profiles on function of *bZIP* genes in plants are far from being understood.

Recently, available complete genome sequences have allowed a comprehensive and phylogenetic annotations for the *bZIP* genes in several species, such as *Arabidopsis* (Riechmann et al. 2000; Jakoby et al. 2002), rice (Nijhawan et al. 2008), soybean (Liao et al. 2008), sorghum (Wang et al. 2011) and maize (Wei et al. 2012). Based on sequence similarity of the basic region and the presence or absence of additional conserved motifs, the characterization of *bZIP* genes has been annotated for *Arabidopsis,* rice and maize. However, limited is known about the holistic profiles of structural and functional annotation in *bZIP* genes in other plants because of the limited availability of complete genome. Castor bean (*Ricinus communis* L., Euphorbiaceae) is one of the most important non-edible oilseed crops and its seed oil is broadly used in industry. In particular, the main composition of its seed oil is ricinoleic acid, which is considered as an ideal and unique feedstock for biodiesel production (Akpan et al. 2006; Ogunniyi 2006; Scholza and da Silva 2008). Due to increasing demand for production of castor bean seed oils in many countries, breeding and improvement of varieties are drawing great attention from breeders (Qiu et al. 2010). Further efforts should be made to elucidate the molecular mechanism underlying the regulation of growth and development. The recent completion of the castor bean genome (Chan et al. 2010) provides a great opportunity to identify and characterize the holistic profile of the bZIP transcription factor family in castor bean, which could provide useful information in understanding the molecular mechanism of the *bZIP* genes that underlies the regulation of growth and development in castor bean. In this study, we identified and characterized bZIP transcription factors based on the complete genome sequences of castor bean. Further, global expression profiles of the castor bean *bZIP* genes among different tissues were examined using high-throughput sequencing for digital gene expression tag profiling (DGE). Results obtained here provide a holistic profile in the understanding of the molecular basis of the bZIP transcription factor family in castor bean and other plants in the family Euphorbiaceae as well.

Materials and methods

Database search for bZIP transcription factors in castor bean

To identify the putative bZIP proteins in castor bean, the BLAST search was performed using the known *Arabidopsis* bZIP sequences downloaded from the TAIR (http://www.arabidopsis.org/), as queries against the castor bean genome database (http://www.jcvi.org/). The proteins with the significant *E*-value (<1.0) were downloaded and then the redundant protein sequences were screened and discarded from our data. In addition, the online programs SMART (http://smart.embl-heidelberg.de/) and pfam (http://pfam.sanger.ac.uk/) were used to check the predicted bZIP domains in these putative bZIP proteins. The proteins which showed the presence of bZIP domain with confidence (*E*-value <1.0) were used for our further analyses.

Protein properties and sequence analyses

The molecular weight (kDa) and isoelectric point (PI) of each protein were calculated by online ExPASy programs (http://www.expasy.org/tools/). The conserved motifs were identified by the online MEME analysis tool (http://meme.ebi.edu.au/meme/intro.html) with parameters

set: optimum motif width ≥6 and ≤200, maximum number of motifs 25. To inspect the exon/intron organizations in the bZIP genes of castor bean, alignments of the genomic and cDNA sequences were performed using the online server GSDS (http://gsds.cbi.pku.edu.cn/). The intron distribution and splicing phases within the genomic sequences were further predicted based on the alignments of the genomic sequences and cDNA full-length sequences obtained by cloning in previous study (Chan et al. 2010) using Spidey (http://www.ncbi.nlm.nih.gov/spidey/). Further, all of the identified castor bean bZIP and known AtbZIP amino acid sequences downloaded from TAIR (http://www.arabidopsis.org) were aligned for phylogenetic analyses using the MUSCLE tool (Edgar 2004). Based on the bZIP protein full-length sequences, phylogenetic trees were generated using the neighbor-joining criteria in MEGA 5.0 (Tamura et al. 2011) with 1,000 bootstrap replicates.

Gene expression analyses

To examine the global expression profiles of castor bean *bZIP* genes identified among different organs or developmental stages, high-throughput sequencing of DGEs for five tissues was conducted, including leaf, root tip, developing seed at the initial stage (seed I) and at the fast oil accumulation stage (seed II), and endosperm. Seeds of castor bean var. ZB306 elite inbred line (kindly provided by Zibo Academy of Agricultural Sciences, Shandong, China) were germinated and grown in the conservatory under natural conditions (11 h light, 13 h dark; 25 °C during the day and 18 °C at night). Mature female flowers were hand-pollinated and tagged. Leaf tissue was collected from a fully expanded young leaves and root tips were collected, washed and dissected for further use. Developing seeds at two different stages, i.e., seed I at the initial stage (developing seeds do not start to accumulate oil, ca. 15 days after pollination) and seed II at the fast oil accumulation stage (developing seeds start to accumulate oil fast, ca. 35 days after pollination), were collected. Immature endosperm tissue was dissected from the immature seeds (ca. 40 days after pollination). All samples were collected from the same individual tagged. Total RNA was isolated from the leaf, root tip, seed I, seed II and endosperm tissues, respectively, using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and purified using an RNeasy Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. The quality of total RNA samples was checked using the NanoDrop Spectrometer (ND-1000 Spectrophotometer, Peqlab) and agarose gel electrophoresis.

High quality RNAs were used to construct tag libraries for illumina sequencing, respectively. Briefly, total RNA (about 6 μg) was enriched by oligo (dT) magnetic beads and oligo (dT) used as the primers to synthesize the first and second-strand cDNA. The cDNA was digested by two types of endonuclease NlaIII or DpnII, acquiring 17 bp tags with different adaptors of both ends to generate a tag library. After 15 cycles of linear PCR amplification, 105 bp fragments were purified by 6 % PAGE gel electrophoresis. After denaturing, the single-chain molecules were fixed onto the illumina sequencing chip (flow cell). Each molecule then grows into a single-molecule cluster sequencing template through in situ amplification. Four colored nucleotides were added for sequencing using the method of sequencing by synthesis (SBS). Millions of raw reads were generated with a sequencing length of 49 bp. Sequencing was performed in BGI Shenzhen (China).

After the raw sequence data from the five tag libraries were pre-processed to filter out low quality reads and clipped adapter sequences, all clean reads were mapped to the castor bean genome (http://castorbean.jcvi.org/index.php) to obtain unique reads and reads abundance using SOAP2 software (Li et al. 2009). To compare the differential expression of genes among tissues, the expression level of each gene in the different tissues was normalized to the number of transcripts per million clean (TPM). Genes with significantly different expression were determined by $P \leq 0.001$ and fold-change ≥2 in two samples.

Results and discussion

Identification of bZIP transcription factors in castor bean

Based on an extensive survey against castor bean genome database using the conserved bZIP domain sequence of *Arabidopsis thaliana* as a BLASTP query, 100 castor bean bZIP candidates were initially obtained with a probability *E*-value threshold of 1.0. Further, 49 non-redundant bZIP transcription factors were identified and designated as RcbZIPs 1–49 according to the nomenclature system proposed in the previous studies (Jakoby et al. 2002; Nijhawan et al. 2008; Wei et al. 2012). The proteins encoded by the 49 transcription factors ranged from 101 to 702 amino acids (aa) in length with an average of 326 aa. The deduced molecular weights of 49 RcbZIP proteins ranged from 11.72 kDa to 76.12 kDa (Suppl Table S1). Compared with *Arabidopsis* (having 75 members with an average of 321 aa in length), rice (having 89 members with an average of 311 aa in length) and maize (having 125 gene loci), castor bean harbors only 49 bZIP transcription factor members with similar average length to other plants. The reason for castor bean having less transcription factors is currently unknown, though previous evidence has showed the higher number of

bZIP members evolved in monocots than in dicots after the divergence of monocots from dicots (Wang et al. 2011).

DNA-binding-site specificity and classification of RcbZIP proteins

The basic and hinge regions in the bZIP domains directly interacting with DNA *cis*-elements are the most conserved regions of the bZIP transcription factor sequences. The position −14 to −17 of the basic region has been identified to play a critical role in DNA specific recognition and binding (Suckow et al. 1993). Therefore, bZIP transcription factors were often annotated and classified into different groups characterized by the diagnostic sequence structure of the basic and hinge regions in *Arabidopsis*, rice and maize (Jakoby et al. 2002; Nijhawan et al. 2008; Wei et al. 2012). Similarly, we designated the first Leu residue in Leu zipper region as the number +1 and the C-terminal amino acid of hinge region as −1 based on Suckow et al. (1993) study. The 49 RcbZIP members were classified into 11 groups (I-XI) (Suppl Fig. S1) according to the classification criteria in rice (Nijhawan et al. 2008). The structural characteristics and DNA-binding specificity of RcbZIP proteins for each group were described in Table 1. Putative DNA-binding sites were identified in nine groups (including groups I–IX). Particularly, six of nine groups with explicit DNA-binding sites (including the groups I–IV, VII and VIII) exhibited a G-box or C-box binding site, suggesting that the G-box or C-box site be the main DNA-binding sites for bZIP members in castor bean. The group VI (covering nine members) displayed an ABRE binding site, a critical binding site responding to ABA signals (Choi et al. 2000). The group IX (including seven members) could bind DNA sequences like CCA/TGG repeats other than the ACGT core (Aeschbacher et al. 1991; Yin et al. 1997). The group V (containing RcbZIP16 only) might form heterodimers by binding to unknown DNA sequence (Shen et al. 2007). In addition, DNA-binding sites within groups X (RcbZIP4) and XI (RcbZIP5) were inexplicit since the conserved Asn at the −18 position was replaced by a Lys residue within RcbZIP4 and a hydrophobic Ile residue was replaced by the conserved Arg/Lys at the −10 position within RcbZIP5.

Gene structure of *RcbZIP* genes

Genomic structure of each gene, in particular the number and distribution of exon/introns, may be a record of key events in evolution and offers insights into understanding the emergence and evolution of a given gene (Betts et al. 2001). To characterize the gene structure of 49 *RcbZIP* genes, their exon/intron organization was analyzed based on the number, distribution and splicing phase of exon/intron. Since the basic and hinge regions of the bZIP domain are the most conserved, the pattern of intron positions and splicing phases within bZIP domain regions were often considered to be more informative than other regions for exploring the homology and evolution of *bZIP* genes (Nijhawan et al. 2008). The splicing phases were designated as three splicing phases: phase 0 (P0), splicing occurred after the third nucleotide of the codon; phase 1 (P1), splicing occurred after the first nucleotide of the codon; and phase 2 (P2), splicing occurred after the second nucleotide.

In total, 39 of 49 *RcbZIP* genes were intronic and the number of introns varied from 1 to 11, of which, 37 had introns in the bZIP domain region (Fig. 1). On the basis of the presence and position of introns and splicing phase, *RcbZIP* genes could be divided into five patterns (*a*, *b*, *c*, *d* and *e*) (Fig. 2; Suppl Fig. S2). The patterns *a* (including 14 genes) and *b* (including 15 genes) were the most prevalent. The pattern *a* had an intron in P2 behind the position −22 (R) within the basic region, whereas the pattern *b* had an intron in P0 before the position −5 (A) within the hinge region. The pattern *c* (including 7 genes) had two introns (each in P0), one behind position −26 (K) within the basic region and the other behind −6 (K) within the hinge region. The pattern *d* (one gene only) had an intron (P2) inserted between −17 (Q) and −16 (Q) within the basic region. The pattern *e* was devoid of any intron in the basic or hinge region. When an intron was presented in the hinge region, both its phase and position were conserved, whereas its positions and splicing phases (P0 or P2) were variable when presented in the basic region. Similarly, the intron distribution and splicing patterns within the bZIP domain regions were observed in rice and maize, indicating that the gene structure and splicing phases of *bZIP* genes were considerably conserved within the basic and hinge regions in plants (Nijhawan et al. 2008; Wei et al. 2012).

After inspecting the gene structure and splicing phase patterns for each group classified above, we found that most members clustered in the same group shared the same or similar gene structure (including intron number and splicing phase), though the introns were variable in length (Fig. 1). For instance, three members within group I shared same gene structure and splicing phases; eleven members within group VI had the same gene structure and splicing phases; seven members within group VII had their own gene structure and splicing phases; and seven members within group XI also displayed similar gene structure and splicing phases. In particular, the seven members within group VII showed the pattern *c*; group I and VI shared the pattern *b*; group II, IV, V, IX, and X shared the pattern *a*; and group IV clustered by most intronless members exhibited the pattern *e*. Correspondingly, the conserved gene structure and splicing phases in each group were in agreement with the classification of *RcbZIP* genes above.

Table 1 Prediction of DNA-binding specificity and the classification of 49 RcbZIP transcription factors (see Fig. S1 for the amino acid positions)

Group	Characteristic features	Putative binding site	Known binding sites
I (3) [RcbZIPs 15,33,46]	Conserved residues in positions −18 (N), −15 (S), −14 (A), −11 (S), −10 (R); conserved sequences RKQS in the basic region; QAEC(/T)E(/D)E in hinge region	G-box (CACGTG) and/or G-box-like sequences	OSBZ8, GBF2 and EmBP-1 (Nakagawa et al. 1996; Niu et al. 1999; Schindler et al. 1992)
II (2) [RcbZIPs 21,25]	Conserved sequences NRVSAQQAR in basic region	G-box-like sequences	ACACGTGG for HY5 (Chattopadhyay et al. 1998)
III (1) [RcbZIP 28]	Conserved sequences RNRESAxxSR in the basic region	G-box and C-box (GACGTC) with equal affinity	Tobacco TGA1b (Katagiri et al. 1989; Niu et al. 1999)
IV (13) [RcbZIPs 3,7,12,14,20,23,24,34, 36,39,40,41,42]	Specific hinge region sequence Q(/K)A(/Q/K/R)H(/Q) L(M/I)T(/N/Q/D/E)E(/D/N)	Hybrid ACGT elements like G/C (GACGTG), G/A(TACGTG) and C/G(CACGTC) boxes	GCCACGT(A/C)AG for REB (Nakase et al. 1997); C/G hybrid for LIP19/OsOBF1 heterodimer (Shimizu et al. 2005)
V (1) [RcbZIP16]	A residue at −19 position of the basic region and conserved sequences QYISE in the hinge region	May form heterodimers to bind to unknown sequences	AtbZIP34 and AtbZIP61 (Shen et al. 2007)
VI (9) [RcbZIPs 13,18,22,27,29,31,35, 43,47,8,49]	Conserved residues MIK in the basic region; conserved residues QAY in the hinge region	ABREs with the core ACGT or others containing GCGT or AAGT	CACGTGG/tC, CGCGTG for ABF1 and TRAB1, respectively (Choi et al. 2000; Hobo et al. 1999)
VII (7) [RcbZIPs 2,6,9,10,26,32,38]	Conserved residues at positions −21 (L), −20 (A), −19 (Q), −18 (N), −15 (A), −14 (A), −12 (K), −11 (S) and −10 (R); a specific sequences KAYVQQ in the hinge region	Specificity C-box sequences	Tobacco TGA1a (Niu et al. 1999)
VIII (2) [RcbZIPs 1,45]	Conserved A residue at position −15	C-box elements	Unknown
IX (7) [RcbZIPs 11,17,19,30,37,44,48]	Conserved K at −10 position of the basic region instead of R	Sequences other than those containing ACGT core, like CCA/TGG repeats	RF2a [CCA(N)nTGG] and *Arabidopsis* PosF21 (Aeschbacher et al. 1991; Yin et al. 1997)
X (1) [RcbZIP 4]	Lack the conserved N at position −18	Unknown	Unknown
XI (1) [RcbZIP 5]	Hydrophobic I residue at position −10 instead of R/K	Unknown	Unknown

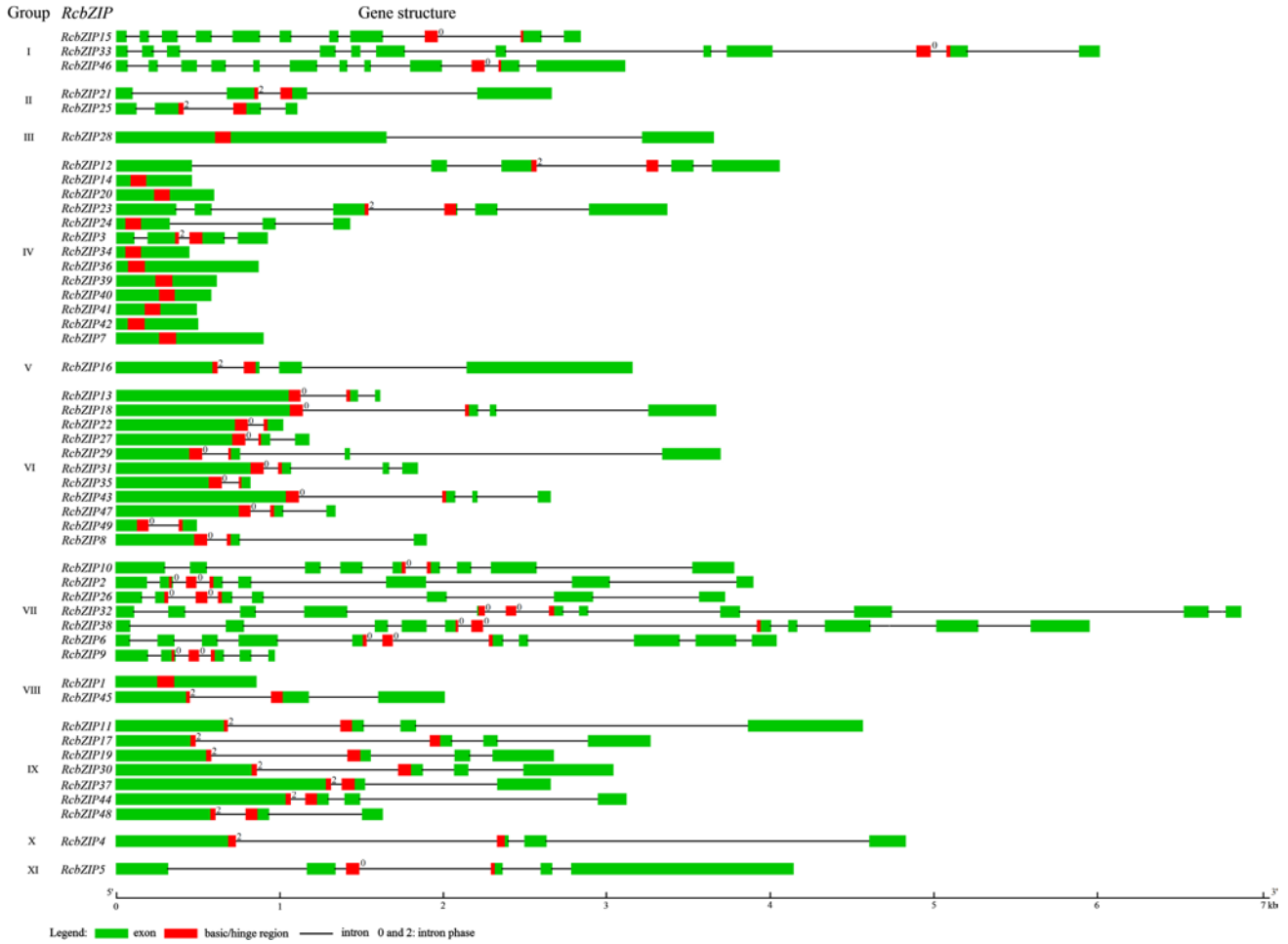

Fig. 1 Exon/intron organization of 49 *RcbZIP* genes was depicted for each group. The exons and introns are represented by *boxes* and *lines*, respectively. The *red box* denotes the bZIP domain region; the numbers above line denote the splicing phase. The numbers "0" and "2" denote different splicing phases, "0" means splicing occurred after the third nucleotide of the codon, and "2" means splicing occurred after the second nucleotide

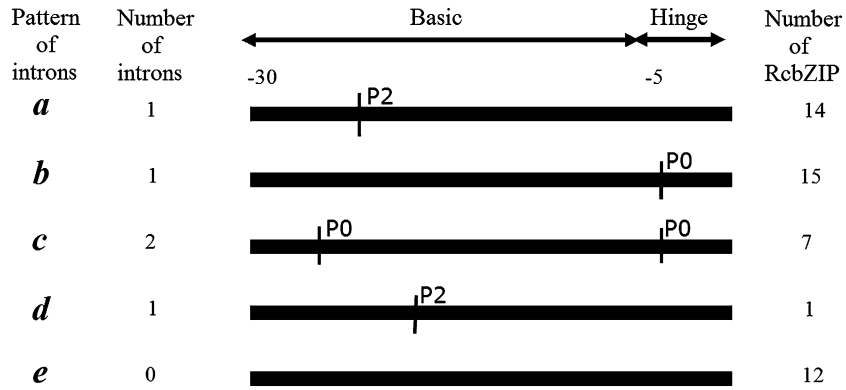

Fig. 2 Intron patterns (**a**–**e**) within the basic and hinge regions of 49 RcbZIP proteins. P0 and P2 stand for the intron splicing phases. P0 means splicing occurred after the third nucleotide of the codon; and P2 means splicing occurred after the second nucleotide. The *black bars* represent the sequence of basic and hinge regions. The *vertical lines* denote the positions of intron splicing phases

Identification of additional motifs in RcbZIP transcription factors

Outside the bZIP domain region, bZIP proteins usually contain additional conserved motifs which might indicate potential function sites or participate in activating the function of bZIP proteins. Diverse conserved motifs outside of the bZIP domain region had been identified in *Arabidopsis*, rice and maize (Jakoby et al. 2002; Nijhawan et al. 2008; Wei et al. 2012). A total of 25 motifs (with *E*-value cutoff

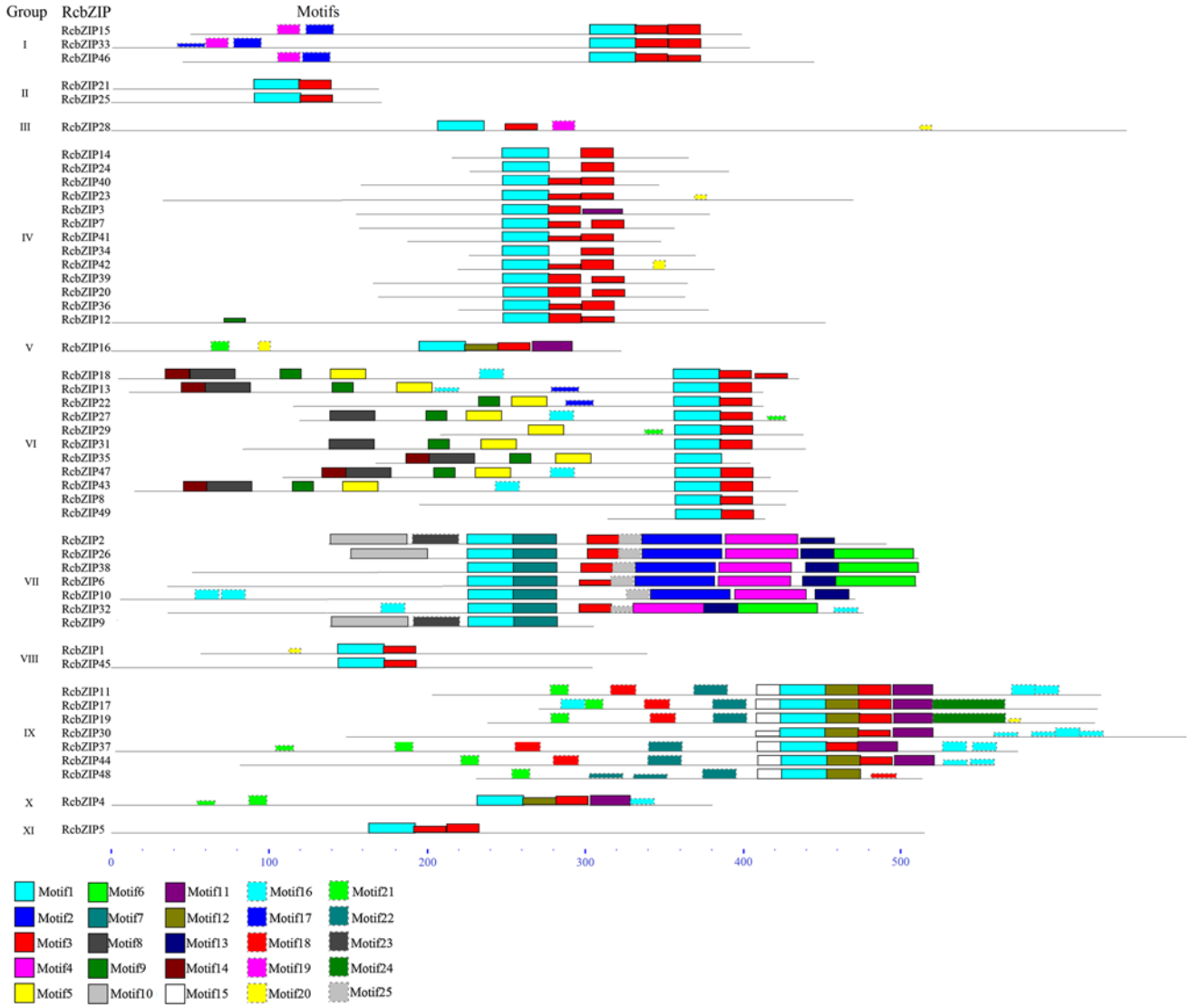

Fig. 3 Summary for the distribution of conserved motifs identified from 49 RcbZIP proteins by each group given separately. Each motif is represented by a number in *colored box*. See Suppl Table S2 for detailed motif information

<e-1.0), including the conserved bZIP domain (motif 1), were identified in castor bean (Suppl Table S2). The distribution of motifs in each amino acid sequence was depicted separately in Fig. 3. It can be observed that the motif 3 are shared by 45 members involving each group, and some motifs are shared by several groups such as the motif 17 present in groups I and VI, the motif 19 present in groups I and III, the motif 20 present in groups III, IV, V and VIII, and the motif 16 present in groups VI, VII, IX and X, respectively. Most of the conserved motifs, however, were presented in specific groups, for example, motifs 5, 8, 9, 14 and 16 in the group VI, motifs 2, 4, 6, 7, 13 and 25 in the group VII, and motifs 11, 12, 15, 18, 21 and 22 in group IX. These observations may indicate that the group-specific motifs aid in determining specific functions for members in each group.

Compared to those conserved motifs identified within bZIP transcription factors in other plants, 14 motifs (including motifs 2, 4, 5, 6, 8, 9, 10, 11, 13, 14, 16, 17, 19 and 25) were shared between castor bean and *Arabidopsis*, eight motifs (including motifs 2, 4, 5, 6, 8, 9, 14 and 19) were shared between castor bean and rice, and 10 motifs (including motifs 2, 4, 5, 6, 7, 8, 9, 14, 19 and 21) were shared between castor bean and maize. In addition, eight motifs (motifs 2, 4, 5, 6, 8, 9, 14 and 19) were commonly shared by castor bean, *Arabidopsis*, rice and maize, indicating these additional motifs outside of bZIP domain might be conserved among plant species. Some of additional

motifs, however, were variable among species and might be species-specific in plants.

The motif 3 was further identified as a part of N-terminal extension of Leu zipper region, tightly associated with motif 1 (within basic region). The motifs 5, 8 and 9 represented potential CKII (casein kinaseII) phosphorylation sites with the presence of TLE [ED], [ST] AE [EA] and TVD [ED] (Pinna 1990), respectively. These CKII phosphorylation sites might be associated with ABA signal (Choi et al. 2000; Uno et al. 2000). The motif 14 contained another phosphorylation site RQ [NGST] S, which might be for Ca^{2+}-dependent protein kinase (R/KxxS/T) (Schworer et al. 1988). The motif 16, a Gln-rich domain, had been proven to have a critical role in transcriptional activation (Emili et al. 1994). The motif 17, a Pro-rich domain, might be involved in mediating protein–protein interactions (Meier and Gruissem 1994). The motif 4 corresponds to the Ras-binding domain of Byr2 kinase in *Schizosaccharomyces* (Scheffzek et al. 2001). Although the presence of these conserved motifs shared among castor bean, *Arabidopsis*, rice and maize might reflect functional similarities of bZIP proteins in plants, the function of most additional conserved motifs identified in RcbZIP transcription factor family is not yet known.

Prediction of dimerization property

Studies have demonstrated that the Leu zipper region of the bZIP domain mediates dimerization between the parallel coiled-coil structures and forms homo and/or hetero-dimerization determined by the amino acid sequences (Oshea et al. 1992). The Leu zipper region, arranged in the form of heptad repeats, is responsible for dimerization. Within each heptad, the amino acid positions are named as g, a, b, c, d, e and f in order (see Suppl Fig. S3; Hodges et al. 1973; Deppmann et al. 2004). Leu zipper dimerization stability and specificity are determined mainly by four amino acids at *a, d, e* and *g* positions, as their locations near the Leu zipper interface. The *a* and *d* residues are hydrophobic and pack in a regular "knobs and holes" pattern along the dimerization interface to create the hydrophobic core that is critical for dimerization. The buried Asparagines (Asns) at *a* position can form a polar pocket in the hydrophobic interface that limits oligomerization when interacting interhelically (Zeng et al. 1997), and cause stable N–N interactions at the $a \leftrightarrow a'$ (the opposite helix corresponding position), but does not interact favorably with other *a* position amino acids (Acharya et al. 2002), tending to form homodimerization Leu zippers. The *d* position contains the conserved Leu, a kind of the most stabilizing aliphatic amino acid (Moitra et al. 1997), and is important to maintain the dimerization stability. However, the *e* and *g* positions that flank the dimerization interface contain a large number of

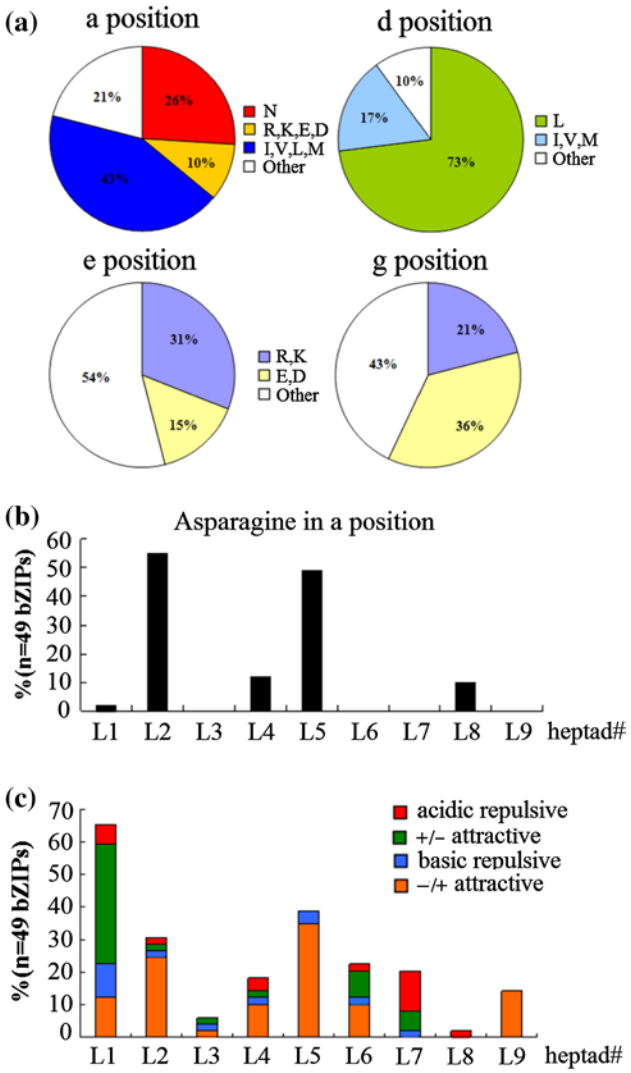

Fig. 4 Prediction of the dimerization properties of 49 RcbZIP proteins. **a** Pie charts depicting the frequency of different amino acids at the *a*, *d*, *e* and *g* positions of the Leu zippers of all RcbZIP proteins (see Suppl Fig. S3 for the amino acid positions within the Leu zipper regions). **b** Histogram of the frequency of Asn residues present at the *a* position of each heptad within Leu zippers for all RcbZIP proteins. **c** Histogram of the frequency of attractive or repulsive $g \leftrightarrow e'$ pairs per heptad within Leu zippers for all RcbZIP proteins

charged amino acids, which are thought to interact electrostatically to form salt bridges between helices (Vinson et al. 1993). Thus, a detailed analysis was carried out to characterize the amino acids present at the *a, d, e* and *g* positions.

At the *a* position, 26 % of amino acids was Asns in 49 RcbZIP proteins (Fig. 4a). For each heptad, the highest of the Asn frequency at *a* position was located in the second heptad followed by the fifth heptad, accounting for 55.1 and 48.98 %, respectively (Fig. 4b), which is identical to *Arabidopsis* and rice. A small number of charged amino acids (R, K, E and D) present at *a* position were observed

which might contribute to form hetero-dimerization. At the *d* position, the frequency of the conserved Leu was of considerable abundance in 49 RcbZIP proteins, accounting for 73 %; the frequency of other hydrophobic amino acids (I, V and M) accounted for 17 % (Fig. 4a). Compared to the AtbZIPproteins (the Leu amino acid accounted for 56 % at the *d* position), RcbZIPs exhibited more abundant Leu content at *d* position, indicating that castor bean bZIP genes may have a stronger stability of Leu zipper dimerization than *Arabidopsis*. For the *e* and *g* positions, the abundant charged amino acids (including acidic amino acids E and D, and basic amino acids R and K) were observed, accounting for 46 and 57 %, respectively (see Fig. 4a). Compared to *Arabidopsis* (charged amino acids accounting for 41 and 53 % at the *e* and *g* positions, respectively), the frequencies of charged amino acids present in the *e* and *g* positions were higher in castor bean. To analyze the contribution of charged residues at the *e* and *g* positions in governing dimerization properties of RcbZIP proteins, we calculated the presence of attractive and repulsive $g \leftrightarrow e'$ pairs in each heptad of castor bean Leu zippers (Suppl Fig. S3), and the histogram of their frequency was presented in Fig. 4c. If both the *g* and the following *e* position amino acids are charged, they were referred as complete $g \leftrightarrow e'$ pairs, classified into four groups (acidic repulsive, basic repulsive, +/−attractive, and −/+attractive) whereas if only the g or e position is charged, it was referred as incomplete $g \leftrightarrow e'$ pairs. It was observed that the maximum frequency of the complete $g \leftrightarrow e'$ pairs was 65 %, appearing in the first heptad, among which, the attractive $g \leftrightarrow e'$ pairs were in the majority. The frequencies of the complete $g \leftrightarrow e'$ pairs in the next three heptads decreased dramatically. In the fifth heptad, the frequency of attractive $g \leftrightarrow e'$ pairs increased. In eighth heptad, only one repulsive $g \leftrightarrow e'$ pair was found. Moreover, only −/+attractive $g \leftrightarrow e'$ pairs were presented in the ninth heptad. When the $g \leftrightarrow e'$ pairs were attractive pairs, they would not prefer homo-dimerization or hetero-dimerization, and for the opposite residues, would form two self-complementary salt bridges. However, the $g \leftrightarrow e'$ pairs were repulsive pairs, they favored the formation of hetero-dimerization, for the interaction of the opposite residues causing two mutually complementary salt bridges (Vinson et al. 1993). Based on the analyses of the dimerization properties above, the 49 RcbZIP proteins could be categorized into 18 types (BZ1-BZ18) (Suppl Fig. S3). These types could be further divided into three classes: class I, containing 2 types (BZ1 and BZ2), favored the formation of homo-dimerization, for the N present in the second heptad at *a* position and the attractive $g \leftrightarrow e'$ pairs in the first and/or second heptads, and lack of any repulsive pairs as well; class III, covering 5 types (BZ14-BZ18), could be considered as having hetero-dimerization specificity on account of no attractive $g \leftrightarrow e'$ pairs; and class II, comprising of the remaining types, could form both homo- and/or hetero-dimerization patterns. The detailed dimerization properties were described in Suppl Table S3.

Phylogenetic and evolutionary relationships of the *RcbZIP* genes

To elucidate the Phylogenetic relationships of the 49 *RcbZIP* genes identified, an unrooted phylogenetic tree generated showed that the 49 *RcbZIP* genes were clustered into nine clades (designating A to I) with well supported bootstrap values (Fig. 5). It was observed that most members, predicted to have similar DNA-binding properties and categorized into the same groups, clustered together. For example, three members categorized in group I were clustered into clade D; seven members categorized in group VII were nested into clade I; and seven members categorized in group IX were embedded in clade G. The 13 members categorized in group IV were separated into three clades (clade A, B and C). The members from group V and X were merged with group IX (clade G) with a well bootstrap support. Combining with the distribution of conserved motifs among clades, most of the members clustered together shared one or more conserved motifs outside the bZIP domain. These observations suggested that more than one RcbZIP motif protein might recognize the same DNA sequence for binding, but they may have different dimerization.

To examine the evolutionary relationships of bZIP transcription factors between castor bean and *Arabidopsis*, another unrooted phylogenetic tree was constructed based on the amino acid sequence similarity of 49 RcbZIPs and 75 AtbZIPs obtained in the previous study (Jakoby et al. 2002). The phylogenetic tree generated 11 distinct clades designated "a" to "k" with well-supported bootstrap values (Suppl Fig. S4). The similar tree topology to Fig. 5 was reconstructed. Ten clades (except for clade k) were clustered by interspecies indicating that *bZIP* genes are homologous between castor bean and *Arabidopsis*. In particular, most members clustered shared one or more conserved motif outside bZIP domain region. For example, most of the members from both castor bean and *Arabidopsis* within Clade *e* shared one or several conserved motifs 5, 8, 9, 14, or 16; most of the members from both castor bean and *Arabidopsis* within Clade *i* shared one or more conserved motifs 2, 4, 6, 7, 13, or 25; and most of the members within Clade *g* shared one or more motifs 11, 12, 15, or 18. Therefore, a majority of the RcbZIP proteins are expected to be orthologs of the AtbZIP proteins. The fact that most of the clades comprised interspecies members sharing similar motifs suggests that the structure and function of most of the *bZIP* genes have probably remained conserved during evolution in plants.

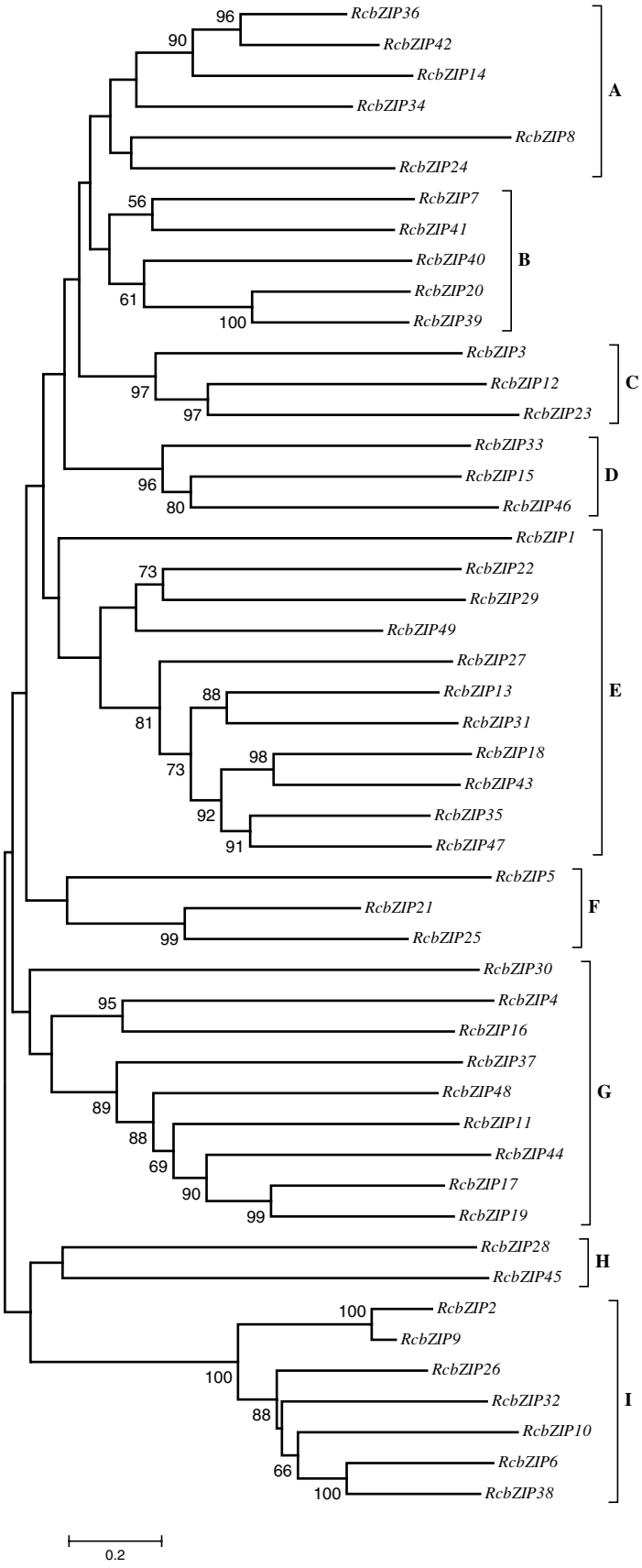

Fig. 5 Phylogenetic relationships of the 49 RcbZIP proteins. The phylogenetic tree was generated using the neighbor-joining criteria in MEGA 5.0 with 1,000 bootstrap replicates. The numbers at each other node indicated the bootstrap values (>50 %). The 9 distinct clades were designated as *A–I*

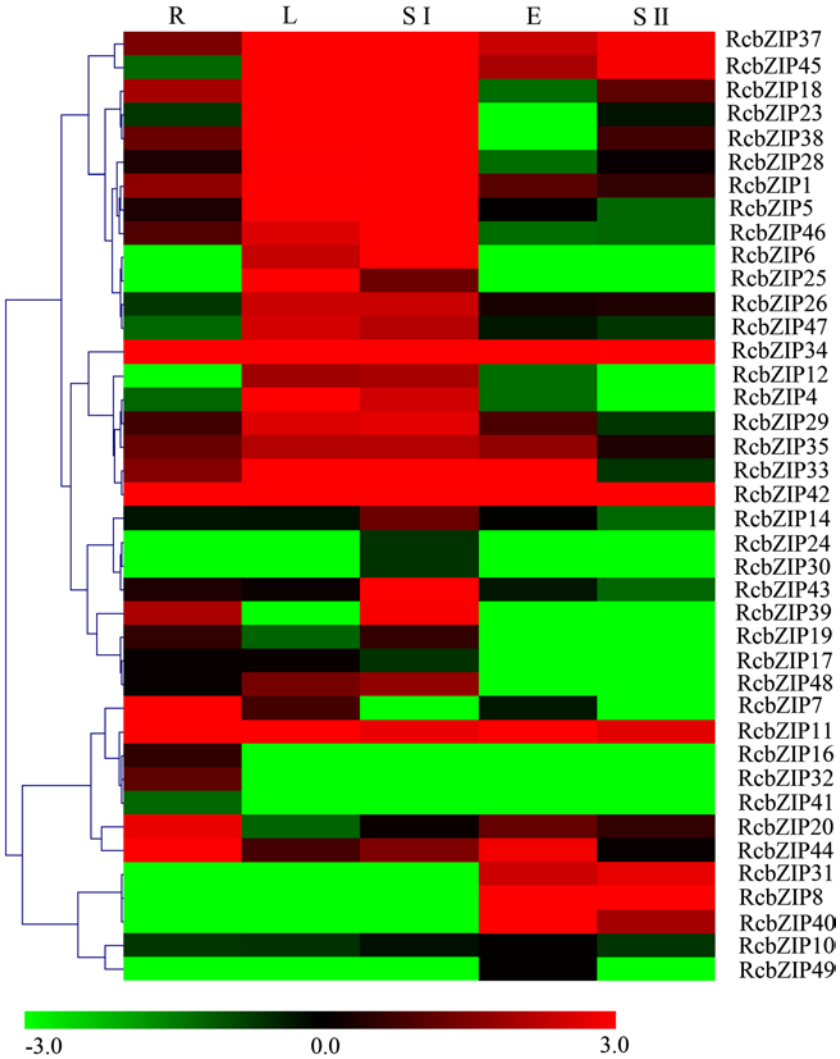

Fig. 6 Heatmap of *RcbZIP* genes expressed among five tissues. This heatmap was generated based on the DGE expression data (normalized log2) using the R software. *Green* indicates low expression, *dark* indicates intermediate expression and *red* indicates high expression. The *R*, *L*, *S I*, *S II* and *E* denote root, leaf, seed I (developing seeds at the initial stage), seed II (developing seeds at the fast oil accumulation stage) and endosperm tissues, respectively

Expression profiles of *RcbZIP* genes among different tissues

The high-throughput sequencing produced 4574301, 4660289, 4543329, 4650533 and 4828665 clean sequence tags for leaf, root, seed I, seed II and endosperm libraries, respectively. After mapping these clean sequence tags to the castor bean genome database, abundance of tags matched to each *bZIP* gene regions in five libraries was 1,407 (leaf), 892 (root), 1,364 (seed I), 736 (seed II), and 981 (endosperm), respectively.

The expression analyses showed that 40 *RcbZIP* genes were expressed in at least one of five tissues tested. As illustrated in Fig. 6, the expression of 31, 30, 32, 25 and 27 genes was detected in root, leaf, seed I, seed II and endosperm tissues, respectively. 20 genes (*RcbZIP*s *33*, *46*, *28*, *14*, *20*, *34*, *42*, *18*, *29*, *35*, *43*, *47*, *10*, *26*, *1*, *45*, *11*, *37*, *44*, and *5*) were expressed in all five tissues tested, but their expression levels are largely varied among tissues, and only four genes (*RcbZIP*s *11*, *34*, *37* and *42*) were highly expressed in all five libraries. Most of *RcbZIP* genes displayed a tissue specific expression pattern. For instance, three genes (*RcbZIP*s *4*, *18* and *25*) were highly expressed in the leaf tissue, five genes (*RcbZIP*s *7*, *20*, *16*, *32* and *11*) were tissue-specifically expressed in root tips, three genes (*RcbZIP*s *40*, *8* and *49*) were tissue-specifically expressed in endosperm, and four genes (*RcbZIP*s *28*, *43*, *37* and *5*) were preferentially expressed in the seed I tissue. Interestingly, the homolog of the *RcbZIP25* is *HYH* (*AT3G17609*) in *Arabidopsis*, which is involved in light regulation during leaf photosynthesis (Brown and Jenkins 2008; Holm et al. 2002), consistent with the high expression of *RcbZIP25* in leaf of castor bean. However, the function of most *bZIP* genes is yet unknown.

To identify potential bZIP transcription factors involved in oil accumulation in developing seeds of castor bean, we purposely analyzed the expressional differences of *RcbZIP* genes detected in seed I and II tissues. Results showed that 21 of 24 *RcbZIP* genes were down-regulated from the

initial developing stage to the fast oil accumulation stage. These down-regulated genes might be negative regulators or expressed early to activate downstream genes and then declined if they were involved in oil accumulation in the developing seeds of castor bean. Also, it is possible that these down-regulated genes might not be involved in oil accumulation in the developing seeds. Only three genes *RcbZIP40* (*30138.m003906*), *RcbZIP31* (*30055.m001548*), and *RcbZIP8* (*29002.m000146*) were significantly up-regulated in the fast oil accumulation stage. Further, we found the homolog of *RcbZIP31* gene is the *AtbZIP67* (*AT3G44460*) in *Arabidopsis*, which has been identified to be a critical transcription factor responsive to ABA signal in the seed filling and storage accumulation of developing seeds (Bensmihen et al. 2005). The expressional up-regulation of *RcbZIP31* gene may imply its involvement in oil accumulation in developing seeds of castor bean. Further studies focusing on the functional analysis of *RcbZIP31* might partially discover the mechanism underlying the regulation of oil accumulation in developing seeds of castor bean. In addition, both *RcbZIP*s *40* and *8* were specifically up-regulated in seed II and immature endosperm, implying their functional involvement in storage material accumulation or endosperm genesis in developing seeds of castor bean, though their function is yet uncertain.

The expression of nine genes (*RcbZIP*s *2*, *3*, *9*, *13*, *15*, *21*, *22*, *27* and *36*) was not detected in any tissue tested, implying their functions might not be involved in these tissues under normal growth conditions. Further, we detected whether the nine genes were expressed under drought stress using RT-PCR techniques. Results from tests showed that four *RcbZIP*s (*2*, *9*, *22* and *36*) were expressed in seedlings of castor bean (data not shown) under drought stress, suggesting that the functions of the four *RcbZIP*s 2, 9, 22 and 36 might be involved in responding to abiotic stresses. Comparing the five genes, whose expression was not detected in this study, with their homologs in *Arabidopsis*, we found that the *HY5* gene (*AT5G11260*), a homolog of *RcbZIP21*, was involved functionally in regulating the photomorphogenic development in seedling (Chattopadhyay et al. 1998); the *ABI5* gene (*AT2G36270*), a homolog of *RcbZIP13*, was involved functionally in mediating embryogenesis in the late embryo development (Finkelstein and Lynch 2000). It is understandable that the expression of some genes (if their functions are involved in regulating the growth and development of other organs specific) is not detected for the limited samples tested in this study.

Conclusions

In total, 49 putative bZIP transcription factors were identified in castor bean, one of the most important non-edible oilseed crops and its seed oil is broadly used for industrial applications. Further, the 49 bZIP transcription factors were characterized according to the conserved amino acid residues within bZIP domain, the conserved motifs and gene organization in structure, phylogenetical analysis, and global expression profiles among different tissues using high-throughput sequencing. Results obtained from this study provide holistic information in understanding the molecular basis of the *bZIP* gene family in castor bean and other plants in the family Euphorbiaceae as well.

Acknowledgments We thank Zibo Academy of Agricultural Sciences, Shandong, China for providing the seeds of castor bean var. ZB306 elite inbred line. This work was supported by the "Hundreds of Talents" program of the Chinese Academy of Sciences (AL).

References

Acharya A, Ruvinov SB, Gal J, Moll JR, Vinson C (2002) A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: amino acids I, V, L, N, A, and K. Biochemistry 41:14122–14131

Aeschbacher RA, Schrott M, Potrykus I, Saul MW (1991) Isolation and molecular characterization of PosF21, an *Arabidopsis thaliana* gene which shows characteristics of a b-Zip class transcription factor. Plant J 1:303–316

Akpan U, Jimoh A, Mohammed A (2006) Extraction, characterization and modification of castor seed oil. Leonardo J Sci 8:43–52

Bensmihen S, Giraudat J, Parcy F (2005) Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during *Arabidopsis thaliana* embryo maturation. J Exp Bot 56:597–603

Betts MJ, Guigo R, Agarwal P, Russell RB (2001) Exon structure conservation despite low sequence similarity: a relic of dramatic events in evolution? EMBO J 20:5354–5360

Brown BA, Jenkins GI (2008) UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature Arabidopsis leaf tissue by requirement for UVR8, HY5, and HYH. Plant Physiol 146:576–588

Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G, Cahoon EB, Gedil M, Stanke M, Haas BJ, Wortman JR, Fraser-Liggett CM, Ravel J, Rabinowicz PD (2010) Draft genome sequence of the oilseed species *Ricinus communis*. Nat Biotechnol 28:951–953

Chattopadhyay S, Ang LH, Puente P, Deng XW, Wei N (1998) Arabidopsis bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. Plant Cell 10:673–683

Choi HI, Hong JH, Ha JO, Kang JY, Kim SY (2000) ABFs, a family of ABA-responsive element binding factors. J Biol Chem 275:1723–1730

Deppmann CD, Acharya A, Rishi V, Wobbes B, Smeekens S, Taparowsky EJ, Vinson C (2004) Dimerization specificity of all 67 B-ZIP motifs in *Arabidopsis thaliana*: a comparison to *Homo sapiens* B-ZIP motifs. Nucleic Acids Res 32:3435–3445

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Ellenberger TE, Brandl CJ, Struhl K, Harrison SC (1992) The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha-helices: crystal structure of the protein-DNA complex. Cell 71:1223–1237

Emili A, Greenblatt J, Ingles CJ (1994) Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. Mol Cell Biol 14:1582–1593

Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. Plant Cell 12:599–609

Foster R, Izawa T, Chua NH (1994) Plant bzip proteins gather at ACGT elements. FASEB J 8:192–200

Hobo T, Kowyama Y, Hattori T (1999) A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. P Natl Acad Sci USA 96:15348–15353

Hodges RS, Sodek J, Smillie LB, Jurasek L (1973) Tropomyosin: amino acid sequence and coiled-coil structure. Cold Spring Harb Sym 37:299–310

Holm M, Ma LG, Qu LJ, Deng XW (2002) Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis*. Gene Dev 16:1247–1259

Hossain MA, Cho JI, Han M, Ahn CH, Jeon JS, An G, Park PB (2010a) The ABRE-binding bZIP transcription factor OsABF2 is a positive regulator of abiotic stress and ABA signaling in rice. J Plant Physiol 167:1512–1520

Hossain MA, Lee Y, Cho JI, Ahn CH, Lee SK, Jeon JS, Kang H, Lee CH, An G, Park PB (2010b) The bZIP transcription factor OsABF1 is an ABA responsive element binding factor that enhances abiotic stress signaling in rice. Plant Mol Biol 72:557–566

Hurst HC (1995) Transcription factors.1. Bzip proteins. Protein Profile 2:105–168

Izawa T, Foster R, Chua NH (1993) Plant bZIP protein DNA binding specificity. J Mol Biol 230:1131–1144

Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F, ZIPRG b (2002) bZIP transcription factors in *Arabidopsis*. Trends Plant Sci 7:106–111

Kang JY, Choi HI, Im MY, Kim SY (2002) *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. Plant Cell 14:343–357

Katagiri F, Lam E, Chua NH (1989) Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature 340:727–730

Landschulz WH, Johnson PF, Mcknight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA-binding proteins. Science 240:1759–1764

Li RQ, Yu C, Li YR, Lam TW, Yiu SM, Kristiansen K, Wang J (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967

Liao Y, Zou HF, Wei W, Hao YJ, Tian AG, Huang J, Liu YF, Zhang JS, Chen SY (2008) Soybean *GmbZIP44*, *GmbZIP62* and *GmbZIP78* genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic *Arabidopsis*. Planta 228:225–240

Lu G, Gao CX, Zheng XN, Han B (2009) Identification of OsbZIP72 as a positive regulator of ABA response and drought tolerance in rice. Planta 229:605–615

Meier I, Gruissem W (1994) Novel conserved sequence motifs in plant G-box binding-proteins and implications for interactive domains. Nucleic Acids Res 22:470–478

Moitra J, Szilak L, Krylov D, Vinson C (1997) Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. Biochemistry 36:12567–12573

Nakagawa H, Ohmiya K, Hattori T (1996) A rice bZIP protein, designated OSBZ8, is rapidly induced by abscisic acid. Plant J 9:217–227

Nakase M, Aoki N, Matsuda T, Adachi T (1997) Characterization of a novel rice bZIP protein which binds to the alpha-globulin promoter. Plant Mol Biol 33:513–522

Nijhawan A, Jain M, Tyagi AK, Khurana JP (2008) Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. Plant Physiol 146:333–350

Niu XP, Renshaw-Gegg L, Miller L, Guiltinan MJ (1999) Bipartite determinants of DNA-binding specificity of plant basic leucine zipper proteins. Plant Mol Biol 41:1–13

Ogunniyi DS (2006) Castor oil: a vital industrial raw material. Bioresource Technol 97:1086–1091

Oh SJ, Song SI, Kim YS, Jang HJ, Kim SY, Kim M, Kim YK, Nahm BH, Kim JK (2005) Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. Plant Physiol 138:341–351

Oshea EK, Rutkowski R, Kim PS (1992) Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell 68:699–708

Pinna LA (1990) Casein kinase 2: an 'eminence grise' in cellular regulation? Biochim Biophys Acta 1054:267–284

Qiu L, Yang C, Tian B, Yang JB, Liu A (2010) Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). BMC Plant Biol 10:278

Qu GZ, Zang LN, Hu XL, Gao CQ, Zheng TC, Li KL (2012) Co-transfer of *LEA* and *bZip* genes from *Tamarix* confers additive salt and osmotic stress tolerance in transgenic tobacco. Plant Mol Biol Rep 30:512–518

Riechmann JL, Heard J, Martin G, Reuber L, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu CL (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105–2110

Rook F, Gerrits N, Kortstee A, van Kampen M, Borrias M, Weisbeek P, Smeekens S (1998) Sucrose-specific signalling represses translation of the *Arabidopsis* ATB2 bZIP transcription factor gene. Plant J 15:253–263

Scheffzek K, Grunewald P, Wohlgemuth S, Kabsch W, Tu H, Wigler M, Wittinghofer A, Herrmann C (2001) The Ras-Byr2RBD complex: structural basis for Ras effector recognition in yeast. Structure 9:1043–1050

Schindler U, Menkens AE, Beckmann H, Ecker JR, Cashmore AR (1992) Heterodimerization between light-regulated and ubiquitously expressed *Arabidopsis* GBF Bzip proteins. EMBO J 11:1261–1273

Scholza V, da Silva JN (2008) Prospects and risks of the use of castor oil as a fuel. Biomass Bioenerg 32:95–100

Schworer CM, Colbran RJ, Keefer JR, Soderling TR (1988) Ca^{2+}/calmodulin-dependent protein kinase II. Identification of a regulatory autophosphorylation site adjacent to the inhibitory and calmodulin-binding domains. J Biol Chem 263:13486–13489

Shen HS, Cao KM, Wang XP (2007) A conserved proline residue in the leucine zipper region of AtbZIP34 and AtbZIP61 in *Arabidopsis thaliana* interferes with the formation of homodimer. Biochem Biophys Res Comm 362:425–430

Shimizu H, Sato K, Berberich T, Miyazaki A, Ozaki R, Imai R, Kusano T (2005) LIP19, a basic region leucine zipper protein, is a fos-like molecular switch in the cold signaling of rice plants. Plant Cell Physiol 46:1623–1634

Suckow M, von Wilcken-Bergmann B, Müller-Hill B (1993) Identification of three residues in the basic regions of the bZIP proteins GCN4, C/EBP and TAF-1 that are involved in specific DNA binding. EMBO J 12:1193–1200

Tak H, Mhatre M (2013) Cloning and molecular characterization of a putative bZIP transcription factor VvbZIP23 from *Vitis vinifera*. Protoplasma 250:333–345

Talanian RV, Mcknight CJ, Kim PS (1990) Sequence-specific DNA-binding by a short peptide dimer. Science 249:769–771

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using

maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K (2000) *Arabidopsis* basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. P Natl Acad Sci USA 97:11632–11637

Vinson CR, Hai TW, Boyd SM (1993) Dimerization specificity of the leucine zipper-containing bzip motif on DNA binding: prediction and rational design. Gene Dev 7:1047–1058

Wang JZ, Zhou JX, Zhang BL, Vanitha J, Ramachandran S, Jiang SY (2011) Genome-wide expansion and expression divergence of the basic leucine zipper transcription factors in higher plants with an emphasis on sorghum. J Integr Plant Biol 53:212–231

Wei KF, Chen J, Wang YM, Chen YH, Chen SX, Lin YN, Pan S, Zhong XJ, Xie DX (2012) Genome-wide analysis of bZIP-encoding genes in maize. DNA Res 19:463–476

Yin YH, Zhu Q, Dai SH, Lamb C, Beachy RN (1997) RF2a, a bZIP transcriptional activator of the phloem-specific rice tungro bacilliform virus promoter, functions in vascular development. EMBO J 16:5247–5259

Zeng X, Herndon AM, Hu JC (1997) Buried asparagines determine the dimerization specificities of leucine zipper mutants. P Natl Acad Sci USA 94:3673–3678

Zou MJ, Guan YC, Ren HB, Zhang F, Chen F (2008) A bZIP transcription factor, OsABI5, is involved in rice fertility and stress tolerance. Plant Mol Biol 66:675–683