1 Genome-wide identification and characterization of Transcription
2 Factors of Basic Leucine Zipper Family in *Malus domestica*
3 Zhengrong Zhang, Haoran Cui, Shanshan Xing, Xin Liu, Xuesen Chen and Xiaoyun
4 Wang*
5 College of Life Science, State Key Laboratory of Crop Biology, Shandong
6 Agricultural University, Shandong Taian 271018, People's Republic of china
7 *Corresponding author:
8 Xiaoyun Wang Tel: +86 538 8242922; Fax: +86 538 8242217
9 E-mail address: xyunwang@sdau.edu.cn
10 Author disclosure statement: The authors declare no conflicts of interests.
11

12 **Abstract**

13 Basic leucine zipper proteins (bZIP) contain a basic DNA-binding region and a
14 leucine zipper region, acting as transcriptional factors in regulation of gene expression
15 exclusively in eukaryotes. In this investigation, total 116 bZIP members were
16 identified in apple genome and mapped on all 17 chromosomes with various densities
17 as *M.bZIPs*. All these members were divided into six groups according to the
18 phylogenetic relationship combining with bZIPs from rice and *Arabidopsis*.
19 Investigating gene structure of *M.bZIPs*, five splicing patterns of intron were found in
20 the DNA-binding region with no splicing position and splicing positions at different
21 nucleotide of codons or different positions. Analyzing of protein structure of *M.bZIPs*,
22 twenty-five motifs were identified with certain characteristic in different phylogenetic
23 groups. To predict dimerization of leucine zipper region, the key positions of amino
24 acid in heptad(s) were investigated. The results showed that most *M.bZIPs* may form
25 hetero-dimer or homo-dimer and some *M.bZIPs* may form both. Expression
26 experiment revealed that *M.bZIP* genes have organ-specific expression and widely
27 expressed in flowers, leaves, and fruits. To investigate the response of *M.bZIPs* to
28 abiotic stresses, the promoter sequences of randomly selected *M.bZIP* genes were
29 analyzed. *Cis*-acting elements related to multiple stresses were found existing widely
30 in promoter sequences. Quantitative real-time PCR results further demonstrated that
31 the expression of some *M.bZIP* genes were quite sensitive to exogenous abscisic acid
32 and osmotic treatments.

33 Keyword: Basic leucine zipper; Genomic survey; Gene expression; Organ-specific

34 expression; Stress response

35

1. Introduction

As a large family, basic leucine zipper proteins (bZIP) are widely present in eukaryotes (Hurst 1994). There are different bZIP members in different species, for example, 17 in *Saccharomyces cerevisiae*, 31 in *Caenorhabditis elegans*, 27 in *Drosophila* (Fassler et al. 2002), 56 in human (Vinson et al. 2002). Generally, bZIP members in plants are much more, 75 in *Arabidopsis* (Jakoby et al. 2002), 89 in rice (Nijhawan et al. 2008), 92 in *sorghum* (Wang et al. 2011) and 131 in soybean (Liao et al. 2008). More bZIP members in plants may suggest that bZIPs are more important in plants than in animals. As transcription factor, bZIP binds to negatively charged DNA in DNA-binding region while the leucine zipper region formed homo- and/or heter-dimer to maintain the stability of the combination. The DNA-binding region is highly conserved, containing about 18 amino acid residues (Asn-X₇-Arg/Lys-X₉). Leucine zipper region is made up of heptad(s) with repeated leucine or similar hydrophobic amino acids. The amino acids at positions nearing interface of leucine zipper determine the dimerization stability and specificity (Vinson et al. 2002, Newman and Keating 2003).

Plant bZIPs is reported to be involved in diverse physiological processes, especially responses to abiotic/biotic stresses. In maize, *ZmbZIP17* functions as a TF by interacting with ABA-responsive *cis*-elements (ABRE) of promoter in gene, acting as an stress transducer in endoplasmic reticulum (Yang et al. 2013). *OsbZIP71* is also strongly induced in ABA-mediated drought and salt tolerance in rice (Liu et al. 2014). Compared to animal, various mechanism of anti-stress is more important, since plants

are non-mobility. Although some plant bZIPs were verified and some investigations on their function have been reported. However, such studies have mostly focused on herbaceous plant species, and the bZIP genes of woody economically important fruit tree species have received less attention.

Apple (*Malus domestica*) is one of the most economically important woody plant and fruit crops in the world. Recently, the divergence of the bZIP gene family in apple, strawberry and peach were reported. It showed that bZIP genes exist multiple modes of gene evolution after duplication (Wang et al. 2015). Using iTAK-Plant Transcription factor and Protein Kinase Identifier and Classifier database, Zhao et al. obtained 112 apple bZIP members and divided all these members into 11 groups. Detailed investigation of some members in one group (group A) indicated that they had organ-specific expression and were responsive to high salinity and drought, as well as to different phytohormones (Zhao et al. 2016).

Studies on bZIP family from different species indicated that gene structure and splicing patterns of intron are important for their functional evolution, since it would affect the coding sequences of protein, especially for the functional basic domain (Liu and Chu 2015). As TFs, bZIP regulate the expression of other genes by binding DNA, which functional motifs are quite important to their function. Dimerization is also an important functional characteristic of bZIP protein which determines the stability of the binding between bZIP transcription factor and DNA.

In the present study, total 116 *M.bZIP* encoding genes were identified from the apple genome and divided them into different groups based on phylogenetic relationship.

80 Splicing patterns of intron in DNA-binding region and characterized dimerization
 81 property within the leucine zipper regions were analyzed. As protein structure,
 82 different functional motifs were identified, some motifs were involved in
 83 ABA-response signal-transduction pathway and abiotic stress response. *Cis*-acting
 84 elements related to abiotic stress tolerance and hormone signaling were massive
 85 existence in the promoters of *M.bZIP* genes. The expressions of some *M.bZIP* genes
 86 which were randomly selected in different phylogenetic groups were quite responsive
 87 to the ABA and osmotic treatments.
 88

2. Results

2.1. Identification, Chromosomal Location and Expansion Patterns of bZIP

Genes in Apple

To identify bZIP proteins in apple genome, two methods were used. For the first method, all of the identified sequences of bZIP proteins from *Arabidopsis* and rice were used as queries to search the Golden Delicious peptide database (GDR <http://www.rosaceae.org/>) using software BLASTp (Gagne et al. 2002). Trying not to miss potential proteins and taking account of credible results, the E value cutoff was set to 0.001 as in references of similar investigations (Li et al. 2011). From this method, about 700 proteins were screened out. For the second method, Hidden Markov Model (HMM) profile of the bZIP domain from characteristics of the amino acid sequences established by the pfam website were used as a query to search in the Golden Delicious peptide database. Total 199 proteins were obtained from this method. To further identify whether all above proteins contains the bZIP domain, the sequences were submitted to the Interpro Database to be verified. Finally, 116 proteins have been demonstrated to have bZIP domain. The detail information of these proteins with their relative encoding genes was listed in Table 1.

For mapping 116 *M.bZIP* genes on the apple chromosomes, their gene ID and location data were retrieved from genome annotations downloaded from the Genome Database for Rosaceae (GDR <http://www.rosaceae.org/>). The chromosome map showing the physical location of all encoding *M.bZIP* genes (for short *M.bZIP* genes) was generated using procedure of MapDraw (Figure 1). Most *M.bZIP* (107 of 116) were

mapped on all 17 apple chromosomes and designated as *M.bZIP* 1-107 according to their top-to-bottom position on chromosomes from 1 to 17, while nine genes could not be mapped on any chromosome and designated as *M.bZIP*108-116. Presumably these nine genes could be localized to unassembled genomic sequence scaffolds (Li et al. 2011). The number of *M.bZIP* genes on different chromosomes was variable. There were four chromosomes harboring low density of *M.bZIP* genes, chromosome 1, 5, 6 and 17 with only two *M.bZIP* genes on each chromosome. Other chromosomes had relatively high density and *M.bZIP* genes usually emerged with gene clusters on each chromosome. Further analyzed the distance between genes which located in the cluster, within 30kb of each other was defined as tandem duplication gene (Du et al. 2013). These tandem duplicates genes were highlighted with green block as shown in Figure 1 and listed in Table S1. Based on this result, we hypothesized that the tandem duplication of genes was one reason for the expansion of the *M.bZIP* genes. In addition to gene tandem duplication, a genome-wide duplication (GWD) event also occurred in apple and formed several intra-genomic homologous regions (Li et al. 2011). Large numbers of *M.bZIP* genes were located in homologous regions (gray shadow region in Figure 1) on different chromosomes, which suggest that genome duplication was another important way for expansion of *M.bZIP* genes.

2.2. Phylogenetic Analysis and Classification of the *M.bZIP* family

Previous studies have shown that the evolution of *bZIP* genes was conservative in plants (Liu and Chu 2015). To obtain the credible phylogenetic relationship of the *M.bZIP* proteins, *bZIP* members from apple (116), *Arabidopsis* (72) and rice (89)

were used together to construct a phylogenetic tree (Figure 2). Comparing evolutionary grouping of *Arabidopsis* and rice from previous reports, the phylogenetic tree constructed in this investigation was credible. All *M.bZIP* members can be divided into six groups, designated as group I to VI, with 25, 35, 13, 24, 8 and 11 members in each group, respectively. The group II was the largest group with 35 members, and the group V was the least with only 8 members. Most of the *M.bZIP* genes exhibited closer relationship with *AtbZIPs* than *OsbZIPs* and they were hierarchical clustered in the same clade in the phylogenetic tree, which suggested that dicots of apple and *Arabidopsis* might had the similar patterns of their bZIP families in evolutionary.

2.3. Gene Structure of *M.bZIP* genes

As a result of key events in evolution, gene structure including the number and distribution of exon/intron, offers insights into understanding the emergence and evolution of a given gene (Patthy 1987). The exon/intron organization of 116 *M.bZIP* was analyzed based on the number, distribution and splicing patterns of intron. In total, 23 of 116 *M.bZIP* genes contained no intron and 93 of 116 *M.bZIP* genes contained varied introns, the number from one to twenty-two (Table 1). The members of *M.bZIP* genes in each group possessed similar gene structure in introns number and distribution. The number of introns in each group was uneven, with intronless genes only found in group IV and VI, dramatically varied from one to twenty-two introns in the group I. In the group II, most genes had less than six introns, and in the group V, most genes had more than ten introns (Table 1).

Further analyzing splicing patterns of DNA sequence encoding the conservative DNA-binding region (basic and hinge region) of bZIPs, five splicing patterns of intron were characterized and designated as pattern S-I to S-V (Figure 3 and Table S2). The pattern S-I was the most prevalent and lacked intron in this regions with 43 *M.bZIPs* from group IV and VI. The pattern S-II had one intron in phase 2 (splicing occurred between the second and the third nucleotide in one codon) in the basic region, while the pattern S-III had one intron in phase 0 (splicing occurred between codons) located between amino acid Gln and Ala. The members in pattern S-III was completely corresponded to that from group III and V except *M.bZIP100*. Pattern S-IV had two introns (each in phase 0), one located between amino acid Lys and Arg within the basic region and the other located between amino acid Gln and Ala within the hinge region, which member was exclusive in group I. The pattern S-V had only one intron in phase 1 (splicing occurred between the first and the second nucleotide in one codon) in the hinge region. Gene structure showed that each group of *M.bZIPs* mostly shared the same intron/exon structural pattern and remained conserved during the course of evolution of *M.bZIP* family genes, which were very consistent with those in other plants.

2.4. Conserved Structural Features of *M.bZIP* Proteins

The bZIP domain is the core of the bZIP proteins, which preferentially binds to the promoter of their downstream target genes on specific *cis*-elements. To analyze the characteristics of the bZIP domain of *M.bZIPs*, the amino acid sequence logo of bZIP domain was generated by submitting all the sequences of 116 members to the MEME

analysis tool (Figure 4). Most bZIP domain of *M.bZIPs* contained 39 amino acids and residues at position 11, 19, 29, 36 were much conserved as Asn, Arg /Lys, Leu and Leu. In all *M.bZIPs*, a conservative sequence could be identified from residue Asn (position 11) as Asn-X₇-Arg/Lys-X₉-Leu-X₆-Leu (Table S3), which was accord with that of bZIP domains in other plants. Besides, some amino acids especially charged (Asp, Glu, Arg and Lys) were also conservative in this region (gray shadow in Figure 4).

Besides the bZIP domain, bZIP proteins usually contain additional conserved motifs which might be related to the function of bZIP proteins as potential sites. A total of 25 motifs, including the conserved bZIP domain (be named motif 1) were identified in 116 *M.bZIPs* (Table S4) and their distribution was shown in Figure 5. Except motif 1, motif 7 was the most widely distributed in 69 *M.bZIP* members in each group. Some motif shared by several groups, motif 9 presented in four groups (group II, III, IV and V), motif 12 presented in three groups (group II, IV and VI), and motif 11 presented in two groups (group IV and V). Most of the conserved motifs, however, were presented in specific group, for example, motif 2, 3, 6, 14, 15 and 19 only existed in the group I, motif 4, 5, 13, 16, 20, 21, 23 and 24 only presented in the group II. Group-specific motifs may indicate their specific functions in each group.

Protein phosphorylation, catalyzed by protein kinase is an important modification and directly related to protein function. Two conserved sequences, R/KxxS/T and S/TxxD/E, were verified as phosphorylated sites by a Ca²⁺ independent protein kinase and casein kinase II, respectively (Choi et al. 2000, Finkelstein and Lynch 2000,

199 Nijhawan et al. 2008). Six motifs (motif 6, 8, 9, 12, 14 and 24) were verified
 200 harboring phosphorylation sites of casein kinase II. Nine motifs (motif 2, 3, 4, 11, 15,
 201 16, 19, 20 and 23) harbored function site of Ca^{2+} independent protein kinase. Motif 3
 202 and motif 18 contained both sites of above two kinases. Other functional sites were
 203 also identified in motifs. Motif 4 and motif 7 were extension of Leucine zipper region,
 204 function for the dimerization of *M.bZIPs*. Motif 6 only presented in the members of
 205 group I, was part of the DOG1 domain, which was related to the dormancy of the
 206 seeds (Bentsink et al. 2006). Motif 17 was found in six members of group V, with the
 207 presence of small stretches of Pro and aromatic amino acids Tyr, Phe and Trp, and is a
 208 part of the Pro-rich domain, similar motifs have been identified in arabidopsis, which
 209 have been shown to have transcription activation potential (Schindler et al. 1992).

210 **2.5. Predicted Dimerization of *M.bZIP* proteins**

211 bZIPs bind to DNA as dimer when acting as transcription factors (Deppmann et al.
 212 2004). The leucine zipper region in bZIPs, arranged in the form of heptad repeats, is
 213 responsible for dimerization. Within each heptad, the amino acid positions are named
 214 as *a*, *b*, *c*, *d*, *e*, *f* and *g* in order in human and *Arabidopsis* (Vinson et al. 2002,
 215 Deppmann et al. 2004). Following this criteria, two to nine heptads were identified in
 216 different *M.bZIPs* (Table S5). Stability and specificity of leucine zipper dimer are
 217 closely related to the amino acids present at the position *a*, *d*, *e* and *g*, since they are
 218 near the interface of leucine zipper (Wei et al. 2012). Amino acids at the position *a*
 219 and *d* are typically hydrophobic, forming a hydrophobic core is essential for
 220 interaction between two monomers. The position *e* and *g* usually present charged

221 amino acids, which mediate the specificity of dimerization as well as stability
222 (Nijhawan et al. 2008).

223 By analyzing all the 116 *M.bZIPs*, less than one tenth of charge amino acids (K, R, E
224 and D) presented at position *a*, while other amino acids especially hydrophobic
225 residues were predominant at the position *a* and *d* (Figure 6A). About 22% of amino
226 acids present at the position *a* were Asn, equivalent amount with that of bZIP from
227 maize (22%). Further analyzing the composition of each heptad, the highest frequency
228 of Asn at the position *a* occurred both in the second and fifth heptads, accounting for
229 59% and 58% respectively (Figure 6B). The high frequency of Asn at the position *a*
230 implied that it will form homo-dimerizing leucine zippers among the bZIP members,
231 because Asn produce more stable N-N interactions than other amino acids (Nijhawan
232 et al. 2008). Leucine present in position *d* was also import for dimerization stability.

233 The frequency of Leu was high to 70% at the position *d* in *M.bZIPs* (Figure 6A), a
234 similar level with that of bZIPs from rice (71%) and greater than that from
235 *Arabidopsis* (56%).

236 Different from the position *a* and *d*, the charged amino acids occupied near half in the
237 position *e* and *g*, with frequencies of 45% and 60% respectively (Figure 6A).The
238 attractive and repulsive *g* and *e* pairs governs dimerization properties of bZIPs
239 (Deppmann et al. 2004). The amino acids present in each heptad of leucine zippers of
240 *M.bZIPs* were characterized (Table S5) and the histogram of their frequency was
241 presented in Figure 6C. Maximum frequency of the complete *g* and *e'* pairs appeared
242 in the first heptad, showing basic repulsive in blue block. High frequency of attractive

243 pairs were found in the second, fifth and ninth heptad, moreover, only -/+ attractive
 244 pairs were found in the ninth heptad.

245 Attractive pairs and presence of Asn in position *a* contribute to homo-dimerization,
 246 while repulsive and incomplete pairs, and charged residues in position *a* may favour
 247 hetero-dimerization (Wei et al. 2012, Liu and Chu 2015). According to these
 248 principles, total 116 *M.bZIPs* were classified into 29 sub-families (Table S5). One
 249 subfamily favoured homo-dimerization (subfamily BZ1 in Table S5), since attractive
 250 pairs presented in the first and second heptads and lacking any repulsive interactions.
 251 Two subfamilies could be considered having hetero-dimerizing specificity because
 252 they contained only repulsive interhelical interactions (subfamily BZ2 and BZ3).
 253 There was a subfamily (subfamily BZ29) was thought not to form dimer because
 254 these proteins lacked leucine zipper region. Other twenty-five subfamilies had both
 255 homo- and hetero-dimerization properties.

256 **2.5. The Organ-Specific Expression of *M.bZIP* Genes**

257 Evidences suggested that bZIPs are widely involved in the integration and
 258 development of many organs in plants, such as seed maturation and germination (Toh
 259 et al. 2012), floral induction and development (Walsh and Freeling 1999, Wigge et al.
 260 2005). To further analyze the organ-specific expression of *M.bZIP* genes, GSE42873,
 261 an expression profile from the GEO DataSets were used. According to the expression
 262 of 116 *M.bZIP* genes in different organs, four expression types were divided and
 263 designated as type I-IV (Figure 7). Generally, most of *M.bZIP* genes were higher
 264 expression in flowers, fruits and leaves, indicating that *M.bZIP* genes widely

participated in the process of flower development and fruit ripening. Type I contained sixteen members with low expression in seedlings. Compared with other types, the gene expression in this type was relatively high in the seeds indicating their potential function in seed germination. In type II, twenty-two genes were clustered together due to their highest expression in the flowers and lowest expression in the seeds. There were sixty-eight genes that were clustered into type III because they showed a low expression in the seeds, seedling, roots and stems compared with other organs. In type IV, all of the ten genes had a low expression in the seeds, but most of them had a high expression in the roots, stems and seedlings, implying that the gene of type IV involved in the vegetative growth of apple.

To verify credibility of expression profile, five *M.bZIP* genes in different types were randomly selected to check the organic expression by qRT-PCR technique using gene-specific primers (Table S6). The expression trend of *M.bZIP* genes in different organs was consistent with the expression data in the GSE42873 except *M.bZIP52* (Figure 8). Compared with other organs, *M.bZIP 52* has high expression in the seeds with checked by qRT-PCR, while low expression was found in expression profile in the seeds. Our results demonstrated that the expression profile of *M.bZIP* genes in different organs was reliable.

2.6. Promoter Analysis and Expression Assay of *M.bZIP* Genes under Abiotic Stress Conditions

Previous studies suggested that some bZIP proteins are involved in signaling and responses to abiotic/biotic stimuli (Kesarwani et al. 2007). To investigate the role of

287 *M.bZIP* in stress response, the promoter sequences of fourteen randomly selected
288 *M.bZIP* genes in different phylogenetic groups were analyzed. Total 163 *cis*-acting
289 elements were detected, which distribution and function annotation were shown in
290 Table S7. Important *cis*-acting elements related stress-response including
291 ABA-response element (ABRE), dehydration-response element (DRE),
292 low-temperature responsive element (LTRE) were widely existed in the promoters of
293 *bZIP* genes as shown in Table 2. These elements have been demonstrated to bind the
294 relative proteins to modulate stress response in other plants (Dai et al. 2007, Maestrini
295 et al. 2009, Zhang et al. 2009, Li et al. 2012). Besides, some elements found in the
296 promoters of *M.bZIPs* function for binding of important transcription factors, which
297 factors were closely related to various abiotic stresses such as transcription factor
298 MYB and MYC (Table 2). High frequency of above elements in the promoters
299 suggested that *M.bZIP* genes were involved in the response of multiple stresses in
300 apple.

301 To further examine whether the expressions of *M.bZIP* gene were induced by
302 abiotic stresses, 20% PEG and 100 μ M exogenous ABA were treated to 1-year-old
303 apple seedlings. The expression levels of seventeen genes randomly selected in
304 different groups were checked by qRT-PCR. Most of the genes were up-regulated or
305 down-regulated under the abiotic stress treatments (Figure 9). Among them, six
306 *M.bZIP* genes were significantly up-regulated by ABA treatment. The relative
307 expressions of three genes (*M.bZIP6*, *M.bZIP45* and *M.bZIP73*) were up-regulated,
308 approximately 3-fold (3 h), 6-fold (1 h) and 3-fold (1 h), respectively. One gene

309 (*M.bZIP101*) exhibited decreased expression at 1 h after treatment, but then increased
310 and reached the initial level after 3 h treatment. While the expression of ten *M.bZIP*
311 genes showed an obvious decrease at all time points after treatment. Under osmotic
312 stress, the expression of sixteen *M.bZIP* genes showed up-regulated and three of these
313 (*M.bZIP46*, *M.bZIP45* and *M.bZIP93*) were very obvious, approximately reaching
314 60-fold (9 h), 11-fold (3 h) and 10-fold (6 h), respectively. There was only one
315 *M.bZIP* gene down-regulated in transcript levels at all time points after treatment with
316 PEG. Combined with the results of the treatment of two kinds of stress, five *M.bZIP*
317 genes increased in response to the two kinds of stress, indicated these genes play
318 multiple roles in the regulation of the abiotic stress signaling pathways.
319

320 3. Discussion

321 In this study, using Local Blast and Hidden Markov Model Profile, total 116 genes
 322 encoded bZIP proteins were identified in the 'Golden Delicious' apple genome, which
 323 was by far the largest compared with the estimates for other plants such as 76 in
 324 *Arabidopsis* (Jakoby et al. 2002) and 89 in rice (Nijhawan et al. 2008). It can be
 325 speculated that the presence of additional *M.bZIP* genes in the apple genome may
 326 reflect the great need for these genes in the complicated transcriptional regulations.
 327 Similar results were also obtained by Wang and his colleagues, when they
 328 investigated the divergence of the bZIP gene family in strawberry, peach and apple
 329 (Wang et al. 2015).

330 DNA-binding region of bZIPs are relatively conservative. Analyzing DNA sequence
 331 encoding this region, different splicing patterns were found. Among 116 *M.bZIP*
 332 genes, 42 members have no intron in this region, while other members with introns
 333 have four splicing patterns. In other plants, *Os*bZIPs contained seven splicing patterns
 334 and *Hvb*bZIPs contained six splicing patterns (Nijhawan et al. 2008, Pourabed et al.
 335 2015). Comprehensive considering the splicing patterns of bZIP gene family of apple,
 336 rice and barley, we find that although there are some differences of the splicing phase
 337 and the interrupted amino acids, the positions of intron are highly conserved. Splicing
 338 patterns in basic and hinge regions of bZIPs indicated that the positions of introns
 339 may be important for the evolution and divergence of the function in the *M.bZIP*
 340 family.

341 Compared to the gene structure, protein structure of *M.bZIPs* is directly related to the

function of transcription factors. bZIP domain is the core domain of *M.bZIP* protein, however, the function diversity of the *M.bZIP* gene family may also be contributed by additional regions in bZIP proteins. In total, 25 motifs were identified in *M.bZIPs*, among them seventeen motifs appeared to contain potential phosphorylation sites (R/KxxS/T and S/TxxD/E). These phosphorylation sites were closely related to ABRE-dependent ABA signaling pathways (Fujita et al. 2013). Motif 4 and motif 7 were found to be extension of leucine zipper regions, which suggested that these two kinds of motifs may function in dimerization of *M.bZIPs*.

Gene expression patterns can provide important clues for gene function. So we also examined the expression of 116 *M.bZIP* genes in the seven organs (seeds, seedling, roots, stems, leaves, flowers and fruits) using the expression profile GSE42873. All of the *M.bZIP* genes exhibited diverse expression patterns, which represented the distinct roles of the individual *M.bZIP* genes. Interestingly, most of *M.bZIP* genes mainly expressed in leaves, flowers and fruits indicated that *M.bZIP* genes widely participated in the process of flower development and fruit ripening. We further randomly selected five *M.bZIP* genes in different types to verify the organ expression, Our results about *M.bZIP* genes in different organs by qRT-PCR were approximately consistent with the expression profile data.

Previous studies have shown that bZIP genes play pivotal roles in developmental processes and responses of multiple stresses (Izawa et al. 1994, Yin et al. 1997, Chen et al. 2012). Promoter analysis revealed that promoter regions of fourteen *M.bZIP* genes contained variety of *cis*-acting elements associated with stresses, including

364 ABRE, DRE, LTRE, MYB and MYC. Combined with the promoter analyze,
365 qRT-PCR was used to further investigate the expression of *M.bZIP* genes under stress.
366 The expression revealed that some *M.bZIP* genes were sensitively responded to stress
367 treatment. Five *M.bZIP* genes were identified to respond to two abiotic stress
368 treatments tested, the results suggested that these genes may be stress-sensitive gene
369 with a potential role in ABA and osmosis acclimation. Some bZIP genes, such as
370 *M.bZIP16* and *M.bZIP24*, exhibited opposing expression patterns under ABA and
371 PEG treatment, which indicate that these genes are involved in the communication
372 between different signal transduction pathways. Except *M.bZIP104*, sixteen of
373 seventeen genes which were randomly selected in different phylogenetic groups
374 shown respond to PEG treatment. It is noteworthy that the expression of *M.bZIP46*,
375 *M.bZIP45* and *M.bZIP93* showed the maximum 60-fold, 11-fold, and 10-fold increase
376 at 9 h, 3 h and 6 h, indicating that these genes play a major role in the response to
377 osmosis conditions. However, in the fourteen genes, the expression results from
378 qRT-PCR did not agree perfectly with *cis*-acting elements analysis. For instance,
379 *M.bZIP46* contained 16 *cis*-acting elements and its relative expression level was more
380 than 60-fold in relation to controls under PEG treatment at 9 h. whereas the value for
381 *M.bZIP24* carrying 67 *cis*-acting elements was only 3-fold (Table 2; Figure 9). One
382 explanation may be that several elements lost their activities or performed in a
383 negative way. Another explanation is that different *cis*-acting elements are different
384 ability to initiate gene expression. Overall, our investigation about bZIP family in the
385 apple may lead to further understanding the relationship of function and structure of

386 bZIP family members.

387

388 **4. Materials and methods**

389 **4.1. Genome-wide identification of bZIP proteins in apple**

390 The genome and proteome sequences of *Arabidopsis* and rice were downloaded from
 391 the NCBI website. Two methods were used. For the first method, BLASTp was
 392 performed by the Bioedit program with an E value cutoff of 0.001 to search the
 393 Golden Delicious peptide database (GDR <http://www.rosaceae.org/>) using bZIP
 394 protein sequences in *Arabidopsis* and rice as queries. For the second method,
 395 Hmmer3.0 software was downloaded from the HMMER website
 396 (<http://hmmer.janelia.org/>) to perform a global search of the apple proteome. The
 397 HMM profile (PF00170) of the bZIP domain was download from the pfam website
 398 (<http://pfam.xfam.org/>) (Finn et al. 2013). All of the proteins that were obtained by
 399 the two methods were submitted to the Interpro Database
 400 (<http://www.ebi.ac.uk/interpro/>) and smart Database
 401 (<http://www.smart.embl-heidelberg.de/>) to ensure the presence of bZIP domain.

402 **4.2. Phylogenetic analysis of *M.bZIP* genes**

403 116 *M.bZIP* protein sequences, 89 *OsbZIP* protein sequences and 72 *AtbZIP* protein
 404 sequences were executed multiple alignments by Clustal X 1.83 program (Larkin et al.
 405 2007). Phylogenetic trees were constructed using phylip-3.695 by the
 406 Neighbor-Joining (NJ), and the bootstrap test carried out with 1000 replications. The
 407 picture of those trees was drawn using FigTree (Version1.4.2).

408 **4.3. Chromosomal locations and gene duplications for *M.bZIP* genes**

409 Location data were retrieved from genome annotations downloaded from the Genome

410 Database for *Rosaceae* (GDR <http://www.rosaceae.org/>). The chromosome map
411 showing the physical location of all *M.bZIP* genes was generated using revised
412 version of MapDraw (Liu and Meng 2003). Tandem duplications were characterized
413 as adjacent genes located within 10 predicted genes apart or within 30kbp of each
414 other.

415 **4.4. Genomic and proteome structures of *M.bZIP* transcription factors**

416 To explore the diverse exon/intron organizations of *M.bZIP* genes, we compared the
417 predicted coding sequences of *M.bZIP* genes with their corresponding genomic
418 sequences using GSDS software (<http://gsds.cbi.pku.edu.cn>) (Hu et al. 2014). The
419 intron distribution and splicing phase within the genomic sequences were further
420 predicted based on the alignments of the genomic sequences and cDNA full-length
421 sequences obtained by Spidey website (<http://www.ncbi.nlm.nih.gov/spidey/>).

422 The isoelectric point (PI) of each protein was calculated by online ExPASy programs
423 (<http://www.expasy.org/tool/>). The conserved motifs were identified by the online
424 MEME analysis tool (<http://meme-suite.org/>). The Interpro Database was used to
425 identify conserved functional domains in apple *M.bZIP* proteins.

426 **4.5. Expression pattern of *M.bZIP* genes in apple**

427 To study the expression pattern of *M.bZIP* genes in different organs, series matrix data
428 from the expression profiles GSE42873 was downloaded from NCBI GEO datasets.
429 Expression data for the identified genes were extracted from the two datasets using
430 their unigene IDs with a Visual Basic (version6.0) script. Clustering was performed
431 by Cluster3.0 (<http://rana.lbi.gov/EisenSoftware.htm>) using the hierarchical clustering

method for the average linkage, a heatmap and clustering tree were constructed and viewed with Java Treeview (http://sourceforge.net/projects/jtreeview/?source=typ_redirect).

4.6. Promoter analysis of *M.bZIP* genes

To investigate *cis*-acting elements in the promoter sequences of *M.bZIP* genes, 2000bp of genomic DNA sequences upstream of the transcriptional start site were obtained from the apple genome. Promoter sequences were submitted to the PLACE database (<http://www.dna.affrc.go.jp/PLACE/>) and *cis*-acting elements were obtained (Higo et al. 1999).

4.7. Sample preparation and total RNA extraction

Golden Delicious plants which were cultivated in the fields were selected as the experimental material for qRT-PCR analysis. The apple (Golden Delicious) seedlings were separately treated in 100 μ M ABA and 20% PEG6000 for 0 h, 1 h, 3 h, 6 h and 9 h. 1g leaf samples from three individual plants of stressed and controlled seedlings were harvested for total RNA extraction. The total RNA was isolated from the leaves using the CTAB procedure (Gasic et al. 2004). RNA concentrations and A_{260}/A_{280} ratios were determined using a NanoDrop Spectrometer (ND-1000 Spectrophotometer, peqlab). The integrity of the RNA was detected by agarose gel electrophoresis. Qualified RNA was used for cDNA synthesis and quantitative real-time PCR.

For organ specificity expression of *M.bZIP* genes, organs of leaf, stem, root, and seed were harvested separately from one year old plants of Golden Delicious and the flowers were picked in summer. All of materials were stored at -80°C after freezing in

454 liquid nitrogen.

455 **4.8. Quantitative real-time PCR (qRT-PCR) analysis**

456 cDNA fragments were synthesized from total RNA using the TransScripTMone-step
 457 gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, Beijing
 458 china).The gene-specific primer pairs were designed based on target gene sequences
 459 using the software Beacon Designers 8.10. The apple *actin* gene was used as an
 460 internal control. The primer sequences were listed in table S6. The qRT-PCR was
 461 carried out with a Stratagene Mx3000P thermocycler (Agilent) in a final volume of 20
 462 µl that contained 1.4 µl cDNA, 10 µl 2*SYBR premix Ex TaqTM (Takara, Shiga,
 463 Japan), 7.8 µl H₂O and 0.8 µl (10 µM) primers. The thermal cycling conditions were
 464 as follows: 44 cycles of 95°C denaturation for 15 s, 55°C annealing for 30 s and 72°C
 465 extension for 15 s. The real-time PCR experiment was carried out at least three times
 466 under identical conditions.

467

468 **Abbreviations:**

469 ABA, abscisic acid; ABRE, abscisic acid responsive elements; bZIP gene, basic
470 leucine zipper encoding gene; DRE, dehydration-response element; LTRE,
471 low-temperature responsive element; qRT-PCR, quantitative real-time PCR;

472

473 **Acknowledgment** This work was supported by Special Research Fund of Public
474 Welfare of China Agricultural Ministry (201303093).

475

476 **Supplementary Material**

477 **Table S1** *M.bZIP* genes involved in tandem duplication

478 **Table S2** Splicing sites and splicing phase in basic and hinge region of *M.bZIP* genes

479 **Table S3** Alignment sequences of bZIP domain of *M.bZIPs*

480 **Table S4** Conserved motifs identified from 116 *M.bZIPs*

481 **Table S5** Amino acid sequences alignment of the leucine zipper region of 116

482 *M.bZIPs*

483 **Table S6** List of primers used in qRT-PCR expression analysis of *M.bZIP* genes

484 **Table S7** *Cis*-acting elements distribution in promoters of six *M.bZIP* genes and

485 function annotation

486

487 **Table captions**

488 **Table 1** The identified *M.bZIP* encoding genes and the relative information.

489 **Table 2** Distribution of ABRE, DRE, LTRE, MYB and MYC *cis*-acting elements in

490 *M.bZIP* promoters.

491

492 **Figure captions**

493 **Figure 1** Distribution of the *M.bZIP* genes in 17 apple chromosomes. The numbers to

494 the left of each chromosome represent a megabase between two *M.bZIP* genes.

495 Tandem duplications are indicated by light green, gray block represent intragenomic

496 homologous regions.

497 **Figure 2** The phylogenetic tree of bZIP proteins from apple, rice and *Arabidopsis*. It

498 was constructed by aligning sequences of bZIP protein from apple (116) *Arabidopsis*

499 (72) and rice (89).

500 **Figure 3** Splicing patterns of intron within the basic and hinge regions of the bZIP

501 domains of the *M.bZIP*s. P₁, P₂ and P₀ means splicing occurred after the first, second

502 and third nucleotide of the codon, respectively.

503 **Figure 4** Sequence logos of bZIP domain from all *M.bZIP*s. Blue shadows represent

504 the most conserved amino acids in the feature sequence of bZIP domain, gray

505 shadows represent the charge amino acids of bZIP domain.

506 **Figure 5** Protein structure analysis and motif identified in *M.bZIP*s. The bZIP domain

507 is shown in purple block, other predicted motifs are indicated in different color blocks

508 with numbers from 2 to 25. The details of predicted conserved motifs are given in

509 table S3.

510 **Figure 6** Prediction of dimerization properties of *M.bZIPs*. (A) The frequency of the
511 amino acids at position *a*, *d*, *e* and *g*, these positions directly related to dimerization of
512 the leucine zipper part of *M.bZIPs*. (B) The frequency of Asn residues present at the
513 position *a* in the nine heptads for *M.bZIPs*. (C) The ratio of attractive or repulsive pair
514 per heptad for *M.bZIPs*.

515 **Figure 7** Expression patterns of *M.bZIP* genes in different organs in apple. This
516 heatmap was generated based on the GSE expression data using cluster3.0 and
517 treeview software. Green indicates low expression, dark indicated intermediate
518 expression and red indicated high expression.

519 **Figure 8** The expression changes of representative *M.bZIP* genes in flowers, seeds,
520 roots, stems and leaves.

521 **Figure 9** The expression changes of representative *M.bZIP* genes in different groups
522 under ABA and PEG treatment.

523

References

- Bentsink, L., Jowett, J., Hanhart, C. J. and Koornneef, M. 2006 "Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis." Proceedings of the National Academy of Sciences. 103, 17042-17047.
- Chen, H., Chen, W., Zhou, J., He, H., Chen, L., Chen, H., et al. 2012 "Basic leucine zipper transcription factor OsbZIP16 positively regulates drought resistance in rice." Plant Sci. 193-194, 8-17.
- Choi, H., Hong, J., Ha, J., Kang, J. and Kim, S. Y. 2000 "ABFs, a family of ABA-responsive element binding factors." J Biol Chem. 275, 1723-1730.
- Dai, X., Xu, Y., Ma, Q., Xu, W., Wang, T., Xue, Y., et al. 2007 "Overexpression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis." Plant Physiol. 143, 1739-1751.
- Deppmann, C. D., Acharya, A., Rishi, V., Wobbes, B., Smeekeens, S., Taparowsky, E. J., et al. 2004 "Dimerization specificity of all 67 B-ZIP motifs in Arabidopsis thaliana: a comparison to Homo sapiens B-ZIP motifs." Nucleic acids research. 32, 3435-3445.
- Du, D., Zhang, Q., Cheng, T., Pan, H., Yang, W. and Sun, L. 2013 "Genome-wide identification and analysis of late embryogenesis abundant (LEA) genes in Prunus mume." Mol Biol Rep. 40, 1937-1946.
- Fassler, J., Landsman, D., Acharya, A., Moll, J. R., Bonovich, M. and Vinson, C. 2002 "B-ZIP proteins encoded by the Drosophila genome: evaluation of potential dimerization partners." Genome Res. 12, 1190-1200.
- Finkelstein, R. R. and Lynch, T. J. 2000 "The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor." Plant Cell. 12, 599-609.
- Finn, R. D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R. Y., Eddy, S. R., et al. 2013 "Pfam: the protein families database." Nucleic acids research, gkt1223.
- Fujita, Y., Yoshida, T. and Yamaguchi-Shinozaki, K. 2013 "Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants." Physiologia plantarum. 147, 15-27.
- Gagne, J. M., Downes, B. P., Shiu, S. H., Durski, A. M. and Vierstra, R. D. 2002 "The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis." Proc Natl Acad Sci U S A. 99, 11519-11524.
- Gasic, K., Hernandez, A. and Korban, S. S. 2004 "RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction." Plant Molecular Biology Reporter. 22, 437-438.
- Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. 1999 "Plant cis-acting regulatory DNA elements (PLACE) database: 1999." Nucleic acids research. 27, 297-300.
- Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J. and Gao, G. 2014 "GSDS 2.0: an upgraded gene feature visualization server." Bioinformatics, btu817.
- Hurst, H. C. 1994 "Transcription factors. 1: bZIP proteins." Protein Profile. 1, 123-168.
- Izawa, T., Foster, R., Nakajima, M., Shimamoto, K. and Chua, N.-H. 1994 "The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development." The Plant Cell. 6, 1277-1287.
- Jakoby, M., Weisshaar, B., Droge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., et al. 2002 "bZIP transcription factors in Arabidopsis." Trends Plant Sci. 7, 106-111.

565 Kesarwani, M., Yoo, J. and Dong, X. 2007 "Genetic interactions of TGA transcription factors in the
566 regulation of pathogenesis-related genes and disease resistance in Arabidopsis." *Plant Physiol.* 144,
567 336-346.

568 Larkin, M. A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., et al. 2007
569 "Clustal W and Clustal X version 2.0." *Bioinformatics.* 23, 2947-2948.

570 Li, W., Cui, X., Meng, Z., Huang, X., Xie, Q., Wu, H., et al. 2012 "Transcriptional regulation of
571 Arabidopsis MIR168a and argonaute1 homeostasis in abscisic acid and abiotic stress responses." *Plant*
572 *Physiol.* 158, 1279-1292.

573 Li, Y., Wu, B., Yu, Y., Yang, G., Wu, C. and Zheng, C. 2011 "Genome-wide analysis of the RING finger
574 gene family in apple." *Mol Genet Genomics.* 286, 81-94.

575 Liao, Y., Zou, H.-F., Wei, W., Hao, Y.-J., Tian, A.-G., Huang, J., et al. 2008 "Soybean GmbZIP44, GmbZIP62
576 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing
577 tolerance in transgenic Arabidopsis." *Planta.* 228, 225-240.

578 Liu, C., Mao, B., Ou, S., Wang, W., Liu, L., Wu, Y., et al. 2014 "OsZIP71, a bZIP transcription factor,
579 confers salinity and drought tolerance in rice." *Plant molecular biology.* 84, 19-36.

580 Liu, R. and Meng, J. 2003 "[MapDraw: a microsoft excel macro for drawing genetic linkage maps based
581 on given genetic linkage data]." *Yi chuan= Hereditas/Zhongguo yi chuan xue hui bian ji.* 25, 317-321.

582 Liu, X. and Chu, Z. 2015 "Genome-wide evolutionary characterization and analysis of bZIP
583 transcription factors and their expression profiles in response to multiple abiotic stresses in
584 *Brachypodium distachyon*." *BMC Genomics.* 16, 227.

585 Maestrini, P., Cavallini, A., Rizzo, M., Giordani, T., Bernardi, R., Durante, M., et al. 2009 "Isolation and
586 expression analysis of low temperature-induced genes in white poplar (*Populus alba*)." *J Plant Physiol.*
587 166, 1544-1556.

588 Newman, J. R. and Keating, A. E. 2003 "Comprehensive identification of human bZIP interactions with
589 coiled-coil arrays." *Science.* 300, 2097-2101.

590 Nijhawan, A., Jain, M., Tyagi, A. K. and Khurana, J. P. 2008 "Genomic survey and gene expression
591 analysis of the basic leucine zipper transcription factor family in rice." *Plant Physiol.* 146, 333-350.

592 Patthy, L. 1987 "Intron-dependent evolution: preferred types of exons and introns." *FEBS Lett.* 214,
593 1-7.

594 Pourabed, E., Ghane Golmohamadi, F., Soleymani Monfared, P., Razavi, S. M. and Shobbar, Z. S. 2015
595 "Basic leucine zipper family in barley: genome-wide characterization of members and expression
596 analysis." *Mol Biotechnol.* 57, 12-26.

597 Schindler, U., Terzaghi, W., Beckmann, H., Kadesch, T. and Cashmore, A. R. 1992 "DNA binding site
598 preferences and transcriptional activation properties of the Arabidopsis transcription factor GBF1."
599 *EMBO J.* 11, 1275-1289.

600 Toh, S., McCourt, P. and Tsuchiya, Y. 2012 "HY5 is involved in strigolactone-dependent seed
601 germination in Arabidopsis." *Plant signaling & behavior.* 7, 556-558.

602 Vinson, C., Myakishev, M., Acharya, A., Mir, A. A., Moll, J. R. and Bonovich, M. 2002 "Classification of
603 human B-ZIP proteins based on dimerization properties." *Mol Cell Biol.* 22, 6321-6335.

604 Walsh, J. and Freeling, M. 1999 "The *liguleless2* gene of maize functions during the transition from the
605 vegetative to the reproductive shoot apex." *Plant J.* 19, 489-495.

606 Wang, J., Zhou, J., Zhang, B., Vanitha, J., Ramachandran, S. and Jiang, S. Y. 2011 "Genome-wide
607 expansion and expression divergence of the basic leucine zipper transcription factors in higher plants
608 with an emphasis on sorghum." *J Integr Plant Biol.* 53, 212-231.

609 Wang, X.-L., Zhong, Y., Cheng, Z.-M. and Xiong, J.-S. 2015 "Divergence of the bZIP Gene Family in
610 Strawberry, Peach, and Apple Suggests Multiple Modes of Gene Evolution after Duplication."
611 International Journal of Genomics. 2015, 17.

612 Wei, K., Chen, J., Wang, Y., Chen, Y., Chen, S., Lin, Y., et al. 2012 "Genome-wide analysis of
613 bZIP-encoding genes in maize." DNA Res. 19, 463-476.

614 Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., et al. 2005 "Integration of
615 spatial and temporal information during floral induction in Arabidopsis." Science. 309, 1056-1059.

616 Yang, Y.-G., Lv, W.-T., Li, M.-J., Wang, B., Sun, D.-M. and Deng, X. 2013 "Maize membrane-bound
617 transcription factor Zmbzip17 is a key regulator in the cross-talk of ER quality control and ABA
618 signaling." Plant and Cell Physiology, pct142.

619 Yin, Y., Zhu, Q., Dai, S., Lamb, C. and Beachy, R. N. 1997 "RF2a, a bZIP transcriptional activator of the
620 phloem-specific rice tungro bacilliform virus promoter, functions in vascular development." EMBO J.
621 16, 5247-5259.

622 Zhang, Y., Chen, C., Jin, X. F., Xiong, A. S., Peng, R. H., Hong, Y. H., et al. 2009 "Expression of a rice
623 DREB1 gene, OsDREB1D, enhances cold and high-salt tolerance in transgenic Arabidopsis." BMB Rep.
624 42, 486-492.

625 Zhao, J., Guo, R., Guo, C., Hou, H., Wang, X. and Gao, H. 2016 "Evolutionary and Expression Analyses of
626 the Apple Basic Leucine Zipper Transcription Factor Family." Frontiers in Plant Science. 7, 376.

627

-27 Basic region -5 Hinge 1

DEKRQKRMLSNRESARRSRERKQAYVEEL
 UACGAAAAGCGACAAAAGCGAAUGCUUUCAAAUAAGGAAUCAUUACGACGAUCACGAGAACGAAAGCAAUUUAAGUAGAAAGAACUU

Numbers
of *M.bZIPs*

Numbers
of introns

S-I

42

0

S-II

31

1

S-III

29

1

S-IV

13

2

S-V

1

1

Group	Motifs	<i>M.bZIPs</i> belonging to this pattern
I		<i>M.bZIP</i> 33
		<i>M.bZIP</i> 9
		<i>M.bZIP</i> 8
		<i>M.bZIP</i> 35, 39, 69, 78, 79, 87, 89, 98, 99. 103, 115
		<i>M.bZIP</i> 2,25, 70, 71, 81, 82, 88, 113
		<i>M.bZIP</i> 13, 14, 31, 32
II		<i>M.bZIP</i> 21
		<i>M.bZIP</i> 18, 19, 20
		<i>M.bZIP</i> 40, 41, 73, 74
		<i>M.bZIP</i> 55, 101, 24, 76, 112
		<i>M.bZIP</i> 15, 16, 61, 62
		<i>M.bZIP</i> 17
		<i>M.bZIP</i> 114
		<i>M.bZIP</i> 54, 64, 22, 53, 107
		<i>M.bZIP</i> 2
		<i>M.bZIP</i> 72
		<i>M.bZIP</i> 111
		<i>M.bZIP</i> 85, 86, 109
		<i>M.bZIP</i> 110
III		<i>M.bZIP</i> 49
		<i>M.bZIP</i> 50, 56, 100, 102
		<i>M.bZIP</i> 7, 36, 59
IV		<i>M.bZIP</i> 26, 28, 48, 67, 68, 93
		<i>M.bZIP</i> 3
		<i>M.bZIP</i> 3
		<i>M.bZIP</i> 83, 84, 94
		<i>M.bZIP</i> 45, 57
V		<i>M.bZIP</i> 10, 11, 12, 34
		<i>M.bZIP</i> 96
		<i>M.bZIP</i> 23, 27, 30, 44, 52, 58, 65, 66, 77, 80, 90, 104, 106
VI		<i>M.bZIP</i> 29, 105, 91, 6
		<i>M.bZIP</i> 1, 37
		<i>M.bZIP</i> 38, 108
VI		<i>M.bZIP</i> 4, 5, 42, 43, 46, 47, 51, 63, 75, 92, 97

A

B

C

ABA-treatment

Relative expression

PEG-treatment

Relative expression

Table 1 The identified *M.bZIP* encoding genes and the relative information.

Group	Gene name	Gene ID	Chromosome location	Number of introns	Protein length (aa)	Isoelectric point	Additional domain
I	<i>M.bZIP8</i>	MDP0000300532	2	6	286	7.03	DOG1
	<i>M.bZIP9</i>	MDP0000174930	2	8	306	8.61	-
	<i>M.bZIP13</i>	MDP0000270677	2	1	755	6.13	-
	<i>M.bZIP14</i>	MDP0000180785	2	1	755	6.13	-
	<i>M.bZIP25</i>	MDP0000834642	4	3	213	9.09	-
	<i>M.bZIP31</i>	MDP0000299504	7	1	741	6.16	-
	<i>M.bZIP32</i>	MDP0000138811	7	1	741	6.16	-
	<i>M.bZIP33</i>	MDP0000378041	7	1	100	9.29	-
	<i>M.bZIP35</i>	MDP0000320524	7	11	526	9.60	DOG1
	<i>M.bZIP39</i>	MDP0000274723	7	7	416	6.15	DOG1
	<i>M.bZIP69</i>	MDP0000262210	12	11	534	9.67	DOG1
	<i>M.bZIP70</i>	MDP0000586302	12	3	233	9.35	-
	<i>M.bZIP71</i>	MDP0000264514	12	4	484	9.49	-
	<i>M.bZIP78</i>	MDP0000145555	13	9	466	5.85	DOG1
	<i>M.bZIP79</i>	MDP0000301884	13	9	466	5.85	DOG1
	<i>M.bZIP81</i>	MDP0000279891	13	4	366	5.13	-
	<i>M.bZIP82</i>	MDP0000219041	13	3	242	9.02	-
	<i>M.bZIP87</i>	MDP0000320322	14	10	542	6.81	DOG1
	<i>M.bZIP88</i>	MDP0000299688	14	22	854	7.94	DUF
	<i>M.bZIP89</i>	MDP0000367943	14	10	494	6.64	DOG1
	<i>M.bZIP98</i>	MDP0000121258	15	11	524	5.70	DOG1
	<i>M.bZIP99</i>	MDP0000277999	5	14	719	8.69	DOG1
	<i>M.bZIP103</i>	MDP0000250967	16	18	1018	8.51	DOG1
	<i>M.bZIP113</i>	MDP0000386314	Unknown	8	433	5.20	-
	<i>M.bZIP115</i>	MDP0000536881	Unknown	10	450	7.16	DOG1
II	<i>M.bZIP2</i>	MDP0000136654	1	3	152	7.95	-
	<i>M.bZIP15</i>	MDP0000293847	3	4	454	5.92	-
	<i>M.bZIP16</i>	MDP0000120158	3	5	472	6.66	-
	<i>M.bZIP17</i>	MDP0000295681	3	2	320	5.62	-
	<i>M.bZIP18</i>	MDP0000222114	3	5	280	9.76	-
	<i>M.bZIP19</i>	MDP0000301399	3	4	340	6.87	-
	<i>M.bZIP20</i>	MDP0000488746	3	4	340	7.75	-
	<i>M.bZIP21</i>	MDP0000772665	3	3	282	5.46	-
	<i>M.bZIP22</i>	MDP0000134936	3	3	353	6.50	-
	<i>M.bZIP24</i>	MDP0000247372	4	4	421	5.86	-
	<i>M.bZIP40</i>	MDP0000282828	8	8	1166	6.43	-
	<i>M.bZIP41</i>	MDP0000300820	8	3	508	5.94	-
	<i>M.bZIP49</i>	MDP0000250947	8	4	528	4.44	-
	<i>M.bZIP53</i>	MDP0000178326	9	3	328	6.46	-
	<i>M.bZIP54</i>	MDP0000306302	9	2	272	6.37	-
	<i>M.bZIP55</i>	MDP0000141948	10	15	917	9.13	-
	<i>M.bZIP60</i>	MDP0000234166	11	3	345	5.83	-
	<i>M.bZIP61</i>	MDP0000297791	11	3	448	5.97	-
	<i>M.bZIP62</i>	MDP0000305387	11	3	448	5.97	-
	<i>M.bZIP64</i>	MDP0000133698	11	3	356	6.22	-
	<i>M.bZIP72</i>	MDP0000636541	12	3	484	9.49	-
	<i>M.bZIP73</i>	MDP0000280559	12	3	580	6.86	PKD
	<i>M.bZIP74</i>	MDP0000479652	12	3	527	6.83	-
	<i>M.bZIP76</i>	MDP0000210251	12	3	361	6.08	-
	<i>M.bZIP85</i>	MDP0000121603	14	6	514	7.20	-
	<i>M.bZIP86</i>	MDP0000147745	14	6	588	8.54	-
	<i>M.bZIP95</i>	MDP0000435971	15	3	449	5.85	-
	<i>M.bZIP101</i>	MDP0000602946	15	3	421	9.13	-
	<i>M.bZIP107</i>	MDP0000555457	Unknown	3	327	6.81	-
	<i>M.bZIP109</i>	MDP0000203904	Unknown	6	543	8.66	-
	<i>M.bZIP110</i>	MDP0000374836	Unknown	3	446	7.56	-
	<i>M.bZIP111</i>	MDP0000267964	Unknown	3	325	10.14	-
	<i>M.bZIP112</i>	MDP0000129203	Unknown	3	361	6.08	-
	<i>M.bZIP114</i>	MDP0000123107	Unknown	1	229	6.06	-
	<i>M.bZIP116</i>	MDP0000120802	Unknown	3	439	5.98	-
	<i>M.bZIP7</i>	MDP0000273211	2	4	535	9.44	-
	<i>M.bZIP26</i>	MDP0000231542	4	10	1017	8.83	-

III	<i>M.bZIP28</i>	MDP0000248567	5	3	431	9.62	-
	<i>M.bZIP36</i>	MDP0000215106	7	6	408	8.80	-
	<i>M.bZIP48</i>	MDP0000701734	8	4	643	8.94	-
	<i>M.bZIP50</i>	MDP0000706379	8	3	299	6.27	-
	<i>M.bZIP56</i>	MDP0000208334	10	3	314	6.07	-
	<i>M.bZIP59</i>	MDP0000740787	10	4	517	8.39	-
	<i>M.bZIP67</i>	MDP0000177486	12	4	470	9.15	-
	<i>M.bZIP68</i>	MDP0000144105	12	4	470	9.15	-
	<i>M.bZIP93</i>	MDP0000296303	15	3	835	8.79	-
	<i>M.bZIP100</i>	MDP0000129112	15	6	489	6.45	-
IV	<i>M.bZIP102</i>	MDP0000198495	15	5	391	5.84	-
	<i>M.bZIP3</i>	MDP0000470928	2	2	240	9.50	-
	<i>M.bZIP10</i>	MDP0000898701	2	0	245	6.79	-
	<i>M.bZIP11</i>	MDP0000893802	2	0	245	6.79	-
	<i>M.bZIP12</i>	MDP0000190186	2	0	245	6.79	-
	<i>M.bZIP23</i>	MDP0000437680	3	0	142	9.09	-
	<i>M.bZIP27</i>	MDP0000275309	5	14	799	7.70	RPBB, bZIP-C
	<i>M.bZIP30</i>	MDP0000917315	6	0	154	6.98	RPBB, bZIP-C
	<i>M.bZIP34</i>	MDP0000772633	7	0	245	6.74	-
	<i>M.bZIP44</i>	MDP0000891108	8	0	192	6.34	-
	<i>M.bZIP45</i>	MDP0000319187	8	1	361	7.92	-
	<i>M.bZIP52</i>	MDP0000407755	9	0	202	5.63	-
	<i>M.bZIP57</i>	MDP0000139670	2	2	280	5.82	-
	<i>M.bZIP58</i>	MDP0000176747	10	14	869	8.08	-
	<i>M.bZIP65</i>	MDP0000234798	11	0	141	6.84	-
	<i>M.bZIP66</i>	MDP0000949327	11	0	141	6.84	-
	<i>M.bZIP77</i>	MDP0000270365	12	5	437	5.80	-
V	<i>M.bZIP80</i>	MDP0000891899	13	0	204	5.93	-
	<i>M.bZIP83</i>	MDP0000431572	13	5	346	4.92	-
	<i>M.bZIP84</i>	MDP0000441891	13	5	346	4.92	-
	<i>M.bZIP90</i>	MDP0000905135	14	0	173	7.07	-
	<i>M.bZIP94</i>	MDP0000680042	15	4	327	5.76	-
	<i>M.bZIP96</i>	MDP0000169473	15	2	286	9.03	-
	<i>M.bZIP104</i>	MDP0000183562	16	3	360	6.60	-
	<i>M.bZIP106</i>	MDP0000863909	17	0	202	5.63	-
	<i>M.bZIP1</i>	MDP0000197219	1	11	457	8.47	MFMR
	<i>M.bZIP6</i>	MDP0000251332	2	11	387	5.78	MFMR
VI	<i>M.bZIP29</i>	MDP0000185553	6	15	661	9.10	MFMR
	<i>M.bZIP37</i>	MDP0000138052	7	11	443	6.12	MFMR
	<i>M.bZIP38</i>	MDP0000493795	7	4	278	6.67	-
	<i>M.bZIP91</i>	MDP0000231274	14	10	351	5.66	MFMR
	<i>M.bZIP105</i>	MDP0000286846	17	11	448	6.85	MFMR
	<i>M.bZIP108</i>	MDP0000545420	Unknown	4	278	6.67	-
	<i>M.bZIP4</i>	MDP0000249561	2	0	157	5.19	-
	<i>M.bZIP5</i>	MDP0000265875	2	0	157	5.73	-
	<i>M.bZIP42</i>	MDP0000521934	8	0	156	6.58	-
	<i>M.bZIP43</i>	MDP0000205823	8	0	159	5.93	-
VII	<i>M.bZIP46</i>	MDP0000140166	8	0	184	9.12	-
	<i>M.bZIP47</i>	MDP0000200822	8	0	152	7.95	-
	<i>M.bZIP51</i>	MDP0000738631	8	0	228	10.62	-
	<i>M.bZIP63</i>	MDP0000239026	11	0	160	6.51	-
	<i>M.bZIP75</i>	MDP0000448715	12	0	157	7.03	-
	<i>M.bZIP92</i>	MDP0000190277	15	0	152	8.52	-
	<i>M.bZIP97</i>	MDP0000261154	15	8	873	9.03	-

Table 2 Distribution of ABRE, DRE, LRTE, MYB and MYC *cis*-acting elements in *M.bZIP* promoters.

Group	Gene	ABRE	DRE	LRTE	MYB	MYC	Total
II	<i>M.bZIP16</i>	1	0	1	24	21	47
II	<i>M.bZIP24</i>	7	0	2	22	36	67
V	<i>M.bZIP29</i>	4	0	1	14	18	37
IV	<i>M.bZIP45</i>	1	0	0	8	20	29
VI	<i>M.bZIP46</i>	3	0	1	6	6	16
II	<i>M.bZIP49</i>	6	0	0	14	14	34
IV	<i>M.bZIP58</i>	2	0	1	20	20	43
III	<i>M.bZIP59</i>	3	0	0	23	2	28
I	<i>M.bZIP69</i>	4	2	2	19	12	39
II	<i>M.bZIP73</i>	2	2	6	20	12	42
I	<i>M.bZIP87</i>	1	4	2	13	18	38
I	<i>M.bZIP98</i>	1	1	2	18	16	38
II	<i>M.bZIP101</i>	4	2	4	24	12	46
V	<i>M.bZIP105</i>	3	0	0	30	10	43