

RESEARCH ARTICLE

Open Access

Genome-wide analysis and expression profile of the bZIP transcription factor gene family in grapevine (*Vitis vinifera*)

Jinyi Liu¹, Nana Chen¹, Fei Chen¹, Bin Cai¹, Silvia Dal Santo², Giovanni Battista Tornielli², Mario Pezzotti² and Zong-Ming (Max) Cheng^{1,3*}

Abstract

Background: Basic leucine zipper (bZIP) transcription factor gene family is one of the largest and most diverse families in plants. Current studies have shown that the bZIP proteins regulate numerous growth and developmental processes and biotic and abiotic stress responses. Nonetheless, knowledge concerning the specific expression patterns and evolutionary history of plant bZIP family members remains very limited.

Results: We identified 55 bZIP transcription factor-encoding genes in the grapevine (*Vitis vinifera*) genome, and divided them into 10 groups according to the phylogenetic relationship with those in *Arabidopsis*. The chromosome distribution and the collinearity analyses suggest that expansion of the grapevine bZIP (*VvbZIP*) transcription factor family was greatly contributed by the segment/chromosomal duplications, which may be associated with the grapevine genome fusion events. Nine intron/exon structural patterns within the bZIP domain and the additional conserved motifs were identified among all *VvbZIP* proteins, and showed a high group-specificity. The predicted specificities on DNA-binding domains indicated that some highly conserved amino acid residues exist across each major group in the tree of land plant life. The expression patterns of *VvbZIP* genes across the grapevine gene expression atlas, based on microarray technology, suggest that *VvbZIP* genes are involved in grapevine organ development, especially seed development. Expression analysis based on qRT-PCR indicated that *VvbZIP* genes are extensively involved in drought- and heat-responses, with possibly different mechanisms.

Conclusions: The genome-wide identification, chromosome organization, gene structures, evolutionary and expression analyses of grapevine bZIP genes provide an overall insight of this gene family and their potential involvement in growth, development and stress responses. This will facilitate further research on the bZIP gene family regarding their evolutionary history and biological functions.

Keywords: bZIP transcription factor family, Grapevine, Gene expression, Drought response, Heat stress response

Background

Transcription factors (TFs) play important roles in regulation and control many crucial biological processes in plants. Functional characterization of TFs is essential for understanding transcriptional regulatory networks and the biological processes they are involved in. Presently, at least 64 families of transcription factors have been

identified in the plant kingdom [1]. The basic leucine zipper (bZIP) transcription factor family is one of the largest and most diverse families [1,2]. The bZIP transcription factors harbor a highly conserved bZIP domain composed of two structural features: a basic region (N-x7-R/K-x9) for sequence-specific DNA binding, and a leucine zipper, which is composed of several heptad repeats of Leu or other bulky hydrophobic amino acids, such as Ile, Val, Phe or Met, for dimerization specificity [2-4]. Outside the bZIP domain, the bZIP transcription factors also contain other conserved domains that function as transcriptional activators. For example, two

* Correspondence: zmc@njau.edu.cn

¹College of Horticulture, Nanjing Agricultural University, 210095 Nanjing, Jiangsu, China

³Department of Plant Sciences, University of Tennessee, 37996 Knoxville, TN, USA

Full list of author information is available at the end of the article

conserved motives, described as R/KxxS/T and S/TxxD/E, have been confirmed as sites phosphorylated by the Ca^{2+} independent protein kinase and casein kinase II, respectively [5]. Proline-rich, glutamine-rich and acidic domains are also thought to play roles in transcription activation of *bZIP* genes [6].

The *bZIP* gene family has been comprehensively identified or predicted in several plant genomes. Seventy-five *bZIP* genes have been found in *Arabidopsis thaliana* [4], 89 in rice (*Oryza sativa*) [2], 131 in soybean (*Glycine max*) [6], 92 in Sorghum [7], and 125 in maize (*Zea mays*) [8]. Among all members in all plant species, just a small number of *bZIP* genes have been functionally characterized. Plant *bZIP* proteins have been found to regulate the integration and development of many organs and tissues, including seed maturation and germination [9,10], floral induction and development [11-18], vascular development [19], embryogenesis [20,21] and photomorphogenesis [22,23]. They have also been found to be involved in responses to a variety of abiotic/biotic stimuli, such as drought [24,25], high salinity [26], heat stress [27], cold stress [28,29], pathogen infection [30], and hormone signaling, such as abscisic acid (ABA) [25,31,32], ethylene [33] and light signaling [34].

Grapevine (*Vitis* spp.) is a popular fruit crop that in cultivated throughout the world and represents one of the most important crops with highly valued products such as juice, liquors and wines. The grapevine (*Vitis vinifera*) genome sequence was released in 2007 [35]. It was suggested that the grapevine genome was composed of three ancestral genomes, and it has retained more ancient imprint of the evolution. However, the notion of grapevine as a “palaeo-hexaploid” organism [35,36] remains to be confirmed. Since *Vitis* is sister to all Rosids (I and II), studying the evolution of a large gene family in grapevine will shed light on our understanding the evolution of other plants [36], and will allow the elucidation of function and divergence of gene families in grapevine. So far, only a limited number of *VvbZIPs* have been isolated and functionally characterized [37]. Among them, two *VvABFs* (homologous of *VvbZIP08* and *VvbZIP45* in our study), isolated from the *V. vinifera* cultivar ‘Thompson Seedless’, were induced by various abiotic stresses including drought, salinity, cold and the exogenous hormone ABA [38], and *VvbZIP45* corresponds to the very recently described *ABSCISIC ACID RESPONSE ELEMENT-BINDING FACTOR2* (*VvABF2*), involved in ABA-dependent grape berry ripening [39]. Another gene, *VvbZIP23*, which was isolated from the cultivar ‘Mango’, was induced by a wide spectrum of abiotic stresses, including drought, salt, cold, and application of abscisic acid, methyl viologen, salicylic acid, jasmonic acid, and ethephon [37]. In addition, the transcriptional analysis of late ripening stages of

grapevine berry detected one *bZIP* transcription factor which exhibited lower expression at the harvest stages (TH) than that at the 7-days before harvest stage (TH-7) [40]. The transcriptomic analysis of grape leaves suggested that some *VvbZIP* transcription factors were involved in stress responses [41].

In the present study, we report the identification of 55 *bZIP* genes from the grapevine genome, and the analysis of the gene family with a focus on the evolution and divergence after multiple duplications in relation to the grapevine genome fusion. We also analyze gene expression of *VvbZIP* genes in different tissues/organs/developmental stages, as well as in response to stresses. Our results provide a prospective for the evolutionary history and general biological involvement of the grapevine *bZIP* transcription factor family.

Results

Identification and nomenclature assignment of *bZIP* transcription factors in grapevine

In order to run a complete search for *VvbZIP* genes in the grapevine genome, all annotated proteins of the latest version (V1) of the grapevine genome from CRIBI (<http://genomes.cribi.unipd.it/>) were considered. The Hidden Markov Model (HMM) profile of the *bZIP* domain (PF00170) (<http://pfam.sanger.ac.uk/>) was then employed as a query to search the database using the program HMMER3.0 with the default E-value. After determining the integrity of the *bZIP* domain using the online program SMART (<http://smart.embl-heidelberg.de/>) with an E-value < 0.1 and sequence alignment, 55 non-redundant genes were assigned as grapevine *bZIP* genes, and named from *VvbZIP01* to *VvbZIP55* based on the coordinate order on grapevine chromosomes from top to bottom. The nomenclature and corresponding information are listed in Table 1.

Phylogenetic analysis and classification of the grapevine *bZIP* gene family

To investigate the phylogenetic relationship of the *bZIP* gene families in grapevine and *Arabidopsis*, the amino acid sequences of the *bZIP* domain of the 55 proteins from grapevine (*Vitis vinifera*), 72 from *Arabidopsis thaliana* (three genes were no longer supported by their annotations comparing to Jakoby et al. [4]) and 99 from poplar (*Populus trichocarpa*) (data from Phytozome v9.1) were used to construct a phylogenetic tree. Three methods, Neighbor-Joining (NJ), Minimal Evolution (ME) and Maximum Parsimony (MP), yielded nearly identical phylogenetic trees (data not shown), therefore, only the NJ tree was used for further analyses (Figure 1). Based on the phylogenetic tree, the *VvbZIPs* can generally be divided into the same 10 subgroups as those in *Arabidopsis*. However, *AtbZIP62/VvbZIP26/Potri.006G114600.1*,

Table 1 The identified VvbZIP genes and their related information

Gene name	Gene ID	Group	Chr.	Locus		Chain	Protein Length	Amount of intron	Intron pattern within bZIP domain	AtbZIP Ortholog (Syntenic orthologs were denoted by overstriking and italics)
<i>WvbZIP01</i>	VIT_01s0011g03230	J	1	2941266	2946275	+	453	10	<i>b</i>	46 , 20, 45, 26, 65, 21, 57, 47, 22, 50
<i>WvbZIP02</i>	VIT_01s0010g00930	C	1	16444033	16444682	-	195	0	<i>h</i>	42, 58 , 48, 43, 3, 8 , 70, 2, 44, 53, 7, 11, 41, 5, 68, 6, 16, 4
<i>WvbZIP03</i>	VIT_02s0025g01020	D	2	989255	996186	+	398	10	<i>a</i>	55(GBF3) , 54(GBF2) , 16, 68, 41
<i>WvbZIP04</i>	VIT_02s0012g02250	D	2	9137736	9153313	-	413	11	<i>a</i>	55, 54, 16, 68 , 41, 44, 42, 43, 58, 3, 70, 8, 48
<i>WvbZIP05</i>	VIT_03s0038g00860	F	3	689247	693308	+	676	5	<i>c</i>	29,30 , 18, 52, 69, 51, 59, 31, 33, 61, 34, 74
<i>WvbZIP06</i>	VIT_03s0038g02420	I	3	1679507	1683223	+	271	1	<i>h</i>	23, 19 , 24
<i>WvbZIP07</i>	VIT_03s0038g04450	C	3	3227416	3228660	-	157	0	<i>h</i>	44, 11, 2, 53, 48, 3, 42, 43, 58, 8, 1, 6, 7, 5, 68, 70
<i>WvbZIP08</i>	VIT_03s0063g00310	A	3	3887073	3890514	-	409	4	<i>a</i>	37(ABF3) , 36, 38(ABF4/AREB2) , 35(ABF1) , 15, 66, 12, 67, 39, 14, 13, 40
<i>WvbZIP09</i>	VIT_04s0008g02750	B	4	2284121	2287115	-	349	5	<i>d</i>	9 , 63, 10, 2, 53, 42, 25, 44, 58, 3, 48, 68, 11
<i>WvbZIP10</i>	VIT_04s0008g05210	G	4	4716191	4719252	+	169	3	<i>f</i>	56(HYS) , 64
<i>WvbZIP11</i>	VIT_04s0069g01150	A	4	9841491	9842587	+	57	1	<i>a</i>	66, 12, 37, 36, 35, 38, 39
<i>WvbZIP12</i>	VIT_04s0023g01360	D	4	17832626	17863952	+	364	10	<i>a</i>	41 , 16, 68, 55, 54, 42, 70, 3, 58, 43, 48, 8, 44, 18
<i>WvbZIP13</i>	VIT_04s0023g02430	C	4	18983616	18984767	+	249	0	<i>h</i>	44, 11, 2, 53, 48, 42, 58, 3, 43, 8, 7, 6, 1, 5
<i>WvbZIP14</i>	VIT_05s0077g01140	C	5	878161	879142	-	154	0	<i>h</i>	53, 2, 44, 11, 58, 3, 5, 48, 7, 1
<i>WvbZIP15</i>	VIT_05s0020g01090	G	5	2860048	2863747	+	210	3	<i>h</i>	64(HYH) , 56
<i>WvbZIP16</i>	VIT_05s0102g01120	F	5	23226450	23231749	-	425	3	<i>c</i>	69, 59, 18, 52, 51, 29, 30, 61, 33, 31, 34, 74
<i>WvbZIP17</i>	VIT_06s0004g08070	F	6	8835674	8841051	-	342	3	<i>c</i>	18,52 , 59, 69, 51, 29, 30, 33, 61, 31, 74, 34
<i>WvbZIP18</i>	VIT_06s0009g01790	A	6	13809459	13815625	-	248	2	<i>a</i>	66, 12, 35, 38, 36, 67, 39, 40, 14, 13, 37, 15
<i>WvbZIP19</i>	VIT_06s0080g00340	A	6	20262007	20263157	+	324	2	<i>a</i>	67(DPBF2) , 39, 66, 12, 35, 36, 38, 15, 37, 40
<i>WvbZIP20</i>	VIT_06s0080g00360	J	6	20304714	20311349	-	500	10	<i>g</i>	21 , 65, 45, 20, 26, 46, 22, 50, 47, 57
<i>WvbZIP21</i>	VIT_07s0141g00170	B	7	116657	126811	-	452	5	<i>d</i>	25, 10, 63, 9
<i>WvbZIP22</i>	VIT_07s0005g01450	C	7	3987031	3988120	+	145	0	<i>h</i>	53 , 44, 2, 11, 42, 3, 6, 48, 7, 58, 43, 8, 16, 1, 68
<i>WvbZIP23</i>	VIT_07s0031g01320	J	7	17404970	17410595	+	349	7	<i>b</i>	47, 57 , 50, 22, 20, 45, 26, 21, 46, 65
<i>WvbZIP24</i>	VIT_08s0040g00870	H	8	11857207	11860728	-	819	1	<i>h</i>	17, 28, 49
<i>WvbZIP25</i>	VIT_08s0007g03420	A	8	17372100	17378057	-	400	0	<i>a</i>	39(AB15) , 35, 67, 66, 38, 37, 36, 12, 13, 15, 40, 14
<i>WvbZIP26</i>	VIT_08s0007g03640	UC	8	17579864	17586189	+	496	5	<i>a</i>	62
<i>WvbZIP27</i>	VIT_08s0007g05170	J	8	19088004	19098613	-	451	10	<i>b</i>	45, 20, 26, 46, 21, 65, 57, 50, 22, 47
<i>WvbZIP28</i>	VIT_08s0007g06160	J	8	19963578	19968190	-	491	10	<i>b</i>	65 , 21, 45, 20, 26, 46, 22, 57, 50, 47
<i>WvbZIP29</i>	VIT_12s0028g02590	E	12	3366293	3367561	-	295	3	<i>e</i>	61, 34, 51, 69, 18, 52, 59, 30, 29, 72, 33, 31
<i>WvbZIP30</i>	VIT_12s0055g00420	A	12	13276799	13300251	+	299	3	<i>a</i>	13, 40, 12, 66, 67, 35, 37, 38, 36, 39, 14
<i>WvbZIP31</i>	VIT_12s0034g00110	A	12	15485471	15491224	+	99	3	<i>a</i>	66, 12, 37, 36, 35, 38, 39
<i>WvbZIP32</i>	VIT_12s0035g00620	E	12	20038502	20045379	+	374	3	<i>e</i>	61, 34, 69, 59, 18, 51, 29, 52, 30

Table 1 The identified VvbZIP genes and their related information (Continued)

<i>WvbZIP33</i>	VIT_13s0067g02900	F	13	1562503	1569804	+	359	3	<i>c</i>	18 , 52, 59, 69, 51, 29, 30, 33, 61, 31, 34, 74, 72
<i>WvbZIP34</i>	VIT_13s0175g00120	A	13	15049338	15072969	+	325	2	<i>a</i>	66, 12, 35, 67, 38, 37, 36, 39, 15, 14
<i>WvbZIP35</i>	VIT_13s0084g00660	J	13	19780668	19805654	-	469	10	<i>b</i>	45, 26, 20, 46, 21, 65, 57, 22, 50, 47
<i>WvbZIP36</i>	VIT_13s0158g00380	I	13	21283084	21284766	-	286	1	<i>h</i>	23, 19, 24
<i>WvbZIP37</i>	VIT_14s0060g01210	C	14	959578	961198	-	145	0	<i>h</i>	53, 11, 2, 44, 1, 48, 58, 3, 8, 7, 42, 43, 70
<i>WvbZIP38</i>	VIT_14s0030g02200	B	14	7402807	7408597	+	423	5	<i>d</i>	63, 10, 25, 9
<i>WvbZIP39</i>	VIT_14s0083g00700	C	14	22839619	22840360	-	198	0	<i>h</i>	42,43 , 58, 3,48 , 8, 70, 44, 2, 53, 6, 11, 41, 7, 5, 68
<i>WvbZIP40</i>	VIT_15s0046g01440	D	15	18416554	18423462	-	430	11	<i>a</i>	55(GBF3) , 54(GBF2) , 16, 68, 41, 3, 42, 43, 70, 58
<i>WvbZIP41</i>	VIT_18s0122g00500	UC	18	420616	423564	+	322	2	<i>f</i>	60
<i>WvbZIP42</i>	VIT_18s0001g04470	J	18	3861987	3867762	-	363	7	<i>b</i>	57, 47, 50, 22 , 26, 20, 65, 45, 21, 46
<i>WvbZIP43</i>	VIT_18s0001g04500	A	18	3875315	3878509	+	311	2	<i>a</i>	12, 66, 35, 36, 39, 38, 37, 14, 15, 67, 40, 13
<i>WvbZIP44</i>	VIT_18s0001g08710	C	18	7185268	7185664	-	81	0	<i>h</i>	7, 6, 4, 5, 2, 3
<i>WvbZIP45</i>	VIT_18s0001g10450	A	18	8761978	8769742	-	447	3	<i>a</i>	36(ABF2/AREB1) , 35, 37(ABF3) , 38, 66, 12, 39, 15 , 67
<i>WvbZIP46</i>	VIT_18s0001g12120	D	18	10309606	10313842	+	396	10	<i>a</i>	41, 16, 68, 55, 54, 43, 58, 42, 3, 48, 70, 8
<i>WvbZIP47</i>	VIT_18s0001g13040	C	18	11138196	11139020	+	154	0	<i>h</i>	44, 11, 2, 53, 48, 3, 42, 58, 43, 1, 7, 6, 8, 5
<i>WvbZIP48</i>	VIT_18s0001g13740	F	18	11745490	11746413	-	187	3	<i>c</i>	30, 29, 51, 69, 52, 18, 59, 33, 31, 61, 34
<i>WvbZIP49</i>	VIT_18s0001g14890	A	18	12936974	12937903	-	214	2	<i>a</i>	14(FD) , 66, 12, 37, 36, 39, 38, 35, 27
<i>WvbZIP50</i>	VIT_18s0076g00330	F	18	16209694	16218061	+	350	3	<i>c</i>	51, 18, 69, 59, 52, 29, 30, 31, 61, 33, 74, 34
<i>WvbZIP51</i>	VIT_18s0072g00470	A	18	19680318	19687522	+	275	3	<i>a</i>	13, 40, 66, 12, 39, 36
<i>WvbZIP52</i>	VIT_19s0014g01780	E	19	1945207	1947498	-	309	3	<i>e</i>	61, 34, 51, 18, 52, 69, 59, 30, 29, 72, 33, 31
<i>WvbZIP53</i>	VIT_19s0015g01020	A	19	9101591	9112651	-	251	2	<i>g</i>	36, 12
<i>WvbZIP54</i>	VIT_18s0001g03010	E	18_random	3068050	3070292	-	272	3	<i>e</i>	61, 34, 69, 51, 59
<i>WvbZIP55</i>	VIT_00s0541g00020	C	Un	31593865	31594595	-	194	0	<i>h</i>	6, 5, 7, 44, 42, 58, 2, 3, 11, 48, 53, 43, 70, 8, 1, 4

Note: AtbZIP Orthologs were identified by local BLAST search with E-value < 1E-10 and the identification of syntenic orthologs (highlight with italics) was performed by using the software MScanX.

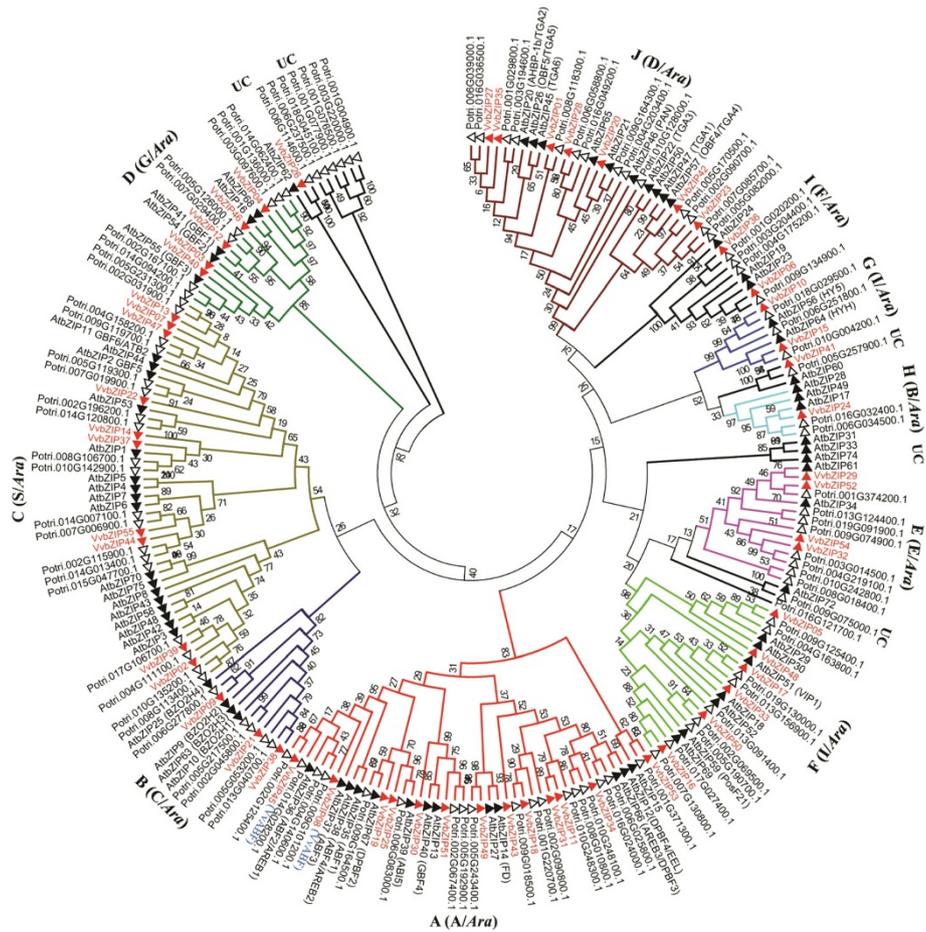


Figure 1 The phylogenetic tree of grapevine and Arabidopsis bZIP genes. The 55 grapevine, 72 Arabidopsis and 99 poplar bZIP domain protein sequences were aligned by Clustal X 1.83 and the phylogenetic tree was constructed using MGA5.0 by the Neighbor-Joining (NJ) method. The Bootstrap value was 1,000 replicates. The gene names of the characterized bZIP genes are shown in parenthesis after the gene numbers, and which were referred to previous research [2,4]. A(A/Ara), B(C/Ara) and so on is presented as the subfamily names of bZIP gene family in grapevine (capital letter outside the parenthesis) and Arabidopsis (capital letter in parenthesis) and Ara represent for Arabidopsis), UC, unclassified genes or clades. The colored branch indicates the different subfamily and the black represent for UC.

AtbZIP60/VvbZIP41/Potri.005G257900.1, and *AtbZIP72/Potri.009G075000.1* (UC, Figure 1) form three small unique clades in the phylogenetic tree and may have had independent evolutionary trajectories from others clades. In addition, *AtbZIP74*, *AtbZIP33* and *AtbZIP31* and *Potri.006G237500.1/Potri.018G045100.1/Potri.001G077900.1/Potri.001G076500.1/Potri.003G220000.1/Potri.001G004900.1*, were distinguished from other bZIP genes as they formed individual clades without any other bZIP genes, suggesting that these individual clades may be specific to Arabidopsis and poplar, respectively. This result, different from the former classification in Arabidopsis [4], was supported by Nijhawan et al. and Wei et al. [2,8] and sustained by the analyses of the gene structures, conserved motifs and DNA-binding site specificity of VvbZIP transcription factors in the present work.

Chromosomal location and expansion patterns of VvbZIP genes

The Vitaceae evolved about 60 million years ago (MYA) and is the earliest diverging lineage of rosids [36,42]. The extant grapevine genome was considered to be formed by fusion of three ancestral genomes following post-fusion shuffling [36] which has not undergone recent genome duplication following the γ WGD. The “palaeo-hexaploid” nature of the genome is supported by many triplicate blocks in the genome, as indicated by the same colors of homeologous blocks adapted from Jaillon et al. [35], and it was hypothesized that the ancestral hexaploid genome of *V. vinifera* was derived from an allopolyploidization of two ancestral genomes, which were named as Va and Vc, respectively [36]. Chromosome triplets consist of two Va and one Vc chromosomes, as expected from the tetraploid

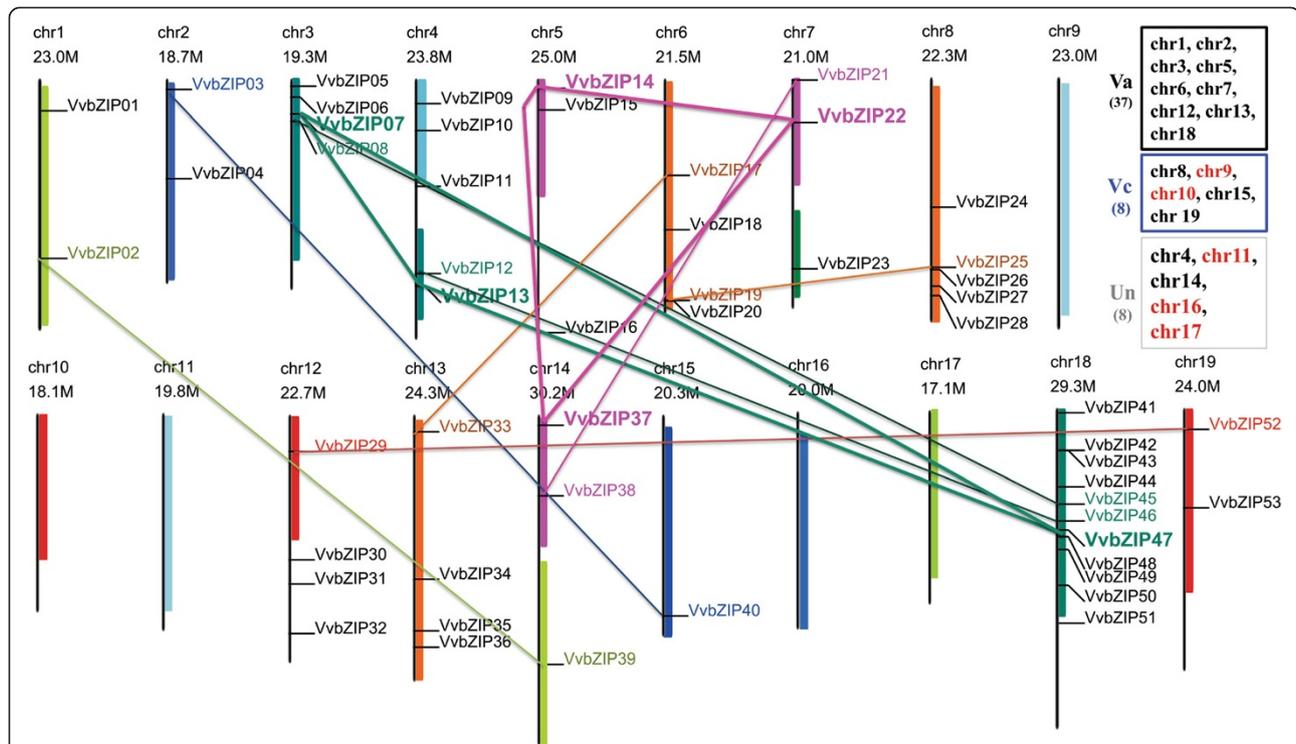


Figure 2 Distributions of *VvbZIP* genes in grapevine Chromosomes. Fifty-three *VvbZIP* genes were mapped to the 14 linkage groups. *VvbZIP54* was located on 18 random chromosomes and *VvbZIP55* was assigned to the Un chromosome. Five chromosomes (chr9, 10, 11, 16 and 17) have no grapevine bZIP gene. The schematic representation of paralogous regions was adapted from Jaillon et al., with the same color [35]. According to the former hypothesis [36], the ancestral hexaploid genome of *V. vinifera* was derived from an allopolyploidization of two ancestral genomes: Va and Vc, which are denoted on the right of the figure, and Un is unassigned chromosomes. The chr names in red underline the chromosomes that harbored no bZIP genes. The gene names with collinearity were performed with the some color of chromosome paralogous region.

and diploid conditions of the two component genomes [36]. To search for imprints of the grapevine genome evolution from the perspective of gene families, the location of each *VvbZIP* gene was given a diagrammatic representation based on the annotation of the 12x V1 assembly grapevine genome available at CRIBI (Figure 2). The results showed that the distribution of *VvbZIP* genes is not even. *VvbZIP54* and *VvbZIP55* are located on the 18 Random and Un chromosome, respectively. Among the other 53 *VvbZIP* genes, 11 (20.75%) genes were situated on chromosome 18; 5 (9.43%) *VvbZIP* genes were each located in the chromosome 4 and 8; 4 *VvbZIP* genes positioned each in chromosomes 3, 6, 12 and 13; 3 *VvbZIP* genes located each in chromosomes 5, 7 and 14; 2 *VvbZIP* genes each in chromosomes 1, 2 and 19; only one gene in chromosome 15; and no *VvbZIP* genes are located in chromosomes 9, 10, 11, 16 and 17.

Furthermore, we analyzed the tandem arrangement and the collinearity by using MCScanX program and mapped the genes to chromosomal locations (Figure 2), and found 14 homologous blocks containing bZIP genes with collinearity that may have resulted from genome fusion or WGD and none of the *VvbZIP* genes was found

to be arranged in tandem. Interestingly, all of those collinearity events occurred within same homologous blocks, and only two triplicate bZIP genes, *VvbZIP7/13/47* and *VvbZIP14/22/37*, were identified, indicating that the expansion patterns of bZIP gene family partially support the triplicate state of the grapevine genome and combined with further genome fusion or WGD events. Further, the rest of the collinearity events were just found within two of the homologous triplets, and the third homologous triplet, such as chromosome 9, 10, 11, 16 and 17, mostly harbored no *VvbZIP* genes, which seem to belong to Vc genome of grapevine described by Malacarne et al. [36], supposing that *VvbZIP* genes in the heterogenous Vc genome may have undergone extensive loss and shuffling after genome fusion.

Gene structure of *VvbZIP* genes

Since the intron/exon organizations and intron types and numbers are typical imprints of the evolution within some gene families [2,43,44], the *VvbZIP* gene structures were examined to obtain further insight into their evolutionary trajectory. The results show that exon/intron structures of *VvbZIPs* were highly conserved within subfamilies.

Genes that clustered together generally possessed a similar gene structure, especially the exon number and length (Additional file 1a). Interestingly, no intron containing genes were found in group C, which accounted for 18.2% (10/55) of the total *VvbZIP* genes. Furthermore, the number of exons in each group (Additional file 1b) was uneven. The numbers of exons in group B, C, E, G and I were much more conserved than that in those in the group A (varying from one to 5). Furthermore, the positions and phases of introns in the bZIP domain (composing of basic region, hinge and leucine zipper) were analyzed (Figure 3). Nine intron/exon structural patterns in the *VvbZIP* gene family were identified based on the intron position, number and splicing phase (details were represented in Additional file 2), similar to those in rice and maize [2,8]. We named these patterns as *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h* and *i*. The patterns *a*, *b* and *i* in *VvbZIP* correspond to the patterns *a*, *b* and *f* in maize and the patterns *b*, *c* and *g* in rice, respectively. The patterns *c*, *d*, *e* and *f* in *VvbZIP* jointly correspond to the pattern *c* in maize or the pattern *a* in rice, because they have the same position of intron in the basic and hinge regions.

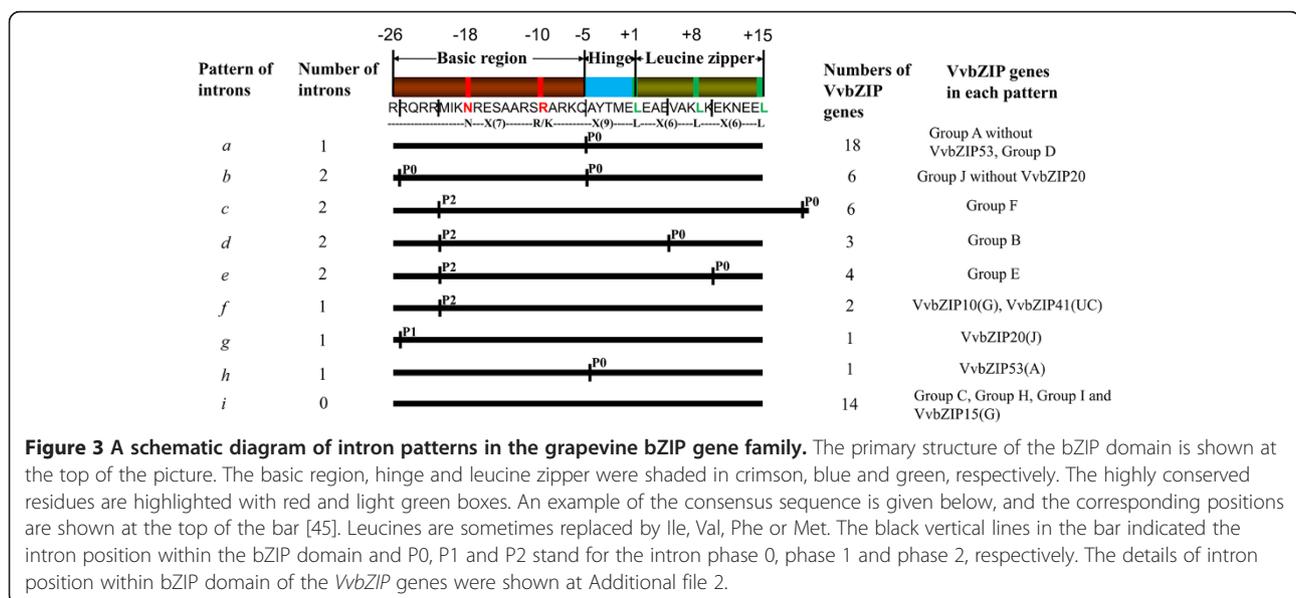
In addition, each subfamily of *VvbZIPs* mostly shared the same intron/exon structural pattern. As illustrated in Figure 3, Pattern *a* has just one intron in phase 0 (P0) in the leucine zipper which inserts between the amino acids Gln and Ala, includes members of group A without *VvbZIP20* and all genes of the group D. Pattern *b* has two introns in phase 0 (P0) in the basic and hinge regions, respectively, and includes the genes of group J without *VvbZIP20*. The pattern *c*, *d*, *e* and *f* share a common intron insert with phase 2 (P2) in the basic region, but for each pattern, differences in intron inserts exist in the leucine zipper region, Pattern *c*, *d* and *e* have

another intron insert in phase 0 (P0) in the leucine zipper region at a different position, whereas the pattern *f* has no intron insert in the leucine zipper region thus different from *c*, *d* and *e*; and the pattern *c*, *d* and *e* are completely corresponded to the group F, B, E, respectively. Pattern *g* has only one intron in phase 1 (P1) in the basic region and pattern *h* has one intron in phase 0 (P0) in the leucine zipper region; both patterns *g* and *h* contained just one member of *VvbZIP* gene each from group J and group A, respectively. Pattern *i* lacks any intron in the basic region and leucine zipper region, and this pattern contains all genes of group C (10 genes), H (one gene), group I (two genes) and *VvbZIP15* in the group G. In summary, the intron position and splicing phase in the basic region and the leucine zipper region of grapevine bZIP gene family have been highly conserved during the course of evolution in each group, which is concordant with the early reports [2,8].

Furthermore, considering the patterns of *bZIP* gene family of rice, maize and grape, we find that although there are some differences of the splicing phase and the interrupted amino acids, the positions of the intron insert are highly conserved. This phenomenon suggested that the position of the intron insert in the conserved domain is important for the evolution and divergence of function in the gene family.

Conserved structural features of *VvbZIP* transcription factors

The bZIP domain is the core of the *bZIP* gene family members, which preferentially binds to the promoter of their downstream target genes on a specific cis-element (e.g. ABREs). However, the functional diversity of the bZIP gene family may also be contributed by additional

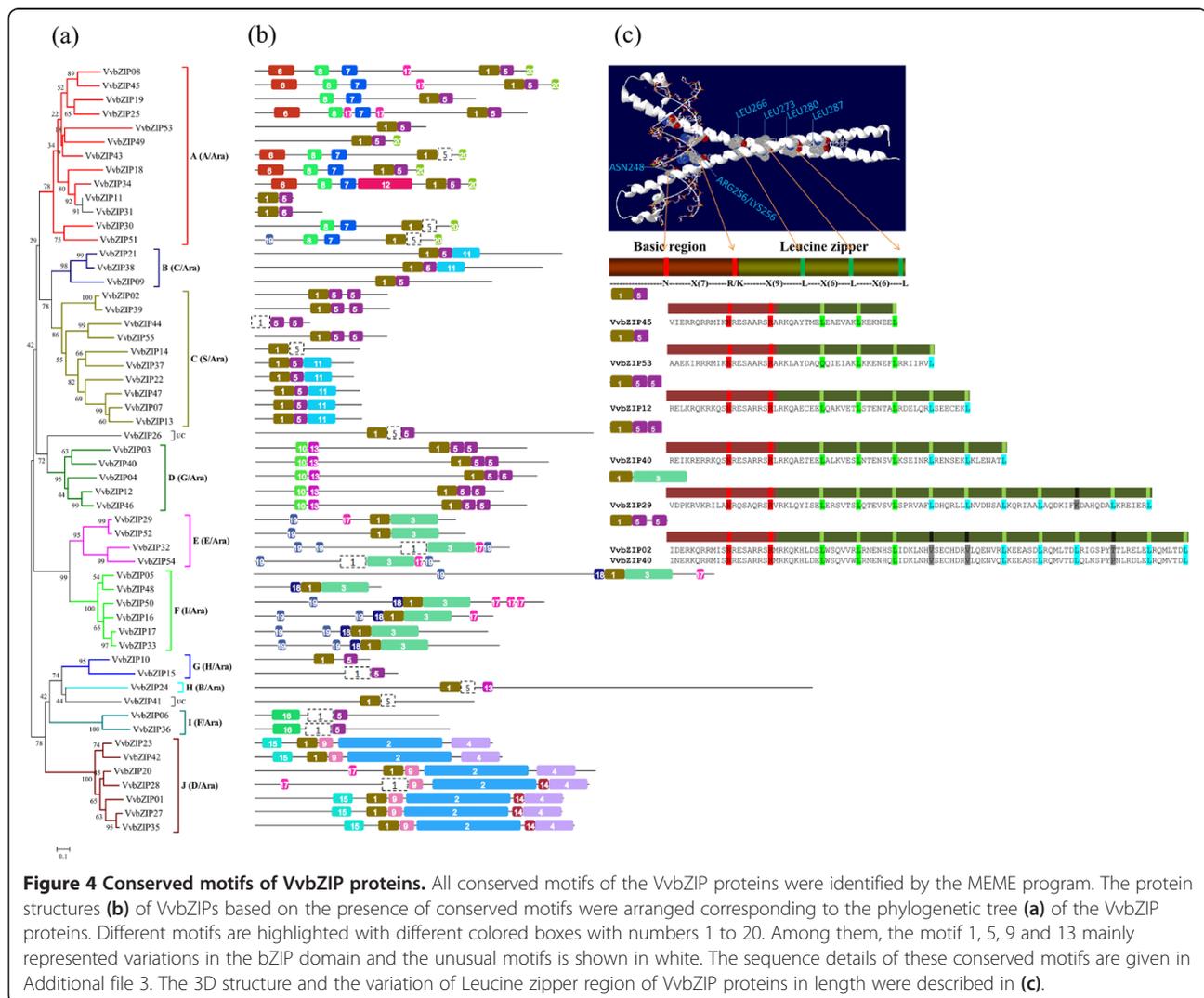


domains in the bZIP proteins [46]. In total, 20 conserved motifs, including the bZIP domain, were identified in the VvbZIP proteins and their multilevel consensus amino acid sequences of motifs are listed in Additional file 3. The motif distribution corresponding to the phylogenetic tree of VvbZIP gene family is shown in Figure 4a,b. Motif 1 is the basic region and the hinge of the bZIP domain; whilst motif 5 (typically), 3 and 9 represent the variable motifs in the leucine zipper region across the bZIP gene family. For example, motif 3 only appeared in the group E and F; motif 9 appeared in the group J exclusively; and group C and D mostly harbored a double motif 5. Further, the number of Leucine units ranged from 3 to 11, and some of them were interrupted by one or two other units which are conserved with the same number of amino acids. Finally, nine gradients (I to IX) were defined (Additional file 4), and some of them showed group specificity, such as the gradient I which harbored the most members of group A, and gradient IX which harbored the

most members of group C. Some typical cases and their corresponding motifs are shown in Figure 4c and may provide a greater understanding of the divergence and the roles of leucine zipper region in the bZIP gene family.

Beside the bZIP domain, each group of VvbZIP shares some motifs. For example, most members of the group A possess motifs 6, 7, 8 and 20; group B and C share motif 11; all members of the group D specifically harbor the motifs 10 and 13; group E shares motif 19 with group F which also possesses motif 18 exclusively. Motif 16 appeared in the group I, motifs 2, 4, 9, 14 and 15 are present in group J exclusively, whereas group G and H lack any additional motif.

Previous evidence suggested that most members of the group A of the bZIP gene family are involved in ABA-responsive signal-transduction pathway and abiotic stress response [4,25,47,48]. Typically, ABF/AREBs are master transcription factors that regulate ABRE-dependent ABA signaling involved in drought stress tolerance, but their



activation requires posttranscriptional modification with ABA-dependent phosphorylation [25]. In the present study, two conserved sites, R/KxxS/T and S/TxxD/E, were verified as the key phosphorylation sites in VvbZIP proteins which are phosphorylated by a Ca²⁺ independent protein kinase and casein kinase II, respectively [5,32]. Furthermore, ABA-dependent multiple phosphorylation sites are also required for their activation. In our study (Figure 4b), motif 6, 7, 8 and 20 appeared in the group A VvbZIP exclusively and they widely represent the potential phosphorylation sites described above. Motif 6 harbors the R/KxxS/T and motif 7 harbors S/TxxD/E, presented as RQ[GNS]S, T[LF]DE and R[QE]PT, TLE[DE], respectively. Motif 8 and 20 contain just one of these sites, presented as TV[DE]E and R[TV][SL][ST], respectively. Moreover, these potential phosphorylation sites were also found in other groups of the VvbZIP gene family, presented as RQQT/ RALS in motif 4, [RK][LV]Q[SM] in motif 11, T[DI][VAM][DEQ] in motif 15, and R[SA]SS in motif 19, respectively. In addition, other conserved motifs in bZIPs that may function as transcriptional activators, have also been found, including proline-rich, glutamine-rich, and acidic domains.

Predicted DNA-binding site specificity of VvbZIP transcription factors

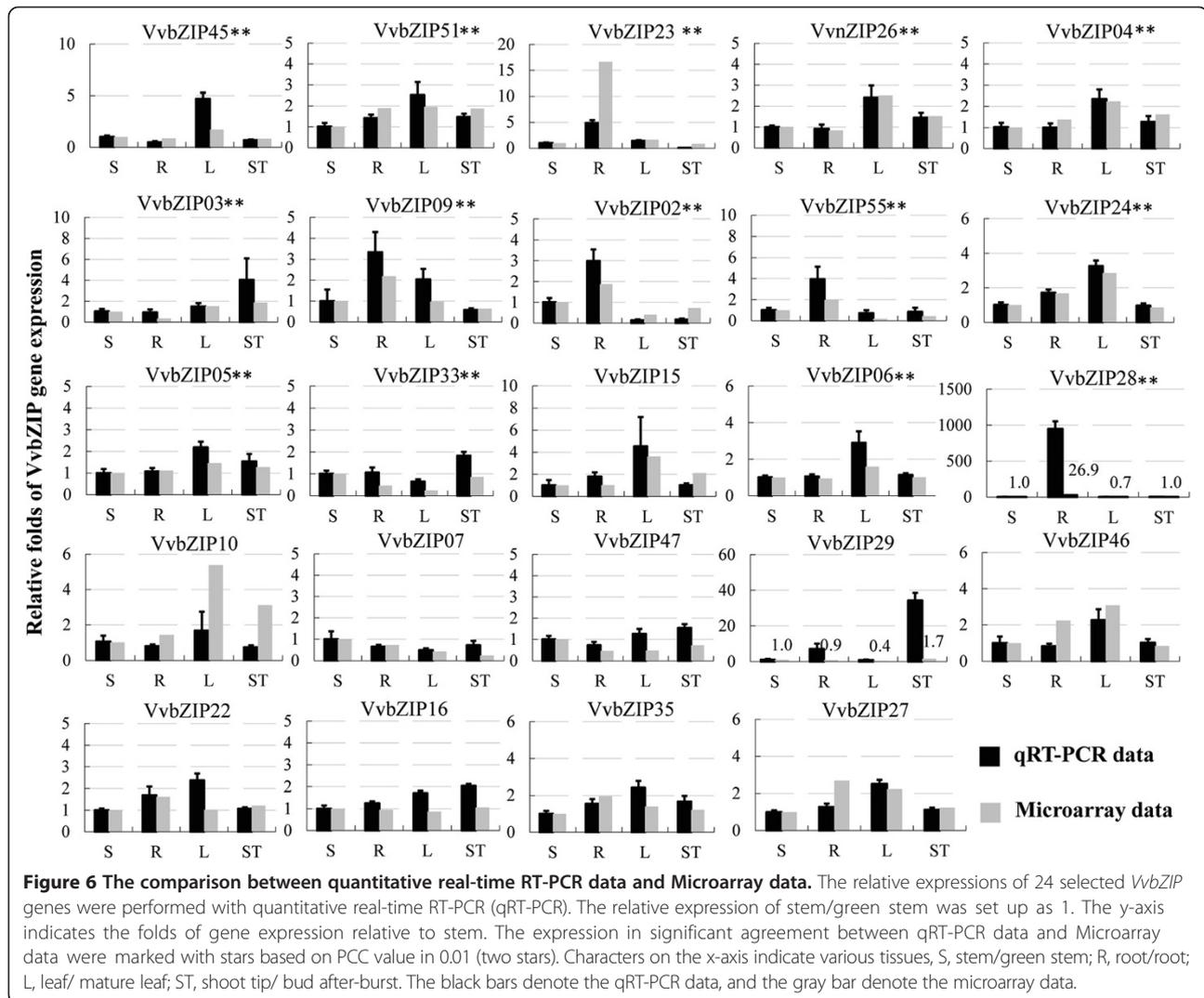
As one of the biggest transcription factor families in eukaryotes, bZIPs owe their DNA-binding ability and specificity to the bZIP domain. We aligned the amino acid sequences of each subfamily and compared them with those in Arabidopsis, rice and other important species in the plant kingdom, including *Physcomitrella patens* (moss), *Selaginella moellendorffii*, *Amborella trichopoda* (Additional file 5). The results show that some amino acid residues are highly conserved within each group across species in the tree of land plant life, indicating that those amino acid residues play crucial roles in the DNA-binding ability of the bZIP domain and have undergone purifying selection during their evolution. Experiments have shown that asparagine (numbered as -18), arginine (numbered as -10) and leucines (numbered as +1, +8, +15), were key and invariant sites of the bZIP domain, and their replacement causes new binding specificities [8,45,49]. In group F and in VvbZIP26 which seemingly belonged to group D in the phylogenetic tree (Figure 4a), the arginine site was replaced by Lys (just occurred in group F) and Ile, respectively, and is consistent with its subfamily members in other species (UC, Additional file 5). These replacements may have contributed to the emergence of a new binding specificity and independent course of evolution of this subfamily. In addition, some leucine sites were also replaced by other amino acids (for example, the third leucine site (+15) of most members in

group C were replaced by Ile, and the second leucine site (+8) were mostly replaced by Met) which may have led to change in the binding specificity of these bZIP domains as well.

Expression analysis of VvbZIP transcription factors at different developmental stages in specific organs in grapevine

Evidence has suggested that bZIP transcription factors are widely involved in the integration and development of many organs and tissues, such as seed maturation and germination [9], floral induction and development [11-18]. To obtain more insight into the temporal and spatial transcription patterns of grapevine bZIP gene family during grapevine development, a heatmap of the global transcription profile of the VvbZIP family was performed based on microarray data of the recent grape gene expression atlas on 54 samples (the sample information is described in Additional file 6) [50]. Fluorescence intensity values were log₂-based to narrow the dynamic range of fluorescence intensities, thus giving a better appreciation of small differences in gene expression. All members of the grapevine bZIP gene family were surveyed. As illustrated in Figure 5, the VvbZIP gene family has a broad expression pattern across various developmental stages of grape organs and tissues. The VvbZIP45, VvbZIP41, VvbZIP24 and VvbZIP05 were abundantly expressed in all examined organs and tissues. In contrast, VvbZIP39, 43, 55, 51 and 44 were expressed at relatively low level. Moreover, VvbZIP53, and VvbZIP19 were mainly expressed in seed at the stage of post-fruit set while VvbZIP25 expression increased during the whole seed development. Notably, those three genes belonged to a subclade of the group A in the phylogenetic tree (Figure 1) harboring Arabidopsis homologous genes (*AtbZIP67/DPBF2*, *AtbZIP39/ABI5* and *AtbZIP15*), which have been confirmed to play an important role in ABA-mediated seed development, germination, and embryo maturation [51,52]. This suggests that VvbZIP19, 25 and 53 may have functions similar to their orthologous Arabidopsis genes during seed development in grapevine. Interestingly, the group E of VvbZIP showed peculiar expression in flower organs. For example, the VvbZIP29 and 32 were mainly expressed in pollen and the VvbZIP54 was highly expressed in Flower-FB, Flower-F, stamen and pollen, and the last member of group E, VvbZIP52, was highly expressed in the carpel but was relatively low in pollen, among others, indicating that the whole group E of grapevine bZIP gene family may be involved in flower development.

To survey the co-expression relationship between VvbZIP genes, Pearson's correlation coefficients (PCC value) were calculated among bZIP genes based on the microarray data of 54 samples (Additional file 7), and showed that

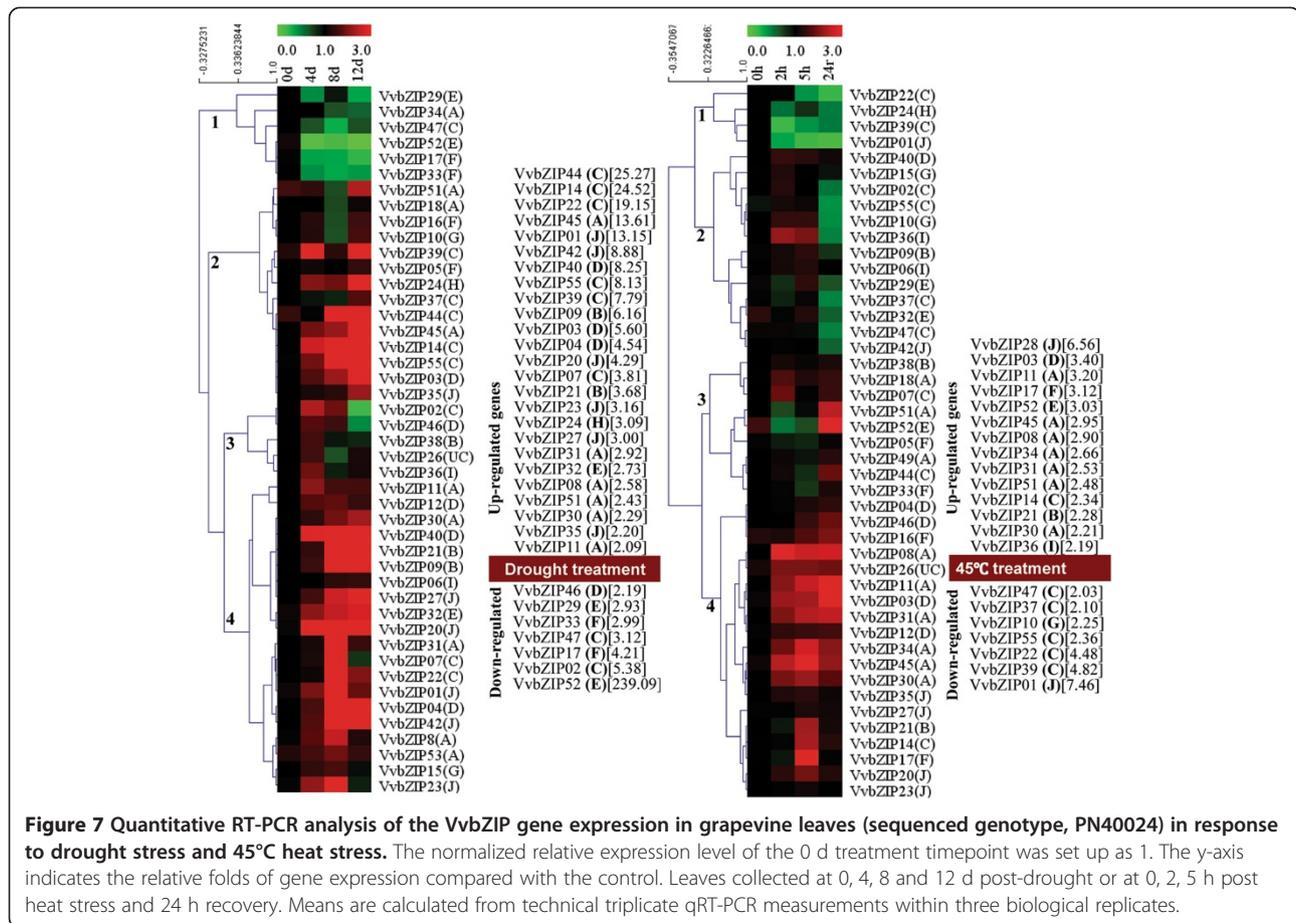


Expression of *VvbZIP* genes under drought and heat stress

Plants and crops are frequently challenged by abiotic stress, such as drought stress and high temperature. Some evidence has suggested that bZIP proteins are widely involved in signaling and responses to abiotic/biotic stimuli [2,26,30,53], but limited information was available on bZIP involvement on stress responses in grapevine. In the present study, long-term drought condition (0, 4, 8 and 12-day) and short-term heat stress (0, 2, 5 hours and 24-hour recovery) on grapevine seedling treatment had marked effect on the expression profile of 47 out of 55 (85.5%) *VvbZIP* genes for which we were able to design primers using for qRT-PCR analysis (Figure 7). As illustrated in Figure 7, both heatmaps (responses to drought and heat) could be divided into four clusters. For the drought treatment, cluster 1 contains 6 (13.3%) members of detectable *VvbZIP* genes, which were widely down-regulated by drought treatment. Cluster 2 contains genes

highly induced at 4 days after drought treatment, but down-regulated at 8 days and 12 days after drought treatment. Cluster 3 mainly consists of genes with the highest expression level at 8 days after drought treatment. Cluster 4 has 15 (33.3%) members which were mainly up-regulated with the increased degree of the drought treatment, and particularly, *VvbZIP44*, *45*, *14*, *55* and *03* were highly induced at 12 days after drought treatment. Overall, 25 *VvbZIP* genes were up-regulated at least two fold after drought treatment relative to the control and ranged from 2 to 25 fold, 7 *VvbZIP* genes were down-regulated at least two fold after drought treatment relative to the control and ranged from 2 to 239 fold.

For the heat stress treatment and recovery, cluster 1 has three (6.7%) members of 45 detectable *VvbZIP* genes, which were widely down-regulated after 45°C treatment and recovery. The members of cluster 2 were mainly down-regulated after 24 hour of recovery. The members



of cluster 3 were mainly down-regulated after 45°C treatment for 2 or 5 hours, but were up-regulated after recovery for 24 hours. Cluster 4 expressed highest across all points of heat stress treatment. Overall, 14 *VvbZIP* genes were up-regulated at least two folds after 45°C treatment or recovery, the fold change ranged from 2 to 7; 7 *VvbZIP* genes were down-regulated at least two fold and ranged from 2 to 7.

Interestingly, some up-regulated genes typically induced by drought were down-regulated during the heat stress or in the recovery process, such as *VvbZIP01*, 39, 22 and 55. On the contrary, some genes down-regulated by the drought treatment were up-regulated by the heat stress or recovery, such as *VvbZIP52* and 17 (Figure 7). This result suggests that different *VvbZIP* genes are involved in these two different biological pathways in response to drought and heat stress.

Furthermore, the Pearson's correlation coefficient (PCC value) among pairs of *VvbZIP* genes based on all of the qRT-PCR expression data in the drought and heat stress conditions (Additional file 9) showed that 66 pairs of genes showed significant correlation at the 0.01 level (2-tailed), including 43 (65.2%) pairs showed significant positive correlation (PCC value ≥ 0.99) and 23

(34.8%) pairs showed significant negative correlation (PCC value ≤ -0.99). The fact that 17 positive and 13 negative pairs of correlations between drought and heat stress conditions suggests that some *VvbZIP* genes may also be involved in some common pathways in responses to drought and heat stresses.

Discussion

Increasing evidence suggests that the bZIP proteins play crucial roles in various developmental and physiological processes in plants, such as seed maturation [9,10], flower development [11-18], stress signaling [25,31,32] and pathogen defense [30]. Much of the research into the functions of the bZIP genes have been characterized in model plants, especially in Arabidopsis, and some crops including rice, barley and Sorghum [2,4,7,54,55]. However, very little is known about this family in other species.

In this report, we exhaustively identified the bZIP genes in grapevine genome, and examined their evolutionary history and high-quality expression profiles across the different developmental stages of various kinds of tissues and under drought and heat stress conditions. The information provides a comprehensive overview of the bZIP gene family in grapevine and lays the foundation for

future functional characterization of individual *VvbZIP* genes.

The distribution and colinearity of grapevine bZIP gene family reflected evolutionary imprint of grapevine genome

Grapevine is evolutionarily the earliest-diverging lineage of rosids and is a sister clade to eurosids I and II [36,42]. The species is believed to have evolved about 60 MYA, and its genome sequence revealed that it has not undergone any whole genome duplication since the gamma genome duplication. Therefore, it serves as a good resource for understanding the evolution and functional divergence of genes in rosids. The bZIP transcription factor family has evolved with a long history, which can be at least traced back to the earliest land plants, such as moss (*Physcomitrella patens*), which contains at least 37 bZIPs [55]. The grapevine, described as a 'palaeo-hexaploid', has homologous triplets in their genome and is believed to have resulted from a genome fusion of a tetraploid and a diploid genome and subsequent diploidization [35]; At the gene family level, the phylogenetic analysis of nucleotide binding site (NBS) gene family members [36], supported the fusion hypothesis, and suggested that the ancestral hexaploid genome of grapevine was derived from an allopolyploidization of two ancestral genomes: Va and Vc. However, the analysis of the calcium-dependent protein kinases (CDPK) gene family showed no indication of such imprint on fusion and divergence [56], and it may be because the gene family (17 members) is not large enough in tracing the evolutionary history, or the retention and loss after fusion was unequal. In this study with a large bZIP gene family (55 members), the distribution and gene duplication events of *VvbZIPs* (Figure 2), are very un-uniform. Five chromosomes (9, 10, 11, 16 and 17) harbor no bZIP gene, and chromosome 15 harbored just one. Interestingly, chromosomes that lack bZIP genes seem to belong to the Vc genome, and this indicates that the grapevine bZIP gene family suffered extensive gene loss after fusion and this loss mainly occurred in Vc genome. The 12 homologous pairs of *VvbZIP* genes plus two triplets (*VvbZIP07/13/47* and *VvbZIP14/22/37*) occurred within the same homologous blocks and this partially supports the fusion hypothesis in the grapevine genome [35,36].

Furthermore, Comparing the number of *VvbZIP* genes with that of other species, it is ~26% less than that in Arabidopsis, and ~38% less than that in rice. This is probably because grapevine did not undergo any whole genome duplication (WGD) events after the gamma WGD as shown in Arabidopsis, or the WGD after split between eudicots and monocots which happened in rice [57].

The colinearity-orthologues of the bZIP genes between grapevine and Arabidopsis

Comparative genomic analyses across different taxa allows the transfer of functional information from a well-

characterized taxon, such as the model plant Arabidopsis, to another less-studied taxon, like grapevine. This is true with bZIP genes, many of which have been well-characterized in Arabidopsis, but few of which have been studied in grapevine. Therefore, the determining orthologous relationships between the Arabidopsis bZIP genes with those in grapevine (Table 1) will allow the grapevine researchers to make a "best guess" of their functions. Most orthologues clustered at the same clade in the phylogenetic tree (Figure 1) share common conserved motifs (data not shown). In Arabidopsis, *AtbZIP36* (*AREB1*), *38* (*AREB2*), and *37* (*ABF3*) have been shown to be master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress response [25]. In grapevine, their orthologues are *VvbZIP8* and *VvbZIP45*, which were also shown to be highly inducible in response to drought stress in this study (Figure 7). Moreover, Nicolas et al. [39] recently demonstrated that *VvbZIP45* transcripts mainly accumulated in the berry during ripening, and were up-regulated by ABA [39]. Comparative genomic analyses between grapevine and Arabidopsis provides a good reference for studying the biological functions of *VvbZIPs*.

The grapevine bZIP gene family may participate in different aspect of organ development

The in-depth analysis on a high-quality microarray data [50] (Figure 5) indicated that bZIP proteins may widely participated in the differentiation and development of many organs and tissues, and this was verified using qRT-PCR (Figure 6). Tissue-specific expression patterns include *VvbZIP19*, *25* and *53* in seed development, *VvbZIP29*, *32* and *54* in flower development, indicating their primary roles in the morphogenesis and organ development. These results may lead to more directed understanding the function of these *VvbZIP* genes in grapevine development biology.

Diverse *VvbZIP* genes are involved in responses to drought and heat stress

Previous studies have shown that bZIP genes play an important role in plants in response to various stresses, such as drought, salt, freezing temperature, and heat [2,4,46]. In this study, the expression profiles of *VvbZIP* genes under drought and heat stresses (Figure 7) revealed that grapevine bZIP genes are widely involved in responding to the drought and heat stress. Interestingly, some drought up-regulated genes were down-regulated in heat stress, which indicates that two sets of *VvbZIP* genes are involved in response to the drought and heat stress, respectively, others are co-expressed in response to both heat and drought stresses, and those genes may participate in shared roles between drought and heat stress. The data of the Pearson's correlation coefficient

(PCC value) among *VvbZIP* gene pairs also supports this notion (Additional file 9).

Conclusions

Grapevine (*Vitis vinifera*) genome contains 55 members in the *bZIP* gene family, and these genes are located in 14 chromosomes. Genome distribution, gene organizations and gene structures suggested a complex evolutionary history of this family in grapevine. The grapevine *bZIP* family is greatly contributed by the segment/chromosomal duplications, which may be associated with the grapevine genome fusion events, their expression profiles in various tissue/organ and their development stages and responses to drought and heat stress condition demonstrated that this gene family is widely involved in grapevine organ development and drought- and heat-responses. The overall picture of this gene family and their potential involvement in growth, development and stress responses will facilitate further research on the *bZIP* gene family regarding their evolutionary history and biological functions.

Methods

Genome-wide identification of *bZIP* transcription factors in grapevine

To identify all *bZIP* transcription factor genes in grapevine (*Vitis vinifera*), all annotated proteins were downloaded from the latest database version (V1) of the 12X assembly of grapevine genome (<http://genomes.cribi.unipd.it/>). The Hidden Markov Model (HMM) profile of the *bZIP* domain (PF00170) downloaded from Protein family (Pfam) (<http://pfam.sanger.ac.uk/>) was used for identification of the *bZIP* genes from the downloaded database of grapevine genome using HMMER3.0 [58]. All output genes with default (<1.0) E-value were collected and the online software SMART (http://smart.embl-heidelberg.de/smart/set_mode.cgi?GENOMIC=1) was used to confirm the integrity of the *bZIP* domain with E-value < 0.1 [59] and the incorrectly predicted genes were rejected. Finally, the non-redundant and confident genes were gathered and assigned as grapevine *bZIP* genes. Furthermore, the same methods were also used for detecting of possible additional *bZIP* signatures in other database, such as Genoscope and NCBI, or identification of poplar *bZIP* genes. The protein sequences of grapevine, poplar and Arabidopsis *bZIP* transcription factors were deposited at Additional file 10.

Phylogenetic analysis and identification of conserved motifs of the *VvbZIP* family

The conserved *bZIP* domain sequences of were extracted from *bZIP* proteins using online program SMART (<http://smart.embl-heidelberg.de/>) for phylogenetic analysis. Multiple alignments of protein sequences were executed by Clustal X 1.83 program [60]. Phylogenetic trees were

constructed using MEGA 5.0 by the Neighbor-Joining (NJ), Minimal Evolution (ME) methods and the bootstrap test carried out with 1000 iterations, but only NJ tree was presented because all trees are consistent. The MEME program (version 4.8.1, <http://meme.nbcr.net/meme/cgi-bin/meme.cgi>) was used for identification of additional conserved motifs outside the *bZIP* domain, with the following parameters: number of repetitions: any; maximum number of motifs: 50; and the optimum motif widths: 6–200 amino acid residues.

Genomic structures, chromosomal locations and gene duplications for *VvbZIP* genes

Each of the *VvbZIP* gene was positioned on grape chromosome derived from the CRIBI (<http://genomes.cribi.unipd.it/>) using the programmed Perl script. The homeologous chromosome segments resulting from whole-genome duplication or fusion events were shown in the same color according to Jaillon et al. [35]. Gene structure display server program (GSDS, <http://gsds.cbi.pku.edu.cn/index.php>) was exploited to illustrate exon/intron organization for the *bZIP* genes by aligning the cDNAs with their corresponding genomic DNA sequences [61]. For detecting the gene duplication events, the MCScanX software (<http://chibba.pgml.uga.edu/mcscan2/>) was used and the E-value was set under 1E-5. The other manipulations followed the operation manual.

Plant materials preparation and treatment

In vitro grapevine plants (*V. vinifera*, the sequenced genotype PN40024) were maintained in vitro on $1/2$ MS medium supplied with 0.3 mg/L indole 3-butyric acid under a 16/8 h photoperiod ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) at 25°C in the culture room.

For drought stress treatment, one-month-old tissue cultured PN40024 plants were acclimatized in the lighted growth chambers [16-h light (25°C)/8-h dark (23°C), relative humidity of 60–70%, and light intensity of $20 \mu\text{mol m}^{-2} \text{s}^{-1}$ for 15 days, then watered regularly for 10 days, and this was set as the 0 d of treatment. One group of plants was then kept watered as the control, the other three groups were stopped being watered (the drought treatment) for a duration of 4, 8 and 12 days, respectively, and every duration of treatment had a corresponding regularly-watered control. For each treatment, three independent plants (biological replications) were sampled.

For heat stress treatment, in vitro-grown one-month old plants were directly placed in the 45°C lighted growth chamber (light intensity of $20 \mu\text{mol m}^{-2} \text{s}^{-1}$), excepting the temperature, the other conditions were the same as in the drought treatment. After heat treatment for 0, 2, 5 hours, plants were kept for 24 hours at 25°C for rapid recovery [27]. Each treatment contained three treated plants, or

biological replicates, and three corresponding non-treated control.

For tissue specificity expression of *VvbZIP* genes, tissues of root, stem, leaf (4-6th leaves from the tip) and shoot tips were harvested separately from three one-month-old tissue in vitro-grown plants of PN40024 (three biological replicates). After freezing in liquid nitrogen, all plant materials were stored at -80°C.

RNA extraction, primer design and quantitative real-time PCR

Total RNA was extracted using Huayueyang Quick RNA isolation Kit (Cat.No.: ZH120; Huayueyang biotechnology, Beijing, China), all protocols followed the manufacturer's procedure. To make sure that there have no DNA contamination, the traditional digestion protocol was used to eliminate DNA in the sample (rDNase[®], Cat. No.: D2270A, TaKaRa Biotechnology, Dalian, China). The qualities and quantities of RNA were determined by agarose gel electrophoresis and Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc.; USA). For cDNA synthesis, 500 ng high quality total RNA was reverse transcribed with oligdT and random primers using SuperScriptIII Reverse Transcriptase (TaKaRa, Dalian, China) according to the manufacturer's instructions.

Gene-specific primers for the *VvbZIP* genes were designed according to the predicted mRNA sequence using Beacon Designer 7.0 software (Premier Biosoft International, USA). To enhance specificity of the primer, most of them were cross-blasted the ORF and 3'UTR of *VvbZIP* genes, all of the primers were tested with PCR amplification, electrophoresis gels, melting curve analysis and the products of PCR amplification were sequenced at least three times, the maximum deviation of which were less than 3 base pairs. The sequences of primers and their products were listed in Additional file 11, and noted that primers for *VvbZIP28*, *49*, *52* and *53* were not very specific, so it was just used in one or two treatment for determining the gene expression.

For quantitative real-time PCR (qRT-PCR) assay, cDNA was diluted to about 100 ng/μl with ddH₂O. The qRT-PCR was carried out on ABI 7300 Real-time PCR System (Applied Biosystems) using SYBR[®] Premix Ex TaqTM (TaKaRa Code: DRR420A, TaKaRa, Dalian, China). The reaction system (total volume of 20 μl) was contained: 10 μl SYBR[®] Premix Ex TaqTM (2×), 0.4 μl of each primer (10 μM), 0.4 μl ROX Reference Dye (50×), 1 μl of template (about 100 ng/ μl), finally, add the ddH₂O to total volume 20 μl. The program were performed as the following steps: 95°C/30 s for pre-denaturation (step 1), 95°C/5 s for denaturation (step 2), 60°C/34 s for primer annealing/extension and gathering the fluorescent signal (step 3), then go to step 2 for 40 circles, each reaction was completed in triplicate for the accuracy of results. At the end,

the melting curve analysis was executed for verifying the specificity of the primer with the following program: 95°C/15 s, 60°C/1 min, 95°C/15 s. Transcript levels were normalized against the average expression of *VvACTIN* gene (VIT_12s0178g00200), and the baseline and threshold cycles (Ct) were automatically determined by the software of the system. For the data analyses, 2^{-ΔΔCt} method [$\Delta\Delta Ct = (Ct_{\text{target gene}} - Ct_{\text{Actin gene}})_{\text{treatment}} - (Ct_{\text{target gene}} - Ct_{\text{Actin gene}})_{\text{control}}$] was used for calculating the relative expression of *VvbZIP* genes. To visualizing the relative fold difference, all data were normalized based on setting up the relative expression level, the expression level of 0-point treatments for drought stress and heat stress and stem for plant tissue specificity was set as 1, "above 1" and "below 1" are considered as upper, or down, regulation in drought and heat treatment, and more, or less abundant in tissue specificity assay, respectively. The data (0 to 3 folds) were displayed and clustered by MeV v4.8 software (<http://www.tm4.org/>), and the genes induced by each stress with their maximum folds (> 2 folds) of relative expression level were accounted and performed in the Figure 7.

Microarray data analysis

To understand the spatial and temporal expression patterns of *VvbZIP* genes during development, a high-throughput microarray data, from recent research [50], was employed for further analysis (Additional file 12). In the data sets, 54 samples, including green and woody tissues and organs at different developmental stages as well as specialized tissues such as pollen and senescent leaves was determined (Additional file 6). The heatmaps were made by MeV v4.8 software (<http://www.tm4.org/>) [62].

Statistics analysis

To determine the co-expression relationship of *VvbZIP* genes in various tissues/organs development stages and drought and heat responses, the Pearson's correlation coefficient (PCC) values among *VvbZIP* gene pairs were calculated based on the microarray data of various tissues/organs development (Additional file 7) and qRT-PCR data of drought and heat stress (Additional file 9) using program Statistical Product and Service Solutions (SPSS v20.0). The significant correlation at the 0.01 level (2-tailed) was determined and denoted with red (significant positive correlation) and blue (significant negative correlation) background.

Additional files

Additional file 1: The gene structure of *VvbZIPs*. (a) The intron-exon arrangement of *VvbZIP* genes; (b) the number of exons in *VvbZIP* genes.

Additional file 2: The position and pattern of introns within bZIP domain.

Additional file 3: Multilevel Consensus Sequence and their logo of VvbZIP proteins as predicted by MEME program.

Additional file 4: The length of Leucine zipper region in bZIP domain of VvbZIP proteins.

Additional file 5: Alignment of bZIP domain of VvbZIP proteins.

Additional file 6: Description of organs and tissues of *Vitis vinifera* cultivar Corvina, from a vineyard in the Verona province (Montorio).

Additional file 7: The Pearson correlation coefficient (PCC value) between VvbZIP genes based on the microarray data of various tissues/organs development.

Additional file 8: Differentially expressed genes of grapevine bZIPs in various kinds of tissues. The microarray data was normalized based on setting up the expression level of first development stage of each tissue as 1, the folds of differential expression was subsequently calculated. The genes above and below the colored tissue bar indicated the up-regulated and down-regulated genes during this tissue development, respectively. The expression fold was presented at bracket along with each VvbZIP gene. The genes fluctuated more than 5 folds was highlight with red characters.

Additional file 9: The Pearson correlation coefficient (PCC value) between VvbZIP genes based on the qRT-PCR data of drought and heat (45°C) treatment.

Additional file 10: Primer sequences used for real-time PCR analysis.

Additional file 11: Protein sequences of grapevine, poplar and Arabidopsis bZIP transcription factors.

Additional file 12: High-throughput microarray data of VvbZIP gene expression.

Abbreviations

ABREs: ABA response elements; WGD: Whole genome duplication; MEME: Multiple expectation maximization for motif elicitation; qRT-PCR: Quantitative reverse transcription polymerase chain reaction.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JYL performed all the bioinformatics analysis and experimental work, and drafted the manuscript; NNC helped in experimental works and data analysis; CF and BC helped in bioinformatics analysis, data mining and management; SDS, GBT and MP shared and analyzed the microarray data and contributed to revision of the manuscript; ZMC conceived the project, provided overall supervision of the study and drafting the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Dr. Anne-Francoise Adam-Blondom (INRA, France) for providing in vitro grapevine plants (*V. vinifera*, the sequenced genotype PN40024). This research is financially supported in part by the Priority Academic of Jiangsu Province, and the Ministry of Education–Nanjing Agricultural University Project # KYT201004.

Author details

¹College of Horticulture, Nanjing Agricultural University, 210095 Nanjing, Jiangsu, China. ²Dipartimento di Biotecnologie, Università degli Studi di Verona, 37134 Verona, Italy. ³Department of Plant Sciences, University of Tennessee, 37996 Knoxville, TN, USA.

Received: 3 August 2013 Accepted: 31 March 2014

Published: 13 April 2014

References

1. Perez-Rodriguez P, Riano-Pachon DM, Correa LG, Rensing SA, Kersten B, Mueller-Roeber B: **PlnTFDB: updated content and new features of the plant transcription factor database.** *Nucleic Acids Res* 2010, **38**(Database issue):D822–D827.

2. Nijhawan A, Jain M, Tyagi AK, Khurana JP: **Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice.** *Plant Physiol* 2008, **146**(2):333–350.
3. Lee SC, Choi HW, Hwang IS, du Choi S, Hwang BK: **Functional roles of the pepper pathogen-induced bZIP transcription factor, CABZIP1, in enhanced resistance to pathogen infection and environmental stresses.** *Planta* 2006, **224**(5):1209–1225.
4. Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F, bZIPRG: **bZIP transcription factors in Arabidopsis.** *Trends Plant Sci* 2002, **7**(3):106–111.
5. Furihata T, Maruyama K, Fujita Y, Umezawa T, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K: **Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1.** *Proc Natl Acad Sci U S A* 2006, **103**(6):1988–1993.
6. Liao Y, Zou HF, Wei W, Hao YJ, Tian AG, Huang J, Liu YF, Zhang JS, Chen SY: **Soybean GmbZIP44, GmbZIP62 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic Arabidopsis.** *Planta* 2008, **228**(2):225–240.
7. Wang J, Zhou J, Zhang B, Vanitha J, Ramachandran S, Jiang SY: **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* 2011, **53**(3):212–231.
8. Wei K, Chen J, Wang Y, Chen Y, Chen S, Lin Y, Pan S, Zhong X, Xie D: **Genome-wide analysis of bZIP-encoding genes in maize.** *DNA Res* 2012, **19**(6):463–476.
9. Toh S, McCourt P, Tsuchiya Y: **HYS is involved in strigolactone-dependent seed germination in Arabidopsis.** *Plant Signal Behav* 2012, **7**(5):556–558.
10. Izawa T, Foster R, Nakajima M, Shimamoto K, Chua NH: **The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development.** *Plant Cell* 1994, **6**(9):1277–1287.
11. Chuang CF, Running MP, Williams RW, Meyerowitz EM: **The PERANTHIA gene encodes a bZIP protein involved in the determination of floral organ number in Arabidopsis thaliana.** *Genes Dev* 1999, **13**(3):334–344.
12. Walsh J, Freeling M: **The liguleless2 gene of maize functions during the transition from the vegetative to the reproductive shoot apex.** *Plant J* 1999, **19**(4):489–495.
13. Strathmann A, Kuhlmann M, Heinekamp T, Droge-Laser W: **BZI-1 specifically heterodimerises with the tobacco bZIP transcription factors BZI-2, BZI-3/TBZF and BZI-4, and is functionally involved in flower development.** *Plant J* 2001, **28**(4):397–408.
14. Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T: **FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex.** *Science* 2005, **309**(5737):1052–1056.
15. Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D: **Integration of spatial and temporal information during floral induction in Arabidopsis.** *Science* 2005, **309**(5737):1056–1059.
16. Muszynski MG, Dam T, Li B, Shirbroun DM, Hou Z, Bruggemann E, Archibald R, Ananiev EV, Danilevskaya ON: **Delayed flowering1 encodes a basic leucine zipper protein that mediates floral inductive signals at the shoot apex in maize.** *Plant Physiol* 2006, **142**(4):1523–1536.
17. Iven T, Strathmann A, Bottner S, Zwafink T, Heinekamp T, Guivarc'h A, Roitsch T, Droge-Laser W: **Homo- and heterodimers of tobacco bZIP proteins counteract as positive or negative regulators of transcription during pollen development.** *Plant J* 2010, **63**(1):155–166.
18. Gibalova A, Renak D, Matczuk K, Dupl'akova N, Chab D, Twell D, Honys D: **AtbZIP34 is required for Arabidopsis pollen wall patterning and the control of several metabolic pathways in developing pollen.** *Plant Mol Biol* 2009, **70**(5):581–601.
19. Yin Y, Zhu Q, Dai S, Lamb C, Beachy RN: **RF2a, a bZIP transcriptional activator of the phloem-specific rice tungro bacilliform virus promoter, functions in vascular development.** *EMBO J* 1997, **16**(17):5247–5259.
20. Guan Y, Ren H, Xie H, Ma Z, Chen F: **Identification and characterization of bZIP-type transcription factors involved in carrot (*Daucus carota* L.) somatic embryogenesis.** *Plant J* 2009, **60**(2):207–217.
21. Shiota H, Ko S, Wada S, Otsu CT, Tanaka I, Kamada H: **A carrot G-box binding factor-type basic region/leucine zipper factor DcBZ1 is involved in abscisic acid signal transduction in somatic embryogenesis.** *Plant Physiol Biochem* 2008, **46**(5–6):550–558.

22. Holm M, Ma LG, Qu LJ, Deng XW: Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis*. *Genes Dev* 2002, **16**(10):1247–1259.
23. Huang X, Ouyang X, Yang P, Lau OS, Li G, Li J, Chen H, Deng XW: *Arabidopsis* FHY3 and HY5 positively mediate induction of COP1 transcription in response to photomorphogenic UV-B light. *Plant Cell* 2012, **24**(11):4590–4606.
24. Chen H, Chen W, Zhou J, He H, Chen L, Deng XW: Basic leucine zipper transcription factor OsbZIP16 positively regulates drought resistance in rice. *Plant Sci* 2012, **193**–194:8–17.
25. Yoshida T, Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K: AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *Plant J* 2010, **61**(4):672–685.
26. Hsieh TH, Li CW, Su RC, Cheng CP, Sanjaya, Tsai YC, Chan MT: A tomato bZIP transcription factor, SJAREB, is involved in water deficit and salt stress response. *Planta* 2010, **231**(6):1459–1473.
27. Liu GT, Wang JF, Cramer G, Dai ZW, Duan W, Xu HG, Wu BH, Fan PG, Wang LJ, Li SH: Transcriptomic analysis of grape (*Vitis vinifera* L.) leaves during and after recovery from heat stress. *BMC Plant Biol* 2012, **12**:174.
28. Shimizu H, Sato K, Berberich T, Miyazaki A, Ozaki R, Imai R, Kusano T: LIP19, a basic region leucine zipper protein, is a Fos-like molecular switch in the cold signaling of rice plants. *Plant Cell Physiol* 2005, **46**(10):1623–1634.
29. Liu C, Wu Y, Wang X: bZIP transcription factor OsbZIP52/RISBZ5: a potential negative regulator of cold and drought stress response in rice. *Planta* 2012, **235**(6):1157–1169.
30. Thurow C, Schiermeyer A, Krawczyk S, Butterbrodt T, Nickolov K, Gatz C: Tobacco bZIP transcription factor TGA2.2 and related factor TGA2.1 have distinct roles in plant defense responses and plant development. *Plant J* 2005, **44**(1):100–113.
31. Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH: ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J* 2002, **32**(3):317–328.
32. Choi H, Hong J, Ha J, Kang J, Kim SY: ABFs, a family of ABA-responsive element binding factors. *J Biol Chem* 2000, **275**(3):1723–1730.
33. Li Z, Zhang L, Yu Y, Quan R, Zhang Z, Zhang H, Huang R: The ethylene response factor ATERF11 that is transcriptionally modulated by the bZIP transcription factor HY5 is a crucial repressor for ethylene biosynthesis in *Arabidopsis*. *Plant J* 2011, **68**(1):88–99.
34. Zander M, Chen S, Imkamp J, Thurow C, Gatz C: Repression of the *Arabidopsis thaliana* jasmonic acid/ethylene-induced defense pathway by TGA-interacting glutaredoxins depends on their C-terminal ALWL motif. *Mol Plant* 2012, **5**(4):831–840.
35. Jaillon O, Aury JM, Noel B, Polcristi A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Huguency P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattanaro F, Anthonard V, Vico V, Del Fabbro C, Alaux M, Di Gasparo G, Dumas V: The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 2007, **449**(7161):463–467.
36. Malacarne G, Perazzolli M, Cestaro A, Sterck L, Fontana P, Van de Peer Y, Viola R, Velasco R, Salamini F: Deconstruction of the (paleo)polyploid grapevine genome based on the analysis of transposition events involving NBS resistance genes. *PLoS One* 2012, **7**(1):e29762.
37. Tak H, Mhatre M: Cloning and molecular characterization of a putative bZIP transcription factor VvbZIP23 from *Vitis vinifera*. *Protoplasma* 2013, **250**(1):333–345.
38. Boneh U, Biton I, Schwartz A, Ben-Ari G: Characterization of the ABA signal transduction pathway in *Vitis vinifera*. *Plant Sci* 2012, **187**:89–96.
39. Nicolas P, Lecourieux D, Kappel C, Cluzet S, Cramer G, Delrot S, Lecourieux F: The basic leucine zipper transcription factor ABSCISIC ACID RESPONSE ELEMENT-BINDING FACTOR2 is an important transcriptional regulator of abscisic acid-dependent grape berry ripening processes. *Plant Physiol* 2014, **164**(1):365–383.
40. Guillaumie S, Fouquet R, Kappel C, Camps C, Terrier N, Moncomble D, Dunlevy JD, Davies C, Boss PK, Delrot S: Transcriptional analysis of late ripening stages of grapevine berry. *BMC Plant Biol* 2011, **11**:165.
41. Cramer GR, Ergul A, Grimplet J, Tillett RL, Tattersall EA, Bohlman MC, Vincent D, Sonderegger J, Evans J, Osborne C, Quilici D, Schlauch KA, Schooley DA, Cushman JC: Water and salinity stress in grapevines: early and late changes in transcript and metabolite profiles. *Funct Integr Genomics* 2007, **7**(2):111–134.
42. Jansen RK, Kaitanis C, Saski C, Lee SB, Tomkins J, Alverson AJ, Daniell H: Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evol Biol* 2006, **6**:32.
43. Boudet N, Aubourg S, Toffano-Nioche C, Kreis M, Lecharny A: Evolution of intron/exon structure of DEAD helicase family genes in *Arabidopsis*, *Caenorhabditis*, and *Drosophila*. *Genome Res* 2001, **11**(12):2101–2114.
44. Wang N, Zheng Y, Xin H, Fang L, Li S: Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. *Plant Cell Rep* 2013, **32**(1):61–75.
45. Suckow M, von Wilcken-Bergmann B, Muller-Hill B: 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* 1993, **12**(3):1193–1200.
46. Fujita Y, Yoshida T, Yamaguchi-Shinozaki K: Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. *Physiol Plant* 2013, **147**(1):15–27.
47. Finkelstein RR, Lynch TJ: The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* 2000, **12**(4):599–609.
48. Ji L, Wang J, Ye M, Li Y, Guo B, Chen Z, Li H, An X: Identification and characterization of the *Populus* AREB/ABF subfamily. *J Integr Plant Biol* 2013, **55**(2):177–186.
49. Suckow M, Schwamborn K, Kisters-Woike B, von Wilcken-Bergmann B, Muller-Hill B: Replacement of invariant bZip residues within the basic region of the yeast transcriptional activator GCN4 can change its DNA binding specificity. *Nucleic Acids Res* 1994, **22**(21):4395–4404.
50. Fasoli M, Dal Santo S, Zenoni S, Torielli GB, Farina L, Zamboni A, Porceddu A, Venturini L, Bicego M, Murino V, Ferrarini A, Delledonne M, Pezzotti M: The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* 2012, **24**(9):3489–3505.
51. Bensmihen S, Giraudat J, Parcy F: Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during *Arabidopsis thaliana* embryo maturation. *J Exp Bot* 2005, **56**(412):597–603.
52. Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T: *Arabidopsis* NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. *Plant J* 2009, **58**(5):843–856.
53. Kesarwani M, Yoo J, Dong X: Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in *Arabidopsis*. *Plant Physiol* 2007, **144**(1):336–346.
54. Onate L, Vicente-Carbajosa J, Lara P, Diaz I, Carbonero P: Barley BL22, a seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of B-hordein promoters in barley endosperm. *J Biol Chem* 1999, **274**(14):9175–9182.
55. Correa LG, Riano-Pachon DM, Schrago CG, dos Santos RV, Mueller-Roeber B, Vincentz M: The role of bZIP transcription factors in green plant evolution: adaptive features emerging from four founder genes. *PLoS One* 2008, **3**(8):e2944.
56. Chen F, Fasoli M, Torielli GB, Dal Santo S, Pezzotti M, Zhang L, Cai B, Cheng ZM: The evolutionary history and diverse physiological roles of the grapevine calcium-dependent protein kinase gene family. *PLoS One* 2013, **8**(12):e80818.
57. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW: Ancestral polyploidy in seed plants and angiosperms. *Nature* 2011, **473**(7345):97–100.
58. Eddy SR: Profile hidden Markov models. *Bioinformatics* 1998, **14**(9):755–763.
59. Letunic I, Doerks T, Bork P: SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 2012, **40**(Database issue):D302–D305.
60. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: The CLUSTAL_X windows interface: flexible strategies for multiple sequence

alignment aided by quality analysis tools. *Nucleic Acids Res* 1997, **25**(24):4876–4882.

61. Guo AY, Zhu QH, Chen X, Luo JC: GSDS: a gene structure display server. *Yi Chuan* 2007, **29**(8):1023–1026.
62. Mar JC, Wells CA, Quackenbush J: Defining an informativeness metric for clustering gene expression data. *Bioinformatics* 2011, **27**(8):1094–1100.

doi:10.1186/1471-2164-15-281

Cite this article as: Liu et al.: Genome-wide analysis and expression profile of the bZIP transcription factor gene family in grapevine (*Vitis vinifera*). *BMC Genomics* 2014 **15**:281.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

