



OPEN ACCESS

EDITED BY

Zhichao Wu,
National Institutes of Health (NIH),
United States

REVIEWED BY

Yang Yu,
Institute of Crop Sciences (CAAS), China
Chang Yuansheng,
Shandong Academy of Agricultural
Sciences Shandong Institute of Pomology,
China
Changwei Bi,
Nanjing Forestry University, China

*CORRESPONDENCE

Xuan Wang
✉ newwxuan@163.com

[†]These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION

This article was submitted to
Plant Bioinformatics,
a section of the journal
Frontiers in Plant Science

RECEIVED 15 February 2023

ACCEPTED 21 March 2023

PUBLISHED 03 April 2023

CITATION

Zhang P, Liu J, Jia N, Wang M, Lu Y,
Wang D, Zhang J, Zhang H and Wang X
(2023) Genome-wide identification
and characterization of the *bZIP* gene
family and their function in starch
accumulation in Chinese chestnut
(*Castanea mollissima* Blume).
Front. Plant Sci. 14:1166717.
doi: 10.3389/fpls.2023.1166717

COPYRIGHT

© 2023 Zhang, Liu, Jia, Wang, Lu, Wang,
Zhang, Zhang and Wang. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Genome-wide identification and characterization of the *bZIP* gene family and their function in starch accumulation in Chinese chestnut (*Castanea mollissima* Blume)

Penglong Zhang^{1,2†}, Jing Liu^{1,2†}, Nan Jia³, Meng Wang¹, Yi Lu¹,
Dongsheng Wang², Jingzheng Zhang², Haie Zhang^{1,2}
and Xuan Wang^{1,2*}

¹Engineering Research Center of Chestnut Industry Technology, Ministry of Education, Qinhuangdao, Hebei, China, ²Hebei Key Laboratory of Horticultural Germplasm Excavation and Innovative Utilization, College of Horticulture Science and Technology, Hebei Normal University of Science and Technology, Changli, Hebei, China, ³Changli Institute of Pomology, Hebei Academy of Agriculture and Forestry Science, Changli, Hebei, China

The transcription factors of basic leucine zipper (bZIP) family genes play significant roles in stress response as well as growth and development in plants. However, little is known about the *bZIP* gene family in Chinese chestnut (*Castanea mollissima* Blume). To better understand the characteristics of *bZIPs* in chestnut and their function in starch accumulation, a series of analyses were performed including phylogenetic, synteny, co-expression and yeast one-hybrid analyses. Totally, we identified 59 *bZIP* genes that were unevenly distributed in the chestnut genome and named them *CmbZIP01* to *CmbZIP59*. These *CmbZIPs* were clustered into 13 clades with clade-specific motifs and structures. A synteny analysis revealed that segmental duplication was the major driving force of expansion of the *CmbZIP* gene family. A total of 41 *CmbZIP* genes had syntenic relationships with four other species. The results from the co-expression analyses indicated that seven *CmbZIPs* in three key modules may be important in regulating starch accumulation in chestnut seeds. Yeast one-hybrid assays showed that transcription factors *CmbZIP13* and *CmbZIP35* might participate in starch accumulation in the chestnut seed by binding to the promoters of *CmISA2* and *CmSBE1_2*, respectively. Our study provided basic information on *CmbZIP* genes, which can be utilized in future functional analysis and breeding studies

KEYWORDS

Castanea mollissima, bZIP, gene family, starch accumulation, yeast one-hybrid, *CmbZIP13*, *CmbZIP35*

1 Introduction

Starch is an important form of carbon storage for the majority of plant species. Throughout the lifecycle of a plant, starch plays roles in development and response to the environment (MacNeill et al., 2017). Short-term storage of starch occurs in the leaves of plants, whereas long-term storage takes place in seeds and tubers, providing material for energy, development, and reproduction (MacNeill et al., 2017). A previous study has reported that starch can be degraded by a kind of α -amylase in response to osmotic stress (Thalmann et al., 2016). In many plants, the starch biosynthesis pathway is catalyzed by enzymes, such as ADP-glucose pyrophosphorylase, starch synthase, starch branching enzyme, and starch de-branching enzyme (Qu et al., 2018), and regulated by transcription factors (TFs) (MacNeill et al., 2017). For example, as an AP2/EREBP TF family member, rice starch regulator 1 negatively regulates the expression of starch synthesis-related genes in rice seeds and is involved in the amylose content of the seed (Fu and Xue, 2010). The endosperm-specific TF TaNAC019 can bind to the promoters of starch metabolism genes, regulate starch accumulation, and improve the quality of wheat grains (Gao et al., 2021).

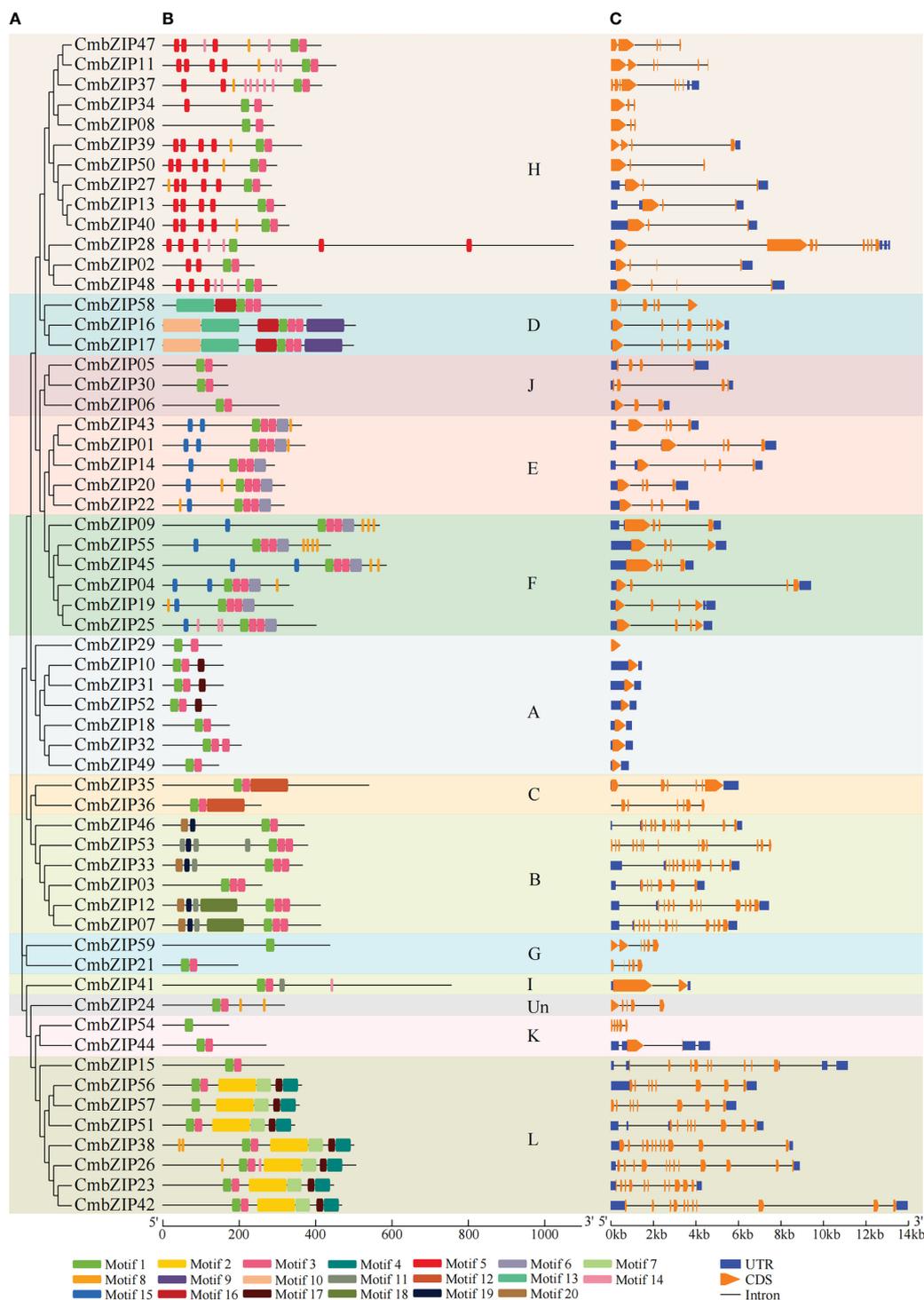
TFs play an indispensable role in plant growth, development, and resistance to biotic and abiotic stresses (Wang et al., 2022). As one of the largest and most diverse TF families, the basic leucine zipper (bZIP) family has been studied extensively (Lee et al., 2006; Schlögl et al., 2008; Wang et al., 2013; An et al., 2018; Song et al., 2020; Duan et al., 2022). The members of the bZIP TF family harbor a highly conserved, 60–80 amino acid long domain, which is composed of a basic region and a leucine zipper region. The sequence of the basic region consists of approximately 20 amino acid residues with a fixed nuclear localization structure N-x7-R/K, which can specifically bind to DNA cis-elements (Lee et al., 2006; Nijhawan et al., 2008). The leucine zipper region, containing the core sequences of L-x6-L-x6-L, consists of various repetitions of leucine or other hydrophobic amino acids, which facilitates hetero- or homo-dimerization of bZIP proteins (Jakoby et al., 2002). The core sequence recognized by bZIP TFs is ACGT, which includes an A-box (TACGTA), C-box (GACGTC), and G-box (CACGTG) (Izawa et al., 1993). Most abscisic acid induced genes have these ACGT cis-elements in their promoter regions. In addition, bZIP TFs can also recognize non-palindrome sequences, such as H-box (CCTACC), GCN4-like motif (GTGAGTCAT), and prolamin box-like (TGAAAA) elements (Kim et al., 2014).

The bZIP TF genes play an important role in many plant biological processes, such as regulating plant morphology and growth. For example, over-expression of the pepper *CabZIP1* gene in *Arabidopsis* slowed plant growth and reduced the number of petals (Lee et al., 2006). The bZIP TFs of tobacco regulate its transition from vegetative growth to reproductive growth (Heinekamp et al., 2002). Plant bZIP transcription factors are induced by plant hormones such as salicylic acid, methyl jasmonate, ethylene, or abscisic acid (Meng et al., 2005; Lee et al., 2006; Schlögl et al., 2008). In addition, many bZIP transcription

factors can regulate abscisic acid synthesis, which regulates gene expression (Finkelstein and Lynch, 2000; Uno et al., 2000). bZIP family genes participate in the regulation of plant resistance to biotic and/or abiotic stresses. Overexpression of bZIP-like proteins in plants under stress conditions can improve the photosynthetic capacity of plants, improving their resistance to salt, cold, herbicides, drought, heat, and other stresses (Kim et al., 2004; Lee et al., 2006; Liao et al., 2008; Zhang et al., 2008). For example, the silencing of the rice endogenous *rT-GA2.1* gene (a member of the bZIP family) mediated by dsRNA can improve the resistance of rice to bacterial pathogens, such as *Xanthomonas oryzae* pv. *oryzae*, indicating that *rTGA2.1* plays a negative role in response to bacterial pathogens (Fitzgerald et al., 2005). Some members of the bZIP family also act as regulators in the starch synthesis pathway. In rice, OsbZIP20, OsbZIP33, and OsbZIP58 TFs interact with *granule-bound starch synthase (GBSS)* and *starch branching enzyme 1 (SBE1)* genes by binding to their promoters, and are capable of regulating starch synthesis (Cai et al., 2002; Wang et al., 2013). In wheat, *TubZIP28* (from *Triticum urartu*) and *TabZIP28* (from *Triticum aestivum*) also participate in the regulation of starch synthesis by interacting with starch synthesis-related genes (Song et al., 2020).

Many bZIP TF families have been identified in important plant species. For example, 75 bZIP genes have been predicted in *Arabidopsis thaliana* (Jakoby et al., 2002). One hundred twenty-five members of the bZIP genes family were identified in maize (*Zea mays*) (Wei et al., 2012), 114 in apple (*Malus domestica*) (Li et al., 2016), 92 in pear (*Pyrus breschneideri*) (Ma et al., 2021), 77 in tobacco (*Nicotiana tabacum*) (Duan et al., 2022), 65 in pomegranate (*Punica granatum*) (Wang et al., 2022), and 227 in wheat (*Triticum aestivum*) (Liang et al., 2022). However, there are no reports on the identification and functionality of bZIP genes in Chinese chestnut (*Castanea mollissima*), despite Chinese chestnut being an economically important dry fruit tree species that is favored for its sweet, fragrant, and waxy characteristics. The waxiness is a critical parameter for chestnut quality and is determined by the starch content in the kernel (Lin et al., 2012; Shi et al., 2021). Over the past four years, several versions of the *C. mollissima* genome have been assembled and published (Xing et al., 2019; Sun et al., 2020; Wang et al., 2020; Hu et al., 2022). These assemblies are resources for the identification of gene families and genetic improvement of chestnut.

In this study, we aimed to identify bZIP genes from Chinese chestnut using the whole genome of N11-1, a seedling Chinese chestnut cultivar line (Wang et al., 2020). We performed phylogenetic, conserved motif, and gene structure analyses to study the relationships between the identified *CmbZIP* family members. Whole genome duplication (WGD) analysis revealed that segmental duplication might be the main factor that led to the expansion of the *CmbZIP* family. Transcriptomic data from different development stages of chestnut seeds showed that *CmbZIP* genes had different expression patterns, and some of them may be related to starch accumulation. These results from this study provide a theoretical reference for *CmbZIP* genes and insight useful for future studies on Chinese chestnut.



intron structure and motif composition were observed for *bZIP* genes of chestnut in the same clade; the structure and motif composition differ between clades, illustrating that the evolution and divergence of *CmbZIPs* might have occurred at an early stage of the evolution of *C. mollisima*.

2.4 Chromosome location and duplication of *bZIPs* in chestnut

Fifty-nine *CmbZIP* genes were distributed unevenly on 11 of the 12 chromosomes (Chr) of chestnut as well as five contigs (Ctg)

(Figure 3; Table S1). Chr02 did not contain any *CmbZIP* genes. There was only one *bZIP* gene on the following chromosomes: Chr07 (*CmbZIP28*), Chr10 (*CmbZIP43*), Ctg1 (*CmbZIP54*), Ctg2 (*CmbZIP55*), Ctg3 (*CmbZIP56*), and Ctg4 (*CmbZIP57*); there were two *bZIP* genes on Ctg5 (*CmbZIP58* and *CmbZIP59*). Eleven *CmbZIP* genes (18.64%) were located on Chr01, which contained the greatest number of *bZIP* genes, with 1 and 10 *CmbZIPs*, respectively, located on the proximate and distal ends of this chromosome. Five *CmbZIPs* were relatively evenly dispersed throughout Chr06. Three to eight *bZIP* genes were located on the remaining eight chromosomes.

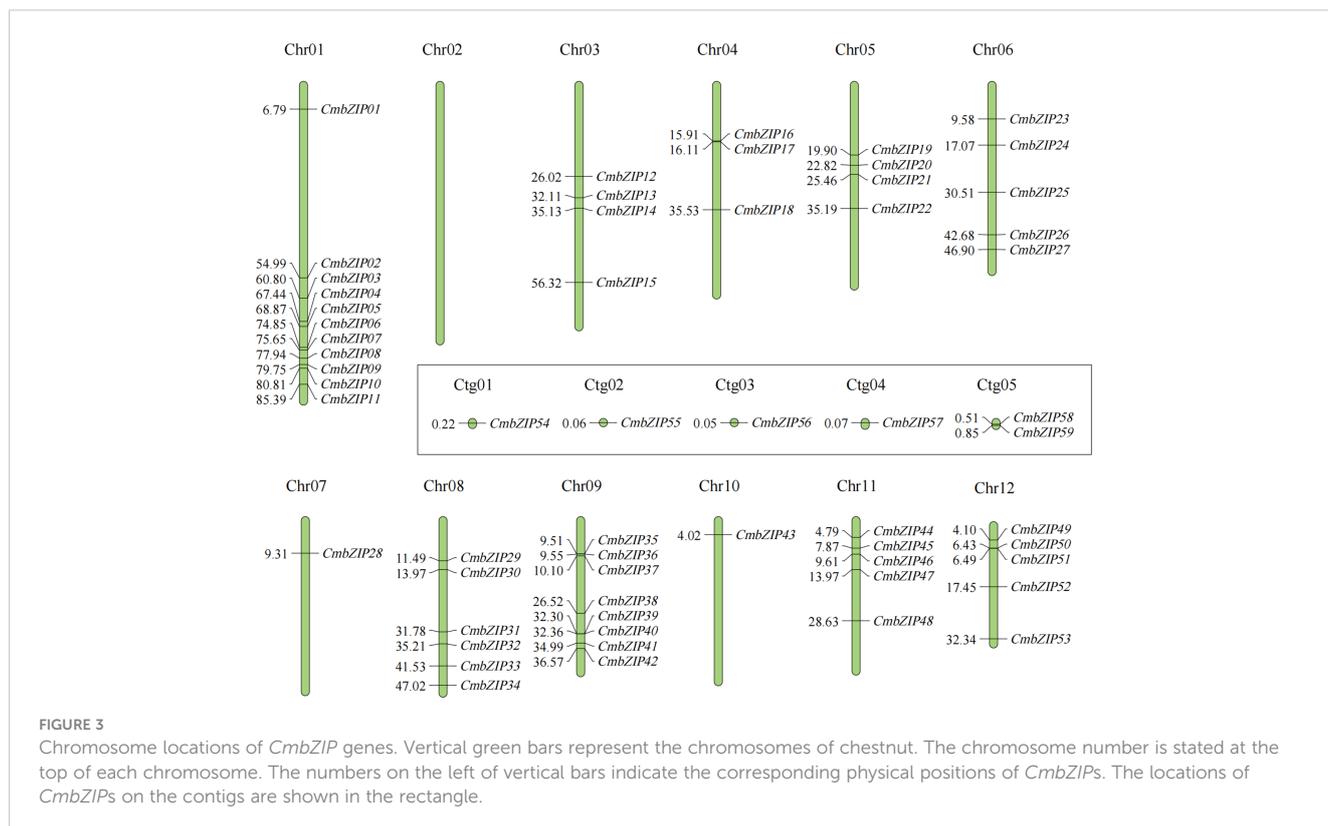
We detected 10 pairs of segmental duplications in *CmbZIP* genes, and no tandem duplications were observed, indicating that segmental duplication events were the major cause of expansion of the *CmbZIP* family (Figure 4; Table 1). Chromosomal distribution analysis revealed that the 20 analogous *CmbZIPs* were unevenly located on the Chinese chestnut genome. We also found that every pair of duplicated *CmbZIPs* were in clades A, B, D, E, F, and H (Table 1). Furthermore, we calculated the synonymous substitution rate (Ks) to estimate the segmental duplication events for *CmbZIPs* (Table 1). The divergence time of all duplicated *CmbZIPs* varied greatly from 9.50–116.09 million years ago (Mya). To predict the selection pressure driving the divergence of *CmbZIPs*, we also calculated the nonsynonymous substitution rate (Ka) and the Ka/Ks ratio. Seven pairs of duplicated *CmbZIPs* might have undergone purifying selection from 27.74–116.09 Mya. The selection pressure on *CmbZIP20/22* was the strongest (Ka/Ks=0.06). Conversely, *CmbZIP09/45* and *CmbZIP19/25* might have recently undergone positive selection from 9.50–11.65 Mya.

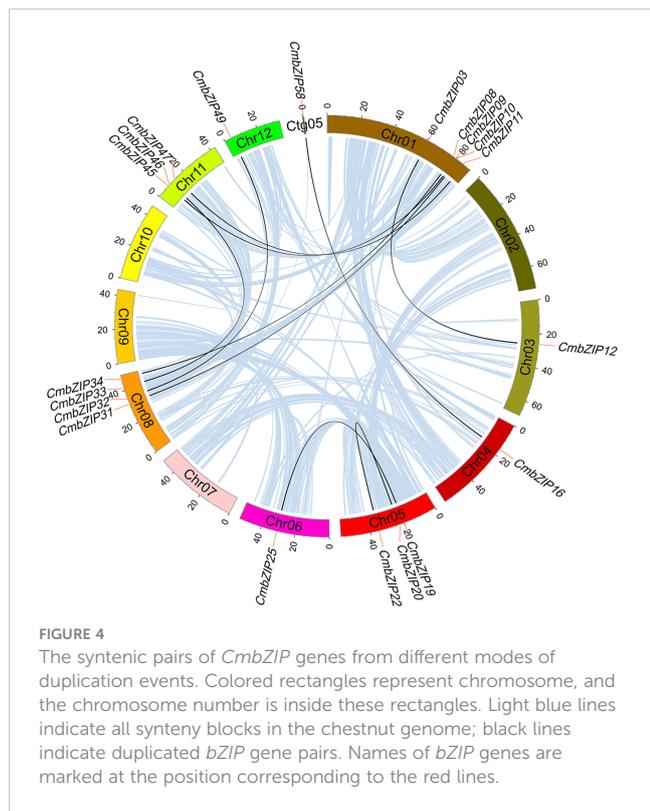
2.5 Synteny analysis of *bZIPs* between genomes

To gain deeper insight into the evolutionary relationships of the *bZIP* genes family among different species, we constructed four comparative syntenic maps of chestnut associated with *Arabidopsis*, rice, wheat, and apple (Figure 5; Table S4). In total, there were 41 orthologous *bZIP* gene pairs between chestnut and the other four species. Further, we found 31 *CmbZIPs* associated with at least two syntenic gene pairs, suggesting that these genes might play an important role in the evolutionary process of the *bZIP* family. Thirty-four *CmbZIPs* showed syntenic relationships with apple *bZIP* genes, 27 with *Arabidopsis*, five with rice, and one with wheat. There was a far greater number of syntenic *bZIP* pairs between chestnut and the two dicots (i.e., apple and *Arabidopsis*) than between chestnut and the two monocots (i.e., rice and wheat), which might indicate that most of these orthologous pairs occurred after the divergence of dicotyledons and monocotyledons.

2.6 Analysis of expression patterns and identification of *CmbZIPs* related to starch accumulation

To confirm the expression patterns of *CmbZIP* genes related to starch synthesis, we used published transcriptome data of all genes from the N11-1 version of the reference genome to determine the fragments per kilobase transcript per million mapped reads (FPKM) (Table S5). All *CmbZIP* genes were clustered to four





subclades by their expression profiles (Figure 6A). The greatest number of *CmbZIPs* were in subclade II, but almost all members (10/19) had very low levels of expression (FPKM < 0.5) in every sample. The 15 members of subclade IV had relatively high expression levels, with an average FPKM ranging from 29.1 to 130.3 (Table S5).

Furthermore, we performed a correlation analysis between expression patterns of 59 *CmbZIPs* and four physiological characteristics: total starch content, amylopectin content, amylose

content, and starch synthase activity (Figures 6B, S2; Table S6). Fourteen highly expressed *CmbZIPs*, namely *CmbZIP33*, *CmbZIP47*, *CmbZIP14*, *CmbZIP38*, *CmbZIP48*, *CmbZIP56*, *CmbZIP35*, *CmbZIP04*, *CmbZIP39*, *CmbZIP40*, *CmbZIP45*, *CmbZIP06*, *CmbZIP13*, and *CmbZIP07*, were identified to be significantly associated ($|r| \geq 0.67$, $p < 0.05$) with starch accumulation (content of total starch or amylopectin) or activity of starch synthase during chestnut seed development (Figure 6B; Table S6).

2.7 Identification of co-expression networks related to starch accumulation

To investigate the co-expression networks related to starch accumulation, eight modules with gene numbers ranging from 921 to 4,890 were identified by weighted gene co-expression network analysis (WGCNA) (Figure S3; Table S5). Analysis of module–trait relationships revealed that three key modules (MEred, MEgreen, and MEbrown) were significantly associated with starch accumulation, with $|r| \geq 0.6$ and $p < 0.05$ (Figure 7A). In detail, MEred module was negatively related to contents of total starch ($r = -0.79$, $p = 0.01$) and amylopectin ($r = -0.8$, $p = 0.01$), and starch synthase activity ($r = -0.82$, $p = 0.006$). In contrast, MEbrown module was positively related to contents of total starch ($r = 0.73$, $p = 0.03$) and amylopectin ($r = 0.76$, $p = 0.02$), and starch synthase activity ($r = 0.76$, $p = 0.02$). MEgreen module was also related to amylopectin content ($r = 0.69$, $p = 0.04$) and starch synthase activity ($r = 0.86$, $p = 0.003$) positively, but not with total starch content. We further identified seven starch accumulation related genes co-expressing with five *CmbZIPs* (*CmbZIP04*, *CmbZIP14*, *CmbZIP33*, *CmbZIP38*, and *CmbZIP56*) in the MEbrown module. Six starch accumulation related genes were found to be co-expressing with one *CmbZIP* (*CmbZIP35*) in the MEgreen module. Two starch accumulation related genes were found to be co-expressing with

TABLE 1 Estimation of the date of segmental duplication events for *CmbZIPs*.

Duplicated gene pairs		Clade		Ka	Ks	Ka/Ks	Divergence time (Mya)
gene 1	gene 2	gene 1	gene 2				
<i>CmbZIP20</i>	<i>CmbZIP22</i>	E	E	0.22	3.48	0.06	116.09
<i>CmbZIP03</i>	<i>CmbZIP12</i>	B	B	0.55	2.42	0.23	80.65
<i>CmbZIP08</i>	<i>CmbZIP34</i>	H	H	0.39	2.29	0.17	76.43
<i>CmbZIP49</i>	<i>CmbZIP32</i>	A	A	0.55	1.60	0.35	53.17
<i>CmbZIP46</i>	<i>CmbZIP33</i>	B	B	0.48	1.43	0.34	47.79
<i>CmbZIP10</i>	<i>CmbZIP31</i>	A	A	0.21	1.35	0.15	45.01
<i>CmbZIP11</i>	<i>CmbZIP47</i>	H	H	0.47	0.83	0.56	27.74
<i>CmbZIP58</i>	<i>CmbZIP16</i>	D	D	1.01	0.96	1.05	32.08
<i>CmbZIP09</i>	<i>CmbZIP45</i>	F	F	0.57	0.35	1.63	11.65
<i>CmbZIP19</i>	<i>CmbZIP25</i>	F	F	0.47	0.29	1.64	9.50

*Ka, the nonsynonymous substitution rate; Ks, the synonymous substitution rate; Mya, million years ago.

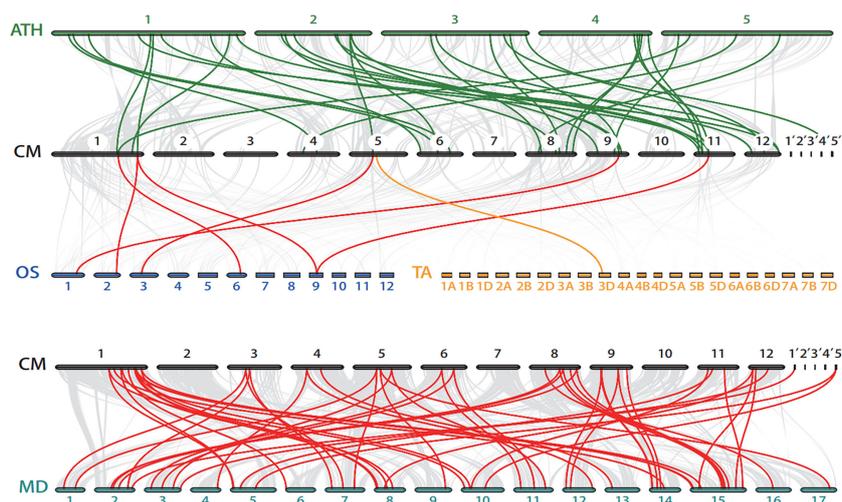


FIGURE 5

Synteny analysis of the *bZIP* genes between chestnut and four other plant species. The grey lines in the background indicate the collinear blocks between the chestnut genome and other genomes. Bold, colored lines highlight the syntenic *bZIP* gene pairs. The colored bars represent chromosomes of different species. 1'-5': Ctg1-Ctg5, contigs in chestnut genome. CM, *Castanea mollissima*; ATH, *Arabidopsis thaliana*; OS, *Oryza sativa*; TA, *Triticum aestivum*; MD, *Malus domestica*.

one *CmbZIP* (*CmbZIP13*) in the MEdred module (Figure 7B; Table S7). The expression profiles of these genes were evaluated by FPKM (Figure 7C).

bZIP transcription factors can recognize ACGT sequences in gene promoter regions, particularly A-box (TACGTA), C-box (GACGTC), and G-box (CACGTG) sequences (Izawa et al., 1993). Interestingly, all starch accumulation related genes from key modules contained at least one ACGT sequence in the promoter region, except for *CmPUL1* (Figure S5; Table S7). A-box occurred at -733 bp and -1,377 bp in promoters of *CmSBE1_2* and *CmAA3*, respectively. C-box was only predicted in the promoter region of *CmBA1_1*, with the positions of -726 bp and -294 bp. G-box was the most frequently identified ACGT sequence in the promoter regions of *CmBA1_1* (-240 bp, -98 bp, and -75 bp), *CmBA4_1* (-289 bp), *CmISA2* (-771 bp), and *CmSSS3* (-648 bp and -595 bp). Based on these findings, we can infer that *CmbZIPs* may regulate the expression of starch accumulation related genes through ACGT cis-elements, and then participate in the accumulation of starch in chestnut seeds.

2.8 Identification of interactions of *CmbZIPs* with promoters of starch accumulation related genes

We selected the only two *bZIP* genes in MEdgreen and MEdred by our interest, *CmbZIP35* and *CmbZIP13*, to identify their potential functions of binding to promoters of starch accumulation related genes. In yeast one-hybrid (Y1H) assays, the Y1H Gold yeast strains containing pCmSBE1_2-AbAi×pGADT7-*CmbZIP35* were able to grow on the screening synthetic dropout medium lacking uracil and leucine (SD-UL) containing 100 ng/mL Aureobasidin A (AbA) (Figure 8A). Similarly, the Y1H Gold yeast strains containing pCmISA2-AbAi×pGADT7-*CmbZIP13* were able to grow under

the same conditions (Figure 8B). These observations suggested that *CmbZIP35* can directly bind to the promoter of *CmSBE1_2*, and *CmbZIP13* can bind to the promoter of *CmISA2*.

3 Discussion

Previous studies have shown that the *bZIP* TF family plays an important role in the regulation of plant growth and development as well as resistance to biotic and abiotic stresses (Lee et al., 2006; Schlögl et al., 2008; Song et al., 2020; Wang et al., 2022). Numerous studies have been conducted on *bZIP* TFs, such as in *Arabidopsis* (Gibalova et al., 2017), wheat (Song et al., 2020), rice (Wang et al., 2013), tobacco (Duan et al., 2022), and apple (An et al., 2018). Although several versions of the Chinese chestnut genome have been published (et al., 2019; Sun et al., 2020; Wang et al., 2020; Hu et al., 2022), to our knowledge, there have been no published studies on the function of *bZIP* TFs in chestnut. In this study, we identified 59 *bZIP* genes in the chestnut genome, which has a complete genome size of 689.98 Mb (Wang et al., 2020), and further analyzed the characteristics of the *bZIP* genes.

The *bZIP* domain consists of a basic region and a leucine zipper region (Jakoby et al., 2002). In the present study, we identified the two structural features *via* analyzing conserved motifs, and they were named motif 1 and motif 3 (Figure S1); this is similar to the *bZIPs* in tobacco (Duan et al., 2022). We also identified another 18 motifs in the 59 *CmbZIP* proteins (Figure 2B). Strong clade-specificity and conservation were detected in the motif distribution and phylogenetic analyses of *CmbZIPs*, and the findings were similar to those reported for pomegranate (Wang et al., 2022), pear (Ma et al., 2021), tobacco (Duan et al., 2022), wheat (Liang et al., 2022), and apple (Li et al., 2016). The *bZIPs* containing the same motifs might have similar functions (Wang et al., 2022). For example, all members in clade D, which contained

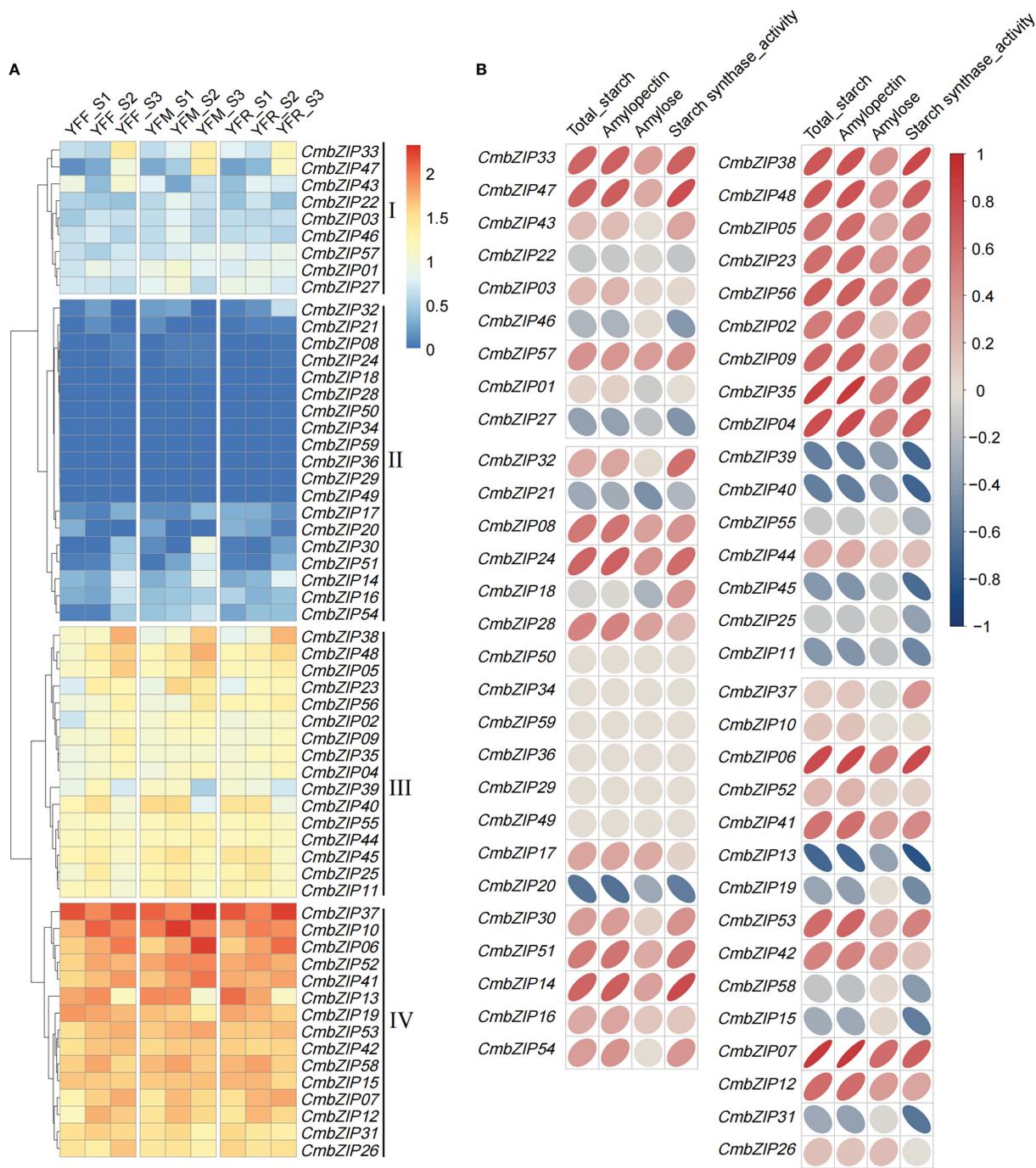


FIGURE 6 Expression pattern of 59 *CmbZIP* genes and correlation analysis of *CmbZIP* expression and physiological characteristics of three chestnut crosses (Li et al., 2021). **(A)** The heatmap shows the expression pattern of *CmbZIP* genes. YFF, YFM, and YFR indicate seeds from crosses of ‘Yongfeng 1’x‘Yongfeng 1,’ ‘Yongfeng 1’x‘Yimen 1,’ and ‘Yongfeng 1’x‘Yongren Zao,’ respectively. S1, S2, and S3 indicate 70, 82, and 94 days after pollination, respectively. The Roman numerals along the right-hand side of the figure indicate \log_{10} FPKM. **(B)** The correlation coefficients between the expression of *CmbZIPs* and four traits are shown by elliptical bubbles. The flatter the bubble, the higher the correlation. Dark red (from top right to bottom left) and dark blue (from top left to bottom right) ellipses represent positive and negative correlations, respectively. The numbers in the legend indicate the correlation coefficient.

motifs 13 and 16, encode the light-inducible protein CPRF2; and most members of clade H, with motifs 5 and 8, may participate in the pathways responding to abiotic stress by encoding ABSCISIC ACID-INSENSITIVE 5-like proteins (Figure 2B; Table S1). In addition, clade-specificity was observed in the number of introns

(Figure 2C; Table S1), similar to previous studies (Duan et al., 2022; Liang et al., 2022; Wang et al., 2022). These analyses suggested that conserved motifs and gene structure were critical for members in same clade during evolution and functional differentiation (Wang et al., 2022).

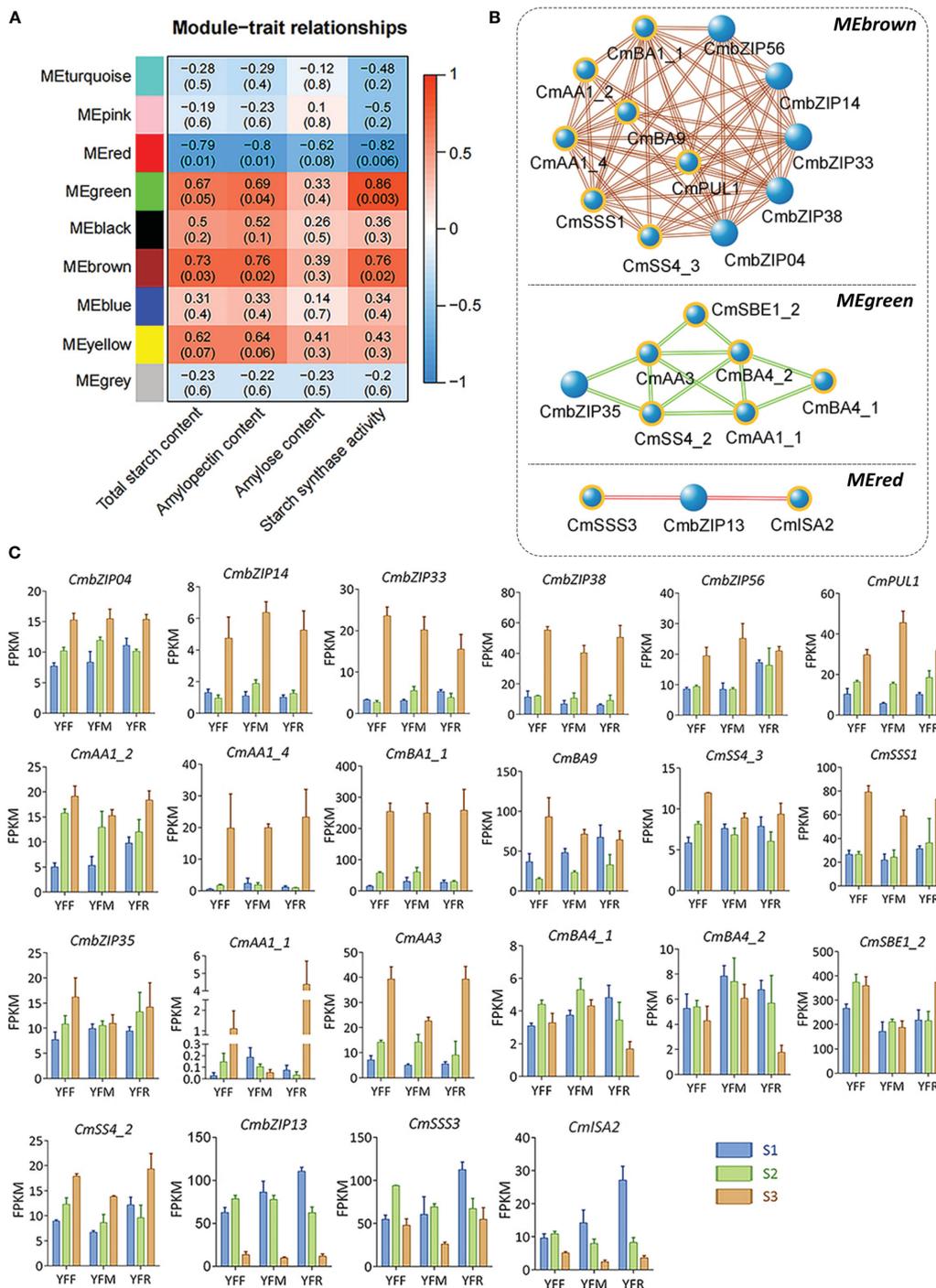
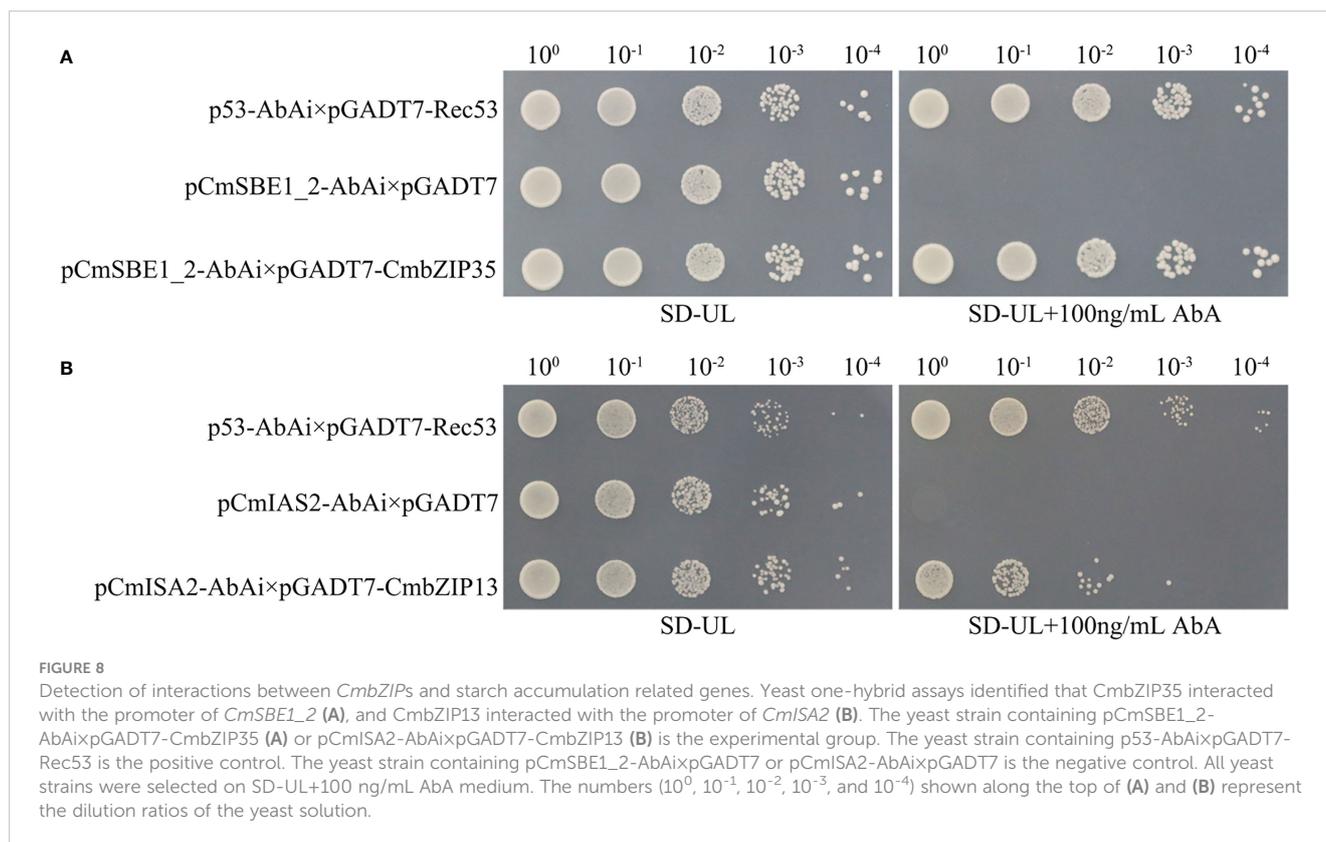


FIGURE 7 Module-trait relationships, co-expression networks, and module-specific gene expression profiles based on results from the WGCNA. **(A)** The heatmap represents relationships between WNCGA modules and traits. The top and bottom numbers in the heat grid represent the correlation coefficients and p-values (shown in parentheses), respectively. **(B)** The networks represent co-expression relationships of *CmbZIPs* and starch accumulation related genes in three key modules. The blue balls highlighted in indicate starch accumulation related genes; the larger balls indicate *CmbZIPs*. Bold and italic characters indicate the names of key modules. **(C)** The column diagrams describe the expression profiles of genes in panel **(B)**. The abbreviations YFF, YFM, YFR, S1, S2, and S3 are the same as those used in Figure 6.

WGD drives the evolution and differentiation of plant genome structure (Paterson et al., 2012). WGD, especially segmental and tandem duplications, drive the expansion of gene families (Li et al., 2019; Sun et al., 2019; Wu et al., 2019; Duan et al., 2022). In this

study, ten pairs of *CmbZIPs* were identified to have derived from segmental duplication events, and these duplicated genes were clustered into the same clade (Table 1). Based on the Ka/Ks ratio, we found that *CmbZIP* genes have undergone purifying and/or



positive selection events. It is likely that seven pairs of *CmbZIPs* have undergone purifying selection pressures before the divergence of *C. mollissima* and *Quercus robur* (18.3 Mya) (Wang et al., 2020). These genes may have maintained similar functions. *CmbZIP09/45* and *CmbZIP19/25* might have undergone positive selection after the formation of the genome of *C. mollissima*, which might play an important role in promoting the evolution of *C. mollissima*.

In the phylogenetic analysis, CmbZIP proteins were grouped according to their homology with AtbZIPs (Jakoby et al., 2002). The bZIPs with highly similar protein sequences were clustered into the same clade, in which the members maintained similar functions. Previous studies have shown that two G-box binding factor (GBF) encoding genes *AtbZIP41* (*GBF1*) and *AtbZIP55* (*GBF3*) might play roles during seed maturation (Chern et al., 1996; Jakoby et al., 2002). Therefore, it is possible that CmbZIPs in clade B, especially CmbZIP33, CmbZIP12, and CmbZIP07 (Figure 1), may be involved in the maturation of the chestnut kernel. Furthermore, AtbZIPs in clades F, H, and L might be involved in abiotic and biotic stress response (Jakoby et al., 2002; Liu et al., 2010; Skubacz et al., 2016; Lapham et al., 2018). CmbZIP16/17/58 and CmbZIP41 were, respectively, closely related to OsbZIP33/58 and TabZIP28/TubZIP28, which have been shown to be involved in starch synthesis (Cai et al., 2002; Song et al., 2020). Therefore, we hypothesize that CmbZIP16/17/58 and CmbZIP41 may participate in the regulation of starch synthesis.

In starchy seeds, such as rice and wheat, starch acts as a sink of carbon allocation. In developing seeds, starch is synthesized from sucrose, which is catalyzed by enzymes and regulated by TFs

(MacNeill et al., 2017). Previous studies found that OsbZIP58 participates in the regulation of starch synthesis in rice by binding to the promoters of *OsAGPL3*, *OsGBSS*, *OsSSIIa*, *OsSBE1*, *OsBEIIb*, and *OsISA2*, which encode enzymes critical to the starch synthesis process (Wang et al., 2013). Similarly, a previous study reported that TubZIP28 and TabZIP28 are both capable of binding to the promoter of cytosolic AGPase encoding gene to enhance the total starch content (Song et al., 2020). In our study, the expression pattern of *CmbZIP* genes in developing chestnut seeds was further analyzed (Figure 6). The results showed that most *bZIP* genes were highly expressed 70–94 DAP, except for *CmbZIP08*, *CmbZIP18*, *CmbZIP24*, *CmbZIP28*, *CmbZIP29*, *CmbZIP34*, *CmbZIP36*, *CmbZIP49*, *CmbZIP50*, and *CmbZIP59*, which showed overall low levels of expression. This suggested that *CmbZIP* genes might participate in the regulation of chestnut seed maturation. In the correlation and co-expression analysis, we identified that the expression of seven *bZIPs* from the modules containing *CmbZIP04*, *CmbZIP13*, *CmbZIP14*, *CmbZIP33*, *CmbZIP35*, *CmbZIP38*, and *CmbZIP56* was closely related to starch (especially amylopectin) accumulation in chestnut seeds (Figure 7; Table S6) (Li et al., 2021). In the analysis of cis-elements, several 'ACGT' elements, including A- and G-boxes, were found in the promoter regions of *CmISA2* (encoding an isoamylase-type starch de-branching enzyme) and *CmsBE1_2* (encoding a starch branching enzyme) (Figure S5). *CmISA2* and *CmsBE1_2* can modify glucan. In the Y1H assays, CmbZIP13 and CmbZIP35 TFs were found to directly bind to the promoters of *CmISA2* and *CmsBE1_2*, respectively (Figure 8). Therefore, we hypothesize that *CmbZIP13* and *CmbZIP35* genes might participate in starch accumulation

in the chestnut seed by interacting with *CmISA2* and *CmSBE1_2*, respectively.

4 Materials and methods

4.1 Identification of *CmbZIP* genes

The protein sequence data from the previously published genome of the N11-1 Chinese chestnut, a seedling chestnut cultivar, were obtained from the Genome Warehouse in BIG Data Center under accession number GWHANWH00000000 (<https://bigd.big.ac.cn/gwh>) (Wang et al., 2020). A hidden Markov model (HMM) was used to identify chestnut bZIP candidates, and the HMM profile of bZIP (PF00170) was downloaded from the Pfam protein database (<http://pfam.xfam.org/>) (Finn et al., 2016). To identify *CmbZIP* genes, the *hmmsearch* tool of HMMER 3.0 software (Potter et al., 2018) was used to retrieve a domain similar to the bZIP domain in chestnut. The *hmmbuild* tool was used to rebuild the new HMM profile to re-identify *CmbZIP* protein sequences. Finally, these protein sequences were confirmed as bZIPs via the conserved domain using Pfam (<http://pfam.xfam.org/>) and Batch CD-Search web tool (<https://www.ncbi.nlm.nih.gov/cdd/>) (Lu et al., 2020).

4.2 Sequence and phylogenetic analyses

All CDS sequences of *CmbZIP* genes were submitted to ExpASy (<https://www.expasy.org>) to determine gene length, amino acid length, relative molecular weight, isoelectric point, hydrophilicity, stability, and other physicochemical properties analysis. The protein sequences of 59 *CmbZIP* genes, 23 *AtbZIP* genes (Jakoby et al., 2002), *OsbZIP20* (Izawa et al., 1994), *OsbZIP33* (Cai et al., 2002), *OsbZIP58* (Wang et al., 2013), *TubZIP28*, and *TabZIP28* (Song et al., 2020) were imported into MEGA X, and ClustalW was used for multiple sequence alignments (Kumar et al., 2018). A Neighbor-Joining (NJ) phylogenetic tree was constructed using MEGA X software with bootstrapping set to 1,000. The optional parameters substitution model and gaps data treatment were set to p-distance and partial deletion, respectively. EvolView (<https://www.evolgenius.info/evolview/>) was used to annotate and visualize the phylogenetic trees (Subramanian et al., 2019). *CmbZIP* proteins were grouped according to their homology with *AtbZIPs* (Jakoby et al., 2002).

4.3 Gene structure and conserved motif analyses

A gene structure analysis of 59 *CmbZIP* genes was performed with general feature format (GFF) file, and visualized using the online software Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/>) (Hu et al., 2015). Conserved motifs were identified using Multiple Expectation Maximization for Motif Elicitation (MEME version

5.1.0) (Bailey et al., 2015) with motif width set to 8–100 and the parameter of maximum motif number set to 20.

4.4 Chromosomal localization and synteny analyses

All *CmbZIP* genes were mapped to Chinese chestnut chromosomes based on physical location information from the GFF file using Mapchart 2.32 software (Voorrips, 2002). Multiple Collinearity Scan toolkit (MCScanX) was used to identify the syntenic gene pairs within the genome, using the results from all-vs-all BLASTP analysis (Wang et al., 2012). Results were displayed with Circos (version 0.69-8) software (Krzywinski et al., 2009). The values of *Ka* and *Ks* were calculated using the *KaKs-calculator 2.0* (Wang et al., 2010). The divergence-times of duplicated gene pairs were estimated using the *Ks* value with the formula $T = Ks/2r$, where *T* is the divergence-time and *r* is the divergence rate of nuclear genes from plants ($r = 1.5 \times 10^{-8}$) (Koch et al., 2000; Huang et al., 2016). We used the Python version of MCScan to analyze the synteny between the genomes of *C. mollissima*, *A. thaliana*, *O. sativa*, *T. aestivum*, and *M. domestica* (Tang et al., 2014).

4.5 Expression profile and co-expression analyses based on RNA-seq

The published RNA-seq data were obtained from the sequence read archive (SRA) in NCBI (accession number PRJNA540079) (Li et al., 2021). All seed samples were collected at 70, 82, and 94 days after pollination from three crosses: ‘Yongfeng 1’×‘Yongfeng 1’, ‘Yongfeng 1’×‘Yimen 1’, and ‘Yongfeng 1’×‘Yongren Zao’ (Li et al., 2021). RNA-seq read-files were converted from SRA to fastq format using *sratoolkit3.0* (Goldberg et al., 2009). The mapping genome used in the RNA-seq analysis was changed from the previous version of the genome assembly to N11-1. The expression profiles of *CmbZIPs* were determined using Tophat2 software (Kim et al., 2013; Wang et al., 2020; Li et al., 2021). The FPKM value was used to evaluate the expression level of each gene. The correlation coefficients and p values between the \log_{10} FPKM of *CmbZIPs* and four physiological characteristics (i.e., total starch content, amylopectin content, amylose content, and starch synthase activity) published in a previous study were estimated using GraphPad Prism version 6.02 for Windows (Li et al., 2021; <https://www.graphpad.com/>). The co-expression modules were identified using WGCNA with the R package (Langfelder and Horvath, 2008). The co-expression networks were generated using Cytoscape software (Otasek et al., 2019).

4.6 Identification of cis-elements in gene promoter regions

We retrieved sequences 1,500 bp upstream of the transcription start site of starch accumulation related genes. These sequences

were submitted to PlantCARE to identify cis-elements, which might affect gene expression and function (Rombauts et al., 1999).

4.7 Sequence cloning and yeast one-hybrid assays

The open reading frames of *CmbZIP35* and *CmbZIP13* were cloned by reverse transcription PCR using RNA from developing seeds of *C. mollissima* cultivar ‘Yanbao.’ We confirmed these sequences using DNAMAN 6.0 software. Subsequently, we fused the two sequences into the pGADT7 vector to construct the pGADT7-*CmbZIP35* and pGADT7-*CmbZIP13* recombinant plasmids, respectively. The *CmSBE1_2* and *CmISA2* promoter fragments were cloned by PCR using DNA from ‘Yanbao’ seeds. The fragments were confirmed and inserted into the pAbAi vector to construct the pCmSBE1_2-AbAi and pCmISA2-AbAi recombinant plasmids. All primer sequences are listed in Table S8.

To determine the optimal AbA concentration, the Y1H Gold yeast strain containing the recombinant pAbAi plasmids were grown on SD-UL screening medium supplemented with different AbA concentrations (Yang et al., 2019). Then, Y1H Gold yeast cells were co-transformed with pGADT7-*CmbZIP35* and pCmSBE1_2-AbAi plasmids, as well as pGADT7-*CmbZIP13* and pCmISA2-AbAi plasmids. Interactions were detected on SD-UL selection medium that was supplemented with 100 ng/mL AbA.

5 Conclusions

A total of 59 *CmbZIP* genes were identified in Chinese chestnut. These *CmbZIPs* were clustered into 13 clades with clade-specific motifs and structures. Segmental duplication was determined as the major driving force of the expansion of the *CmbZIP* gene family. *CmbZIP04*, *CmbZIP13*, *CmbZIP14*, *CmbZIP33*, *CmbZIP35*, *CmbZIP38*, and *CmbZIP56* were confirmed to be highly correlated with starch accumulation in chestnut seeds. We demonstrated that *CmbZIP13* and *CmbZIP35* may regulate starch accumulation in the chestnut seed by binding to the promoters of *CmISA2* and *CmSBE1_2*, respectively. These results indicated that *CmbZIP* genes contained information related to starch accumulation in chestnut seeds, which can be used in future functional analysis and breeding studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found in the article/Supplementary Material.

Author contributions

XW and HZ designed the experiments. MW and YL collected data of Chinese chestnut genome and RNA-seq. NJ and JL drew all figures of this study. PZ and JL carried out the yeast one-hybrid experiments. PZ and XW performed the bioinformatic analysis and wrote the paper. DW and JZ directed and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by a subproject of Hebei Collaborative Innovation Center of Chestnut Industry (202202) and Scientific Research Foundation of Hebei Normal University of Science and Technology (2022YB002).

Acknowledgments

The authors thank the Hebei Normal University of Science and Technology, and Engineering Research Center of Chestnut Industry Technology, Ministry of Education for the experimental materials and technical assistance provided.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1166717/full#supplementary-material>

References

- An, J., Yao, J., Xu, R., You, C., Wang, X., and Hao, Y. (2018). Apple bZIP transcription factor MdbZIP44 regulates abscisic acid-promoted anthocyanin accumulation. *Plant Cell Environ.* 41 (11), 2678–2692. doi: 10.1111/pce.13393
- Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME suite. *Nucleic Acids Res.* 43 (W1), W39–W49. doi: 10.1093/nar/gkv416
- Cai, Y., Xie, D., Wang, Z., and Hong, M. (2002). Interaction of rice bZIP protein REB with the 5'-upstream region of both rice *sbe1* gene and waxy gene. *Chin. Sci. Bulletin.* 47, 310–314. doi: 10.1360/02tb9074
- Chern, M. S., Bobb, A. J., and Bustos, M. M. (1996). The regulator of MAT2 (ROM2) protein binds to early maturation promoters and represses PvALF-activated transcription. *Plant Cell.* 8 (2), 305–321. doi: 10.1105/tpc.8.2.305
- Duan, L., Mo, Z., Fan, Y., Li, K., Yang, M., Li, D., et al. (2022). Genome-wide identification and expression analysis of the bZIP transcription factor family genes in response to abiotic stress in *Nicotiana glauca* L. *BMC Genomics* 23 (1), 318. doi: 10.1186/s12864-022-08547-z
- 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 (4), 599–609. doi: 10.1105/tpc.12.4.599
- Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* 44 (D1), D279–D285. doi: 10.1093/nar/gkv1344
- Fitzgerald, H. A., Canlas, P. E., Chern, M. S., and Ronald, P. C. (2005). Alteration of TGA factor activity in rice results in enhanced tolerance to *Xanthomonas oryzae* pv. *oryzae*. *Plant J.* 43 (3), 335–347. doi: 10.1111/j.1365-3113.2005.02457.x
- Fu, F., and Xue, H. (2010). Coexpression analysis identifies rice starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol.* 154 (2), 927–938. doi: 10.1104/pp.110.159517
- Gao, Y., An, K., Guo, W., Chen, Y., Zhang, R., Zhang, X., et al. (2021). The endosperm-specific transcription factor TaNAC019 regulates glutenin and starch accumulation and its elite allele improves wheat grain quality. *Plant Cell.* 33 (3), 603–622. doi: 10.1093/plcell/koaa040
- Gibalova, A., Steinbachova, L., Hafidh, S., Blahova, V., Gadiou, S., Michailidis, C., et al. (2017). Characterization of pollen-expressed bZIP protein interactions and the role of ATbZIP18 in the male gametophyte. *Plant Reprod.* 30 (1), 1–17. doi: 10.1007/s00497-016-0295-5
- Goldberg, D. H., Victor, J. D., Gardner, E. P., and Gardner, D. (2009). Spike train analysis toolkit: Enabling wider application of information-theoretic techniques to neurophysiology. *Neuroinformatics* 7, 165–178. doi: 10.1007/s12021-009-9049-y
- Heinekamp, T., Kuhlmann, M., Lenk, A., Strathmann, A., and Droge-Laser, W. (2002). The tobacco bZIP transcription factor BZI-1 binds to G-box elements in the promoters of phenylpropanoid pathway genes *in vitro*, but it is not involved in their regulation *in vivo*. *Mol. Genet. Genomics* 267, 16–26. doi: 10.1007/s00438-001-0636-3
- Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31 (8), 1296–1297. doi: 10.1093/bioinformatics/btu817
- Hu, G., Cheng, L., Cheng, Y., Mao, W., Qiao, Y., and Lan, Y. (2022). Pan-genome analysis of three main Chinese chestnut varieties. *Front. Plant Sci.* 13:916550. doi: 10.3389/fpls.2022.916550
- Huang, Z., Duan, W., Song, X., Tang, J., Wu, P., Zhang, B., et al. (2016). Retention, molecular evolution, and expression divergence of the auxin/indole acetic acid and auxin response factor gene families in *Brassica rapa* shed light on their evolution patterns in plants. *Genome Biol. Evol.* 8 (2), 302–316. doi: 10.1093/gbe/evv259
- Izawa, T., Foster, R., and Chua, N.-H. (1993). Plant bZIP protein DNA binding specificity. *J. Mol. Biol.* 230 (4), 1131–1144. doi: 10.1006/jmbi.1993.1230
- 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. *Plant Cell.* 6 (9), 1277–1287. doi: 10.1006/jmbi.1993.1230
- Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., et al. (2002). bZIP transcription factors in arabidopsis. *Trends Plant Sci.* 7 (3), 106–111. doi: 10.1016/S1360-1385(01)02223-3
- Kim, S., Kang, Jy, Cho, D. I., Park, J. H., and Kim, S. Y. (2004). ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *Plant J.* 40 (1), 75–87. doi: 10.1111/j.1365-3113.2004.02192.x
- Kim, D., Perthe, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptsomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14 (4), 1–13. doi: 10.1186/gb-2013-14-4-r36
- Koch, M. A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in arabidopsis, arabis, and related genera (*Brassicaceae*). *Mol. Biol. Evol.* 17 (10), 1483–1498. doi: 10.1093/oxfordjournals.molbev.a026248
- Krzywinski, M., Schein, J., Biról, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19 (9), 1639–1645. doi: 10.1101/gr.092759.109
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35 (6), 1547. doi: 10.1093/molbev/msy096
- Langfelder, P., and Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf.* 9 (1), 1–13. doi: 10.1186/1471-2105-9-559
- Lapham, R., Lee, L.-Y., Tsugama, D., Lee, S., Mengiste, T., and Gelvin, S. B. (2018). VIP1 and its homologs are not required for agrobacterium-mediated transformation, but play a role in botrytis and salt stress responses. *Front. Plant Sci.* 9, 749. doi: 10.3389/fpls.2018.00749
- Lee, S. C., Choi, H. W., Hwang, I. S., Choi, D. S., and Hwang, B. K. (2006). Functional roles of the pepper pathogen-induced bZIP transcription factor, CABZIP1, in enhanced resistance to pathogen infection and environmental stresses. *Planta* 224, 1209–1225. doi: 10.1007/s00425-006-0302-4
- Li, Y., Meng, D., Li, M., and Cheng, L. (2016). Genome-wide identification and expression analysis of the bZIP gene family in apple (*Malus domestica*). *Tree Genet. Genomes* 12, 1–17. doi: 10.1007/s11295-016-1043-6
- Li, S., Shi, Z., Zhu, Q., Tao, L., Liang, W., and Zhao, Z. (2021). Transcriptome sequencing and differential expression analysis of seed starch accumulation in Chinese chestnut *metaxenia*. *BMC Genomics* 22, 1–14. doi: 10.1186/s12864-021-07923-5
- Li, M., Wang, R., Liu, Z., Wu, X., and Wang, J. (2019). Genome-wide identification and analysis of the WUSCHEL-related homeobox (WOX) gene family in allotetraploid *Brassica napus* reveals changes in WOX genes during polyploidization. *BMC Genomics* 20 (1), 1–19. doi: 10.1186/s12864-019-5684-3
- Liang, Y., Xia, J., Jiang, Y., Bao, Y., Chen, H., Wang, D., et al. (2022). Genome-wide identification and analysis of bZIP gene family and resistance of TaABI5 (TabZIP96) under freezing stress in wheat (*Triticum aestivum*). *Int. J. Mol. Sci.* 23 (4), 2351. doi: 10.3390/ijms23042351
- Liao, Y., Zou, H.-F., Wei, W., Hao, Y.-J., Tian, A.-G., Huang, J., et al. (2008). Soybean GmbZIP44, GmbZIP62 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic arabidopsis. *Planta* 228, 225–240. doi: 10.1007/s00425-008-0731-3
- Lin, S., Pang, L., and Zhu, M. (2012). Correlation between chestnut starch complex and glutinous characteristics. *Food Sci.* 33, 308–311.
- Liu, Y., Kong, X., Pan, J., and Li, D. (2010). VIP1: Linking agrobacterium-mediated transformation to plant immunity? *Plant Cell Rep.* 29, 805–812. doi: 10.1007/s00299-010-0870-4
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* 48 (D1), D265–D268. doi: 10.1093/nar/gkz991
- Ma, M., Chen, Q., Dong, H., Zhang, S., and Huang, X. (2021). Genome-wide identification and expression analysis of the bZIP transcription factors, and functional analysis in response to drought and cold stresses in pear (*Pyrus breschneideri*). *BMC Plant Biol.* 21 (1), 583. doi: 10.1186/s12870-021-03356-0
- MacNeill, G. J., Mehrouyan, S., Minow, M. A., Patterson, J. A., Tetlow, I. J., and Emes, M. J. (2017). Starch as a source, starch as a sink: The bifunctional role of starch in carbon allocation. *J. Exp. Bot.* 68 (16), 4433–4453. doi: 10.1093/jxb/erx291
- Meng, X., Zhao, W., Lin, R., Wang, M., and Peng, Y. (2005). Identification of a novel rice bZIP-type transcription factor gene, OsbZIP1, involved in response to infection of *Magnaporthe grisea*. *Plant Mol. Biol. Reporter.* 23, 301–302. doi: 10.1007/BF02727262
- Nakase, M., Aoki, N., Matsuda, T., and Adachi, T. (1997). Characterization of a novel rice bZIP protein which binds to the α -globulin promoter. *Plant Mol. Biol.* 33, 513–522. doi: 10.1023/A:1005784717782
- Nijhawan, A., Jain, M., Tyagi, A. K., and Khurana, J. P. (2008). Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. *Plant Physiol.* 146 (2), 333. doi: 10.1104/pp.107.112821
- Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., and Demchak, B. (2019). Cytoscape automation: Empowering workflow-based network analysis. *Genome Biol.* 20, 1–15. doi: 10.1186/s13059-019-1758-4
- Paterson, A. H., Wendel, J. F., Gundlach, H., Guo, H., Jenkins, J., Jin, D., et al. (2012). Repeated polyploidization of gossypium genomes and the evolution of spinnable cotton fibres. *Nature* 492 (7429), 423–427. doi: 10.1038/nature11798
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46 (W1), W200–W204. doi: 10.1093/nar/gky448
- Qu, J., Xu, S., Zhang, Z., Chen, G., Zhong, Y., Liu, L., et al. (2018). Evolutionary, structural and expression analysis of core genes involved in starch synthesis. *Sci. Rep.* 8 (1), 12736. doi: 10.1038/s41598-018-30411-y
- Rombauts, S., Déhais, P., Van Montagu, M., and Rouzé, P. (1999). PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* 27 (1), 295–296. doi: 10.1093/nar/27.1.295
- Schlögl, P. S., Nogueira, F. T. S., Drummond, R., Felix, J. M., De Rosa, V. E., Vicentini, R., et al. (2008). Identification of new ABA- and MEJA-activated sugarcanes bZIP genes by data mining in the SUCEST database. *Plant Cell Rep.* 27, 335–345. doi: 10.1007/s00299-007-0468-7

- Shi, L., Wang, J., Liu, Y., Ma, C., Guo, S., Lin, S., et al. (2021). Transcriptome analysis of genes involved in starch biosynthesis in developing Chinese chestnut (*Castanea mollissima* blume) seed kernels. *Sci. Rep.* 11 (1), 1–13. doi: 10.1038/s41598-021-82130-6
- Skubacz, A., Daszkowska-Golec, A., and Szarejko, I. (2016). The role and regulation of ABI5 (ABA-insensitive 5) in plant development, abiotic stress responses and phytohormone crosstalk. *Front. Plant Science.* 7, 1884. doi: 10.3389/fpls.2016.01884
- Song, Y., Luo, G., Shen, L., Yu, K., Yang, W., Li, X., et al. (2020). TubZIP28, a novel bZIP family transcription factor from triticum urartu, and TabZIP28, its homologue from triticum aestivum, enhance starch synthesis in wheat. *New Phytol.* 226 (5), 1384–1398. doi: 10.1111/nph.16435
- Subramanian, B., Gao, S., Lercher, M. J., Hu, S., and Chen, W.-H. (2019). Evolvview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* 47 (W1), W270–W275. doi: 10.1093/nar/gkz357
- Sun, W., Ma, Z., Chen, H., and Liu, M. (2019). MYB gene family in potato (*Solanum tuberosum* L.): Genome-wide identification of hormone-responsive reveals their potential functions in growth and development. *Int. J. Mol. Sci.* 20 (19), 4847. doi: 10.3390/ijms20194847
- Sun, Y., Lu, Z., Zhu, X., and Ma, H. (2020). Genomic basis of homoploid hybrid speciation within chestnut trees. *Nat. Commun.* 11 (1), 3375. doi: 10.1038/s41467-020-17111-w
- Tang, J., Wang, F., Hou, X., Wang, Z., and Huang, Z. (2014). Genome-wide fractionation and identification of WRKY transcription factors in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) reveals collinearity and their expression patterns under abiotic and biotic stresses. *Plant Mol. Biol. Reporter.* 32, 781–795. doi: 10.1007/s11105-013-0672-2
- Thalmann, M., Pazmino, D., Seung, D., Horrer, D., Nigro, A., Meier, T., et al. (2016). Regulation of leaf starch degradation by abscisic acid is important for osmotic stress tolerance in plants. *Plant Cell.* 28 (8), 1860–1878. doi: 10.1105/tpc.16.00143
- Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000). Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc. Natl. Acad. Sci.* 97 (21), 11632–11637. doi: 10.1073/pnas.190309197
- Voorrips, R. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Heredity.* 93 (1), 77–78. doi: 10.1093/jhered/93.1.77
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40 (7), e49–e4e. doi: 10.1093/nar/gkr1293
- Wang, J., Tian, S., Sun, X., Cheng, X., Duan, N., Tao, J., et al. (2020). Construction of pseudomolecules for the Chinese chestnut (*Castanea mollissima*) genome. *G3: Genes Genomes Genet.* 10 (10), 3565–3574. doi: 10.1534/g3.120.401532
- Wang, J., Xu, H., Zhu, Y., Liu, Q., and Cai, X. (2013). OsbZIP58, a basic leucine zipper transcription factor, regulates starch biosynthesis in rice endosperm. *J. Exp. Bot.* 64 (11), 3453–3466. doi: 10.1093/jxb/ert187
- Wang, S., Zhang, X., Li, B., Zhao, X., Shen, Y., and Yuan, Z. (2022). Genome-wide identification and characterization of bZIP gene family and cloning of candidate genes for anthocyanin biosynthesis in pomegranate (*Punica granatum*). *BMC Plant Biol.* 22 (1), 1–18. doi: 10.1186/s12870-022-03560-6
- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinf.* 8 (1), 77–80. doi: 10.1016/S1672-0229(10)60008-3
- Wei, K., Chen, J., Wang, Y., Chen, Y., Chen, S., Lin, Y., et al. (2012). Genome-wide analysis of bZIP-encoding genes in maize. *DNA Res.* 19 (6), 463–476. doi: 10.1093/dnares/dss026
- Wu, A., Hao, P., Wei, H., Sun, H., Cheng, S., Chen, P., et al. (2019). Genome-wide identification and characterization of glycosyltransferase family 47 in cotton. *Front. Genet.* 10, 824. doi: 10.3389/fgene.2019.00824
- Xing, Y., Liu, Y., Zhang, Q., Nie, X., Sun, Y., Zhang, Z., et al. (2019). Hybrid *de novo* genome assembly of Chinese chestnut (*Castanea mollissima*). *Gigascience* 8 (9), giz112. doi: 10.1093/gigascience/giz112
- Yang, G., Chao, D., Ming, Z., and Xia, J. (2019). A simple method to detect the inhibition of transcription factor-DNA binding due to protein-protein interactions *In vivo*. *Genes* 10 (9), 684. doi: 10.3390/genes10090684
- Zhang, X., Wollenweber, B., Jiang, D., Liu, F., and Zhao, J. (2008). Water deficits and heat shock effects on photosynthesis of a transgenic arabidopsis thaliana constitutively expressing ABP9, a bZIP transcription factor. *J. Exp. Bot.* 59 (4), 839–848. doi: 10.1093/jxb/erm364