Genome-wide identification of the bZIP transcription factor genes related to starch synthesis in barley (*Hordeum vulgare* L.)

Journal:	<i>Genome</i>
Manuscript ID	gen-2020-0195.R1
Manuscript Type:	Article
Date Submitted by the Author:	21-Apr-2021
Complete List of Authors:	<p>Zhong, XiaoJuan; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Feng, Xiuqin; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Li, Yulong; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Guzmán, Carlos; Universidad de Córdoba Escuela Técnica Superior de Ingeniería Agronómica y de Montes, Departamento de Genética</p> <p>Lin, Na; Yibin University, College of Sichuan Tea</p> <p>Xu, Qiang; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Zhang, Yazhou; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Tang, Huaping; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Qi, Peng Fei; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Deng, Mei; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Ma, Jian; Sichuan Agricultural University - Chengdu Campus, Triticeae Research Institute</p> <p>Wang, Ji Rui; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Chen, Guo Yue; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Lan, Xiu Jin; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Wei, Yu Ming; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Zheng, Youliang; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Jiang, Qiantao; Sichuan Agricultural University, Triticeae Research Institute</p>
Keyword:	Barley, Basic leucine zipper, Transcription factors, Gene interaction, Starch synthesis
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

**Genome-wide identification of the bZIP transcription factor genes
related to starch synthesis in barley (*Hordeum vulgare* L.)**

Xiaojuan Zhong^{1,2+}, Xiuqin Feng^{1,2+}, Yulong Li^{1,2}, Carlos Guzmán³, Na Lin⁴, Qiang Xu^{1,2},
Yazhou Zhang^{1,2}, Huaping Tang^{1,2}, Pengfei Qi^{1,2}, Mei Deng^{1,2}, Jian Ma^{1,2}, Jirui Wang^{1,2},
Guoyue Chen^{1,2}, Xiujin Lan^{1,2}, Yuming Wei^{1,2}, Youliang Zheng^{1,2}, Qiantao Jiang^{1,2*}

¹*State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China,
Sichuan Agricultural University, Chengdu, Sichuan, 611130, China*

²*Triticeae Research Institute, Sichuan Agricultural University, Chengdu, Sichuan
611130, China*

³*Departamento de Genética, Escuela Técnica Superior de Ingeniería Agronómica y de
Montes, Edificio Gregor Mendel, Campus de Rabanales, Universidad de Córdoba,
Cordoba, 14071, Spain*

⁴*College of Sichuan Tea, Yibin University, Yibin, Sichuan 644000, China*

The first and second authors contributed equally to this paper.

*Corresponding author

Dr. Qiantao Jiang

E-mail: qiantaojiang@sicau.edu.cn

Tel: +86 28 86290958; Fax: +86 28 82650350

20 **Abstract**

21 The basic leucine zipper (bZIP) family genes encode transcription factors which play
22 key roles in plant growth and development. In this work, a total of 92 *HvbZIP* genes
23 were identified and compared with previous studies, using the recent released barley
24 genome data. Two novel genes were characterized in this study, and some mis-
25 annotated and duplicated genes in previous studies have been corrected.
26 Phylogenetic analysis results showed that 92 *HvbZIP* genes were classified into 10
27 groups and three unknown groups. The gene structure and motif distribution of the
28 three unknown groups implied that the genes of the three groups may be functionally
29 different. Expression profiling indicated that *HvbZIP* genes exhibited different patterns
30 of spatial and temporal expression. Using qRT-PCR, more than 10 *HvbZIP* genes were
31 identified with expression patterns similar to those of starch synthase genes in barley.
32 Yeast one-hybrid analysis found that two of the *HvbZIP* genes exhibited in vitro binding
33 activity to the promoter of *HvAGP-S*. The two *HvbZIP* genes may be candidate genes
34 for further study, to explore the mechanism by which they regulate the synthesis of
35 barley starch.

36 **Keywords**

37 Barley, Basic leucine zipper, Transcription factors, Gene interaction, Starch synthesis

39 Introduction

40 Transcription factors (TFs) play key roles in the regulation of gene transcription,
41 by binding to cis-acting elements of their target genes. The basic region/leucine zipper
42 (bZIP) TFs are one of the largest and most conserved gene families in plants. The bZIP
43 TFs are named and classified according to their highly conserved DNA-binding
44 domains: a basic region that binds DNA, and a leucine zipper dimerization motif (Hurst
45 1995). The bZIP domain consists of 60–80 amino acids, of which the highly conserved
46 basic region contains about 16 amino acid residues with an N-x7-R/K motif, and is
47 responsible for nuclear localization and DNA binding. The less conserved leucine
48 zipper motif is composed of a heptad repeat of leucines or other bulky hydrophobic
49 amino acids, creating an amphipathic helix. The leucine zipper motif is primarily
50 responsible for forming homo- and heterodimers (Hurst 1995). In plants, bZIP TFs bind
51 preferentially to the A-box (TACGTA), C-box (GACGTC), and G-box (CACGTG) which
52 have core ACGT sequences (Izawa et al. 1993). In *Arabidopsis*, the bZIP gene family is
53 categorized into 10 groups according to the sequence similarity of the basic region
54 and the presence of additional conserved motifs. Member of a same group share a
55 similar DNA-binding basic region (Jakoby et al. 2002).

56 Although members of the bZIP TFs family have been characterized in many plant
57 species: 125 in maize (Wei et al. 2012), 187 in wheat (Li et al. 2015), 89 in rice
58 (Nijhawan et al. 2008), 75 in *Arabidopsis* (Jakoby et al. 2002), 92 in sorghum (Wang et
59 al. 2011), 55 in grapevine (Liu et al. 2014) and 64 in cucumber (Baloglu et al. 2014),
60 only a small part of bZIP genes are functionally characterized in plants. These studies
61 show that bZIP TFs are involved in regulating diverse developmental processes, such

as seed maturation (Li et al. 2018; Song 2020) and flower development (Abe et al. 2005; Muszynski et al. 2006), as well as multiple biotic and abiotic stress responses (Gupta et al. 2012; Lim et al. 2015; Stracke et al. 2010; Cao et al. 2012).

Barley (*Hordeum vulgare* L.) is the fourth most abundant cereal in both area and tonnage harvested. It is widely used for animal feed and beer making (Mayer et al. 2012). The International Barley Sequencing Consortium (IBSC) released the genome sequencing map of barley cultivar Morex, a US spring six-row malting barley, for the first time in 2012 (Mayer et al. 2012). In 2017, the IBSC combined hierarchical shotgun sequencing of bacterial artificial chromosomes (BAC) with novel technologies such as optical mapping and chromosome-scale scaffolding with chromosome conformation capture sequencing (Hi-C) to assemble a highly contiguous reference genome sequence for Morex (Mascher et al. 2017). In 2019, Monat et al. (2019) used TRITEX, an open-source computational workflow, to generate an improved annotated reference genome assembly for Morex as an important resource for the barley research community (Monat et al. 2019). The improvement of barley genome provides a good opportunity for us to make full use of barley germplasm resources and carry out the cloning and functional identification of unknown genes. In previous studies, Pourabed et al. (2014) and Li et al. (2015) identified 65 and 107 HvZIP members from the barley genome (2012 version), respectively (Pourabed et al. 2014; Li et al. 2015). The different results of these two studies might be caused by incomplete splicing and mis-annotation of the early barley genome sequence. The second version of the reference genome provided the basis for re-analysis of the ZIP family genes in barley.

Starch is the main reserve substance of cereal crops, and has been an important dietary energy source for humans. The type and physicochemical characteristics of starch directly affect the quality of crops (Hebelstrup et al. 2017; Hsieh et al. 2019). Starch synthesis in cereals require a suite of well-characterized enzymes, including ADP-glucose pyrophosphorylase (AGPase), granule bound starch synthase (GBSS), starch synthase (SS), starch branching enzyme (SBE) and starch de-branching enzyme (DBE) (Hannah and James 2008; James et al. 2003; Jeon et al. 2010). The first step of starch biosynthesis is the conversion of glucose-1-phosphate into adenosine diphosphoglucose (ADP-glucose), and is catalyzed by AGPase (Atta et al. 2014). GBSS encoded by *waxy* gene is responsible for amylose synthesis by catalyzing the transfer of ADP-glucose to linear chains of α -(1,4)-linked glucose residues (Guzmán and Alvarez 2016). SS elongates glucan chains using ADP-glucose, and then SBE and DBE catalyze the branched structure to form amylopectin (Sonnewald and Kossmann 2013). The isoamylase 1 (ISA1) is one of the starch debranching enzyme and plays an indispensable role in amylopectin biosynthesis by trimming the glucan cluster (Ball et al. 1996).

There have been some studies into the transcriptional regulation of the starch synthesis pathway in plants, which found that TFs are involved in starch synthesis. In rice, the transcription factor genes such as *Rice Starch Regulator 1 (RSR1)*, *OsbZIP58* and *OsbZIP76* were proved to regulate starch synthesis directly or indirectly (Fu and Xue 2010; Niu et al. 2020; Wang et al. 2013). In maize, *ZmNAC36*, *ZmbZIP91*, *ZmbZIP22*, *Opaque 2*, *ZmEREB156*, *ZmNAC128*, and *ZmNAC130* are regulators of starch synthesis (Chen et al. 2016; Dong et al. 2019; Huang et al. 2016; Zhang et al. 2014; Zhang et al.

2016; Zhang et al. 2019). In wheat, *TubZIP28* and *TabZIP28* are transcriptional activators of starch synthesis, while *TaNAC019-A1* and *TaRSR1*, a gene homologous to *OsRSR1*, negatively regulate the expression of many starch synthesis genes (Liu et al. 2016; Liu et al. 2020; Song 2020). Unlike rice, maize and wheat, just WRKY family transcription factor gene, *SUSIBA2* (sugar signaling in barley), was identified in barley (Sun et al. 2003). The product of *SUSIBA2* gene can bind directly to the promoter of *ISA1* (encoding isoamylase1) gene to regulate its expression, thereby influencing starch synthesis in barley. Previous studies had shown that bZIP family members regulating starch synthesis could be identified in rice, maize, and wheat. However, no regulator of starch synthesis in the bZIP family was identified in barley.

Accurate analysis of bZIP gene family members is valuable for screening for bZIP genes associated with starch synthesis. In the work presented here, all members of the bZIP family were identified using the most recent assembly of the Morex barley genome. The differences between previous studies and the present study were analyzed, and the structural features, phylogenetic relationships, chromosomal location, and expression patterns of all bZIP genes identified were characterized. We reported the expression pattern of *HvbZIP* genes in different tissues and developmental stages with a focus on the coexpression in relation to the starch synthase gene. Our results provided more accurate and a new insight into the bZIP family of barley and identified some candidate *HvbZIP* genes might play a key role in the regulation of starch synthesis in barley.

Materials and Methods

Identification of bZIP family transcription factor genes in Morex

Two methods were used to comprehensively identify the bZIP domain-containing sequences in barley. The first method involved searching for the bZIP gene family members in the Morex genome on IPK (<https://apex.ipk-gatersleben.de/apex/f?p=284:49:::NO:::>) by inputting the keyword “bZIP” (in Morex v2 Gene Models-2019). Secondly, the genome sequences including DNA fasta and GFF3 files of Morex were obtained from the e!DAL database (<https://doi.ipk-gatersleben.de/DOI/83e8e186-dc4b-47f7-a820-28ad37cb176b/64687610-b20c-4702-9698-7dc2401a80e5/0>). The bZIP transcription factor protein sequences of *Arabidopsis thaliana* and *Hordeum vulgare* downloaded from the PlantTFDB v5.0 database (<http://planttfdb.gao-lab.org>) were used as the query to do the first BLAST search (e-value $\leq 1e^{-10}$) in the Morex genomes, with the biological data handling toolkit TBtools (Chen et al. 2020). Redundant sequences were removed manually. Based on the TBtools searches, the protein sequences obtained from the first BLAST search were extracted and used as the query for second BLAST in the NCBI-BLASTP searches (e-value $\leq 1e^{-10}$). All the alignment results were downloaded and the non-bZIP members were deleted manually. Proteins thus identified from the two BLAST searches were further searched for presence of conserved bZIP domain, using the NCBI Conserved Domain Database server (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) (Marchler-Bauer et al. 2011). Proteins confirmed by domain searches were regarded as putative bZIP TFs.

Comparison of *HvbZIP* genes identified in new and old barley genome versions

The sequences of *HvbZIP* genes identified in previous studies (Li et al. 2015; Pourabed et al. 2014) were searched and downloaded from the Morex WGS Gene Models (2012) in IPK database (http://apex.ipk-gatersleben.de/apex/f?p=284:27::::RP:P27_GENE_CHOICE:1). To compare the result with the previous studies, the *bZIP* genes identified from previous studies were used as the query sequences to do BLAST search in the Morex V2 (all Morex V2 in this article represented the barley genome published in 2019) database. The gene with the Expect=0 value was the same gene with the query sequence. Secondly, the *bZIP* genes identified from the Morex V2 genome were used to do BLAST search (e-value $\leq 1e^{-10}$) in the IPK HC_genes_CDS_Seq_2012, LC_genes_CDS_Seq_2012 and full length cDNA databases (http://webblast.ipk-gatersleben.de/barley_ibsc/), respectively. The different gene members of *HvbZIP* family between new and old barley genome versions were able to be determined through BLAST. The Golden promise genome (<http://ics.hutton.ac.uk/gmapper/index.html>) were used to do BLAST search for verifying the different genes between previous studies and our search.

Gene structure and phylogenetic analysis

To construct an exon/intron map, the coding sequence of each *HvbZIP* gene was aligned with its genomic sequence using TBtools. To identify the evolutionary relationships among the barley, rice and *Arabidopsis thaliana* *bZIP* proteins, all the amino acid sequences were aligned using the ClustalW program implemented in MEGAX (<http://www.megasoftware.net/>). The phylogenetic tree was constructed using the neighbor-joining method based on the JTT matrix-based model with 1,000 bootstrap replications. The *bZIP* members of rice and *Arabidopsis* were obtained from

the previous studies (Nijhawan et al. 2008; Jakoby et al. 2002) and downloaded from the RAP (The Rice Annotation Project, <https://rapdb.dna.affrc.go.jp/>) and TAIR (The *Arabidopsis* Information Resource, <https://www.arabidopsis.org/>) databases, respectively. The amino acid and coding sequences of *HvbZIP*, *OsbZIP* and *AtbZIP* genes used for phylogenetic analysis were listed in the supplementary text files Text S1 and Text S2.

Conserved motif analysis and localization of bZIP genes on Morex chromosomes

The motifs of the *HvbZIP* protein sequences were searched using the online software MEME 5.1.1 (<http://meme-suite.org/tools/meme>), with the following parameters: number of repetitions—any, maximum number of motifs—20, optimum motif width ≥ 6 and ≤ 200 (Ma et al. 2017). The chromosomal locations of all bZIP genes identified were analyzed by mapping the sequences back to the corresponding genome annotation GFF3 file using TBtools (Chen et al. 2020).

Expression analysis

Firstly, all the CDS sequences of *HvbZIP* genes were compared to the transcriptomic database in the Barley Reference Transcript (BaRTv1.0) Dataset (http://ics.hutton.ac.uk/barleyrtd/blast_page.html) (Rapazote-Flores et al. 2019). The RNA-seq data of 13 tissues (Root (10cm seedlings), Root 2 (4 week seedling), Shoot (10cm seedlings), Rachis (5 weeks pa), Senescing leaf (2 months), Tillers (3rd internode), Inflorescence-1 (0.5cm), Inflorescence-2 (1-1.5cm), Embryo (germinating), Palea (6 weeks pa), Epidermis (4 weeks), Grain (5 DPA) and Grain (15 DPA)) of Morex were retrieved from The James Hutton Institute

(http://ics.hutton.ac.uk/barleyrtd/about_material.html), and the log₂ (TPM) (TPM, transcripts per million) value of each bZIP gene was visualized as a heat map with a blue-yellow-red gradient.

Plant growth, RNA extraction and quantitative reverse transcription PCR analysis

The expression levels of key starch synthase genes in barley and 15 *HvbZIP* genes with a high expression level in the endosperm were quantified using quantitative reverse transcription PCR (qRT-PCR) to verify the RNA-seq data and screen out candidate *HvbZIP* genes. Barley accession “Golden Promise” were grown in a phytotron chamber under 16 h light/8 h dark and 24°C day/18°C night temperatures. Grains were harvested at 5, 10, 15, 20, and 25 days post-anthesis (DPA), transferred immediately into liquid nitrogen, and stored at -80°C until RNA extraction. Total RNA was isolated using Plant RNA kits (Biofit, Chengdu, China) according to the manufacturer’s protocol. Three biological replicates were prepared and tested for each developmental stage. First-strand cDNA was synthesized using PrimeScript™ RT reagent kits with gDNA Eraser (TaKaRa, Dalian, China), and qRT-PCR was carried out with SYBR® Premix Ex Taq™ II (TaKaRa) on a CFX 96 Real-Time System (Bio-Rad, Hercules, USA). The CFX Manager software (Bio-Rad, Hercules, USA) was used to analyze the qRT-PCR data and to calculate relative expression using the 2^{-ΔΔCt} method. The barley β-actin and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) genes were used as the internal reference genes to normalize the relative expression of candidate genes.

Yeast one-hybrid assay

The yeast one-hybrid assay was conducted according to the protocol of the

Matchmaker Gold Yeast One-Hybrid Library Screening System (Clontech, Palo Alto, CA). Three copies of the G-box fragments from the *HvAGP-S* gene promoter were subcloned into the pAbAi vector to produce a bait construct. The construct was then linearized by digestion with *BstBI*, and integrated into the *URA3–52* locus of the Y1HGold yeast genome to generate a Y1H bait strain. The coding sequences of *HvbZIP10* and *HvbZIP70* were cloned into the pGADT7 vector to produce the pGADT7-TFs construct. These constructs or empty vector were separately transformed into the Y1H bait strain and selected on a synthetic dropout (SD)/-Leu plate containing 100 ng/mL AbA (Aureobasidin A).

Results

Identification of the bZIP genes in Morex

The characterization of *HvbZIP* family members is more accurate using the recently released barley genome data (Morex V2) (Monat et al. 2019), than it was previously (Pourabed et al. 2014). We identified a total of 92 non-redundant *HvbZIP* genes in Morex, based on a genome-wide search and confirmation (Table 1). Since there was no standard nomenclature, each gene was named with a generic name (*HvbZIP01*-*HvbZIP92*) based on their order of original sequence ID. The sequence analysis indicated that all the *HvbZIP* genes possessed the bZIP domain, and the amino acid (aa) length of these *HvbZIP* genes ranged from 132 (*HvbZIP18*) to 881 (*HvbZIP62*), with an average size of 329 aa (Table 1). All the characterized bZIP genes except for *HvbZIP90* were extracted based on protein sequences of high-confidence barley genes. The *HvbZIP02*, *HvbZIP27* and *HvbZIP62* genes identified in barley (including in this

study and previous studies) did not match any bZIP family genes in *Arabidopsis* or rice, indicating that these three genes may be specific *bZIP* genes in barley. Chromosomal distribution analysis showed that chromosome 5 contained 22 bZIP genes, 23.9% of the total, and a higher number of bZIP genes than the other chromosomes, which contained from 10 to 16 bZIP genes (Fig. S1).

Comparison of *HvbZIP* genes identified in different barley genome versions

In this study, we found that the previous studies (Pourabed et al. 2014; Li et al. 2015) listed duplicate accessions for the same gene as separate genes, overestimated the size of the family and misnamed genes through the comparative analysis of *HvbZIP* genes in different barley genome versions. Eight-nine and 107 *HvbZIP* family genes were identified in two different studies from the barley genome (2012 version). To further verify our result, the *HvbZIP* genes identified in the Morex V2 genome were compared with the *HvbZIP* genes identified previously (Table S1 and Table S2). Compared with previous studies, two different *HvbZIP* genes (*HvbZIP57* and *HvbZIP74*) were found in the Morex V2 genome. Three sequences (*AK361769*, *MLOC_15674.1*, and *MLOC_20530*) identified in the previous studies (including Pourabed et al. 2014 and Li et al. 2015) were not found in the Morex V2 genome. *AK361769* was not able to match any genes in the Morex V2 and Golden promise genome. The conserved domain of *MLOC_15674* was searched in the NCBI-CDD database and the result showed that the *MLOC_15674* did not contain the bZIP domain. Although *MLOC_20530* contained a bZIP domain and could be matched in Golden promise genome, the sequence of *MLOC_20530* was too short to match any genes in the genome of Morex and Golden promise. In addition, 14 *HvbZIP* genes identified in

Morex V2 were same or highly similar with two or three sequences identified from the 2012 genome version. The results of multiple sequence alignment (Fig. S2) showed that this phenomenon might be caused by incorrect splicing of the 2012 genome version. The results of comparison with the previous studies showed that the number of *HvbZIP* genes identified from the Morex V2 was larger, the gene sequences were more complete, and the results were more accurate than the previous studies.

Phylogenetic analysis of *HvbZIP* genes

The *Arabidopsis thaliana* genome contains typical bZIP family genes, which are well classified according to the structural characteristics of protein sequences (Jakoby et al. 2002). Therefore, a total of 89 *OsbZIP* and 75 *AtbZIP* genes were selected as a standard to classify the types of *HvbZIP* genes in barley. The phylogenetic analysis was performed with all identified *HvbZIP*, *OsbZIP* and *AtbZIP* proteins (Fig. S3). *HvbZIP* were distributed in groups A, B, C, D, E, F, G, H, I, S and three unknown groups (Group U1, Group U2 and Group U3). The two *HvbZIP* members of Group U1 were clustered with the *OsbZIP80* which was not classified any group in rice. The Group U2 and Group U3 contained three *HvbZIP* genes that were not clustered with any rice and *Arabidopsis* bZIP gene, respectively. Groups A, D, I and S contained 17, 16, 11 and 12 *HvbZIP* genes, respectively, and had more member genes than the other groups. The remaining genes were distributed in nine groups with varied numbers of genes. Group B and U2 were the smallest groups which contained only two genes, respectively. With the phylogenetic tree, we found that three members of Group A, one member of Group H and the Group U of *Arabidopsis* were not clustered together. And the Group S was divided into three subgroups by the members in Group C and Group G. A similar

situation also appeared in the phylogenetic analysis of wheat bZIP gene family (Li et al. 2015). We speculated that the reason for this phenomenon may be that the gene classification of *Arabidopsis* is based on the sequence similarities of base regions in AtbZIP proteins, while the phylogenetic tree of this study was made on the basis of all amino acid sequence alignment results.

Gene and protein structural features of the barley bZIP genes

Combined with the data from the phylogenetic analysis (Fig. S3), the gene structural characteristics of *HvbZIP* genes were analyzed with the comparison of *OsbZIP* and *AtbZIP* genes (Fig. 1). Of the 92 genes, 75 possessed introns, with the number varying from 1 to 12, while the remaining 17 genes had no introns (Fig. 1). Most genes in each cluster group had highly similar gene structures, and intron and exon lengths and distributions. However, some close gene pairs did differ in intron/exon arrangements. For example, the Group S genes *HvbZIP55* and *HvbZIP57* had two exons, whereas its close homologs had one, and the same situation were existed in same group of rice and *Arabidopsis*. The genes in Group D had more introns than the other groups. The characterized barley bZIP family members possessed a total of 10 conserved motifs, of which only motif 1 was presumably the bZIP core motif, was shared by all members of the family (Fig. 2 and Fig. S4). The position of the bZIP domain was different among the *HvbZIP* family members. *HvbZIPs* in the same group shared similar conserved motif composition, except for Group A which contained two types of motif composition and Group U3. Some *HvbZIPs* only contained the bZIP domain (Groups B, C, G, H, S, U1 and U2), but there were differences in gene structures among these groups (Fig 1). Group D had the largest number of motif types. We also found that the motif 4 and bZIP

domains of Group U2 overlapped each other. The function of these given motifs were mostly unknown except for the motif 3 which regarded as the DOG1 (Delay Of Germination 1) domain. This domain of plant proteins appears to be a highly specific controller of seed dormancy (Bentsink et al. 2006).

Expression profiles of *HvbZIP* genes during vegetative and reproductive development

To check whether the identified *HvbZIP* genes were able to expressed normally in barley, they were compared to the Barley Reference Transcript (BaRTv1.0) Dataset. The result showed that each of 92 *HvbZIP* genes can well match to a normal transcript, proved that the 92 *HvbZIP* genes were true genes and could normally perform function. To understand the temporal and spatial transcription patterns of the *HvbZIP* genes in barley, data from publicly available RNA-seq databases were used to examine the expression profiles of *HvbZIP* genes in different tissues (Fig. 3 and Table S3). The *HvbZIP* genes were expressed in at least one organ, and could be divided into five groups according to their expression patterns (Fig. 3). The transcripts of group I, containing 31 *HvbZIP* genes, were widely involved in various tissues of barley growth and development. The genes of group II were primarily expressed in inflorescences, with a total of 12 genes, and suggested that these genes were mainly involved in inflorescence development. Almost all the genes of group III were weakly expressed in grains. However, the genes of group IV were specifically expressed in grains with high expression. The remaining 27 genes belonging to group V were primarily expressed in vegetative tissues except for four genes (*HvbZIP12*, *HvbZIP70*, *HvbZIP15* and *HvbZIP65*) which were also highly expressed in grains.

Two bZIP genes with high expression in endosperm bound specifically to the promoters of *HvAGP-S* in yeast

In order to screen out the candidate *HvbZIP* genes involved in starch synthesis, we selected 15 genes highly expressed in the endosperm from different groups (Fig. 3) and detected the expression of these genes and some key starch synthase genes at different developmental stages (5 DPA to 25 DPA), using qRT-PCR (Fig. 4a and b). More than 10 *HvbZIP* genes were highly expressed in the middle stage of grain development, and their expression patterns were similar to those of the starch synthase gene and the starch accumulation rate during grain filling stages (Fig. 4a, b). Therefore, we speculated that these genes might be involved in starch synthesis in barley, and two *HvbZIP* members, *HvbZIP10* and *HvbZIP70*, were selected as candidate genes for further study.

Yeast one-hybrid assays were conducted to investigate whether *HvbZIP10* and *HvbZIP70* could interact with the starch synthetase genes. Both *HvbZIP10* and *HvbZIP70* could bind to the promoter fragments containing the cis-element 5'-CACGTC-3' of *HvAGP-S* (Fig. 4c). These results implied that both *HvbZIP10* and *HvbZIP70* had DNA-binding activity, and might directly regulate the expression of the target gene *HvAGP-S*.

Discussion

Starch accumulation in barley endosperm requires coordinated regulation of gene expression. The bZIP family genes regulating starch synthesis have been identified in wheat, maize and rice (Chen et al. 2016; Dong et al. 2019; Song 2020; Wang et al.

2013), while no *bZIP* genes involved in starch synthesis have been identified in barley. Accurate identification and analysis of *bZIP* family genes of barley could provide a reliable basis for screening for candidate *HvbZIP* genes involved in the regulation of starch synthesis. In previous reports (Li et al. 2015; Pourabed et al. 2014), 89 and 107 *bZIP* genes were identified in barley. However, these works were carried out based on the obsolete version of the barley genome full-length cDNA, HC_genes_CDS_Seq, and LC_genes_CDS_Seq-2012 sequence data. Even in the BAC-based reference genome data (Beier et al. 2017), there were some gaps in the physical map or failed BAC assemblies that might contain important genes. The new version (Barley all CDS Morex v2.0 2019) of the barley genome sequence was recently released, and provides more comprehensive and reliable gene information than the V1 annotation, making it possible to conduct a more accurate characterization of *bZIP* family (Monat et al. 2019). In this study, we identified 92 *bZIP*s from the new barley genome data, and two (*HvbZIP57* and *HvbZIP74*) of these genes were not identified in the previous studies which might be because the Morex V2 genome data had a more complete gene sequence and better annotation. *AK361769* and *MLOC_20530* identified in previous studies could not match any gene in Morex V2 and Golden promise genome database. The BLAST result of *AK361769* in NCBI database showed that the *AK361769* sequence is 100% similar to an *Rhopalosiphum maidis* gene and is likely a contaminant mRNA from *Rhopalosiphum maidis* that was misannotated as a *Hordeum* sequence. And we speculated that the *MLOC_20530* gene sequences might be caused by incorrect sequence splicing or mis-annotation in the old genome. In addition, the results of multiple sequence alignment showed that there were some repeat genes in previous

378 studies. Therefore, our results are more reliable and provide a complete
379 understanding on barley *bZIP* genes.

380 *Arabidopsis* genome contain a rather ancient diversity of *bZIP* genes that are
381 representative of other flowering plants (Jakoby et al. 2002). The *AtbZIP* genes were
382 divided into 10 groups according to the structural features and functional information
383 of *bZIPs*. This classification method has become the classification standard of the *bZIP*
384 family members, and is applied to other plants (Baloglu et al. 2014; Li et al. 2015; Liu
385 et al. 2014; Wang et al. 2011; Wei et al. 2012). In our study, the *bZIP* gene number of
386 Groups A, D, I, and S were more than 10, higher than those of other groups. This
387 number is consistent with that of *Arabidopsis* (Jakoby et al. 2002), but different from
388 those identified from previous barley data (Pourabed et al. 2014). We identified two
389 members more in Group A and ten more in Group S, than were previously reported
390 by Pourabed et al. (2014). In the *Arabidopsis* genome, three *AtbZIP* genes (*At1g42990*,
391 *At1g19490*, and *At5g07160*) did not fit into any group (Jakoby et al. 2002). In this study,
392 based on the classification of the *AtbZIP* genes, we identified eight *HvbZIP* genes
393 belonging to three different clades, Group U1, Group U2 and Group U3, which could
394 not be fitted into any group (Fig. S3). According to the result of gene structure, motifs
395 and *bZIP* domain distribution, the gene structures of Group U2 were highly similar and
396 complete, and the *bZIP* domain of two members in Group U2 overlapped with that of
397 motif 4 (Fig. 2). This situation was also consistent with the classification criteria of the
398 *bZIP* gene family in *A. thaliana*. The classification was not based purely on phylogeny,
399 but emphasized conserved motifs that reflect functional similarities. This approach
400 should aid in determining specific functions for each *bZIP* (Jakoby et al. 2002). The

genes in Group U2 were matched to the group U bZIP genes of Pourabed et.al (2014) study. In the Pourabed et.al (2014) study, the conserved –10 arginine in the DNA-binding domain of group U bZIP genes was replaced with a hydrophobic isoleucine, which might affect their DNA-binding specificity. It was speculated that the U group members either have a different DNA-binding specificity from other plant bZIP factors, or they do not bind to DNA (Pourabed et al. 2014).

Expression profiling indicated that the genes in the bZIP family played important roles in plant growth and development. A previous study identified the one *HvbZIP* gene with the highest expression level during seed germination, development, and other tissues and treatments, suggesting that it could be a housekeeping gene in barley (Pourabed et al. 2014). However, we did not find any gene that was highly expressed in all tissues. Our expression results indicated that *HvbZIP* genes were expressed in at least one barley tissue, and showed tissue-specific differential expression. Most genes of group IV were specifically expressed in grains, which indicated these genes may be mainly involved in the development of grain and the accumulation of storage substances.

Previous work found that *bZIP* TF genes can regulate the expression of genes involved in starch synthesis. For example, Rice Starch Regulator1 (*RSR1*) for starch biosynthesis was identified using gene coexpression analysis in rice (Fu and Xue 2010). Our results indicated that more than 10 *HvbZIP* genes had an expression pattern similar to that of the starch synthetase genes (Fig. 4). Of the starch synthetase genes, *AGPase* is responsible for the first key step of starch synthesis, and is the rate-limiting

enzyme (Preiss 1993). AGPase is a heterotetramer composed of two large subunits (AGPL) and two small subunits (AGPS) in higher plants. AGPL and AGPS have complementary roles in AGPase function (Qu et al. 2018). In wheat, *TabZIP28* and *TabZIP28* which are two orthologous genes from two related species *Triticum urartu* and *Triticum aestivum* regulate starch synthesis by binding to the promoter of cytosolic *AGPase* and enhancing its transcription and activity (Song 2020). In our study, the yeast one-hybrid analysis showed that two *HvbZIP* genes, *HvbZIP10* and *HvbZIP70* which were not orthologous to *TabZIP28*, could bind to the promoter of *HvAGP-S* (Fig.4c). *HvbZIP10* is the *BLZ2* gene identified in previous study in barley (*Hordeum vulgare*) cv. *Bomi* (Onate et al. 1999) and homologous to the *Opaque2* gene of maize (Zhang et al. 2016). The *BLZ2* is an endosperm-specific protein and as an important transcriptional activator of seed storage protein genes containing the GLM (GCN4-like motif) in their promoters (Onate et al. 1999). However, the study on the involvement of *BLZ2* gene in starch synthesis has not been reported. In maize, the *Opaque2* (*O2*) encodes an endosperm-specific bZIP transcription factor, and primarily regulates the expression of α - and β -zein genes by recognizing the *O2* box in their promoters (Holdsworth et al. 1995). The endosperm-specific transcription factor *O2* could interact with other TFs or proteins to regulate starch synthesis (Zhang et al. 2016). The other gene, *HvbZIP70*, is the homologous gene of *OsbZIP20*. In rice, yeast one-hybrid analysis indicated that *OsbZIP20* could bind to the promoters of both *Waxy* and *SBE1* (Wang et al. 2013). These studies suggested that these two candidate genes, as homologs of *Opaque2* and *OsbZIP20*, were highly likely to be involved in the regulation

of starch synthesis in barley. The two candidate *HvbZIP* genes might be the first transcriptional regulators of starch synthesis identified in barley, and could be valuable subjects for further study.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2017YFD0100900), the Key Research and Development Program of Sichuan Province (2018NZDZX0002), the International Science & Technology Cooperation project of Sichuan Province, China (2020YFH0140), the International Science & Technology Cooperation project of Chengdu, Sichuan Province, China (2019-GH02-00078-HZ). Carlos Guzmán gratefully acknowledges the European Social Fund and the Spanish State Research Agency (Ministry of Science, Innovation and Universities) for financial funding through the Ramon y Cajal Program (RYC-2017-21891).

Conflict of Interest: We declare that we do not have any conflicts of interest.

Author contribution

XJZ carried out analysis of gene members, qRT-PCR and wrote this article. XQF and YLL downloaded sequences and did the yeast one-hybrid assay. CG and NL analyzed the expression data. HPT contributed to the material planting. QTJ conceived and designed the experiments. QX, YZZ, PFQ, MD, JM, JRW, GYC, XJL, YMW, and YLZ partially participated in its design and revised the manuscript. All authors read and approved the final manuscript.

References

- Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., Araki, T. 2005. FD, a bZIP Protein Mediating Signals From the Floral Pathway Integrator FT at the Shoot Apex. *Science* **309**: 1052-1056.
- Atta, S., Ayele, B.T., Fouad, D., and Permyakov, E.A. 2014. Biochemical and Molecular Characterization of Barley Plastidial ADP-Glucose Transporter (HvBT1). *PloS one* **9**(6): e98524.
- Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., Colonna, P. and Preiss, J. 1996. From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. *Cell* **86**: 349–352. doi: 10.1016/S0092-8674(00)80107-5.
- Baloglu, M.C., Eldem, V., Hajyzadeh, M., and Unver, T. 2014. Genome-wide analysis of the bZIP transcription factors in cucumber. *PloS one* **9**(4): e96014. doi:10.1371/journal.pone.0096014.
- Beier, S., Himmelbach, A., Colmsee, C., Zhang, X.Q., Barrero, R.A., Zhang, Q., Li, L., Bayer, M., Bolser, D., Taudien, S., Groth, M., Felder, M., Hastie, A., Simkova, H., Stankova, H., Vrana, J., Chan, S., Munoz-Amatriain, M., Ounit, R., Wanamaker, S., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Sampath, D., Heavens, D., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J.K., Tan, C., Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard, M.I., Houben, A., Dolezel, J., Ayling, S., Lonardi, S., Langridge, P., Muehlbauer, G.J., Kersey, P., Clark, M.D., Caccamo, M., Schulman, A.H., Platzer, M., Close, T.J., Hansson, M., Zhang, G., Braumann, I., Li, C., Waugh, R., Scholz, U., Stein, N., and Mascher, M. 2017.

- 486 Construction of a map-based reference genome sequence for barley, *Hordeum vulgare* L. Sci
487 Data **4**: 170044. doi:10.1038/sdata.2017.44.
- 488 Bentsink, L., Jowett, J., Hanhart, C. J., Koornneef, M. 2006. Cloning of DOG1, a quantitative trait locus
489 controlling seed dormancy in Arabidopsis. Proceedings of the National Academy of Sciences
490 of the United States of America **103**(45): 17042-7. doi: 10.1073/pnas.0607877103.
- 491 Cao, X.Y., Chen, M., Xu, Z.S., Chen, Y.F., Li, L.C., Yu, Y.H., Liu, Y.N., Ma, Y.Z. (2012) Isolation and
492 functional analysis of the bZIP transcription factor gene TaABP1 from a Chinese wheat landrace.
493 Journal of Integrative Agriculture **11** (10):1580-1591
- 494 Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. 2020. TBtools: An
495 Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant **13**(8):
496 1194-1202. doi:10.1016/j.molp.2020.06.009.
- 497 Chen, J., Yi, Q., Cao, Y., Wei, B., Zheng, L., Xiao, Q., Xie, Y., Gu, Y., Li, Y., Huang, H., Wang, Y.,
498 Hou, X., Long, T., Zhang, J., Liu, H., Liu, Y., Yu, G., and Huang, Y. 2016. ZmbZIP91 regulates
499 expression of starch synthesis-related genes by binding to ACTCAT elements in their promoters.
500 Journal of experimental botany **67**(5): 1327-1338. doi:10.1093/jxb/erv527.
- 501 Dong, Q., Xu, Q., Kong, J., Peng, X., Zhou, W., Chen, L., Wu, J., Xiang, Y., Jiang, H., and Cheng, B.
502 2019. Overexpression of ZmbZIP22 gene alters endosperm starch content and composition in
503 maize and rice. Plant science : an international journal of experimental plant biology **283**: 407-
504 415. doi:10.1016/j.plantsci.2019.03.001.
- 505 Fu, F.F., and Xue, H.W. 2010. Coexpression analysis identifies Rice Starch Regulator1, a rice
506 AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. Plant
507 physiology **154**(2): 927-938.

- 508 Gupta, S.K., Rai, A.K., Kanwar, S.S., Chand, D., Singh, N.K., Sharma, T.R. (2012) The single
509 functional blast resistance gene Pi54 activates a complex defence mechanism in rice. *Journal of*
510 *Experimental Botany* **63** (2):757-772
- 511 Guzmán, C., and Alvarez, J.B. 2016. Wheat waxy proteins: polymorphism, molecular characterization
512 and effects on starch properties [journal article]. *Theoretical and Applied Genetics* **129**(1): 1-
513 16. doi:10.1007/s00122-015-2595-9.
- 514 Hannah, L.C., and James, M. 2008. The complexities of starch biosynthesis in cereal endosperms.
515 *Curr.opin.biotechnol* **19**(2): 160-165.
- 516 Hebelstrup, K.H., Nielsen, M.M., Carciofi, M., Andrzejczak, O., Shaik, S.S., Blennow, A., and Palcic,
517 M.M. 2017. Waxy and non-waxy barley cultivars exhibit differences in the targeting and
518 catalytic activity of GBSS1a. *Journal of experimental botany* **68**(5): 931-941.
519 doi:10.1093/jxb/erw503.
- 520 Holdsworth M.J., Muoz-Blanco J., Hammond-Kosack M., Colot V., Schuch W. and Bevan M.W. 1995.
521 The maize transcription factor Opaque-2 activates a wheat glutenin promoter in plant and yeast
522 cells. *Plant molecular biology* **29**: 711-720.
- 523 Hsieh, C.F., Liu, W.C., Whaley, J.K., and Shi, Y.C. 2019. Structure and functional properties of waxy
524 starches. *Food Hydrocolloids* **94**: 238-254. doi:10.1016/j.foodhyd.2019.03.026.
- 525 Huang, H., Xie, S., Xiao, Q., Wei, B., Zheng, L., Wang, Y., Cao, Y., Zhang, X., Long, T., Li, Y., Hu,
526 Y., Yu, G., Liu, H., Liu, Y., Huang, Z., Zhang, J., and Huang, Y. 2016. Sucrose and ABA
527 regulate starch biosynthesis in maize through a novel transcription factor, ZmEREB156.
528 *Scientific reports* **6**: 27590. doi:10.1038/srep27590.
- 529 Hurst, H.C.(1995) Transcription factors 1: bZIP proteins. *Protein Profile* **2** (2):101-168

- 530 Izawa, T., Foster, R., Chua, N.H. (1993) Plant bZIP protein DNA binding specificity. *Journal of*
531 *Molecular Biology* 230 (4):1131-1144
- 532 Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., Parcy, F.,
533 bZIP Research Group (2002) bZIP transcription factors in Arabidopsis. *Trends in Plant*
534 *Science* 7 (3):106-111
- 535 James, M.G., Denyer, K., and Myers, A.M. 2003. Starch synthesis in the cereal endosperm. *Current*
536 *opinion in plant biology* 6(3): 215-222.
- 537 Jeon, J.S., Ryoo, N., Hahn, T.R., Walia, H., and Nakamura, Y. 2010. Starch biosynthesis in cereal
538 endosperm. *Plant physiology and biochemistry : PPB* 48(6): 383-392.
- 539 Li, C., Yue, Y., Chen, H., Qi, W., and Song, R. 2018. The ZmbZIP22 Transcription Factor Regulates 27-
540 kD gamma-Zein Gene Transcription during Maize Endosperm Development. *The Plant cell*
541 30(10): 2402-2424. doi:10.1105/tpc.18.00422.
- 542 Li, X., Gao, S., Tang, Y., Li, L., Zhang, F., Feng, B., Fang, Z., Ma, L., and Zhao, C. 2015. Genome-wide
543 identification and evolutionary analyses of bZIP transcription factors in wheat and its relatives
544 and expression profiles of anther development related TabZIP genes. *BMC genomics* 16: 976.
545 doi:10.1186/s12864-015-2196-7.
- 546 Lim, C.W., Back, W., Lim, S., Han, S.W., Lee, S.C. (2015) Expression and functional roles of the
547 pepper pathogen-induced bZIP transcription factor, CabZIP2, in enhanced disease resistance to
548 bacterial pathogen infection. *Molecular Plant–Microbe Interactions* 28 (7):825-833
- 549 Liu, G., Wu, Y., Xu, M., Gao, T., Wang, P., Wang, L., Guo, T., and Kang, G. 2016. Virus-Induced Gene
550 Silencing Identifies an Important Role of the TaRSR1 Transcription Factor in Starch Synthesis
551 in Bread Wheat. *Int J Mol Sci* 17(10). doi:10.3390/ijms17101557.

- 552 Liu, J.Y., Chen, N.N., Chen, F., Cai, B., Santo, S.D., Tornielli, G.B., Pezzotti, M., and Cheng, Z.M. 2014.
- 553 Genome-wide analysis and expression profile of the bZIP transcription factor gene family in
- 554 grapevine (*Vitis vinifera*). *BMC genomics* **15**.
- 555 Liu, Y., Hou, J., Wang, X., Li, T., Majeed, U., Hao, C., and Zhang, X. 2020. The NAC transcription
- 556 factor NAC019-A1 is a negative regulator of starch synthesis in wheat developing endosperm.
- 557 *Journal of experimental botany*. doi:10.1093/jxb/eraa333.
- 558 Ma, J., Yang, Y., Luo, W., Yang, C., Ding, P., Liu, Y., Qiao, L., Chang, Z., Geng, H., Wang, P., Jiang,
- 559 Q., Wang, J., Chen, G., Wei, Y., Zheng, Y., and Lan, X. 2017. Genome-wide identification and
- 560 analysis of the MADS-box gene family in bread wheat (*Triticum aestivum* L.). *PloS one* **12**(7):
- 561 e0181443. doi:10.1371/journal.pone.0181443.
- 562 Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong,
- 563 J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z.,
- 564 Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L.,
- 565 Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S.H.
- 566 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic*
- 567 *Acids Res* **39**(Database issue): D225-229. doi:10.1093/nar/gkq1189.
- 568 Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radchuk, V.,
- 569 Dockter, C., Hedley, P.E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X.Q.,
- 570 Zhang, Q., Barrero, R.A., Li, L., Taudien, S., Groth, M., Felder, M., Hastie, A., Simkova, H.,
- 571 Stankova, H., Vrana, J., Chan, S., Munoz-Amatriain, M., Ounit, R., Wanamaker, S., Bolser, D.,
- 572 Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Chailyan, A.,
- 573 Sampath, D., Heavens, D., Clissold, L., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X.,

- 574 Lin, C., McCooke, J.K., Tan, C., Wang, P., Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard,
575 M.I., Borisjuk, L., Houben, A., Dolezel, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P.,
576 Muehlbauer, G.J., Clark, M.D., Caccamo, M., Schulman, A.H., Mayer, K.F.X., Platzer, M.,
577 Close, T.J., Scholz, U., Hansson, M., Zhang, G., Braumann, I., Spannagl, M., Li, C., Waugh,
578 R., and Stein, N. 2017. A chromosome conformation capture ordered sequence of the barley
579 genome. *Nature* **544**(7651): 427-433. doi:10.1038/nature22043.
- 580 Mayer, K.F., Waugh, R., Brown, J.W., Schulman, A., Langridge, P., Platzer, M., Fincher, G.B.,
581 Muehlbauer, G.J., Sato, K., Close, T.J., Wise, R.P., and Stein, N. 2012. A physical, genetic and
582 functional sequence assembly of the barley genome. *Nature* **491**(7426): 711-716.
583 doi:10.1038/nature11543.
- 584 Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C.,
585 Muehlbauer, G.J., Schulman, A.H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer,
586 K.F.X., Spannagl, M., Stein, N., and Mascher, M. 2019. TRITEX: chromosome-scale sequence
587 assembly of Triticeae genomes with open-source tools. *Genome Biol* **20**(1): 284.
588 doi:10.1186/s13059-019-1899-5.
- 589 Muszynski, M.G., Dam, T., Li, B., Shirbourn, D.M., Hou, Z., Bruggemann, E., Archibald, R.L., Ananiev,
590 E.V., and Danilevskaya, O.N. 2006. *delayed flowering1* Encodes a Basic Leucine Zipper
591 Protein That Mediates Floral Inductive Signals at the Shoot Apex in Maize. *Plant physiology*
592 **142**(4): 1523-1536.
- 593 Niu, B., Deng, H., Li, T., Sharma, S., Yun, Q., Li, Q., Zhiguo, E., and Chen, C. 2020. OsbZIP76 interacts
594 with OsNF-YBs and regulates endosperm cellularization in rice (*Oryza sativa*). *J Integr Plant*
595 *Biol.* doi:10.1111/jipb.12989.

- 596 Pourabed E., Ghane Golmohamadi F., Soleymani Monfared P., Razavi S M., Shobbar Z-S. 2014. Basic
597 Leucine Zipper Family in Barley: Genome-Wide Characterization of Members and Expression
598 Analysis. *Molecular Biotechnology* **57**: 12-26. doi:10.1007/s12033-014-9797-2).
- 599 Preiss, J. 1993. Biosynthesis of Starch: ADPglucose Pyrophosphorylase, the Regulatory Enzyme of
600 Starch Synthesis: Structure-Function Relationships. *Journal of Applied Glycoence* **40**(2): 117-
601 131.
- 602 Qu, J., Xu, S., Zhang, Z., Chen, G., Zhong, Y., Liu, L., Zhang, R., Xue, J., and Guo, D. 2018.
603 Evolutionary, structural and expression analysis of core genes involved in starch synthesis.
604 *Scientific reports* **8**(1): 12736. doi:10.1038/s41598-018-30411-y.
- 605 Rapazote-Flores, P., Bayer, M., Milne, L., Mayer, C.D., Fuller, J., Guo, W., Hedley, P.E., Morris, J.,
606 Halpin, C., Kam, J., McKim, S.M., Zwirek, M., Casao, M.C., Barakate, A., Schreiber, M.,
607 Stephen, G., Zhang, R., Brown, J.W.S., Waugh, R., and Simpson, C.G. 2019. BaRTv1.0: an
608 improved barley reference transcript dataset to determine accurate changes in the barley
609 transcriptome using RNA-seq. *BMC genomics* **20**(1): 968. doi:10.1186/s12864-019-6243-7.
- 610 Song, Y.L., Luo, G.B., Shen, L., Yu, K., Yang, W., Li, X., Sun, J.Z., Zhan, K.H., Cui, D.Q., Liu, D.C.,
611 Zhang, A. 2020. TubZIP28, a novel bZIP family transcription factor from *Triticum urartu*, and
612 TabZIP28, its homologue from *Triticum aestivum*, enhance starch synthesis in wheat. *New*
613 *Phytologist* **226**: 1384-1398. doi:10.1111/nph.16435.
- 614 Sonnewald, U., and Kossmann, J. 2013. Starches--from current models to genetic engineering. *Plant*
615 *biotechnology journal* **11**(2): 223-232. doi:10.1111/pbi.12029.

- 616 Stracke, R., Favory, J.J., Gruber, H., Bartelniewoehner, L., Bartels, S., Binkert, M., Funk, M.,
617 Weisshaar, B., Ulm, R. (2010) The Arabidopsis bZIP transcription factor HY5 regulates
618 expression of the PFG1/MYB12 gene in response to light and ultraviolet-B radiation. *Plant*,
619 *Cell and Environment* 33 (1):88-103
- 620 Sun, C., Palmqvist, S., Olsson, H., Borén, M., Ahlandsberg, S., and Jansson, C. 2003. A novel WRKY
621 transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-
622 responsive elements of the iso1 promoter. *THE PLANT CELL ONLINE* 15(9): 2076-2092.
- 623 Wang, J., Zhou, J., Zhang, B., Vanitha, J., Ramachandran, S., and Jiang, S.Y. 2011. Genome-wide
624 expansion and expression divergence of the basic leucine zipper transcription factors in higher
625 plants with an emphasis on sorghum. *J Integr Plant Biol* 53(3): 212-231. doi:10.1111/j.1744-
626 7909.2010.01017.x.
- 627 Wang, J.C., Xu, H., Zhu, Y., Liu, Q.Q., and Cai, X.L. 2013. OsbZIP58, a basic leucine zipper
628 transcription factor, regulates starch biosynthesis in rice endosperm. *Journal of experimental*
629 *botany* 64(11): 3453-3466. doi:10.1093/jxb/ert187.
- 630 Wei, K., Chen, J., Wang, Y., Chen, Y., Chen, S., Lin, Y., Pan, S., Zhong, X., and Xie, D. 2012. Genome-
631 wide analysis of bZIP-encoding genes in maize. *DNA research : an international journal for*
632 *rapid publication of reports on genes and genomes* 19(6): 463-476. doi:10.1093/dnares/dss026.
- 633 Zhang, J., Chen, J., Yi, Q., Hu, Y., Liu, H., Liu, Y., and Huang, Y. 2014. Novel role of ZmaNAC36 in
634 co-expression of starch synthetic genes in maize endosperm. *Plant molecular biology* 84(3):
635 359-369. doi:10.1007/s11103-013-0153-x).
- 636 Zhang, Z., Zheng, X., Yang, J., Messing, J., and Wu, Y. 2016. Maize endosperm-specific transcription
637 factors O2 and PBF network the regulation of protein and starch synthesis. *Proceedings of the*

638 National Academy of Sciences of the United States of America **113**(39): 10842-10847.
639 doi:10.1073/pnas.1613721113.

640 Zhang, Z., Dong, J., Ji, C., Wu, Y., and Messing, J. 2019. NAC-type transcription factors regulate
641 accumulation of starch and protein in maize seeds. Proceedings of the National Academy of
642 Sciences of the United States of America **116**(23): 11223-11228. doi:10.1073/pnas.1904995116.

Draft

643

Table 1 The related information on HvbZIP transcription factors.

No.	Original_ID	Length(aa)	Molecular weight (average)	Theoretical PI	Confidence ¹	Group	Best <i>Arabidopsis</i> hit	Best rice hit
HvbZIP09	HORVU.MOREX.r2.1HG0061450.1	244	25879.94	6.07	HC1	A	AT5G44080	Os05t0437700
HvbZIP13	HORVU.MOREX.r2.2HG0091060.1	278	29951.83	9.21	HC1	A	AT2G41070	Os07t0686100
HvbZIP18	HORVU.MOREX.r2.2HG0161640.1	132	14476.23	10.26	HC2	A	AT4G35900	Os06t0719500
HvbZIP19	HORVU.MOREX.r2.2HG0167560.1	204	22389.06	9.8	HC1	A	AT2G17770	Os01t0859300
HvbZIP21	HORVU.MOREX.r2.2HG0172870.1	201	21885.47	9.96	HC1	A	AT4G35900	Os07t0686100
HvbZIP30	HORVU.MOREX.r2.3HG0246690.1	337	36320.96	9.35	HC1	A	AT3G56850	Os01t0813100
HvbZIP31	HORVU.MOREX.r2.3HG0249890.1	246	26061.17	8.96	HC1	A	AT1G03970	Os01t0867300
HvbZIP33	HORVU.MOREX.r2.3HG0250900.1	399	42043.65	6.57	HC1	A	AT3G56850	Os01t0859300
HvbZIP39	HORVU.MOREX.r2.4HG0316240.1	220	24036.78	9.47	HC1	A	AT3G56850	Os03t0322700
HvbZIP44	HORVU.MOREX.r2.4HG0342890.1	206	21991.51	9.69	HC1	A	AT2G17770	Os12t0601800
HvbZIP52	HORVU.MOREX.r2.5HG0383360.1	314	34865.3	6.49	HC1	A	AT3G56850	Os05t0489700
HvbZIP56	HORVU.MOREX.r2.5HG0402830.1	368	39332.7	9.07	HC1	A	AT3G19290	Os09t0456200
HvbZIP61	HORVU.MOREX.r2.5HG0415650.1	187	19849.49	9.85	HC2	A	AT2G17770	Os09t0540800

HvbZIP76	HORVU.MOREX.r2.6HG0513990.1	358	38590.95	9.7	HC1	A	AT1G49720	Os02t0766700
HvbZIP77	HORVU.MOREX.r2.7HG0550580.1	333	36056.19	8.84	HC1	A	AT3G56850	Os06t0211200
HvbZIP92	HORVU.MOREX.r2.7HG0620710.1	161	17618.92	9.26	HC2	A	AT4G35900	Os06t0719500
HvbZIP04	HORVU.MOREX.r2.1HG0051850.1	646	68319.73	8.52	HC1	B	AT2G40950	Os05t0411300
HvbZIP15	HORVU.MOREX.r2.2HG0098040.1	579	60953.84	6.25	HC1	B	AT2G40950	Os07t0644100
HvbZIP10	HORVU.MOREX.r2.1HG0063740.1	410	44697.83	5.53	HC2	C	AT5G28770	Os07t0182000
HvbZIP46	HORVU.MOREX.r2.5HG0359930.1	378	40343.88	5.21	HC2	C	AT5G28770	Os12t0601800
HvbZIP64	HORVU.MOREX.r2.5HG0431770.1	397	41763.39	5.2	HC1	C	AT5G28770	Os03t0796900
HvbZIP70	HORVU.MOREX.r2.6HG0473090.1	313	32843.5	5.29	HC2	C	AT5G24800	Os02t0266800
HvbZIP86	HORVU.MOREX.r2.7HG0613070.1	244	26486.58	7.75	HC2	C	AT5G24800	Os02t0266800
HvbZIP06	HORVU.MOREX.r2.1HG0055970.1	322	35289.72	5.6	HC1	D	AT5G06950	Os01t0808100
HvbZIP08	HORVU.MOREX.r2.1HG0060430.1	551	60898.52	6.16	HC1	D	AT1G08320	Os05t0443900
HvbZIP12	HORVU.MOREX.r2.2HG0090690.1	337	37489.36	9.09	HC1	D	AT3G12250	Os07t0687700
HvbZIP20	HORVU.MOREX.r2.2HG0172010.1	424	46934.95	6.2	HC1	D	AT5G10030	Os04t0637000
HvbZIP24	HORVU.MOREX.r2.3HG0216190.1	332	36823.31	7.79	HC1	D	AT3G12250	Os01t0279900
HvbZIP29	HORVU.MOREX.r2.3HG0246180.1	477	51849.08	6.53	HC1	D	AT5G06950	Os01t0808100
HvbZIP32	HORVU.MOREX.r2.3HG0250850.1	570	61532.81	6.41	HC1	D	AT1G08320	Os01t0859500
HvbZIP37	HORVU.MOREX.r2.4HG0296360.1	567	62324.57	6.47	HC1	D	AT1G08320	Os11t0152700

HvbZIP40	HORVU.MOREX.r2.4HG0316820.1	335	37268.18	7.87	HC1	D	AT3G12250	Os03t0318600
HvbZIP54	HORVU.MOREX.r2.5HG0390100.1	469	51394.2	8.37	HC1	D	AT5G06839	Os09t0280500
HvbZIP58	HORVU.MOREX.r2.5HG0406840.1	526	58522.73	6.64	HC1	D	AT5G06839	Os09t0489500
HvbZIP71	HORVU.MOREX.r2.6HG0475320.1	461	50155.42	7.28	HC1	D	AT5G06839	Os02t0194900
HvbZIP79	HORVU.MOREX.r2.7HG0556450.1	451	48698.94	7.1	HC1	D	AT5G06950	Os06t0265400
HvbZIP23	HORVU.MOREX.r2.3HG0204170.1	308	33295.89	6.3	HC1	E	AT3G58120	Os01t0211800
HvbZIP26	HORVU.MOREX.r2.3HG0240630.1	267	28754.84	5.74	HC1	E	AT3G58120	Os01t0756200
HvbZIP34	HORVU.MOREX.r2.4HG0288270.1	349	38515.87	8.64	HC2	E	AT1G58110*	Os11t0218000
HvbZIP49	HORVU.MOREX.r2.5HG0373090.1	361	40283.85	6.46	HC2	E	AT1G58110*	Os12t0194600
HvbZIP53	HORVU.MOREX.r2.5HG0383680.1	241	26208.55	6	HC1	E	AT2G42380	Os01t0756200
HvbZIP74	HORVU.MOREX.r2.6HG0487850.1	323	35396.85	7.59	HC2	E	AT5G04840*	Os02t0538900
HvbZIP05	HORVU.MOREX.r2.1HG0055460.1	226	24010.77	5.44	HC1	F	AT2G16770	Os01t0801901
HvbZIP28	HORVU.MOREX.r2.3HG0245710.1	214	23011.87	5.23	HC2	F	AT2G16770	Os01t0801901
HvbZIP45	HORVU.MOREX.r2.5HG0352300.1	261	27538.93	5.97	HC1	F	AT4G35040)	Os06t0716800
HvbZIP66	HORVU.MOREX.r2.5HG0440220.1	215	23946.05	6.21	HC1	F	AT4G35040	Os06t0716800
HvbZIP85	HORVU.MOREX.r2.7HG0603710.1	232	25692.03	5.63	HC1	F	AT4G35040	Os06t0716800
HvbZIP89	HORVU.MOREX.r2.7HG0619460.1	252	27243.48	5.59	HC1	F	AT4G35040	Os06t0716800
HvbZIP90	HORVU.MOREX.r2.7HG0619470.1	256	28356.23	8.26	LC2	F	AT4G35040	Os06t0716800

HvbZIP91	HORVU.MOREX.r2.7HG0619630.1	202	21914.52	7.25	HC1	F	AT4G35040	Os06t0716800
HvbZIP11	HORVU.MOREX.r2.1HG0074290.1	406	43014.49	9.24	HC1	G	AT4G01120	Os05t0569300
HvbZIP16	HORVU.MOREX.r2.2HG0121940.1	381	39622.99	6.42	HC1	G	AT2G35530	Os07t0209800
HvbZIP42	HORVU.MOREX.r2.4HG0327070.1	382	40706.05	9.17	HC1	G	AT2G35530	Os03t0239400
HvbZIP48	HORVU.MOREX.r2.5HG0372090.1	383	41015.22	8.33	HC1	G	AT2G35530	Os12t0233800
HvbZIP65	HORVU.MOREX.r2.5HG0434960.1	349	35959.56	7.19	HC1	G	AT2G35530	Os07t0209800
HvbZIP67	HORVU.MOREX.r2.6HG0459560.1	352	37155.61	5.96	HC1	G	AT2G35530	Os02t0128200
HvbZIP22	HORVU.MOREX.r2.3HG0200390.1	187	19637.66	9.95	HC2	H	AT5G11260	Os01t0174000
HvbZIP73	HORVU.MOREX.r2.6HG0479210.1	157	17392.07	9.68	HC1	H	AT5G11260	Os02t0203000
HvbZIP83	HORVU.MOREX.r2.7HG0592070.1	162	17358.94	9.8	HC1	H	AT5G11260	Os02t0203000
HvbZIP84	HORVU.MOREX.r2.7HG0600140.1	294	31867.94	4.73	HC1	H	AT1G42990	Os06t0622700
HvbZIP03	HORVU.MOREX.r2.1HG0040670.1	381	41043.73	6.39	HC1	I	AT4G38900	Os10t0531900
HvbZIP14	HORVU.MOREX.r2.2HG0091980.1	369	39686.4	6.94	HC1	I	AT2G40620	Os03t0336200
HvbZIP17	HORVU.MOREX.r2.2HG0147450.1	390	41811.29	5.84	HC2	I	AT1G06070	Os09t0516200
HvbZIP35	HORVU.MOREX.r2.4HG0291940.1	324	34754.53	6.46	HC1	I	AT1G43700	Os12t0162500
HvbZIP38	HORVU.MOREX.r2.4HG0314880.1	356	38302.64	6.42	HC1	I	AT2G40620	Os03t0336200
HvbZIP43	HORVU.MOREX.r2.4HG0342270.1	391	41755.54	6.22	HC2	I	AT4G38900	Os03t0127500
HvbZIP50	HORVU.MOREX.r2.5HG0378980.1	323	34887.85	6.42	HC1	I	AT1G43700	Os12t0162500

HvbZIP59	HORVU.MOREX.r2.5HG0411590.1	376	40854.6	6.74	HC1	I	AT1G06070	Os09t0516200
HvbZIP60	HORVU.MOREX.r2.5HG0412520.1	508	54005.96	9.91	HC2	I	AT4G38900	Os08t0543900
HvbZIP81	HORVU.MOREX.r2.7HG0569750.1	366	39441.48	6.34	HC1	I	AT1G06070	Os09t0516200
HvbZIP82	HORVU.MOREX.r2.7HG0569990.1	472	50787.18	6.51	HC1	I	AT4G38900	Os08t0543900
HvbZIP01	HORVU.MOREX.r2.1HG0012580.1	153	16339.25	9.38	HC1	S	AT3G62420	Os05t0129300
HvbZIP25	HORVU.MOREX.r2.3HG0219550.1	156	16396.44	8.49	HC1	S	AT1G75390	Os01t0542700
HvbZIP41	HORVU.MOREX.r2.4HG0318130.1	169	18662.36	6.6	HC2	S	AT3G62420	Os05t0129300
HvbZIP47	HORVU.MOREX.r2.5HG0367280.1	159	17454.59	9.02	HC1	S	AT3G62420	Os05t0129300
HvbZIP55	HORVU.MOREX.r2.5HG0390950.1	155	16935.04	9.16	HC1	S	AT1G75390	Os09t0306400
HvbZIP57	HORVU.MOREX.r2.5HG0404670.1	207	23066.59	6.67	HC1	S	AT2G04038	Os09t0474000
HvbZIP63	HORVU.MOREX.r2.5HG0426550.1	186	20201.94	11.04	HC2	S	AT1G13600	Os03t0770000
HvbZIP68	HORVU.MOREX.r2.6HG0460660.1	173	18002.34	7.87	HC1	S	AT4G34590	Os02t0132500
HvbZIP72	HORVU.MOREX.r2.6HG0476540.1	185	19781.87	6.52	HC2	S	AT3G30530	Os02t0191600
HvbZIP75	HORVU.MOREX.r2.6HG0507030.1	197	21463.08	9.4	HC2	S	AT5G38800	Os02t0728001
HvbZIP78	HORVU.MOREX.r2.7HG0552060.1	223	23818.27	6.24	HC2	S	AT5G38800	Os08t0487100
HvbZIP80	HORVU.MOREX.r2.7HG0557340.1	195	21607.2	6.31	HC2	S	AT3G30530	Os02t0728001
HvbZIP36	HORVU.MOREX.r2.4HG0293150.1	203	21687.27	7.86	HC2	U1	AT1G19490	Os11t0154900
HvbZIP51	HORVU.MOREX.r2.5HG0379500.1	402	42721.83	7.08	HC2	U1	AT1G19490	Os11t0154900

HvbZIP69	HORVU.MOREX.r2.6HG0471050.1	251	26628.75	9.63	HC2	U2	AT4G36730	Os02t0182700
HvbZIP87	HORVU.MOREX.r2.7HG0616920.1	251	25888.98	9.65	HC2	U2	AT4G36730	Os02t0182700
HvbZIP88	HORVU.MOREX.r2.7HG0617000.1	264	27020.03	9.97	HC2	U2	AT4G36730	Os02t0182700
HvbZIP02	HORVU.MOREX.r2.1HG0033880.1	847	91552.39	5.8	HC1	U3	---	---
HvbZIP27	HORVU.MOREX.r2.3HG0245040.1	682	73911.30	8.47	HC1	U3	---	---
HvbZIP62	HORVU.MOREX.r2.5HG0422280.1	881	94014.87	6.12	HC1	U3	---	---

¹ Members which obtained from high confidence (HC) or low confidence (LC) data

(-) The *HvbZIP* gene didn't match any bZIP gene in *Arabidopsis* and rice.

* The *AtbZIP* genes didn't be identified in Jakoby et al. 2002.

Draft

Fig.1 Gene structures of barley, rice and Arabidopsis bZIP genes with full-length coding sequences. The lengths of the exons and introns of each gene were displayed proportionally. Exons were represented by black boxes and introns by black lines. The 13 groups were represented by different colors.

865x500mm (96 x 96 DPI)

Fig.2 Protein motifs and bZIP domains of barley bZIP proteins. Each motif was represented by a colored box.

307x500mm (96 x 96 DPI)

Fig.3 Heatmaps of expression profiles for HvbZIP genes at different developmental stages of eight tissues in barley. The color scale represented expression values. * represented the 15 genes selected for qRT-PCR verification. The arrows represented the genes selected for yeast one-hybrid verification. Root (10cm seedlings), Root 2 (4 weeks seedling), Shoot (10cm seedlings), Rachis (5 weeks pa), Senescing leaf (2 months), Tillers (3rd internode), Inflorescence-1 (0.5cm), Inflorescence-2 (1-1.5cm), Embryo (germinating), Palea (6 weeks pa), Epidermis (4 weeks), Grain (5 DPA) and Grain (15 DPA), DPA: days post-anthesis.

1295x2343mm (72 x 72 DPI)

Fig.4 Candidate HvZIP genes involved in starch synthesis were screened based on coexpression analysis and interaction verification. (a) Relative expression levels of key starch synthase genes. AGP-L: ADP-glucose pyrophosphorylase-large subunit; AGP-S: ADP-glucose pyrophosphorylase-small subunit; SBE1: starch-branching enzyme 1; SBE2a: starch-branching enzyme IIa; SBE2b: starch-branching enzyme IIb; SS1: starch synthase 1; SS3: starch synthase 3; waxy: the gene encodes GBSS (granule-binding starch synthase); ISA1: isoamylase 1; 5-25D: 5-25 days post-anthesis. (b) Relative expression levels of 15 HvZIP genes. (c) Interactions between two HvZIP genes and the promoter of HvAGP-S by yeast one-hybrid assay. The black triangle indicated that the yeast concentration was diluted with 10-fold gradient.

1000x1114mm (72 x 72 DPI)