Comparison of the Agronomic, Cytological, Grain Protein Characteristics, as Well as Transcriptomic Profile of Two Wheat Lines Derived From Wild Emmer

Fangyi Gong¹, Tiangang Qi¹, Tian Zhang¹, Yufen Lu¹, Jia Liu¹, Xiaoying Zhong¹, Jingshu He¹, Yunfang Li², Youliang Zheng¹,³, Dengcai Liu¹,³, Lin Huang¹* and Bihua Wu¹,³*

¹State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Triticeae Research Institute, Sichuan Agricultural University, Chengdu, China, ²Chengdu Institute of Biology, Chinese Academy of Science, Chengdu, China, ³Key Laboratory of Crop Genetic Resources and Improvement, Ministry of Education, Sichuan Agricultural University, Chengdu, China

Two advanced wheat lines BAd7-209 and BAd23-1 without the functional gene GPC-B1 were obtained from a cross between common wheat cultivar Chuannong 16 (CN16) and wild emmer wheat accession D97 (D97). BAd7-209 showed superior quality parameters than those of BAd23-1 and CN16. We found that the components of glutenins and gliadins in BAd7-209 and BAd23-1 were similar, whereas BAd7-209 had higher amount of glutenins and gliadins than those of BAd23-1. RNA sequencing analysis on developing grains of BAd7-209 and BAd23-1 as well as their parents revealed 382 differentially expressed genes (DEGs) between the high-grain protein content (GPC) (D97 + BAd7-209) and the low-GPC (CN16 + BAd23-1) groups. DEGs were mainly associated with transcriptional regulation of the storage protein genes, protein processing in endoplasmic reticulum, and protein export pathways. The upregulated gluten genes and transcription factors (e.g., NAC, MYB, and bZIP) may contribute to the high GPC in BAd7-209. Our results provide insights into the potential regulation pathways underlying wheat grain protein accumulation and contribute to make use of wild emmer for wheat quality improvement.

Keywords: wild emmer wheat, GPC, transcriptome, processing quality, gluten

INTRODUCTION

Wheat (Triticum aestivum L.) is one of the most important staple crops, which can be processed into a wide range of products such as bread, noodle, and biscuit. Grain protein content (GPC) is an important quality trait in wheat and determines the nutritional value and processing quality (Shewry et al., 1995, 2009). However, GPC and grain yield-related traits are usually negative correlated (Groos et al., 2003), which hampered their simultaneous improvement in conventional wheat breeding program. The characterization and transferring of gene(s)/quantitative trait loci (QTLs) from wheat wild relatives is an effective strategy in the development of elite varieties with high GPC and/or yield (Krugman et al., 2018; Liu et al., 2019; Xiang et al., 2019; Gong et al., 2021).

Wild emmer (T. turgidum ssp. dicoccoides, 2n = 4x = 28, AABB), the tetraploid ancestor of common wheat, provides a valuable reservoir of genetic variation for GPC (Uauy et al., 2006; Liu
et al., 2019). A number of QTLs affecting GPC were reported in wild emmer wheat (Joppa et al., 1997; Gonzalez-Hernandez et al., 2004; Uauy et al., 2006; Fatikhab et al., 2019; Liu et al., 2019). To date, however, only Gpc-B1 on chromosome 6BS has been cloned. The introgression of Gpc-B1 in wheat breeding programs can significantly improve GPC while reduce grain-yield related traits in some wheat lines and environments (Uauy et al., 2006; Chen et al., 2017). Therefore, it is desirable to identify GPC-QTLs that are not negative correlated or less correlated with grain yield-related traits.

GPC was regulated by a plethora of genes and easily affected by environment (Liu et al., 2019). In forward genetics, identification of candidate genes related to GPC is a time-consuming and laborious process. RNA sequencing (RNA-seq) provides high-resolution methods for deciphering quality traits and quantifying expression levels of candidate genes on a genome-wide scale (Furtado et al., 2015; Rangan et al., 2017). Currently, RNA-seq has been utilized to study the differentially expressed genes (DEGs) and regulation networks that associated with wheat grain protein accumulation (Cantu et al., 2011; Gong et al., 2021).

Grain protein synthesis in cereal crops was determined by several pathways, mainly including transcriptional regulation of the storage protein genes (glutenin and gliadin) and protein processing in endoplasmic reticulum (ER) and Golgi apparatus. Transcription factors (TFs) belong to bZIP, Dof, MYB, and NAC families have been widely reported in transcriptional regulation of grain protein genes in rice, maize, and barley (Suzuki et al., 1998; Onodera et al., 2001; Diaz et al., 2002; Shewry and Halford, 2002; Zhang et al., 2019). Previous studies demonstrated that the protein processing in ER and Golgi apparatus had crucial role in grain protein synthesis in rice (Takemoto et al., 2002; Wang et al., 2016; Ren et al., 2020), whereas their function for GPC was less reported in wheat.

In our previous studies, the agronomically stable advanced wheat lines were developed from a cross between common wheat cultivar Chuanrong 16 (CN16) and wild emmer accession D97 (D97) followed by successive selfing (Jiang et al., 2017; Wang et al., 2018; Liu et al., 2019). Some advanced lines with simultaneous improvement of GPC and thousand-kernel weight (TKW) were obtained. In the present study, two sister lines, BAd7-209 and BAd23-1, with desirable agronomic traits were obtained from CN16×D97. These two lines showed contrasting GPC, while both of them did not contain the functional Gpc-B1. A comparison of the transcriptomes of developing grains from BAd7-209, BAd23-1, CN16, and D97 revealed candidate genes and regulation pathways that may be contributed to wheat grain protein accumulation.

### MATERIALS AND METHODS

#### Plant Materials

Two sister lines BAd7-209 and BAd23-1 as well as their parents D97 and CN16 were kept at the Triticeae Research Institute, Sichuan Agricultural University, Chengdu, China. CN16 and BAd23-1 were characterized as low-gluten wheat lines, whereas BAd7-209 and D97 were high-GPC lines (Table 1). Wheat plants were grown in the field with three biological replicates (10 rows each replicate) at the experimental field of the Sichuan Agricultural University over two wheat growing seasons (2016 and 2017) at Wenjiang (2016WJ and 2017WJ) and Chongzhou (2016CZ and 2017CZ). Individuals were planted 10-cm apart in a 2-m row with 30 cm between rows. A compound fertilizer [N: P₂O₅: K₂O (15: 15: 15)] was applied before sowing at a rate of 450 kg per hectare. Developing grains were sampled at 10, 14, 18, 22, 26, 30, 34, and 38 days after pollination (DAP) and snap-frozen in liquid nitrogen and then kept at −80°C for RNA-seq.

#### Characterization of Agronomic Traits and Karyotype

Agronomic traits of BAd7-209 and BAd23-1 and their parents were measured as listed in Table 1. The traits of plant height (PH), number of spikelets per spike (NSp), and spike number per plant (SN) were averaged by 10 plants. The weight of 300 grain samples (GB/T 5519-2008, 2008) was recorded as 1,000-kernel weight (TKW) as described by Wang et al. (2018).

At least 30 root-tip cells of BAd7-209 and BAd23-1 and their parents D97 and CN16 were observed for detection the chromosome number using the methods as described by Zhang et al. (2007).

#### Single-Nucleotide Polymorphism (SNP) Genotyping

Genomic DNA from leaves was isolated using a plant genomic DNA kit (Tiangen Biotech, Beijing Co. Ltd. Beijing, China). Chip-based genotyping was conducted using the wheat 55 K SNP array (www.capit albio.com). The flanking sequence of each SNP was used to map onto the bread wheat reference
sequence (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v1.0/), using BLASTN with E-value threshold of $10^{-10}$ and a maximum mismatch of one base. SNP markers showed homozygous genotype among the parents, and advanced wheat lines were used to estimate D97 introgressions. The ratios of same SNP to the total SNPs scored between D97 and two advanced lines were computed using a sliding window of 10 Mb and step length of 1 Mb as described by Hao et al. (2019). Graphical representations were constructed using the R package ggplot2 (Wickham 2016).

### Quality Parameters Determination

Mature seeds were harvested for measurement of wheat quality parameters (Supplementary Table S1). Seeds of D97, CN16, BAd7-209, and BAd23-1 were conditioned to 14% moisture and milled using the Chopin CD1 AUTO (Renault, Boulogne-Billancourt, France) (Wang et al., 2018). Wet gluten content (WGC) was determined following the American Association of Cereal Chemists (AACC) method 38-12A. The sedimentation volume (SV) was measured following the AACC 44-15A (FOSS, Denmark). GPC (0.5 g) was recorded according to the AACC 39-10. Hygroscopic rate (HR) was measured following the AACC 54-21. All of the experiments were repeated three times.

### Grain Protein Component Measurement

Albumins, globulins, glutenins, and gliadins were extracted from mature grains using the methods as described by Duan and Zhao, 2004 and Wan et al. (2000) with slight modifications. Protein content was tested using Kjeldahl method (Kjeltec TM 8400). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate the glutenins following the method of Hu et al. (2010) using the protein extraction buffer consisting of 62.5 mM Tris-HCl, pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 0.002% (w/v) bromophenol blue, and 1.5% (w/v) dithiothreitol (DTT).

### RNA-Seq Analysis

Total RNA was extracted using a plant RNA extraction kit v1.5 (Biofit Biotechnologies, Chengdu). The RNA concentration and integrity were assessed using an Qubit®RNA assay kit on a Qubit® 2.0 Flurometer (Life Technologies, CA, United States) and an RNA Nano 6000 assay kit, respectively. RNAs from different grain developing stages were pooled in equi-molar concentrations (1 µg of RNA per sample) for cDNA library construction (three biological replicates per library). The sequencing libraries were completed using the NEBNext Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States) and then sequenced using the
Illumina Hiseq platform (Novogene Bioinformatics Technology Co. Ltd. Beijing, China).

The adaptors, reads with more than 10% N, and reads with phred quality scored Q < 20 from RNA-seq raw data were removed using Trimmomatic (Bolger et al., 2014). The generated clean reads from each library were assessed using the Q20, Q30, and GC contents and aligned to Chinese Spring reference genome (https://urgi.versailles.inra.fr/blast_iwgsc/) using Hisat2 (Kim et al., 2019). FPKM (Elowitz et al., 2002; Jiang and Wong, 2009) was estimated to represent the expression level. DESeq was used for differential expression analysis as described by Wang et al. (2010). Genes with adjusted FDR < 0.001 and |log2 FC| ≥ 2 found by EBSeq (Leng et al., 2013) were considered as DEGs. Short Time-series Expression Miner (STEM) software (v1.3.11) was used to explore expression patterns of DEGs as described by Ernst and Bar-Joseph (2006) with log2 standardization, p-value ≤ 0.05, and |log2 FC|≥2. Functional annotation of the DEGs was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) databases (Kanehisa et al., 2004).

**Statistical Analysis**

Statistical analyses were conducted using SPSS version 22 software (SPSS Inc. Chicago, IL, United States). The morphological traits and quality parameters were compared by analysis of variance (ANOVA) complemented by Tukey’s test.

**RESULTS**

**Agronomic and Karyotype Characteristics**

The NSp of BAd7-209 and BAd23-1 were less than that of CN16 (p < 0.05), while higher than D97 (p < 0.05). The BAd7-209 resembled the BAd23-1 with respect to spike and SN. We have observed significantly differences in PH and TKW among BAd7-209, BAd23-1, and CN16 (Figure 1; Table 1). The D97 had significantly less NSp and TKW, while significantly higher PH than that of advanced wheat lines (Table 2).

Cytological observations indicated that both BAd23-1 and BAd7-209 have 42 chromosomes (Figure 1C), which reached the genetic background of common wheat. SNP genotyping analyses revealed 150 and 108 potential D97 segments on chromosomes of BAd23-1 and BAd7-209. The number of CN16 segments on BAd7-209 and BAd23-1 chromosomes was 212 and 232 (Supplementary Figure S1).

**Grain Protein Component and Glutenin Subunits**

The total amount of glutenins, gliadins, albumins, and globulins in D97 was higher than those of CN16. There were no differences in albumin and globulin contents between BAd23-1 and BAd7-209. The glutenin and gliadin contents of BAd7-209 were significantly higher than those of BAd23-1 (Table 2).

SDS-PAGE analysis showed different high-molecular-glutenin subunits (HMW-GSs) in CN16 (1Ax1, 1Dx5, 1Dy10, and 1Bx+1By20) and D97 (1Ax2.2, 1Bx, 1By8.1, and 1Ay). The BAd7-209 possessed five HMW-GSs (1Ax1.2, 1Bx+1By8, and 1Dx5 + 1Dy10), which were consistent with previous reports (Jiang et al., 2017; Xiang et al., 2019). The BAd23-1 had five HMW-GSs, including 1Bx7 and 1Dx5+1Dy10 that are...
identical to BAd7-209, and one different 1Ax subunit (1Ax1) and one 1By subunit (1By8.1) (Figure 2).

The expression levels of 1Ax and 1By in high-GPC group (D97 + BAd7-209) were higher than those of low-GPC group (CN16 + BAd23-1). The composition of low-molecular-glutenin subunits (LMW-GSs) in CN16, BAd7-209, and BAd23-1 was similar. The wild emmer D97 showed less LWM-GS subunits probably due to the absence of the D genome. The expression levels of LMW-GSs in high-GPC group (D97 + BAd7-209) were higher than those of low-GPC group (CN16 + BAd23-1) (Figure 2).

**Analysis of RNA-Seq Data**

RNA-seq generated 189.93 million raw reads from developing grains of BAd7-209, BAd23-1, CN16, and D97 pools. A total of 94.96 million clean reads were retained and further mapped to the *T. aestivum* cDNA database (IWGSC1.0). The mapping ratios ranged from 79.42% to 84.33%, 86.92% to 88.12%, 85.76% to 89.33%, and 79.15% to 80.89% in BAd7-209, BAd23-1, CN16, and D97, respectively. The GC contents among replicates were almost identical, and the Q30 was over 90% in each library (Supplementary Table S2). These results demonstrate that the RNA-seq data were qualified for subsequent analysis.

**Differential Expression Analysis**

In total, 62,404 genes mapped to IWGSC1.0 genome were expressed (FPKM≥1) in D97, CN16, BAd7-209, and BAd23-1 libraries. A total of 24,700, 2,332, 21,555, and 2,459 DEGs (|log2 FC|≥2) were found in D97 vs. CN16, BAd7-209 vs. CN16, D97 vs. BAd23-1, and BAd7-209 vs. BAd23-1, respectively (Figure 3A). The comparison between the high-GPC group (D97 + BAd7-209) and low-GPC group (CN16 + BAd23-1) revealed 382
DEGs (Figure 3B), of which 148 genes were common expressed, 34 genes only expressed in D97 and BAd7-209, whereas 56 genes only expressed in CN16 and BAd23-1. Nine DEGs encoding HWM-GS, alpha-gliadin, nodulin protein, trypsin inhibitor, purothionin, and pre-mRNA-splicing factor proteins were selected to perform qRT-PCR (Supplementary Table S3). The expression changes of seven DEGs were quite consistent with those obtained from the RNA-seq (Figure 4, Supplementary Table S3).

All genes were further analyzed with the STEM software (v1.3.11) (Ernst and Bar-Joseph, 2006) to obtain the temporal expression patterns. Nineteen expression profiles were clustered, and seven profiles (9, 10, 0, 19, 4, 7, and 6) were regarded as significantly changed ($p \leq 0.05$) (Figure 5A). The profile 19 had similar tendency to the GPC in advanced lines and the parents (Figure 5B), indicating the positive role in regulation of grain protein accumulation.

**The Functional Annotation of DEGs**

A total of 459 DEGs in profile 19 were subjected to GO and KEGG analyses. On the basis of GO terms for those genes, three categories can be classified: biological process, cellular component, and molecular function. The GO term metabolic process (GO: 0008152, 172 DEGs), cellular process (GO: 0009987, 131 DEGs), and single-organism process (GO: 0044699, 93 DEGs) were highly enriched in biological process; those of cell (GO: 0005623, 137 DEGs), cell part (GO: 0044464, 136 DEGs), and organelle (GO: 0043226, 105 DEGs) were dominant in cellular component; and those of binding (GO:
0005488, 138 DEGs), catalytic activity (GO: 0003824, 145 DEGs), and nutrient reservoir activity (GO: 0045735, 19 DEGs) were primary in molecular function (Figure 6A, Supplementary Table S4).

KEGG pathway enrichment analysis showed that starch and sucrose metabolism (ko00500, 10 DEGs), protein processing in ER (ko04141, eight DEGs), fructose and mannose metabolism (ko00051, five DEGs), glycolysis/gluconeogenesis (ko00010, five DEGs), protein export (ko03060, four DEGs), inositol phosphate metabolism (ko00562, three DEGs), pentose and glucuronate interconversions (ko00040, three DEGs), valine, leucine, and isoleucine degradation (ko002800, one DEG) were significantly enriched pathways (Figure 6B, Supplementary Table S5). Eight DEGs in pathway of protein processing in ER are related to protein transport protein Sec61 (Sec61, 3 genes), heat shock protein (HSP, 3 genes), SKP1-like protein (SKP1, 1 gene), and E3 ubiquitin-protein ligase RMA3 (RMA3, 1 gene). Four genes in pathway of protein export are related to Sec61 (3 genes) and signal recognition particle 54 kDa protein 2-like (SRP, 1 gene).

Characterization of DEGs Encoded Gluten and TFs

Twelve out of 459 genes in profile 19 were related to α/β-gliadins (six genes), γ-gliadins (three genes), LMW-GSs (two genes), and HWM-GSs (one gene) (Table 3). All these genes were upregulated in the high-GPC group compared to the low-GPC group. Twenty-eight genes belonging to 16 TF families, such as...
TABLE 4 | The list of TFs identified from the 459 DEGs in profile 19.

| GeneName      | GeneID                  | TF family | CPKM | BAd23-1  | BAd7-209 | D97  |
|---------------|-------------------------|-----------|------|----------|----------|------|
| TaALX0996     | Triticum_aestivum_newGene_30.996 | AUX/IAA  | 0    | 0        | 0.03     | 1.13 |
| TaB30200      | TraesCSB02G330200        | B3        | 0.01 | 0.54     | 0.14     | 3.37 |
| SPA           | TraesCS1A02G3299900      | bZIP      | 5.32 | 9.01     | 12.48    | 23.15|
| TabZIP8700    | TraesCS3A02G3787000      | bZIP      | 1.08 | 1.55     | 1.93     | 4.86 |
| TaC2H60600    | TraesCS2B02G4060000      | C2H2      | 0.31 | 1.79     | 2.43     | 5.27 |
| TaC3H8000     | TraesCS2B02G1380000      | C3H       | 0.8  | 1.99     | 4.5      | 8.65 |
| TaCSD9000     | TraesCS2A02G3865000      | CSD       | 23.33| 57.2     | 84.36    | 115.58|
| TaDDP8500     | TraesCSB02G3985000       | D2P       | 8.3  | 7.93     | 27.02    | 37.26|
| TaFA0079      | Triticum_aestivum_newGene_56.079 | FAR1      | 0.26 | 0.52     | 0.84     | 1.62 |
| TaGAG8000     | TraesCS6B02G1382000      | GARP-G2-like | 0.03 | 0.04    | 0.13     | 0.97 |
| TaGA1000      | TraesCS7A02G4110000      | GARP-G2-like | 0.04 | 0.19    | 0.94     | 1.16 |
| TaGIR9000     | TraesCS4A02G1659000      | GRAS      | 0.44 | 0.88     | 1.28     | 2    |
| TaGIR6200     | TraesCS3A02G1662000      | Junoniji  | 0    | 0.84     | 1.44     | 1.59 |
| TaM1500       | TraesCS4B02G3515000      | MADS-M-type | 1.41 | 1.36    | 2.8      | 5.67 |
| TaMYB0300     | TraesCS2A02G3703000      | MYB       | 7.12 | 12.29    | 17.97    | 28.99|
| TaMYB2000     | TraesCS2B02G2520000      | MYB       | 1.45 | 1.63     | 2.49     | 4.86 |
| TaMYB7800     | TraesCS2B02G3878000      | MYB       | 0.75 | 1.96     | 4.3      | 5.94 |
| TaGAMyb       | TraesCS2A02G3965000      | MYB       | 3.54 | 8.04     | 12.93    | 16.15|
| TaNAC7700     | TraesCS2A02G3777700      | NAC       | 0.19 | 0.26     | 0.5      | 1.31 |
| TaNAC2900     | TraesCS3B02G0929000      | NAC       | 0.43 | 1.77     | 4.14     | 6.29 |
| TaNAC2400     | TraesCS7A02G1524000      | NAC       | 0.52 | 2.63     | 4.18     | 6.16 |
| TaNAC4700     | TraesCS7A02G1947000      | NAC       | 6.33 | 32.32    | 37.62    | 56.61|
| TaNAC9500     | TraesCS7A02G3495000      | NAC       | 0.08 | 0.25     | 0.18     | 0.55 |
| TaNAC9100     | TraesCS7A02G3691000      | NAC       | 5.76 | 23.67    | 34.2     | 56.1 |
| TaNAC8500     | TraesCS7A02G4896000      | NAC       | 0.04 | 0.54     | 1.25     | 3.03 |
| TaNAC8300     | TraesCS7B02G0663000      | NAC       | 3.18 | 11.97    | 30.07    | 64.68|
| TaNAC6700     | TraesCS7A02G3367000      | NAC       | 15.73| 113.72   | 199.94   | 192.12|
| Ta3500        | TraesCU02G1935000       | Others    | 0    | 0        | 0.03     | 1.82 |

FIGURE 7 | The synthesis and regulation pathways of grain storage protein.
NAC, MYB, GARP-G2-like, bZIP, and C2H2, were identified (Table 4). To understand the expression patterns, the expression levels of TFs in different tissues sampled at different time points were retrieved from public available RNA-seq data of IWGSC (Thomas et al., 2014). We found that seven genes (e.g., TabZIP8700, TaC2H20600, and TaMYB0300) were expressed in different tissues at different time points. Ten genes (e.g., SPA, TaC3H8000, TaGAMYb, and TaNAC019) were only expressed in grains. Two genes (TaMYB7800 and TaNAC9500) were root specific. The others were not detected in all tissues (Supplementary Table S6).

**DISCUSSION**

GPC is an important quality trait in common wheat. The wild emmer gene GPC-B1 positively impacts protein, Zn, and Fe in wheat grain (Uauy et al., 2006). In the current study, two sister wheat lines BAd7-209 and BAd23-1 derived from wild emmer showed contrasting GPC, while both of them did not contain the Gpc-B1. Processing quality parameters of BAd7-209 were significantly higher than those of BAd23-1. These two lines had genetic background of common wheat and introgression segments from wild emmer D97. Our results indicate the presence of other wild emmer gene(s) that contributed to high GPC in BAd7-209.

Previous reports revealed that the components and expression levels of glutenins and gliadins could affect GPC (Xi and Zheng, 2011; Jiang et al., 2017; Wang et al., 2018; Xiang et al., 2019; Shen et al., 2020). In this study, similar glutenin subunits were identified in BAd7-209 and BAd23-1, whereas BAd7-209 had significantly higher amount of glutenins and gliadins than BAd23-1. Nineteen DEGs including gliadin (nine genes), glutenin (three genes), globulin (three genes), avenin-like (three genes), and serpin (two genes) were enriched in the GO term nutrient reservoir activity (GO: 0045735). These genes were upregulated in the high GPC (D97 + BAd7-209) vs. the low-GPC (CN16 + BAd23-1) groups. Previous reports revealed that overexpression gliadin and glutenin genes could increase GPC in wheat (Guo et al., 2015; Li et al., 2019) and rice (Xiang et al., 2019). In the present study, we have identified 25 TFs, such as NAC, MYB, GARP-G2-like, bZIP, and C2H2, from the high-GPC and low-GPC groups. Some TFs (e.g., TaC3H8000, TaNAC2400, TaNAC4700, and TaNF6700) were specifically expressed in grains. In addition, we have identified either NAC (ACGCAA/G), MYB (AACA), bZIP (GCN4), or DoF (P-box) motifs in the promoters of gluten genes that were differentially expressed in the high-GPC vs. low-GPC groups. Taken together, our results indicate that some TFs are involved in regulation of gluten and gliadin genes that contributed to the GPC accumulation in BAd7-209 (Figure 7).

**CONCLUSION**

In the present study, we have characterized the agronomic, cytological, grain protein characteristics, and transcriptomic profile of two advanced wheat lines (BAd7-209 and BAd23) from a cross between CN16 and D97. We found that BAd7-209 and BAd23 had genetic background of common wheat and introgression segments from wild emmer D97. The two advanced lines had contrasting GPC, while both absence of the functional GPC-B1. BAd7-209 had superior processing quality parameters and higher amount of glutenins and gliadins than those of BAd23-1, while their glutenin and gliadin subunits were similar. RNA-seq revealed that the contrasting GPC in BAd7-209 and BAd23-1 may be closely associated with the expression levels of glutenin and gliadin.
genes which regulated by TFs. The protein processing in ER and protein export pathways and related DEGs are important in wheat grain protein accumulation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available in NCBI under accession number PRJNA777016.

AUTHOR CONTRIBUTIONS

FG: formal analysis and writing—original draft preparation. TQ: investigation. YLu: investigation. JL: investigation. XZ: investigation. JH: investigation. YLi: investigation. YZ: supervision. DL: supervision. LH: writing and editing. BW: conceptualization, project administration, funding acquisition.

FUNDING

This work was financially supported by the Key Research and Development Program of Sichuan Province, China (2021YFYZ0002) and the International Cooperation Program of Science & Technology Department of Sichuan Province (2021YFH0110).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.804481/full#supplementary-material

REFERENCES

Albani, D., Hammond-Kosack, M. C., Smith, C., Conlan, S., Colot, V., Holdsworth, M., et al. (1997). The Wheat Transcriptional Activator SPA: A Seed-specific bZIP Protein that Recognizes the GCN4-like Motif in the Bifactorial Endosperm Box of Prolamin Genes. Plant Cell 9, 171–184. doi:10.1105/tpc.9.2.171

Benbow, H. R., Jermiin, L. S., and Doohan, F. M. (2019). Serpins: Genome-Wide Characterisation and Expression Analysis of the Serine Protease Inhibitor Family In Triticum aestivum. G3. 9, 2709–2722. doi:10.1534/g3.119.40444

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. Bioinformatics 30, 2114–2120. doi:10.1093/bioinformatics/btu170

Canu, D., Pearce, S. P., Distelfeld, A., Christiansen, M. W., Uauy, C., Akhunov, E., et al. (2019). Effect of the Down-Regulation of the High Grain Protein Content (GPC) Genes on the Wheat Transcriptome during Monocarpic Senescence. BMC Genomics 12, 492. doi:10.1186/s12864-016-2644-8

Chen, X.-y., Song, G.-q., Zhang, S.-j., Li, Y.-l., Gao, J., Shahidul, I., et al. (2017). The Allelic Distribution and Variation Analysis of the NAM-BI Gene in Chinese Wheat Cultivars. J. Integr. Agric. 16, 1294–1303. doi:10.1007/s13218-016-1017-2

Chen, Y.-x., Song, G.-q., Zhang, S.-j., Li, Y.-l., Gao, J., Shahidul, I., et al. (2017). The Allelic Distribution and Variation Analysis of the NAM-BI Gene in Chinese Wheat Cultivars. J. Integr. Agric. 16, 1294–1303. doi:10.1007/s13218-016-1017-2

Cho, K., Jo, Y. M., Lim, S. H., Kim, J. Y., Han, O., and Lee, J. Y. (2019). Overexpressing Wheat Low-Molecular-Weight Glutenin Subunits in Rice (Oryza Sativa L. Japonica Cv. Koami) Seeds. Biotech. 9, 49. 3. doi:10.1007/s13205-019-1579-x

Crofts, A. J., Leborgne-Castel, N., Pesca, M., Vitale, A., and Denecke, J. (1998). BiP and Calreticulin Form an Abundant Complex That Is Independent of Endoplasmic Reticulum Stress. The Plant Cell 10, 813–824. doi:10.1105/tpc.10.5.813

D’Arcangelo, J. G., Stahmer, K. R., and Miller, E. A. (2013). Vesicle-mediated export from the ER: COPII Coat Function and Regulation. Biochim. Biophys. Acta (Bba) - Mol. Cell Res. 1833, 2464–2472. doi:10.1016/j.bbamcr.2013.02.003

Diaz, I., Vicente-Carbajosa, J., Abraham, Z., Martinez, M., Isabel-La Moneda, I., Carbonero, P., et al. (2002). The GAMYB Protein from barley Interacts with the DOF Transcription Factor BBFV and Activates Endosperm-specific Genes during Seed Development. Plant J. 28, 453–464. doi:10.1046/j.1365-3104.2000.01230.x

Dong, K., Zhen, S., Cheng, Z., Hui, C., Ge, P., and Yan, Y. (2015). Proteomic Analysis Reveals Key Proteins and Phosphoproteins Upon Seed Germination of Wheat (Triticum aestivum L.). Front. Plant Sci., 6, 1017. doi:10.3389/fpls.2015.01017

Duan, S. E., and Zhao, W. M. (2004). Rapid Separation and SDS-PAGE Analysis of Wheat Glutenin Subunits. J. Shaxin Normal Univ. 32, 77–79.
Nitrogen Fertilization. Plant J. 96, 1160–1177. doi:10.1111/tpj.14096
Zhang, L.-Q., Yen, Y., Zheng, Y.-L., and Liu, D.-C. (2007). Meiotic Restriction in Emmer Wheat Is Controlled by One or More Nuclear Genes that Continue to Function in Derived Lines. Sex. Plant Reprod. 20, 159–166. doi:10.1007/s00497-007-0052-x
Zhang, Z., Dong, J., Ji, C., Wu, Y., and Messing, J. (2019). NAC-type Transcription Factors Regulate Accumulation of Starch and Protein in maize Seeds. Proc. Natl. Acad. Sci. USA 116, 11223–11228. doi:10.1073/pnas.1904995116
Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., et al. (2016). iTAK: a Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol. Plant 9, 1667–1670. doi:10.1016/j.molp.2016.09.014
Zhu, J., Hao, P., Chen, G., Han, C., Li, X., Zeller, F. J., et al. (2014). Molecular Cloning, Phylogenetic Analysis, and Expression Profiling of Endoplasmic Reticulum Molecular Chaperone BiP Genes from Bread Wheat (Triticum aestivum L.). BMC Plant Biol. 14, 260. doi:10.1186/s12870-014-0260-0
Zhu, J., Fang, L., Yu, J., Ying, Z., and Xia, G. (2018). 5-Azacytidine Treatment and Tapbf-D Over-Expression Increases Glutenin Accumulation Within the Wheat Grain by Hypomethylating the Glu-1 Promoters. Theoretical and Applied Genetics 131, 735–746. doi:10.1007/s00122-017-3032-z

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.

Copyright © 2022 Gong, Qi, Zhang, Lu, Liu, Zhong, He, Li, Zheng, Liu, Huang and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.