Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

QTL Mapping and Candidate Gene Analysis of Grain Size and Weight in Synthetic Hexaploid Wheat-Derived Lines

Dongxia Wang
Department of Agriculture and Forestry, College of Agriculture and Animal Husbandry, Qinghai University

Chengkai Liu
Department of Agriculture and Forestry, College of Agriculture and Animal Husbandry, Qinghai University

Ming Hao
Triticeae Research Institute, Sichuan Agricultural University

Yunlong Liang
Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences

Yun Li
Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences

Baolong Liu
Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences

Dong Cao
Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences

Guangrui Liu
liuguangrui@nwipb.cas.cn

Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences

---

Research Article

Keywords: synthetic hexaploid wheat, grain size, QTL, haplotype analysis

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11141334/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

QTL Mapping and Candidate Gene Analysis of Grain Size and Weight in Synthetic Hexaploid Wheat-Derived Lines

Dongxia Wang[1], Chengkai Liu[1], Ming Hao[4], Yunlong Liang [2], Yun Li [2], Baolong Liu [2], Dong Cao [2], Guangrui Liu [2, 3*]

[1] Department of Agriculture and Forestry, College of Agriculture and Animal Husbandry, Qinghai University, Xining, Qinghai, 810016, China

[2] Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, China

[3] University of Chinese Academy of Sciences, Beijing, China

[4] Sichuan Agricultural University, Sichuan, China

[*]Correspondence:liuguangrui@nwipb.cas.cn(G.L)

Abstract Synthetic hexaploid wheat (SHW) contains abundant genetic variation, and is the important breeding resource for enhancing genetic diversity of common wheat. Grain size, including thousand-grain weight (TGW), grain length (GL), grain width (GW), grain area (GA), aspect ratio (AR), and factor from density (FFD), is closely related with grain yield and quality in wheat. In the research, Quantitative Trait Loci (QTL) analyses were carried out for a mapping population of 284 F_2 plants derived from a cross between two sister lines derived from synthetic hexaploid wheat in order to identify genomic regions affecting grain size. The linkage map based on 1671 polymorphic Diversity arrays technology (DArT) markers detected 17 QTLs for six grain size traits on chromosomes 1A, 1B, 1D, 3B, 4A, 5A, 5B, and 7D. These QTL explained 3.8–12.4% of the phenotypic variation. For the thousand-grain weight (TGW) loci, *TraesCS1B01G144800* and *TraesCS1B01G145800* were prioritized within *QTgw.nwipb-1B*, while *TraesCS5A01G111700*, encoding a WD40 repeat-containing protein, was identified as a high-confidence candidate within *QTgw.nwipb-5A*. Notably, *TraesCS1B01G144800* showed pronounced grain-preferential expression, supporting its potential involvement in grain development. For the grain-length QTL *QGl.nwipb-3B*, *TraesCS3B01G232800*, encoding an IAA-amino acid hydrolase ILR1, and

TraesCS3B01G227700, encoding a cyclin family protein, were prioritized based on their potential roles in auxin homeostasis and cell-cycle regulation, respectively. Haplotype analysis further revealed favorable haplotypes of *TraesCS1B01G144800*, *TraesCS1B01G145800*, *TraesCS5A01G111700*, and *TraesCS3B01G232800* that were significantly associated with increased TGW or GL. These results provide new insights into the genetic regulation of wheat grain morphology and identify valuable alleles for marker-assisted selection and molecular breeding.

Key words synthetic hexaploid wheat; grain size; QTL; haplotype analysis

Introduction

Wheat is one of the most important food crops and yield improvement is a priority objective as the global population increases. It is projected that the demand for wheat will increase by 50% by 2030 (Ogbonnaya et al. 2013). Common wheat has evolved from two spontaneous hybridization events. The first one involved the diploid wheat *Triticum urartu* (AA, 2n = 2x = 14) and an *Aegilops* species (carrying the B genome), and formed tetraploid wheat (AABB, 2n = 4x = 28). The second one occurred between tetraploid wheat and the diploid goat grass *Aegilops tauschii* (DD, 2n = 2x = 14), and resulted in hexaploid common wheat (AABBDD, 2n = 6x = 42). During the process, an increase of grain size was considered as one of the domestication syndromes (Doebley et al. 2006; Fuller 2007).

Grain size is crucial to breed high-yield cultivar in wheat. Wheat grain size can be quantified as three factors, including grain length (GL), grain width (GW), and grain thickness (Cheng et al. 2017). A larger grain is considered as one of important breeding targets, and has contributed notably to grain yield of wheat (Zhang et al. 2012b). Grain weight, measured as thousand-grain weight (TGW), is key component of wheat grain yield. Therefore, the understanding of the genetic mechanism of these grain size traits is helpful to quickly cultivate higher-yielding wheat cultivars with good grain morphology.

The majority of QTLs for grain size and weight have been mapped. QTLs for TGW almost existed in 21 wheat chromosomes(Campbell et al. 1999; Dholakia et al. 2003; Kato et al. 2000; Shah et al. 1999; Sun et al. 2009). Many QTLs for grain size were detected and distributed on most wheat chromosomes (Cheng et al. 2017; Maphosa et al. 2014; Xie et al. 2015). The genes controlling grain size and weight have been cloned in

Arabidopsis and rice (Heang and Sassa 2012; Li et al. 2012; Li et al. 2011; Wang et al. 2008; Yu et al. 2014). Meanwhile, some homologous genes also were identified in wheat, including *TaCwi-A1* (Ma et al. 2012), *TaSAP1-A1* (Chang et al. 2013), *TaCYP78A3* (Ma et al. 2015); *TaGw2-6A* (Su et al. 2011), *TaSus2-2B* (Jiang et al. 2011), *TaCKX6-D1* (Zhang et al. 2012a), *TaGS1a-6D* (Guo et al. 2013) and *TaLSU1* (Kang et al. 2013).

With the limited genetic variation in wheat itself, it is imperative that a wider spectrum of variability be examined. However, Scientists have proposed an effective approach to increase yield of modern cultivars by using wild relatives of wheat and synthetic wheat (Gororo et al. 2002; Swamy and Sarla 2008). Recently Over 1,500 SHW have been developed globally and about 1300 SHW lines have been produced from crosses of more than 900 *Aegilops tauschii* accessions stored at the International Maize and Wheat Improvement Center (CIMMYT-Mexico) (Ogbonnaya et al. 2013). The hexaploid wheat yield could be significantly improved using SHW and SHW-derived varieties (Li et al. 2014). Grain yield and size of SHW lines had better performance than the common wheat (Calderini and Reynolds 2000; Li et al. 2014). Scientific researchers have identified plenty of QTLs associated with important traits from SHW varieties (lines). For TGW, the introgressed QTLs derived from SHW were detected across 21 chromosomes (Díaz De León et al. 2011; Huang et al. 2004; Li et al. 2008; Liu et al. 2006; Pshenichnikova et al. 2008). But, a few studies have reported QTLs for grain size traits from the SHW varieties (lines) (Rasheed et al. 2014; Yan et al. 2017).

In this paper, the objective was to detect potential QTLs associated with six grain size traits by combing grain phenotypes and DArT markers in a F_2 population constructed by two sister lines derived from synthetic hexaploid wheat.

Materials and Methods

Plant materials and field experiments

Two SHW-derived lines, B2037 and B2039, were selected as parental materials in this study. B2037 exhibited large grains and high thousand-grain weight (TGW), whereas B2039 showed small grains and low TGW (Table 1). Both parental lines were developed by the Triticeae Research Institute of Sichuan Agricultural University and were derived from the same cross combination: SHW-L1/SW8188//Chuanyu 18/3/Chuanmai 42. SHW-L1 is a synthetic hexaploid wheat (SHW) line derived from the hybridization between durum wheat

(*Triticum turgidum*) accession AS2255 and *Aegilops tauschii* accession AS60 (Zhang et al. 2004). SHW-L1 is characterized by late heading and longer spikes (Ma et al. 2014), whereas common wheat progenitors and elite advanced lines generally show superior performance in thousand-grain weight, spikelet number, and kernel width.

In the summer of 2012, B2037 was crossed with B2039 to generate F1 seeds at Xining, Qinghai, China (36° 30′ N, 101° 74′ E). The F_1 plants were grown in Xining and self-pollinated to produce 284 F_2 individuals in 2014. The F_2 plants were grown in 20 rows (1 m in length per row, with 20 cm spacing between rows) and self-pollinated to obtain $F_{2:3}$ seeds in 2015. In 2016, the $F_{2:3}$ families were evaluated using a randomized complete block design (RCBD) with three replicates. Each family was planted in 2-m rows with 20 cm row spacing. Field management practices followed local agricultural practices.

Phenotypic data

The grain traits evaluated in this study included grain length (GL), grain width (GW), grain area (GA), and thousand-grain weight (TGW). The 284 F_2 individuals were evaluated based on measurements from individual plants. For the $F_{2:3}$ families, 15 plants from each row were randomly selected for phenotypic evaluation, and the final phenotypic value was calculated as the mean across three replicates. Grain length, grain width, and grain area were measured using a Marvin grain size analyzer (Marvin, Germany). Thousand-grain weight was determined by weighing three samples of 500 kernels from each replicate. Additional grain-related parameters were calculated using the following formulas: aspect ratio (AR) = GL/GW; factor from density (FFD) = individual grain weight/(GL × GW).

Genotyping for DArT markers and QTL analysis

Genomic DNA was extracted from 284 F_2 plants and their parental lines and subsequently submitted to Diversity Arrays Technology Pty. Ltd. (Yarralumla, Canberra, Australia) for whole-genome genotyping. QTL analysis was performed using the inclusive composite interval mapping (ICIM) method implemented in IciMapping software (version 4.1) (Meng et al. 2015). The analysis was conducted using a 1-cM scanning step and a 0.001 probability level for entering variables in the stepwise regression. The LOD threshold for QTL detection was determined by a permutation test with 1,000 permutations at the 5% significance level, following

the method described by Li *et al* (Li et al. 2015). QTLs with LOD scores exceeding the permutation-derived threshold were considered significant. The proportion of phenotypic variance explained (PVE) and additive effect (A) were estimated for each detected QTL.

Annotation of the DArT markers related with QTLs

DArT marker sequences were obtained from the marker database. For DArT-seq markers significantly associated with the target traits, their corresponding physical positions on the wheat genome were determined by sequence alignment. The marker sequences were then aligned against the Chinese Spring (CS) reference genome (IWGSC RefSeq v1.0) to identify their genomic locations. All annotated genes located within the confidence intervals of the QTL regions were retrieved from the JBrowse database (https://urgi.versailles.inra.fr/jbrowseiwgsc). Candidate genes potentially associated with grain traits were subsequently identified based on their functional annotations and previously reported biological functions.

Haplotype analysis of candidate genes associated with grain traits

Candidate gene coordinates were obtained from the GFF3 annotation file of the gaoyuan448 v3.0(https://gy448.tcuni.com/#/)(Cao et al. 2026). The genomic intervals corresponding to each candidate gene were extracted, and biallelic SNPs located within these intervals were retrieved from the WES-derived VCF file. Both coding and non-coding SNPs within the candidate gene regions were considered for haplotype construction. For each accession, the alleles of the selected SNPs were ordered according to their physical positions along the gene and combined to define the haplotype pattern of the corresponding candidate gene.

Haplotype frequencies were calculated across the 400 wheat accessions. Haplotypes with a frequency of ≥5% were retained for subsequent phenotypic analysis, whereas rare haplotypes with frequencies <5% were excluded because of their limited sample sizes. The haplotype information was then integrated with phenotypic data according to accession IDs. Differences in grain-related traits among haplotype groups were evaluated using one-way analysis of variance (ANOVA). When significant differences were detected, post hoc multiple-comparison tests were performed to determine differences between individual haplotype groups. The phenotypic distributions of the major haplotypes were visualized using boxplots.

Results

Phenotype variations and correlation

Phenotypic variations of the grain traits are presented in Table 1. The two parental lines, B2037 and B2039, exhibited significant differences in grain length (GL), grain width (GW), grain area (GA), aspect ratio (AR), factor from density (FFD), and thousand-grain weight (TGW). In the F2 population, all six traits displayed continuous variation, approximately normal distributions, and transgressive segregation, indicating that these traits were quantitatively inherited Correlation analysis among grain-related traits showed that TGW was significantly and positively correlated ($P < 0.01$) with GL, GW, GA, AR, and FFD. Grain length was significantly and positively correlated ($P < 0.01$) with GA, AR, and FFD. Grain width exhibited a significant positive correlation ($P < 0.01$) with GA but a significant negative correlation with AR. Furthermore, AR was significantly and positively correlated ($P < 0.01$) with FFD (Figure S1).

Table 1. Phenotypic analysis of grain traits in two parents and 284 F_2 individuals

Grain traits	Parents		F_2 population		
	B2037	B2039	Max.	Min.	Average
Thousand-grain weight (TGW) (g)	44.44 a	39.03 b	49.04	34.03	42.32
Grain length (GL) (mm)	6.95 a	6.00 b	7.13	6.25	6.73
Grain width (GW) (mm)	3.10 a	2.76 b	3.63	3.15	3.38
Grain area (GA) (mm^2)	17.51 a	14.42 b	21.18	16.90	19.22
Aspect ratio (AR)	2.24a	2.17a	2.38	1.90	1.99
Factor from density (FFD) (%)	0.21a	0.24a	0.25	0.15	0.19

QTL analysis

Among the 13,553 DArT markers screened for polymorphism between B2037 and B2039, 1,671 polymorphic markers were selected for construction of the linkage map. A total of 17 significant QTLs associated with six grain size traits were identified (Table S1). These QTLs were distributed across eight chromosomes, including 1A, 1B, 1D, 3B, 4A, 5A, 5B, and 7D (Fig. 1).

For thousand-grain weight (TGW), two QTLs (*QTgw.nwipb-1B* and *QTgw.nwipb-5A*) were detected. Both QTLs showed positive additive effects, with LOD scores of 9.3, and explained 2.6% and 12.4% of the phenotypic variation, respectively. Four QTLs were identified for grain length (GL). *QGl.nwipb-3B*,

QGl.nwipb-4A and *QGl.nwipb-5A* accounted for 6.8%, 9.2% and 8.2% of the phenotypic variation, respectively, with the favorable alleles increasing GL contributed by B2037. In contrast, the QTL *QGl.nwipb-1A* on chromosome 1A explained 10.3% of the phenotypic variation and the allele from B2039 reduced GL.

For grain area (GA), eight significant QTLs were detected on chromosomes 1A, 1D, 3B, 5A, 5B, and 7D. Among these, *QAr.nwipb-1D*, *QAr.nwipb-5A* and *QAr.nwipb-5B* were positively associated with GA which explained 3.8%–7.8% of the phenotypic variation. In contrast, the alleles contributed by B2039 at five QTLs on chromosomes 1A, 3B, 5B, and 7D were associated with reductions in GA ranging from 1.0% to 7.0%.

Three additional QTLs were identified for grain width (GW), grain area (GA) and factor from density (FFD), and these QTLs were named *QGw.nwipb-5B*, *QGa.nwipb-3B* and *QFfd.nwipb-5A*. These QTLs were located on chromosomes 5B, 3B, and 5A and explained 5.5%, 7.4%, and 7.9% of the phenotypic variation, respectively. In all three cases, the favorable alleles associated with increased GW, GA, and FFD were contributed by B2037.

Overall, these results demonstrate that grain size-related traits in the B2037 × B2039 population are controlled by multiple loci with diverse genetic effects. Notably, B2037 contributed favorable alleles at the majority of the identified loci, highlighting its potential value as a donor parent for improving grain size-related traits in wheat.

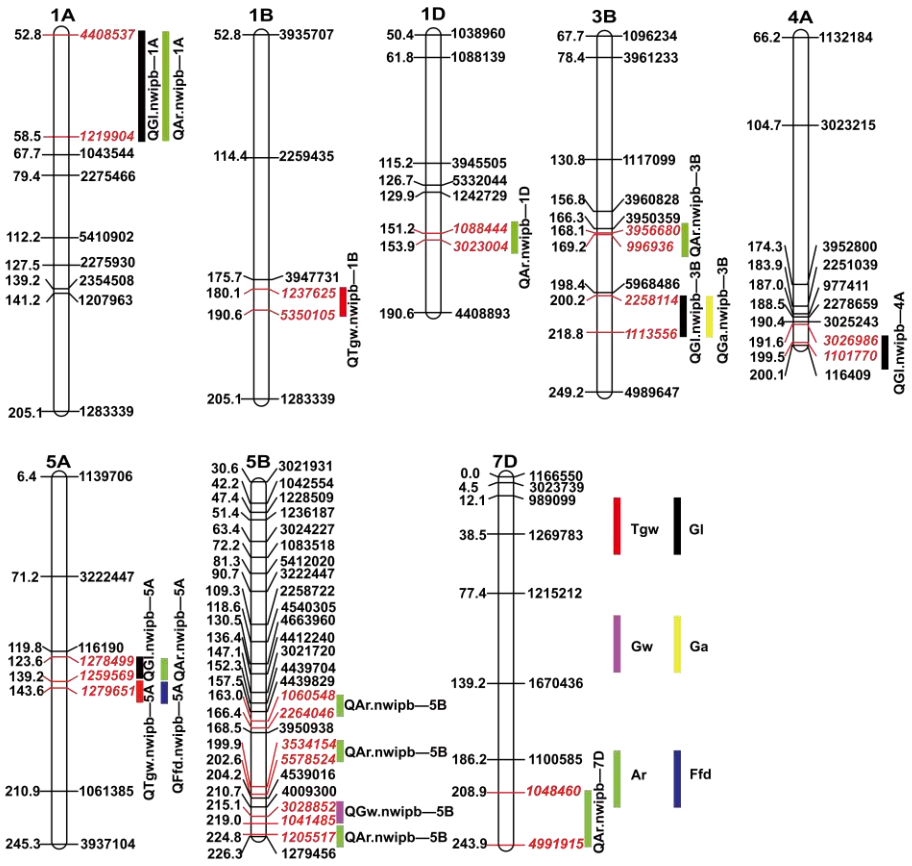


Fig.1 Chromosomal locations of QTLs for grain size and shape detected by DArT marker analysis in lines derived from synthetic hexaploid wheat

Annotation of DArT markers associated with QTLs

A total of 24 DArT markers associated with 17 QTLs were employed for candidate gene identification. The sequences of these markers were aligned to the Chinese Spring reference genome (IWGSC RefSeq v1.0) to determine their physical genomic positions. Based on the genomic locations of the flanking markers, the confidence interval of each QTL was delineated, and all annotated genes within these intervals were retrieved from the JBrowse database. Following functional annotation of the identified genes, candidate genes were prioritized by integrating functional annotations with publicly available gene expression profiles (WheatOmics 2.0) according to their corresponding gene IDs (Table S2).

Based on the confidence interval of *QTgw.nwipb-1B* and the functional annotation of genes within the corresponding chromosome 1B region, *TraesCS1B01G144800* and *TraesCS1B01G145800* were prioritized as

positional candidate genes for thousand-grain weight (TGW). Their transcript abundance was examined across five tissues, with particular attention to grain tissue. *TraesCS1B01G144800*, which is annotated as an AGAMOUS MADS-box transcription factor, exhibited the highest transcript abundance in grains (47.5 TPM), compared with 1.6, 0.1, 0.1, and 13.2 TPM in roots, stems, leaves, and spikes, respectively. In contrast, *TraesCS1B01G145800*, encoding an abscisic acid receptor, showed a lower transcript abundance of 2.7 TPM in grains. The pronounced grain expression of *TraesCS1B01G144800*, together with its predicted function as a MADS-box transcription factor, suggests a potential role in the regulation of grain development. Although *TraesCS1B01G145800* displayed relatively low expression in grains, its annotation as an abscisic acid receptor implicates it as a potential component of hormone-mediated regulatory pathways associated with grain development and yield formation.

Within the confidence interval of *QTgw.nwipb-5A*, only one annotated gene, *TraesCS5A01G111700*, was identified, making it a high-confidence positional candidate. This gene spans 215,475,309–215,484,896 bp and encodes a WD40 repeat-containing protein. Transcriptome profiling across roots, stems, leaves, spikes, and grains revealed that *TraesCS5A01G111700* was expressed in all five tissues, with transcript abundances of 4.1, 3.0, 0.6, 1.5, and 2.7 TPM, respectively. Although no grain-preferential expression pattern was observed, its detectable expression in grains suggests that this gene may be involved in grain development. Given that *TraesCS5A01G111700* was the sole annotated gene within the *QTgw.nwipb-5A* confidence interval, together with its expression in grains and predicted WD40 repeat-containing protein function, it represents a strong positional candidate underlying the *QTgw.nwipb-5A* locus and warrants further functional investigation.

For the grain-length QTL *QGl.nwipb-3B*, candidate genes were investigated within the QTL interval delimited by DArT markers 2258114 and 1113556, with the maximum LOD score detected at 218 cM. Candidate prioritization was based on functional annotation and tissue-specific expression profiles. Among the annotated genes in this interval, *TraesCS3B01G232800* and *TraesCS3B01G227700* showed comparatively strong functional relevance to grain growth and development. *TraesCS3B01G232800*, annotated as an IAA-amino acid hydrolase ILR1, exhibited relatively high transcript abundance in grains (11.8 TPM), suggesting a potential role in auxin homeostasis and, consequently, grain development. *TraesCS3B01G227700*, which encodes a cyclin family protein, was also prioritized based on its predicted function in cell-cycle regulation, a

biological process closely associated with cell proliferation and organ growth. Taken together, the functional annotations and expression profiles support *TraesCS3B01G232800* and *TraesCS3B01G227700* as promising positional candidate genes within *QGl.nwipb-3B*.

Haplotype analysis of candidate genes associated with TGW and GL

Haplotype analysis of the thousand-kernel weight (TGW) candidate gene *GY448v3.1B.HC.G00000014454* (*TraesCS1B01G145800*) 206,692,231-206,693,822 bp identified four haplotypes, among which Hap2 showed the highest TGW and represented the most favorable haplotype. Hap3 exhibited the second-highest TGW and was also classified as a favorable haplotype associated with increased TGW (Fig.2A). *GY448v3.1B.HC.G00000014438* (*TraesCS1B02G144800*) 204,577,896-204,582,456 bp haplotype analysis revealed that Hap2 was favorable haplotypes for TGW, although Hap3 shows better phenotypic values, but the sample size of natural populations needs to be expanded to further verify the stability and statistical significance of its effects (Fig.2B). For *GY448v3.5A.HC.G00000092375* (*TraesCS5A01G111700*) 221,071,051-221,080,828 bp haplotype analysis revealed that Hap2 was favorable haplotypes for TGW, and showed an extremely significant difference from the major haplotype Hap1 (P <0.01). With a sufficient sample size (n=79), it is a superior haplotype regulating 1000-grain weight at this locus and can be applied for molecular marker-assisted selection (Fig.2C).

For the grain length candidate *geneGY448v3.3B.HC.G00000059449* (*TraesCS3B01G232800*) 390,321,733-390,327,149 bp haplotype analysis showed that Hap1 was the favorable haplotype for grain length. Hap1 exhibits the highest peak for grain-length phenotype and has the potential to become a superior haplotype. If significant differences are verified by expanding the sample size of natural populations in follow-up studies, its favorable regulatory effect on grain length can be further confirmed (Fig.2D).

Collectively, these results identified several favorable haplotypes associated with TGW and grain length, providing valuable genetic resources for further wheat improvement.

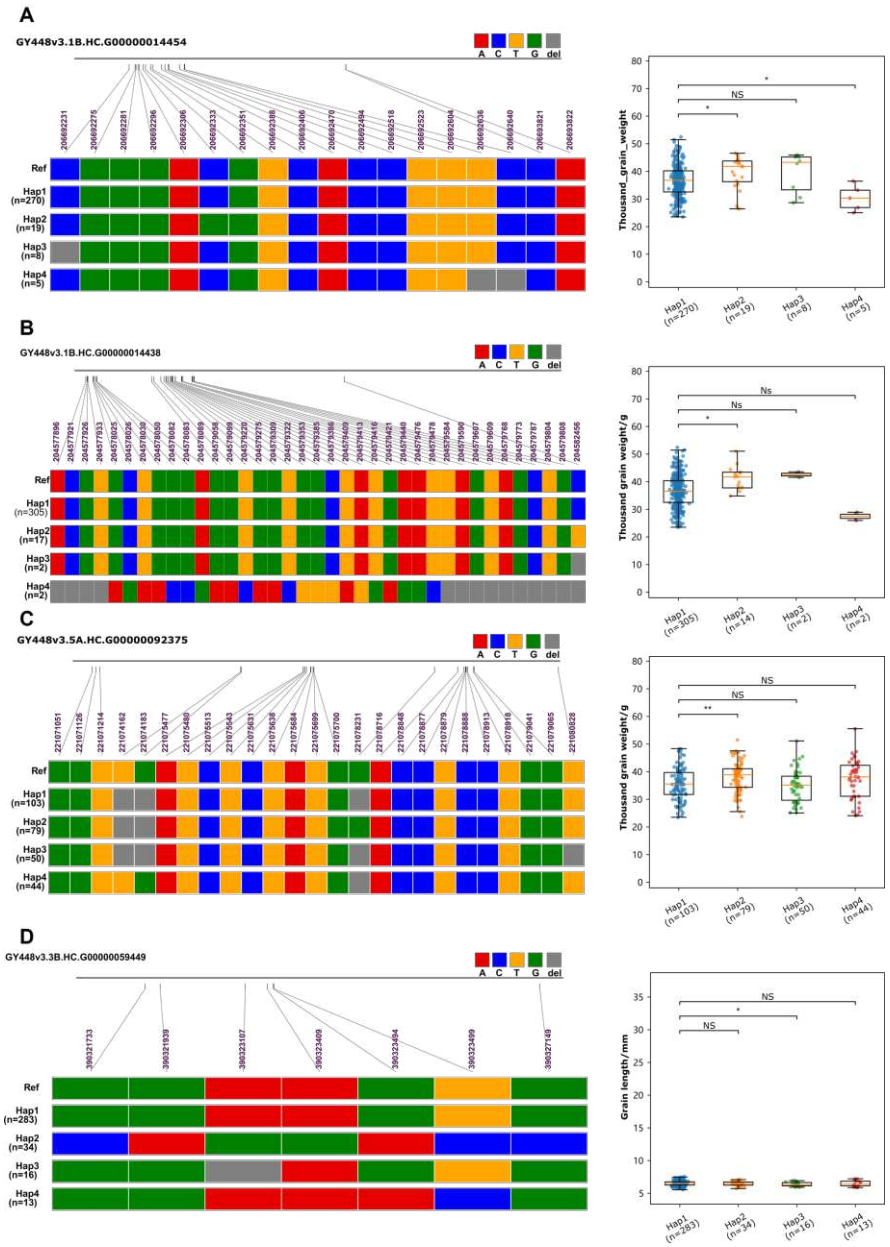


Fig.2 Haplotype analysis of candidate QTLs genes

Discussion

Grain morphology has long been an important target of wheat breeding, and domestication and modern selection have driven a transition from primitive long and narrow grains toward wider and shorter grains in modern cultivars (Gegas et al. 2010). Grain size and shape are major determinants of wheat yield, particularly through their contribution to thousand-grain weight (TGW) variation (Breseghello and Sorrells 2007). Although numerous QTLs controlling grain-related traits have been identified, the complex genetic architecture underlying

grain size and weight remain incompletely understood, particularly in wheat with its large and polyploid genome.

In this study, QTL analysis using an F_2 population derived from two sister lines originating from synthetic hexaploid wheat (SHW) identified loci associated with grain size, shape, and weight on eight chromosomes. Several detected QTLs overlapped with previously reported genomic regions, including TGW-related loci on chromosomes 1B and 5A, suggesting the reliability of these regions in regulating grain weight variation. The chromosome 1B region is particularly noteworthy, as *TaAGP-S1-1B*, previously reported to influence TGW variation (Hou et al. 2017), is located within this region, indicating its potential contribution to grain weight regulation.

Multiple QTLs associated with grain length (GL), grain width (GW), grain area (GA), aspect ratio (AR), and factor from density (FFD) were also identified. Several loci corresponded to previously reported genomic regions, whereas some may represent novel genetic factors contributing to grain morphology. The identification of pleiotropic QTLs further highlighted the complex genetic relationships among grain traits. For example, QTLs on chromosomes 5A and 3B simultaneously affected multiple traits, suggesting coordinated regulation of grain development. Correlation analysis revealed a significant relationship between GL and TGW, supporting previous observations that grain length contributes substantially to grain weight variation (Campbell et al. 1999). However, the absence of a significant association between GW and TGW in this population suggests that the contribution of individual grain dimensions to yield-related traits may depend on specific genetic backgrounds.

For *QTgw.nwipb-1B*, *TraesCS1B01G144800*, encoding an AGAMOUS MADS-box transcription factor, showed pronounced grain-preferential expression (47.5 TPM), supporting its potential role in grain developmental regulation. Class-C MADS-box gene *AGAMOUS* (*AG*) specifies stamen and carpel identity, represses class-A gene expression, and maintains floral meristem identity in *Arabidopsis thaliana*. In rice, two class-C paralogs *OsMADS3* and *OsMADS58* arose from a duplication event predating rice-maize divergence and have undergone partial sub functionalization(Yamaguchi et al. 2006). *osmads3* knockout mutants display homeotic transformation of stamens into lodicules and ectopic second-whorl lodicules, but retain normal carpel development. By contrast, *OsMADS58* silencing causes severe floral meristem determinacy defects with reiterated formation of lodicules, stamens and carpel-like structures. The *mads3 mads58* double mutant

completely loses reproductive organ identity, with abundant lodicules accumulating in the third and fourth whorls; together with *OsMADS13*, both genes are essential for floral meristem fate (Dreni et al. 2011). Epigenetically, *OsMADS58* is silenced via H3K27 trimethylation, and CCP1 maintains this modification to regulate rice palea development (Yan et al. 2015). *TraesCS1B01G145800*, annotated as an abscisic acid receptor, represents a potential regulatory candidate linking ABA signaling with grain development. Although its transcript abundance in grain was relatively low (2.7 TPM), PYL-type ABA receptors function as central components of the ABA–PP2C–SnRK2 signaling pathway, which regulates plant development, seed maturation, and stress responses (Ma et al. 2009; Park et al. 2009; Tuan et al. 2018). For *QTgw.nwipb-5A*, only *TraesCS5A01G111700*, encoding a WD40 repeat-containing protein, was identified within the confidence interval. Although it was not grain-preferentially expressed, its detectable expression in grains and its status as the sole annotated gene within the interval provide strong positional support for its candidacy. WD40 proteins are widely involved in developmental and regulatory processes, although functional validation will be required to establish its contribution to TGW.

For the grain-length QTL *QGl.nwipb-3B*, *TraesCS3B01G232800* and *TraesCS3B01G227700* were prioritized. *TraesCS3B01G232800*, encoding IAA-amino acid hydrolase ILR1, showed relatively high grain expression (11.8 TPM), suggesting a possible role in auxin homeostasis and grain growth. *TraesCS3B01G227700*, encoding a cyclin family protein, provides a complementary candidate because cell-cycle regulation and cell proliferation are fundamental processes underlying organ growth. Together, these candidates suggest that auxin signaling and cell proliferation may contribute to grain-length variation at *QGl.nwipb-3B*. This is consistent with studies demonstrating that grain size in cereals is regulated by multiple developmental pathways involving hormonal signaling and cell proliferation(Gegas et al. 2010; Sreenivasulu and Schnurbusch 2012).

Overall, the candidate genes identified in this study point to distinct but interconnected mechanisms underlying wheat grain traits, including reproductive development, ABA signaling, auxin homeostasis, and cell-cycle regulation. However, candidate-gene prioritization based on positional information and transcript abundance does not establish causality. Further validation using developmental-stage-specific expression analysis, sequence and haplotype analysis, and functional assays will be necessary to determine whether these

genes directly contribute to natural variation in TGW and grain length.

Haplotype analysis further linked the identified candidate genes with phenotypic variation in wheat grain traits and highlighted their potential breeding value. For TGW, *TraesCS1B01G145800* showed clear haplotypic effects, with Hap2 exhibiting the highest TGW and Hap3 the second-highest, suggesting that these haplotypes may contribute to increased grain weight. For *TraesCS1B01G144800*, Hap2 was identified as a favorable haplotype, whereas the higher TGW observed for Hap3 requires validation in a larger population to determine its stability and statistical significance. Notably, *TraesCS5A01G111700* Hap2 showed an extremely significant difference from the major Hap1 ($P < 0.01$) in a relatively large population (n = 79), providing stronger evidence that it is a superior haplotype for TGW and a promising target for marker-assisted selection.

For grain length, *TraesCS3B01G232800* Hap2 displayed the highest phenotypic value, indicating its potential favorable effect on grain elongation. However, further validation using larger and genetically diverse populations is required to confirm the robustness of this association. Overall, these results demonstrate that haplotype analysis can refine candidate-gene associations by identifying favorable allelic combinations with distinct phenotypic effects. In particular, the validated or strongly supported favorable haplotypes identified here could provide useful genetic resources for marker-assisted and haplotype-based breeding. Further validation across independent populations and environments, together with functional characterization, will be important for confirming their effects and facilitating their application in wheat yield improvement.

Conclusion

In this study, QTL mapping was conducted in an F_2 population derived from two synthetic hexaploid wheat (SHW)-derived sister lines with contrasting grain phenotypes. A total of 17 QTLs associated with grain size, shape, and weight were identified on eight chromosomes, demonstrating the complex genetic architecture underlying wheat grain-related traits. Several QTLs were located in genomic regions previously reported to influence grain traits, while others may represent additional genetic factors contributed by SHW-derived germplasm. Candidate gene analysis further prioritized *TraesCS1B01G144800* and *TraesCS1B01G145800* within *QTgw.nwipb-1B*, *TraesCS5A01G111700* within *QTgw.nwipb-5A*, and *TraesCS3B01G232800* and *TraesCS3B01G227700* within *QGl.nwipb-3B*. Their functional annotations and expression profiles suggest that

reproductive development, ABA signaling, auxin homeostasis, and cell-cycle regulation may contribute to the genetic regulation of grain weight and grain length. Haplotype analysis additionally identified favorable haplotypes associated with increased TGW or GL, particularly the Hap2 of *TraesCS5A01G111700*, which showed a significant association with increased TGW and represents a promising target for marker-assisted selection. Collectively, this study expands our understanding of the genetic basis of grain morphology in SHW-derived wheat and provides favorable alleles and candidate genes with potential value for wheat improvement. Further validation across independent populations and environments, together with developmental expression analysis and functional characterization, will be necessary to confirm the causal roles of these candidate genes and to facilitate their application in molecular breeding.

Author Contributions Guangrui Liu: designed the research; Dongxia Wang: conducted the experiments; Chengkai Liu, Ming Hao and Yunlong Liang: performed data analysis and prepared the figures; Yun Li and Baolong Liu: supervised the research; Dong Cao: reviewed and edited the manuscript and acquired funding. All authors have read and agreed to the published version of the manuscript.

Funding This work was supported by the Qinghai Provincial Department of Science and Technology (grant number: 2025-QY-216).

Data Availability Data are contained within the article and Supplementary Materials.

Supplementary Materials Table S1: QTLs detected for grain size traits in 284 F_2 population; Table S2: Candidate Genes and Expression Levels of QTLs; Table S3: Candidate gene haplotype analysis interval (gaoyuan448 v3.0 as the reference genome); Figure S1. Phenotypic correlation coefficients among grain size traits in 284 F_2 individuals.

Declarations

Conflicts of Interest The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QTL	Quantitative trait locus
LOD	Logarithm of Odds
PVE	Proportion of phenotypic variance explained
DArT	Diversity Arrays Technology
SNP	Single nucleotide polymorphism
GL	grain length
GW	grain width
GA	grain area
TGW	thousand-grain weight
AR	aspect ratio = GL/GW
FFD	factor from density= individual grain weight/(GL × GW).
TPM	Transcripts per kilobase million

Reference

Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field crops research 101:172-179

Calderini DF, Reynolds MP (2000) Changes in grain weight as a consequence of de-graining treatments at pre-and post-anthesis in synthetic hexaploid lines of wheat (*Triticum durum* x *T. tauschii*). Australian Journal of Plant Physiology 27:183-191

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft× hard wheat cross. Crop science 39:1184-1195

Cao D, Liu G, Liang Y, Yao Y, Wang H, Li Y, Liu B (2026) Construction of a high-quality indexed EMS mutant library and a pipeline for identifying causal mutations in wheat by whole-exome sequencing. Journal of genetics and genomics= Yi chuan xue bao:S1673-8527 (1626) 00215-00218

Chang J, Zhang J, Mao X, Li A, Jia J, Jing R (2013) Polymorphism of *TaSAP1-A1* and its association with agronomic traits in wheat. Planta 237:1495-1508

Cheng R, Kong Z, Zhang L, Xie Q, Jia H, Yu D, Huang Y, Ma Z (2017) Mapping QTLs controlling kernel dimensions in a wheat inter-varietal RIL mapping population. Theoretical and Applied Genetics 130:1405-1414

Dholakia B, Ammiraju J, Singh H, Lagu M, Röder M, Rao V, Dhaliwal H, Ranjekar P, Gupta V, Weber W (2003) Molecular marker analysis of kernel size and shape in bread wheat. Plant Breeding 122:392-395

Díaz De León JL, Escoppinichi R, Geraldo N, Castellanos T, Mujeeb-Kazi A, Röder MS (2011) Quantitative trait loci associated with salinity tolerance in field grown bread wheat. Euphytica 181:371-383

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309-1321

Dreni L, Pilatone A, Yun D, Erreni S, Pajoro A, Caporali E, Zhang D, Kater MM (2011) Functional analysis of all AGAMOUS subfamily members in rice reveals their roles in reproductive organ identity determination and meristem determinacy. The Plant Cell 23:2850-2863

Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. Annals of Botany 100:903-924

Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape variation in wheat. The Plant Cell 22:1046-1056

Gororo NN, Eagles HA, Eastwood RF, Nicolas ME, Flood RG (2002) Use of Triticum tauschii to improve yield of wheat in low-yielding environments. Euphytica 123:241-254

Guo Y, Sun J, Zhang G, Wang Y, Kong F, Zhao Y, Li S (2013) Haplotype, molecular marker and phenotype effects associated with mineral nutrient and grain size traits of TaGS1a in wheat. Field Crops Research 154:119-125

Heang D, Sassa H (2012) An atypical bHLH protein encoded by POSITIVE REGULATOR OF GRAIN LENGTH 2 is involved in controlling grain length and weight of rice through interaction with a typical bHLH protein APG. Breeding science 62:133-141

Hou J, Li T, Wang Y, Hao C, Liu H, Zhang X (2017) ADP-glucose pyrophosphorylase genes, associated with kernel weight, underwent selection during wheat domestication and breeding. Plant Biotechnology Journal 15:1533-1543

Huang XQ, Kempf H, Ganal MW, Roder MS (2004) Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (*Triticum aestivum* L.). Theor Appl Genet 109:933-943

Jiang Q, Hou J, Hao C, Wang L, Ge H, Dong Y, Zhang X (2011) The wheat (*T. aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is associated with yield traits. Functional & integrative genomics 11:49-61

Kang G, Liu G, Peng X, Wei L, Wang C, Zhu Y, Ma Y, Jiang Y, Guo T (2013) Increasing the starch content and grain weight of common wheat by overexpression of the cytosolic AGPase large subunit gene. Plant Physiology and Biochemistry 73:93-98

Kato K, Miura H, Sawada S (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theoretical and Applied Genetics 101:1114-1121

Li J, Chu H, Zhang Y, Mou T, Wu C, Zhang Q, Xu J (2012) The rice HGW gene encodes a ubiquitin-associated (UBA) domain protein that regulates heading date and grain weight. PloS one 7:e34231

Li J, WAN HS, YANG WY (2014) Synthetic hexaploid wheat enhances variation and adaptive evolution of bread wheat in breeding processes. Journal of Systematics and Evolution 52:735-742

Li S, Wang J, Zhang L (2015) Inclusive composite interval mapping of QTL by environment interactions in biparental populations. PLoS One 10:e0132414

Li XJ, Li LQ, Wang H, Sorrells ME (2008) QTL Mapping for Kernel Weight Using *GW3-1* and *IND109* Markers in Common Wheat (*Triticum aestivum* L .). Acta Bot Boreal Occident Sin 28:1106-1111

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nature genetics 43:1266-1269

Liu SB, Zhou RH, Dong YC, Li P, Jia JZ (2006) Development utilization of introgression lines using a synthetic wheat as donor. Theor Appl Genet 112:1360-1373

Ma D, Yan J, He Z, Wu L, Xia X (2012) Characterization of a cell wall invertase gene *TaCwi-A1* on common wheat chromosome 2A and development of functional markers. Molecular Breeding 29:43-52

Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, Song W, Appels R, Zhao H (2015) Expression of *TaCYP78A3*, a gene encoding cytochrome P450 CYP 78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. The Plant Journal 83:312-325

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E (2009) Regulators of PP2C phosphatase activity function as abscisic acid sensors. Science 324:1064-1068

Ma Y, ZHANG L-q, ZHANG L, ZHENG Y-l (2014) QTL mapping for important agronomic traits in synthetic hexaploid wheat derived from *Aegiliops tauschii* ssp. *tauschii*. Journal of Integrative Agriculture 13:1835-1844

Maphosa L, Langridge P, Taylor H, Parent B, Emebiri LC, Kuchel H, Reynolds MP, Chalmers KJ, Okada A, Edwards J (2014) Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. Theoretical and applied genetics 127:1607-1624

Meng L, Li H, Zhang L, Wang J (2015) QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. The Crop Journal 3:269-283

Ogbonnaya FC, Abdalla O, Mujeeb-Kazi A, Kazi AG, Xu SS, Gosman N, Lagudah ES, Bonnett D, Sorrells ME, Tsujimoto H (2013) Synthetic hexaploids: harnessing species of the primary gene pool for wheat improvement. Plant Breed Rev 37:35-122

Park S-Y, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow T-fF (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. science 324:1068-1071

Pshenichnikova TA, Ermakova MF, Chistyakova AK, Shchukina LV, Berezovskaya EV, Lochwasser U, Röder M, Börner A (2008) Mapping of the quantitative trait loci (QTL) associated with grain quality characteristics of the bread wheat grown under different environmental conditions. RUSS J GENET+ 44:74-84

Rasheed A, Xia X, Ogbonnaya F, Mahmood T, Zhang Z, Mujeeb-Kazi A, He Z (2014) Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. BMC plant biology 14:128

Shah M, Gill K, Baenziger P, Yen Y, Kaeppler S, Ariyarathne H (1999) Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. Crop Science 39:1728-1732

Sreenivasulu N, Schnurbusch T (2012) A genetic playground for enhancing grain number in cereals. Trends in plant science 17:91-101

Su Z, Hao C, Wang L, Dong Y, Zhang X (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 122:211-223

Sun X-Y, Wu K, Zhao Y, Kong F-M, Han G-Z, Jiang H-M, Huang X-J, Li R-J, Wang H-G, Li S-S (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615-624

Swamy BP, Sarla N (2008) Yield-enhancing quantitative trait loci (QTLs) from wild species. Biotechnol Adv 26:106-120

Tuan PA, Kumar R, Rehal PK, Toora PK, Ayele BT (2018) Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals. Frontiers in plant science 9:668

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nature genetics 40:1370-1374

Xie Q, Mayes S, Sparkes DL (2015) Carpel size, grain filling, and morphology determine individual grain weight in wheat. Journal of experimental botany 66:6715-6730

Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano H-Y (2006) Functional diversification of the two C-class MADS box genes *OSMADS3* and *OSMADS58* in *Oryza sativa*. The Plant Cell 18:15-28

Yan D, Zhang X, Zhang L, Ye S, Zeng L, Liu J, Li Q, He Z (2015) CURVED CHIMERIC PALEA 1 encoding an EMF 1-like protein maintains epigenetic repression of *OsMADS58* in rice palea development. The Plant Journal 82:12-24

Yan L, Liang F, Xu H, Zhang X, Zhai H, Sun Q, Ni Z (2017) Identification of QTL for grain size and shape on the D genome of natural and synthetic allohexaploid wheats with near-identical AABB genomes. Frontiers in Plant Science 8:1705

Yu F, Li J, Huang Y, Liu L, Li D, Chen L, Luan S (2014) FERONIA receptor kinase controls seed size in *Arabidopsis thaliana*. Molecular plant 7:920-922

Zhang L, Liu D, Yan Z, Lan X, Zheng Y, Zhou Y (2004) Rapid changes of microsatellite flanking sequence in the allopolyploidization of new synthesized hexaploid wheat. Science in China Series C: Life Sciences 47:553-561

Zhang L, Zhao YL, Gao LF, Zhao GY, Zhou RH, Zhang BS, Jia JZ (2012a) *TaCKX6-D1*, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. New Phytologist 195:574-584

Zhang X, Wang J, Huang J, Lan H, Wang C, Yin C, Wu Y, Tang H, Qian Q, Li J (2012b) Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. Proceedings of the National Academy of Sciences 109:21534-21539

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable.xlsx
- SupplementaryFigure.docx