QTL localisation of seed-related traits in Tibetan hulless barley based on a high-density single-nucleotide polymorphism genetic map

Yan $Wang^{1,2}$, Xiaohua Yao^{1,2}*, Youhua Yao^{1,2}, Yixiong Bai^{1,2}, Likun $An^{1,2}$, Xin $Li^{1,2}$, Yongmei $Cui^{1,2}$, Kunlun $Wu^{1,2}$ *

Citation: Wang Y., Yao X.H., Yao Y.H., Bai Y.X., An L.K., Li X., Cui Y.M., Wu K.L. (2023): QTL localisation of seed-related traits in Tibetan hulless barley based on a high-density single-nucleotide polymorphism genetic map. Czech J. Genet. Plant Breed., 59: 95–108.

Abstract: The effective use of high-quality and high-yielding germplasm resources is of great importance for the development of hulless barley varieties. Therefore, the search for such resources has long been a goal of the breeding community. In this study, a genotyping-by-sequencing (GBS) analysis was performed on an F2 (Nierumuzha × Kunlun 10) population. A high-density genetic map of hulless barley was constructed, which contained 1 475 efficient single-nucleotide polymorphism markers with 7 052 bin markers. The total length of the seven chromosomes was 1 139.4 cM, with an average marker distance of 0.59 cM. Based on this high-density linkage map, a total of 54 quantitative trait loci (QTLs) related to the seed traits were detected, including seed colour (SC), thousand kernel weight (TKW), seed average area (SAA), seed perimeter (SP), seed length (SL), seed width (SW), seed length/width (SLW), seed diameter (SD), and seed circular degree (SCD). These QTLs explained 3.32-35.73% (mean = 11.45%) of the phenotypic interpretation, including 24 main QTLs and 30 epistatic QTLs. A total of 24 candidate genes were identified within the QTL region, including one SC-associated transcription factor (ANT1), two TSW-related genes, a transcription factor (NAC021), a gene associated with the non-homologous end joining (NHEJ) pathway (ku70), three SAA-associated genes (LOL2, NAC021, TSK), two SL-associated genes (MADS21, MADS4), six SW-associated genes (FIP1, NAC021, DREB 1A, HVA22A, CYP78A6, SAUR71), five LW-related genes (NAM-B2, CRY1, LHY, CYP710A1, WRKY72), two SP-related genes (SKIP11, TCP18), two SD-related genes (NAC021, SKIP8), and three SCD-related genes (MYB1R1, RAX3, NAC100). These genes are involved in the regulation of the cell development, material transport, signal transduction, and plant morphogenesis and play an important role in the regulation of agronomic traits in hulless barley. The high-density genetic mapping and QTL identification of the seed traits in hulless barley provide a valuable genetic resource and the basis for further molecular marker-assisted selection and genomic studies.

Keywords: genotyping-by-sequencing (GBS); linkage analysis; quantitative trait loci; seed traits

Hulless barley (*Hordeum vulgare* L. var. *nudum* Hook. f.), one of the most important crops on the Tibetan plateau, is a common barley species be-

longing to the grass family. It is also referred to as naked barley due the ease by which the lemmas are shed when the glumes mature (Yao et al. 2022). The

Supported by the Natural Science Foundation of China (31960427), the Construction Project for Innovation Platform of Qinghai province (2022-ZJ-Y01) and the Agriculture Research System of China (CARS-05).

¹Qinghai University, Xining, P.R. China

²Laboratory for Research and Utilization of Qinghai Tibet Plateau Germplasm Resources, Qinghai Subcenter of National Hulless Barley Improvement, Qinghai Key Laboratory of Hulless Barley Genetics and Breeding, Qinghai Academy of Agricultural and Forestry Sciences, Xining, P.R. China

^{*}Corresponding authors: yaoxiaohua009@126.com; wklqaaf@163.com

[©] The authors. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

hulless barley grain has the characteristics of "three highs and two lows" (high in protein, high in fibre, and high in vitamins, and low in fat and low in sugar) and has high nutritional value and medicinal health effects (Han & Hong 2020). It is also an important raw material for brewing, food processing, and livestock farming (Wang et al. 2022). Research has found that compared to common hulless barley, coloured hulless barley is richer in anthocyanins, phenolic compounds, proteins, and some trace elements, which has led to increased focus on coloured hulless barley (Bellido & Beta 2009).

The majority of the agronomic traits are contained in the QTLs, and studying these loci are of great importance for determining the genetic basis of the agronomic traits and for molecular breeding and improvement (Zhang & Zhang 2017). The construction of high-quality ultra-high density molecular genetic linkage maps are the basis for fine QTL localisation and molecular marker-assisted selection for important traits such as the yield, quality, and resistance in plants (Liu et al. 2017). Importantly, the higher the number and density of the markers, the more accurate and useful the genetic maps are. Genotyping-by-sequencing (GBS) has been successfully applied to the development of single nucleotide polymorphism (SNP) markers and the construction of high-density molecular genetic maps in wheat, maize, rice, and other important crops (Qiao et al. 2019; Wang et al. 2020b; Yang et al. 2021). The greatest advantage of this technique is that the construction of genetic maps is not limited by the reference genome, and thus plants without a reference genome can be genotyped and SNP markers can be developed on a large scale (Xue et al. 2020). The number of molecular markers determines the quality of the genetic map, thus the substantial amount of work required for genotyping populations using traditional molecular markers has hindered the development of genetics. Bin markers can greatly reduce the cost of the analysis compared to SNP markers while still maintaining the map quality (Schafleitner et al. 2016). By contrast, plant genetic linkage maps constructed with traditional molecular markers, such as rapid fragment length polymorphisms, amplified fragment length polymorphisms, and simple sequence repeats, often have large linkage gaps or more linkage breakpoints due to the limited number of markers, and these methods are prone to manual errors when counting bands based on gel electrophoresis (Chen et al. 2014). In addition, these methods are time-consuming and costly to map, and also require the selection of large populations, making it difficult to meet the needs of subsequent studies such as fine QTL localisation (Song et al. 2022). New SNP markers developed with the GBS technology can compensate for these traditional molecular markers; GBS has been successfully used to construct high-density genetic maps in many plants (Liang et al. 2022; Hussain et al. 2017). However, there have been few reports on the use of GBS for genetic map construction in hulless barley. Previous studies on hulless barley have mainly focused on the physiological and biochemical analysis as well as the nutritional quality, but little genetic mapping was performed. Yao et al. (2021) isolated arabinoxylan HBAX-60 from the whole grains of hulless barley by ethanol precipitation and investigated its physicochemical properties and structure. Nie et al. (2021) purified a homogeneous water-soluble polysaccharide from Tibetan hulless barley 25 and investigated its structure and in vitro fermentation profile. However, some scholars have carried out a genetic mapping analysis on the seed coat colour, ripening-stage related traits, Yao et al. (2018) constructed a high-density barley genetic map by simplified genome sequencing (GBS-SEQ) using doubled haploid (DH) lines, then obtained six intervals associated with purple grains of hulless barley. Li et al. (2021) constructed a high-density linkage map containing 1 243 recombinant bin markers using F₂ lines containing 146 populations to map the early heading date of hulless barley (Li et al. 2021). It can be seen that using GBS to construct a genetic map in hulless barley to obtain some genes associated with agronomic traits is desirable.

To date, there have been few reports on the use of the GBS technology to construct a high-density genetic map of hulless barley, and even fewer reports on the use of the GBS technology to construct a high-density genetic map of coloured hulless barley. In the present study, an F_2 segregating population was constructed based on a cross between Nierumuzha (purple) and Kunlun 10 (white) barley, and the F_2 segregating population was used as the experimental material to construct a genotyping library where genetic linkage maps were used to locate the QTLs for certain agronomic traits in hulless barley. A high-density genetic linkage map of the SNPs in hulless barley was constructed, and QTL analyses of the major seed phenotypic traits were carried out using the F_2 population.

MATERIAL AND METHODS

Plant materials. The F_2 cross between the purple-grain hulless barley variety Nierumuzha and the white-grain hulless barley variety Kunlun 10 was used as the material, and the material was planted at an experimental field at the Academy of Agricultural and Forestry Sciences, Qinghai University, China. In the F_2 population, the mean trait values were averaged from up to three repeats from each genotype. The means were used for the QTL identification of the seed traits.

Seed-related traits investigation and data analysis

The seed traits (seed colour, SC; thousand kernel weight, TKW; seed average area, SAA; seed length, SL; seed width, SW; seed length/width, SLW; seed perimeter, SP; seed diameter, SD; seed circular degree, SCD) were assessed in 2017 using the Wseen seed test system SC-G (Wseen, Hangzhou, China). Excel 2007 and SPSS software were used to calculate the frequency distribution, standard errors and make normal distribution graphs for the above data.

DNA extraction. The total gDNA was extracted from the young leaves of Nierumuzha and Kunlun 10 using a plant DNA kit (TIANGEN DP320, Beijing, China). The absorbance of the DNA was measured by a miniature nucleic acid protein instrument, and the bands were detected by gel electrophoresis with 1.2% agarose.

Library construction. For the GBS library construction, the genome was first digested using a combination of MseI+NlaIII+EcoRI digests, followed by amplification of each sample with barcode connectors, following which the samples were mixed to select the desired fragments for the library construction.

Sequencing and sequencing data quality assessment. Paired-end 142 sequencing was performed using the Illumina HiSeq sequencing platform. The sequencing data (i.e., raw data or raw reads) were then rigorously filtered to obtain high-quality clean data using the following three methods: the pair of reads containing splice sequences was filtered out; paired reads were removed when the N content of a singleend sequencing read exceeded 10% of the length of the read; and paired reads were removed when the number of low-quality $(Q \le 5)$ bases in a single-end sequencing read exceeded 50% of the length of the read. The final 148 offspring and parental sequencing data were counted, including the sequencing data volume, data yield, sequencing error rate, quality (Phred) score of 20 (Q20), quality score of 30 (Q30), and GC content. Comparative analysis. Sequencing data from two parents (GBS) and 140 offspring (GBS) were aligned to the reference genome. The paired-end (PE) reads of the parental and offspring clean data were matched against the reference genome using Burrows-Wheeler Aligner (BWA) matching software (Li & Durban 2009). The results were formatted and converted to SAM/BAM files. Perl scripts were used to calculate the matching rate and coverage. SAMtools was used to sort the results (parameter: sort) for the variant detection.

Population SNP detection. The parents and 140 offspring were tested for SNPs using a populationbased SNP assay based on the BWA comparisons. The analysis steps were as follows: filtering of the BWA alignment results, whereby reads that were uniquely positioned on the cotton genome were selected for the subsequent analysis; SNP detection, whereby the Genome Analysis Toolkit (GATK) (Unified Genotyper) was used to detect population SNPs in the filtered bam files (Aaron et al. 2010); SNP filtering, whereby the false-positive SNPs caused by sequencing errors were reduced, the number of parental SNP bases supported was restricted to at least 4, and at least four supported SNP bases were required for the offspring; and SNP-related information statistics, which included calculating the number of heterozygous SNPs, number of pure SNPs, and the heterozygous SNP ratio.

SNP marker development and genetic marker screening. Based on the results of the genotype testing of the barley parents, inter-parental polymorphic marker development was carried out. Loci with polymorphisms between parents were screened, i.e., loci with differences between the parents. After the interparental marker development, 140 offspring were genotyped at the parental polymorphic marker loci. The typed markers were then checked for abnormal bases, abnormal genotypes and, in addition, the markers were filtered for completeness. Markers with genotypes covering at least 75% of all the individuals in the offspring were screened, i.e., for a single polymorphic marker locus, at least 75 individuals out of 100 offspring had a definite genotype.

Linkage map construction and QTL analysis. For each cluster, Join-Map 4.0 was used to sort the markers in each cluster, calculate the genetic distance between the markers (regression algorithm), and remove markers that were seriously unlinkable. The Perl Scalable Vector Graphics module was used to draw the chain map. The limit of detection (LOD) value threshold for each phenotype was determined

Table 1. Sequence depth and coverage statistics

Sample	Clean reads ^a	Mapping reads ^b	Mapping rate ^c	Coverage1X ^d	Coverage4X ^e
Kunlun 10	7 566 250	7 332 637	96.91	6.68	3.08
Nierumuzha	7 695 118	7 366 910	95.73	6.73	3.09

^aNo. of reads used for the alignment; ^bNo. of clean reads that mapped to the reference genome; ^cthe percentage of reads that mapped to the genome; ^dpercentage of genomes with sites covered by at least 1 base in the reference genome; ^epercentage of the genome covered by at least 4 bases in the reference genome

using the permutation test (PT) in MapQTL (Van Ooijen 2009). The composite interval mapping (CIM) algorithm in WinQTL software was used for the QTL localisation, and the QTL segment corresponding to each phenotype was determined based on the threshold value obtained from the permutation test.

RESULTS

GBS sequencing of parental and F₂ offspring.

A total of 7 332 637 reads from Kunlun 10 and 7 366 910 reads from Nierumuzha were mapped to the barley genome. The mapping rates for Kunlun 10 and Nierumuzha were 96.91% and 95.73%, respectively (Table 1). Ultimately, 308 828 polymorphic loci were identified. The marker genotype data are summarised in Table 2. Only the genotype aa × bb, consisting of 34 065 markers between Kunlun10 and Nierumuzha, was used for the subsequent analyses (Table 3).

The F_2 lines were then genotyped using the GBS technology. The average number of raw individual reads of the 150 hulless barley lines was about 622 M. This yielded 93 Gb of high-quality sequence reads (Q20 \geq 90%, Q30 \geq 85%), and a normal guaninecytosine (GC) content was detected among the reads. The average MseI enzyme capture rate was \geq 98% across the GBS of the 142 F_2 lines, indicating that the enzyme digestion had high quality. The map-

Table 2. Marker types

Marker type	Numbers
efxeg	5 865
nnxnp	65 308
abxcc	55
aaxbb	34 065
abxcd	19
lmxll	64 194
hkxhk	139 237
ccxab	85
Total markers	308 828

ping rate of the 140 offspring was 96.32% of the whole genome, with 5.38% (coverage at least $1\times$) and 1.96% (coverage at least $4\times$) of sites. The screened genotypes contained all the markers in more than 75% of the individual lines. Therefore, at least 111 of the 140 progeny lines contained all the markers.

Genotyping of the progeny and selection of genetic markers. The low-coverage sequences of the F_2 lines (coverage under 75%) were filtered out. This left 7 052 markers out of the original 34 065.

Genetic linkage map. The unlinked markers were filtered out, and 7 052 SNPs were mapped to seven linkage maps using Joinmap 4.0. A high-density genetic map was constructed after these 7 052 markers were mapped onto the seven chromosomes of hulless barley. The genetic maps were 1 139.4 cM in length with an average distance of 0.59 cM between the markers. 1H was the largest group among the seven linkage groups, consisting of 285 markers and a genetic length of 271.2 cM. 5H was the smallest group, with 79 markers and a genetic length of 54.5 cM (Table 4, Figure 1). A total of 1 469 gaps were detected between the markers. Among these, 1 432 gaps were < 5 cM, 24 gaps were between 5 and 10 cM in size, 11 gaps were between 10 and 20 cM in size, and only two gaps over 20 cM in size were observed in any of the chromosomes (Table 5).

QTL analysis of the seed colour. An F₂ population of 142 offspring from Nierumuzha × Kunlun 10

Table 3. Progeny genotyping

Chr ^a	Position ^b	Ref ^c	A1 ^d	A2e	10-1 ^f	13-1	14-1	15-1
1	61 271 190	С	TT	CC	CC	CC	CC	CC
1	61 271 394	G	GG	TT	TT	TT	TT	TT
1	61 283 007	C	CC	AA			AA	
1	61 380 972	G	GG	CC	GG		GG	GG

^aScaffold No. where the marker is located; ^blocation on the Scaffold where the marker is located; ^creference genome base type; ^dparental A1 genotype; ^eparental A2 genotype; ^fgenotype of offspring 10-1 individuals at marker loci, other offspring by analogy; -- – indicates missing

Table 4. Genetic linkage group statistics

Chr	SNP	Length	Avg. gap	Max. gap
	SINP	(bp)	(c)	M)
1	285	271.2	0.95	24.88
2	203	190.5	0.94	21.13
3	194	77.6	0.40	8.73
4	157	189.7	1.21	16.57
5	79	54.5	0.69	12.19
6	279	124.5	0.45	9.59
7	278	124.1	0.45	17.26
overall	1475	1139.4	0.59	24.88

Chr – chromosome No.; SNP – single nucleotide polymorphism; lLength – total physical length of the chromosomes; avg. gap – average genetic distance between markers; max. gap – maximum gap between markers

Table 5. Linkage map gap statistics

Chr	< 5	5–10	10-20	> 20	- Ratio			
Chr		(cM)						
lg1	274	7	3	1	96.14			
lg2	194	3	4	1	96.04			
lg3	191	2	0	0	98.964			
lg4	147	8	1	0	94.231			
lg5	77	0	1	0	98.718			
lg6	275	2	1	0	98.921			
lg7	274	2	1	0	98.917			

Chr - chromosome No. + lg-linkage group

was used in this study. We divided them artificially into three categories according to the result of the Wanseen seed test system SC-G (purple seeds, light purple seeds and white seeds) (Table 6), the results for the other phenotypic parameters show normal distribution as shown in Figure 2, where they are all normally distributed. A total of three motifs were identified: one motif on 1H, accounting for 15.78%, and two motifs on 7H (Figure 3), accounting for 4.59% and 10.36% of the phenotypic interpretation (Table 7). In particular, we focused on two main effect QTLs with >10% phenotypic interpretation on 1H and 7H (Table 7). The locus peak positions were located on 70.31 cM (1H), 91.01 cM (7H), and 96.51 cM (7H), and were located within the marker

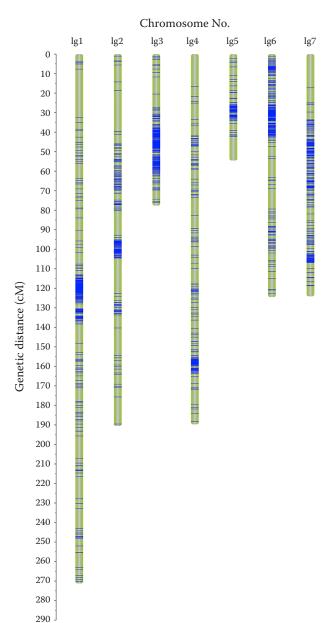


Figure 1. Distribution map of the linkage groups Left end of the horizontal line: marker genetic distance, right end of horizontal line: marker name + lg-linkage group

intervals of 1H-bin1404 to 1H-bin1400 (63 genes), 7H-bin941 to 7H-bin938 (29 genes), and 7H-bin903 to 7H-bin934 (three genes) (Table 7).

QTL analysis of the thousand kernel weight. Two QTLs were identified for the TKW: one QTL on 5H, accounting for 23.18% of the phenotypic interpre-

Table 6. Segregation for the colour of the lemma and seed coat in the F2 populations

Traits	No. of purple colour	No. of purple and white intermediate	No. of white colour	$\chi^2_{0.05}$ (5.99)	P
Seed coat colour (C)	35	68	37	0.17	0.05

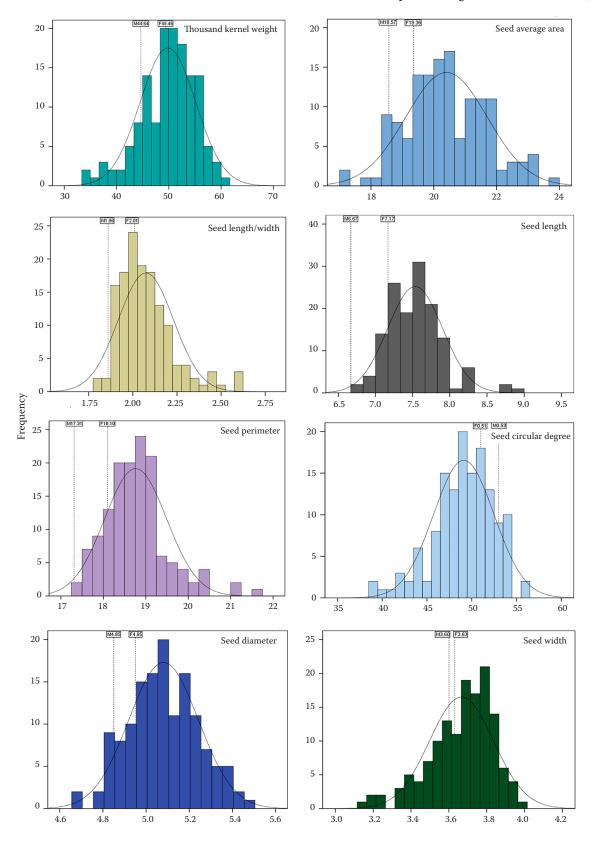


Figure 2. Distribution of the phenotypic data for the Nierumuzha and Kunlun 10 inbred F_2 population The top of each histogram indicates the values of the two parents, M and F represent Kunlun 10 (male parent) and Nierumuzha (female parent), respectively

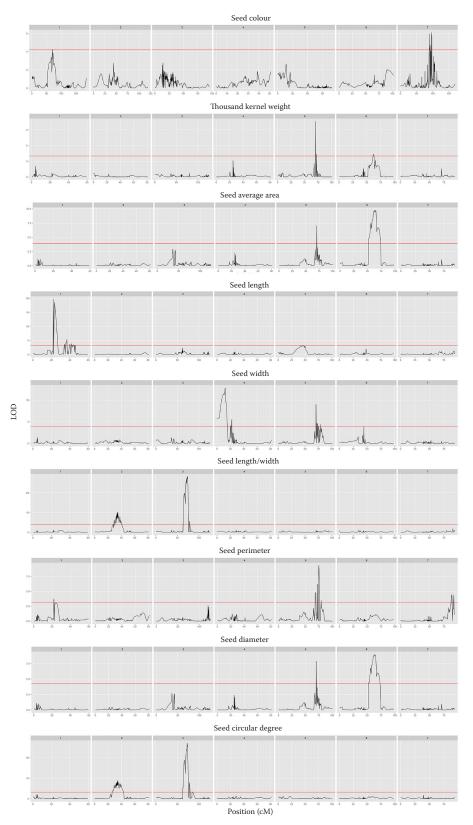


Figure 3. Linkage analysis of the agronomic traits associated with hulless barley in all the linkage groups
The x-axis is the genetic distance of each linkage group, the y-axis is the physical position of the bin marker versus the limit of detection (LOD) score, and the red line is the threshold LOD value determined by the substitution test; 1 to 7, seven linkage groups

Table 7. Linkage analysis of the related agronomic traits

						-			
Trait	Chr	LOD peak	Position	99% CI (cM)	$-\frac{R^2}{(\%)}$	Additive effect	Left marker	Right marker	Gene No.
	lg1	4.24	70.31	70.3–71.7	15.78	0.49	bin1404	bin1400	63
Seed colour	lg7	5.95	91.01	90-91.3	4.59	-0.47	bin941	bin938	29
(SC)	lg7	6.11	96.51	96.4-97.7	10.36	-0.53	bin903	bin934	3
Thousand kernel	lg5	10.64	70.31	70.2–70.5	23.18	-6.9394	bin907	bin906	16
weight (TKW)	lg6	4.34	62.41	59.6-63.4	6.47	2.071	bin1296	bin1296	7
<u>- </u>	lg5	7.03	70.31	70.2-70.5	12.81	-2.03	bin907	bin906	16
Seed average	lg6	7.45	56.01	55.3-57.4	10.13	0.90	bin1303	bin1302	38
area (SAA)	lg6	9.77	63.41	62.4-64.3	13.27	1.00	bin1295	bin1295	17
	lg6	9.73	65.31	60.2-62.4	14.81	1.04	bin1294	bin1294	20
	lg1	3.75	22.01	21.9-22.2	4.41	0.35	bin80	bin9	1 581
	lg5	4.84	69.61	68.1-69.7	6.21	-0.36	bin1157	bin1155	174
Seed perimeter	lg5	6.20	73.01	72.9-73.2	7.85	-0.48	bin1153	bin1152	82
(SP)	lg7	3.20	79.41	78.7-79.8	4.33	-0.36	bin1113	bin1117	22
	lg7	4.46	88.91	87.8-89.4	5.44	-0.32	bin1370	bin1370	18
	lg7	4.40	92.11	91.6-92.7	5.29	-0.30	bin1369	bin1368	37
-	lg2	4.13	27.91	26.8-28.4	4.70	-0.06	bin484	bin397	166
	lg2	6.28	29.11	28.8-30.0	6.93	-0.08	bin329	bin321	108
	lg2	5.33	30.91	30.3-31.0	6.38	-0.07	bin392	bin254	238
	lg2	9.29	33.31	33.2-33.4	10.60	-0.10	bin256	bin471	77
Seed length/	lg2	8.86	33.91	33.8-34.0	9.51	-0.10	bin243	bin352	209
width (SLW)	lg2	8.52	35.01	34.8-35.1	9.11	-0.10	bin388	bin372	680
	lg3	19.99	66.01	65.5–66.6	27.02	0.15	bin528	bin541	134
	lg3	19.19	67.21	66.6–67.7	26.55	0.14	bin544	bin546	123
	lg3	28.27	73.81	73.5–74.2	35.73	0.15	bin517	bin518	79
	lg3	5.73	77.91	77.7–79.1	10.63	-0.08	bin672	bin672	262
	lg1	18.55	23.31	23-27.5	25.71	0.44	bin12	bin12	311
	lg1	3.654	34.11	33.1–34.6	3.79	-0.13	bin58	bin99	156
	lg1	5.34	36.81	35.6-37.4	6.50	-0.17	bin100	bin137	97
Seed length	lg1	3.73	40.21	39.7-40.5	4.23	0.12	bin38	bin39	101
(SL)	lg1	3.89	41.61	41.2-42.8	5.76	0.13	bin141	bin63	196
	lg1	3.50	44.01	42.8-44.5	3.65	0.11	bin64	bin69	89
	lg1	3.38	45.41	45.1–46.3	3.72	0.112	bin36	bin71	103
	lg5	3.32	46.11	43.4-4.5	3.45	-0.13	bin1050	bin1046	33
	lg4	12.84	12.31	10.5–12.7	17.16	-0.15	bin857	bin861	110
	lg5	9.01	70.31	70.2–70.5	16.23	-0.27	bin907	bin906	293
Seed width	lg5	4.63	72.31	71.9–71.9	11.83	-0.18	bin912	bin912	475
(SW)	lg5	4.58	73.51	73.1–73.1	10.81	-0.18	bin1152	bin911	725
	lg5	3.98	79.21	78.3–78.3	5.26	-0.10	bin1132	bin1109	987
	lg5	7.88	70.31	70.2–70.6	13.85	-0.13	bin907	bin906	16
Seed diameter	lg6	6.79	56.01	54.1-57.4	9.30	0.09	bin1303	bin1302	47
(SD)	lg6	9.01	63.41	62.4–66.3	12.30	0.10	bin1295	bin1302	25
Seed circular	lg2	4.27	27.91	25.9–28.4	4.61	0.10	bin484	bin397	212
degree (SCD)	lg2	6.62	29.11	28.8–29.8	6.91	0.01	bin329	bin322	151
degree (SCD)	182	0.02	47.11	20.0-27.0	0.71	0.01	D111347	DIIIDZZ	131

Table 7 to be continued

Trait	Chr	LOD	Position	99% CI	R^{2} (%)	Additive	Left	Right	Gene No.
		peak	((cM)		effect	marke	marker	
	lg2	6.73	33.61	33.5-33.7	7.00	0.01	bin349	bin472	279
	lg2	7.80	33.91	33.8-34.0	8.03	0.01	bin243	bin352	133
	lg2	7.99	34.21	34.1-34.3	8.52	0.01	bin440	bin371	372
	lg2	7.79	34.71	34.6-34.7	7.98	0.01	bin494	bin489	200
~	lg2	7.76	35.01	34.9-35.1	7.93	0.01	bin415	bin372	157
Seed circular degree (SCD)	lg3	17.05	65.31	64.5-65.5	21.92	-0.02	bin548	bin534	99
degree (SCD)	lg3	18.70	66.81	65.5-67.7	23.71	-0.02	bin528	bin546	135
	lg3	23.92	71.31	70.3-71.5	29.24	-0.02	bin524	bin523	95
	lg3	26.99	72.91	72.4-73.5	32.18	-0.02	bin501	bin519	53
	lg3	7.84	77.91	77.7-79.0	11.16	0.01	bin672	bin672	293
	lg3	3.40	84.11	83.9-89.3	3.32	0.01	bin745	bin745	185

Chr – chromosome No.; LOD – limit of detection; CI – confidence interval; R^2 – proportion of phenotypic interpretation explained by the QTL; left marker – left boundary markers for each site; right marker – right boundary markers for each site + lg-linkage group

tation, and another on 6H, accounting for 6.47% of the phenotypic interpretation, where we were mainly interested in the main effect QTL on 5H. The genetic distances were approximately 70.31 cM and 62.41 cM, and the QTLs were located within the marker intervals 5H-bin907 to 5H-bin906 (16 genes) and 5H-bin1296 to 5H-bin1296 (seven genes) (Figure 3) (Table 7).

QTL analysis of the seed average area. A total of four QTLs for the SAA were identified: all four QTLs were located on 5H and 6H, accounting for 12.81%, 10.13%, 13.27%, and 14.81% of the phenotypic interpretation, and all these QTLs have phenotypic interpretation rates greater than 10%, so they are all of interest to us. The genetic distances were approximately 70.31, 56.01, 63.41, and 65.31 cM, and the QTLs were located in 5H-bin907 to 5H-bin906 (16 genes), 6H-bin1303 to 3H-bin1302 (38 genes), 6H-bin1295 to 3H-bin1295 (17 genes), and 6H-bin1294 to 6H-bin1294 (20 genes) within the marker interval (Figure 3) (Table 7).

QTL analysis of the seed length, seed width and length/width ratio. In total, eight SL QTLs were identified: seven QTLs on 1H, accounting for 25.71%, 3.79%, 6.50%, 4.23%, 5.76%, 3.65%, and 3.72% of the phenotypic interpretation; and one QTL on 5H, accounting for 3.45% of the phenotypic interpretation. We used the main effect QTL with > 10% phenotypic interpretation on 1H as the main study. The genetic distances were approximately 23.31, 34.11, 36.81, 40.21, 41.61, 44.01, 45.41 and 46.11 cM, and the QTLs

were located within the marker intervals of 1H-bin12 to 1H-bin12 (311 genes), 1H-bin58 to 1H-bin99 (156 genes), 1H-bin100 to 1H-bin137 (97 genes), 1H-bin38 to 1H-bin39 (101 genes), 1H-bin141 to 1H-bin63 (196 genes), 1H-bin64 to 1H-bin69 (89 genes), 1H-bin36 to 1H-bin71 (103 genes), and 5H-bin1050 to 5H-bin1046 (33 genes) (Figure 3) (Table 7).

Five SW QTLs were identified: one QTL on 4H, accounting for 17.16% of the phenotypic interpretation; and four QTLs on 5H, accounting for 16.23%, 11.83%, 10.81%, and 5.26% of the phenotypic interpretation, there is only one epistatic QTL on 5H accounting for 5.26% of the phenotypic interpretation rate. The genetic distances were approximately 12.31, 70.31, 72.31, 73.51, and 79.21 cM, and the QTLs were located within the marker intervals 4H-bin857 to 4H-bin861 (110 genes) and 5H-bin907 to 5H-bin906 (293 genes), 5H-bin912 to 5H-bin912 (475 genes), 5H-bin1152 to 5H-bin911 (725 genes), and 5H-bin1115 to 5H-bin1109 (987 genes) (Figure 3) (Table 7).

Ten SLW QTLs were identified: the QTLs were located both on 2H and 3H, accounting for 4.70%, 6.93%, 6.38%, 10.60%, 9.51%, and 9.11% on 2H, and 27.02%, 26.55%, 35.73%, and 10.63% on 3H of the phenotypic interpretation, where we focused only on the QTLs with the phenotypic interpretation rate of > 10%, four QTLs on 3H and one QTL on 2H, with genetic distances of approximately 27.91, 29.11, 30.91, 33.31, 33.91, 35.01, 66.01, 67.21, 73.81, and 77.91 cM. These QTLs were located from 2H-bin484 to 2H-bin397 (166 genes), 2H-bin329 to 2H-bin321 (108 genes),

2H-bin392 to 2H-bin254 (238 genes), 2H-bin256 to 2H-bin471 (77 genes), 2H-bin243 to 2H-bin352 (209 genes), 2H-bin388 to 2H-bin372 (680 genes), 3H-bin528 to 3H-bin541 (134 genes), 3H-bin544 to 3H-bin546 (123 genes), 3H-bin517 to 3H-bin518 (79 genes), and 3H-bin672 to 3H-bin672 (262 genes) within the marker interval (Figure 3) (Table 7).

QTL analysis of the seed perimeter and diameter. A total of six SP QTLs were identified: one QTL on 1H, accounting for 4.41% of the phenotypic interpretation; two QTLs on 5H, accounting for 6.21% and 7.85%; and three QTLs on 7H, accounting for 4.33%, 5.44%, and 5.29% of the phenotypic interpretation, we did not get the main effect QTL here (Figure 3) (Table 7). The six perimeter QTLs spanned genetic distances of approximately 22.01, 69.61, 73.01, 79.41, 88.91, and 92.11 cM and were located within the marker intervals 1H-bin80 to 1H-bin9 (1581 genes), 5H-bin1157 to 5H-bin1155 (174 genes), 5H-bin1153 to 5H-bin1152 (82 genes), 7H-bin1113 to 5H-bin1117 (22 genes), 7H-bin1370 to 7H-bin1370 (18 genes), and 7H-bin1369 to 7H-bin1368 (37 genes) (Figure 3) (Table 7). In addition, three SD QTLs were identified: one QTL on 5H, accounting for 13.85% of the phenotypic interpretation, and two QTLs on 6H, accounting for 9.30% and 12.30% of the phenotypic interpretation, we focused on the QTLs with the phenotypic interpretation rate greater than 10% (Figure 3) (Table 7). Their genetic distances were approximately 70.31, 56.01, and 63.41 cM, and the QTLs were located within the marker intervals 5H-bin907 to 5H-bin906 (16 genes), 6H-bin1303 to 6H-bin1302 (47 genes), and 6H-bin1295 to 6H-bin1294 (25 genes) (Figure 3) (Table 7).

QTL analysis of the seed circular degree. Thirteen SCD QTLs were identified: the QTLs were located on 2H and 3H, accounting for 4.61%, 6.91%, 7.00%, 8.03%, 8.52%, 7.98%, and 7.93% of the phenotypic interpretation rate on 2H and 21.92%, 23.71%, 29.24%, 32.18%, 11.16%, and 3.32% on 3H, we identified the QTLs with the phenotypic interpretation rate greater than 10% as the dominant QTLs and the rest as epistatic QTLs. The genetic distances were approximately 27.91, 29.11, 33.61, 33.91, 34.21, 34.71, 35.01, 65.31, 66.81, 71.31, 72.91, 77.91, 84.11, and 77.91 cM, and the QTLs were located from 2H-bin484 to 2H-bin397 (212 genes), 2H-bin329 to 2H-bin322 (151 genes), 2Hbin349 to 2H-bin472 (279 genes), 2H-bin243 to 2Hbin352 (133 genes), 2H-bin440 to 2H-bin371 (372 genes), 2H-bin494 to 2H-bin489 (200 genes), 2H-bin415 to 2H-

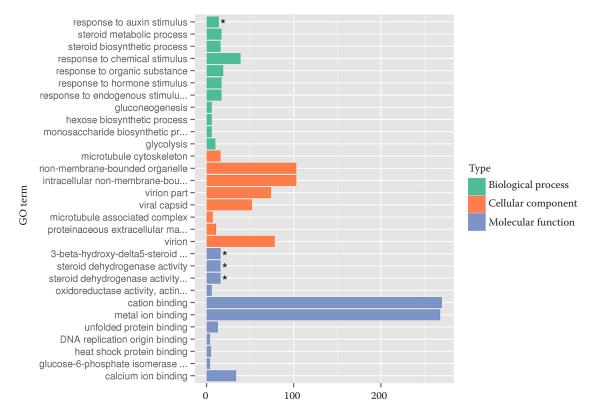


Figure 4. Gene ontology (GO) enrichment pathway for the width

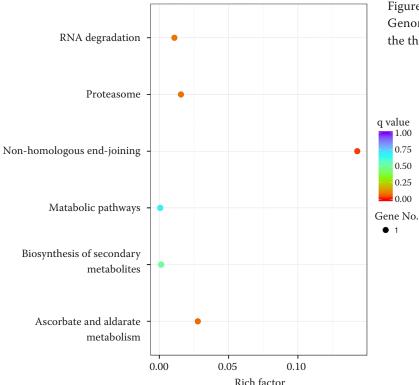


Figure 5. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway for the thousand seed weight

bin372 (157 genes), 3H-bin548 to 3H-bin534 (99 genes), 3H-bin528 to 3H-bin546 (135 genes), 3H-bin524 to 3H-bin523 (95 genes), 3H-bin501 to 3H-bin519 (53 genes), 3H-bin672 to 3H-bin672 (293 genes), and 3H-bin745 to 3H-bin745 (185 genes) within the marker interval (Figure 3) (Table 7)

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis within the candidate genetic intervals. We conducted an enrichment analysis of the gene functions of all the trait-related genes. Notably, there were significantly enriched pathways in the gene ontology (GO) pathway analysis of the SW (Figure 4), and a total of 30 genes were enriched in four pathways, including the response to the auxin stimulus, 3-beta-hydroxydelta5-steroid dehydrogenase activity, steroid dehydrogenase activity, and steroid metabolic process. In addition, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the TKW showed that only one gene was significantly enriched in the non-homologous end-joining pathway (Figure 5), and no other traits were significantly enriched.

DISCUSSION

Using a high-throughput sequencing technology to develop a large number of SNP markers, Yao et al.

(2018) performed a GBS-SNP analysis on the doubled haploid locus population Nierumuzha × Kunlun 10 to locate the purple seed coat colour gene (Psc), where significant progress has been made in the high-density genetic mapping of hulless barley grain colour (Yao et al. 2018). In this study, we studied the seed colour of barley as well as several other important agronomic traits in order to provide a reference for the selection of good hulless barley varieties. We used the hulless barley material Nierumuzha (purple) and Kunlun 10 (white) for the construction of an F₂ segregating population and used the F₂ segregating population as the experimental material for the construction of a genotyping library. In this study, we identified a total of 24 main effect QTLs and 30 epistatic QTLs, we focused on the main effect QTLs, including two main effect QTLs associated with the SC, located on 1H and 7H, and one main effect QTL associated with the TKW on 5H, one QTL associated with the SAA on 1H and three QTLs on 5H, one main effect QTL associated with the SLW on 2H and four main effect QTLs on 3H, and only one main effect QTL associated with the SLW on 3H. There was only one main effect QTL associated with the SL which was located on 1H with 25.71% phenotypic interpretation, four main effect QTLs on the SW, two main effect QTLs on the SD,

and finally five main effect QTLs on the SCD. Five main-effect QTLs were obtained on the SCD, all of which were located on 3H (Table 7). The high precision of the genetic map enabled us to finely localise the QTLs for agronomic traits. Moreover, the gene related to the SC was found to overlap with ANT1, a transcription factor related to anthocyanin synthesis that was previously reported by our group (Yao et al. 2018). The SW is localised on chromosomes 5 and 6, with a high phenotypic interpretation rate of 23.18% on chromosome 5, a region containing 47 genes. The analysis revealed that NAC021 (MLOC_64240) was possibly associated with grain weight traits, and studies have shown that a large number of NAC transcription factors play a role in plant growth and development, signal transduction, abiotic stresses, and pathogenic stresses (Shahnejat-Bushehri et al. 2016; Wang et al. 2020a; Wang 2021). The SAA was localised on chromosomes 5 and 6, both with phenotypic interpretation rates exceeding 10%. Two genes potentially related to the SAA were localised to chromosome 5, namely LOL2 (MLOC_64242) and NAC021 (MLOC_64240). The NAC021 transcription factor was previously found to be associated with the SW, and one potentially related gene TSK (MLOC_22596) on chromosome 6 plays an important role in plant cell division and growth and development (He et al. 2019). The SL was localised on chromosomes 1 and 5, with a 25.71% phenotypic interpretation rate within the chromosome 1 segment, and the two transcription factors MADS21 (MLOC_65843) and MADS4 (MLOC_74587) may possibly be associated with the length. Studies have shown that MADSbox genes play a role in floral organ development, flower initiation, fruit development, differentiation of meristematic tissues, embryo, root, and vascular tissue development (Lai et al. 2021). One gene, FIP1 (MLOC_58516), possibly related to the width, was localised on chromosome 4, but its function is currently unclear. A total of five genes were localised on chromosome 5, including an NAC family transcription factor NAC021 (MLOC_64240) associated with the SW; the DREB 1A (MLOC_74460) transcription factor and HVA22A (MLOC_44915), which are two resistance genes; CYP78A6 (MLOC_64838), a Cytochrome P450 gene that has been found to play an important role in the synthesis of phenylpropanoids, alkaloids, terpenoids, glycosides, and phytohormones (He et al. 2008); and lastly SAUR71 (MLOC_60693). In Arabidopsis thaliana, AtSAUR71 and AtSAUR72 were shown to belong to the same subfamily, were strongly associated with growth hormone transport, and were expressed in the mid-column of the young roots and hypocotyls (Qiu et al. 2013). In addition, the SAUR71 gene (MLOC_60693) was also found in the GO enrichment pathway data results, which is a noteworthy gene. Two genes localised on chromosome 2 may be related to the SLW, including the NAM-B2 (MLOC_11017) transcription factor and CRY1 (MLOC_64083), which is a member of the NAC transcription factor family - a class of transcription factors unique to plants that has a variety of important physiological functions during growth, development, and senescence. Studies have shown that CRY1 has the function of regulating plant growth and development (Canamero et al. 2006), and three genes were localised to chromosome 3 as MYB transcription factors, including LHY (MLOC_59912), CYP710A1 (MLOC_57804), and WRKY72 (MLOC_7939). The MYB transcription factors play a regulatory role in secondary metabolism, cell morphogenesis, the response to the environment, and the cell cycle (Wang et al. 2003), and WRKY transcription factors – a class of transcription factors unique to plants - are involved in regulating growth and development, secondary metabolism, senescence, and the response to biotic and abiotic stresses (Shan et al. 2018). The SP was localised on chromosomes 1, 5 and 7 and we analysed the segments with high phenotypic interpretation, locating a total of two genes, the SKIP11 (MLOC_72534) and TCP18 (MLOC_4614) transcription factors, to chromosome 5. SKIP has many functions in plants, including growth, development and stress defence (Wang et al. 2013; Cui et al. 2017), and TCP transcription factors are involved in plant growth and development, hormone signalling, and regulation of plant stress responses (Wang et al. 2021). The SD was localised to chromosomes 5 and 6, and we localised one NAC gene NAC021 (MLOC_64240) on chromosome 5 that was previously found to be associated with weight and width, as well as a SKIP gene SKIP8 (MLOC_51023) on chromosome 6. The SCD was localised to chromosomes 2 and 3. We analysed the high-contributing segments and localised a total of three genes, all located on chromosome 3: MYB1R1 transcription factor (MLOC_76271), RAX3 (MLOC_67874), and NAC100 (MLOC_63121), with the RAX gene encoding a MYB-like transcription factor. Finally, a gene significantly enriched in the KEGG pathway associated with the TKW was identified as the ku70 gene (MLOC_62977), which is a key gene in the non-homologous end joining (NHEJ)

pathway. The inactivation of ku70 could theoretically lead to the inactivation of the NHEJ pathway, allowing exogenous DNA fragments to be integrated into the genome only through the homology-directed repair pathway, thereby improving the targeting efficiency (Fell & Schild-Poulter 2015). Whether it is related to the TKW deserves further study.

The high-density genetic linkage map constructed based on GBS sequencing and sliding window methods greatly improved the accuracy and localisation of hulless barley genetic mapping. In addition, the *NAC021* gene (MLOC_64240) was located several times in this study, and this gene could be a key research target in the future, potentially leading to significant discoveries. Therefore, the high-density genetic linkage map constructed in this study provides an important research basis for the localisation of QTLs for important agronomic traits in hulless barley, and further validation of the localised QTL loci is expected to provide some reference for the breeding improvement of the hulless barley grain quality.

CONCLUSION

Based on this high-density linkage map, a total of 54 QTLs associated with seed traits were detected, including 24 main QTLs and 30 epistatic QTLs. A total of 24 candidate genes were identified within the QTL regions, including one SC-related transcription factor (ANT1), two TKW-related genes (NAC021, ku70), three SAA-related genes (LOL2, NAC021, TSK), two SL-related genes (MADS21. MADS4), six SW-associated genes (FIP1, NAC021, DREB1A, HVA22A, CYP78A6, SAUR71), five SLWassociated genes (NAM-B2, CRY1, LHY, CYP710A1, WRKY72), two SP-associated genes (SKIP11, TCP18), two SD-associated genes (NAC021, SKIP8), and three SCD-related genes (MYB1R1, RAX3, NAC100). The high-density genetic mapping and QTL identification of seed traits in hulless barley provides a valuable genetic resource and basis for molecular marker-assisted selection and genomic studies in hulless barley.

Acknowledgement: We would like to thank Accdon (www. accdon.com) for editing this manuscript. We are grateful to the reviewers for their valuable comments.

REFERENCES

Aaron M., Matthew H., Eric B. (2010): The genome analysis toolkit: A MapReduce framework for analyzing next-

- generation DNA sequencing data. Genome Research, 20: 1297–1303.
- Bellido G.G., Beta T. (2009): Anthocyanin composition and oxygen radical scavenging capacity (ORAC) of milled and pearled purple, black, and common barley. Journal of Agricultural and Food Chemistry, 57: 1022–1028.
- Canamero R.C., Bakrim N., Bouly J.P. (2006): Cryptochrome photoreceptors cry1 and cry2 antagonistically regulate primary root elongation in *Arabidopsis thaliana*. Planta, 224: 995–1003.
- Chen Z.L., Wang B.B., Dong X.M., Liu H., Ren L.H., Chen J., Hauck A., Song W.B., Lai J.S. (2014): An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F₂ maize population. BMC Genomics, 15: 433.
- Cui Z., Tong A., Huo Y. (2017): SKIP controls flowering time via the alternative splicing of SEF pre-mRNA in *Arabidopsis*. BMC Bioinformatics, 15: 80.
- Fell V.L., Schild-Poulter C. (2015): The Ku heterodimer: Function in DNA repair and beyond. Mutation Research, 763: 15–29.
- Han J.Q., Hong Q.Y. (2020): Research progress of bioactive ingredients and physiological functions in hulless barley. Journal of Food Science and Technology, 38: 11–20.
- He L.H., Zhao S.J., Hu Z.B. (2008): Advances in plant cytochrome P450 genes and functions. Drug Biotechnology, 15: 142–147.
- He M.Z., Huang J., Du L.E., Wen Q.S., Chen J., Tian M.L. (2019): Progress in the study of *TONSOKU* gene in *Arabidopsis thaliana*. Molecular Plant Breeding, 17: 1175–1185.
- Hussain W., Baenziger P.S., Belamkar V. (2017): Genotyping-by-sequencing derived high-density linkage map and its application to QTL mapping of flag leaf traits in bread wheat. Scientific Reports, 7: 16394.
- Lai X., Vega-Léon R., Hugouvieux V. (2021): The intervening domain is required for DNA-binding and functional identity of plant MADS transcription factors. Nature Communications, 12: 4760.
- Li H., Durbin R. (2009): Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25: 1754–1760.
- Li X., Yu X.L., Yao X.H., Yao Y.H., Bai Y.X., An L.K., Wu K.L. (2021): Mapping the major quantitative trait loci of the heading date trait in Qingke barley (*Hordeum vulgare* L.) from the Qinghai-Tibetan Plateau via genotyping by sequencing. All Life, 14: 882–893.
- Liang X., Gao M., Sikandar Amanullah., Yu G., Xiu J.L., Hong G.X., Ji X.L., Yue G., Cheng Z.Y., Fei S.L. (2022): Identification of QTLs linked with watermelon fruit and seed traits using GBS-based high-resolution genetic mapping, Scientia Horticulturae, 303: 111237.

- Liu N., Li M., Hu X. (2017): Construction of high-density genetic map and QTL mapping of yield-related and two quality traits in soybean RILs population by RAD-sequencing. BMC Genomics, 18: 466.
- Nie C., Yan X., Xie X., Zhang Z., Zhu J., Wang Y., Li J. (2021): Structure of β -glucan from Tibetan hull-less barley and its in vitro fermentation by human gut microbiota. Chemical and Biological Technologies in Agriculture, 8: 1–14.
- Qiao Y., Sun J.Q., Wei R.Q., Liu Y. (2019): Research progress of GBS technology in plants. China Agricultural Science and Technology Herald, 21: 47–55.
- Qiu T., Chen Y., Li M. (2013): The tissue-specific and developmentally regulated expression patterns of the SAUR41 subfamily of SMALL AUXIN UP RNA genes: Potential implications. Plant Signaling & Behavior, 8: e25283.
- Schafleitner R., Huang S.M., Chu S.H., Yen J.Y., Lin C.Y., Yan M.R., Krishnan B., Liu M.S., Lo H.F., Chen C.Y., Chen L.F., Wu D.C., Bui T.G., Ramasamy S., Tung C.W., Nair R. (2016): Identification of single nucleotide polymorphism markers associated with resistance to bruchids (*Callosobruchus* spp.) in wild mungbean (*Vigna radiata* var. *sublobata*) and cultivated *V. radiata* through genotyping by sequencing and quantitative trait locus analysis. BMC Plant Biology, 16: 159.
- Shahnejat-Bushehri S., Tarkowska D., Sakuraba Y. (2016): *Arabidopsis* NAC transcription factor JUB1 regulates GA/BR metabolism and signalling. Nature Plants, 2: 16013.
- Shan R.Y., Chen C.S., Zhong Q.S., Lin Z.H., Chen Z.H., You X.M. (2018): Cloning and differential expression characterization of CYP710A1, a cytochrome P450 gene from tea tree. Journal of Tea, 59: 181–185.
- Song B.W., Dang X.X., Zhao Z., Chen C., Huang M., Chen W.X., Liang K.Q., Xiao W.M. (2022): Localization and analysis of grain shape QTL in rice based on high-density genetic mapping. Journal of Crop Science, 48: 1–17.
- Van Ooijen J.W. (2009): MapQTL 6.0.: Software for the Mapping of Quantitative Trait Loci in Experimental Populations of Diploid Species. Wageningen, Kyazma.
- Wang J.C. (2021): Progress in the study of TCP family genes in plants. Agriculture and Technology, 41: 63–66.
- Wang J., Zheng C., Shao X. (2020a): Transcriptomic and genetic approaches reveal an essential role of the NAC transcription factor SlNAP1 in the growth and defense response of tomato. Horticulture Research, 7: 209.

- Wang J., Wang Y., Zhang J. (2021): The NAC transcription factor ClNAC68 positively regulates sugar content and seed development in watermelon by repressing ClINV and ClGH3.6. Horticulture Research, 8: 214.
- Wang N., Yuan Y., Wang H. (2020b): Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding. Scientific Reports, 10: 16308.
- Wang X., Li Z., Yan F. (2013): ZmSKIP, a homologue of SKIP in maize, is involved in response to abiotic stress in tobacco. Plant Cell, Tissue and Organ Culture, 112: 203–216.
- Wang X.Q., Chen B.J., Yin L.P. (2003): MYB transcription factors in plants. Biotechnology Bulletin, 2: 22–25.
- Wang X.R., Li Y., Zhang Y.J., Li Y.S., Wang J.C., Xu Y.P., Qi X.S. (2022): Identification of drought resistance of barley germplasm resources at the establishment stage and screening of drought tolerance indicators. Journal of Crop Science, 48: 1279–1287.
- Xue X.J., Du X.Y., Gai Y., Tang Y., Sun Y.X., Song L.Q., Jiang Z.W. (2020): Progress of developing SNPs in plants based on GBS sequencing. Jiangsu Agricultural Science, 48: 62–68.
- Yang D.S., Yu Z., Yu X.X., Li J.Q., Li J.W., Wu G.F., Lu Q.Q. (2021): Construction of an ultra-high density genetic linkage map of tetraploid hybrid ice grass based on GBS technology. Journal of Cereal Crops, 41: 1197–1210.
- Yao X.H., Wu K.L., Yao Y.H. (2018): Construction of a highdensity genetic map: Genotyping by sequencing (GBS) to map purple seed coat color (Psc) in hulless barley. Hereditas, 155: 37.
- Yao H., Wang Y., Yin J., Nie S., Xie, M. (2021): Isolation, physicochemical properties, and structural characteristics of arabinoxylan from hull-less barley. Molecules, 26: 3026.
- Yao X.H., Wang Y., Yao Y.H., An L.K., Wang Y., Wu K.L. (2022): Cloning and expression of a new gene *HvMEL1 AGO* in hulless barley under stripe disease stress. Journal of Crop Science, 48: 1181–1190.
- Zhang L., Zhang B.J. (2017): Localization and cloning of genes for quantitative traits in plants. Botany Bulletin, 24: 553–560.

Received: September 7, 2022 Accepted: November 10, 2022 Published online: January 31, 2023