

# Functional Gene Assessment of Wheat: Breeding Implication in Ningxia Province

**Weijun Zhang**

Ningxia Academy of Agriculture and Forestry Sciences

**Junjie Zhao**

Chinese Academy of Agricultural Sciences Cotton Research Institute <https://orcid.org/0000-0002-9827-6798>

**Jinshang He**

Ningxia Academy of Agriculture and Forestry Sciences

**Ling Kang**

Ningxia Academy of Agriculture and Forestry Sciences

**Xiaoliang Wang**

Ningxia Academy of Agriculture and Forestry Sciences

**Fuguo Zhang**

Ningxia Academy of Agriculture and Forestry Sciences

**Chenyang Hao**

Chinese Academy of Agricultural Sciences

**Xiongfeng Ma**

Chinese Academy of Agricultural Sciences

**Dongsheng Chen** (✉ [cdsnky@163.com](mailto:cdsnky@163.com))

Ningxia Academy of Agriculture and Forestry Sciences

---

## Research article

**Keywords:** adaptation genes, founder parents, KASP, functional markers, quality, genes, stress resistance genes, yield genes

**Posted Date:** December 4th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-117878/v1>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on February 18th, 2021. See the published version at <https://doi.org/10.1186/s12870-021-02870-5>.

1 **Functional gene assessment of wheat: breeding implication in Ningxia province**

2 Weijun Zhang<sup>1†</sup>, Junjie Zhao<sup>2†</sup>, Jinshang He<sup>1</sup>, Ling Kang<sup>1</sup>, Xiaoliang Wang<sup>1</sup>, Fuguo  
3 Zhang<sup>1</sup>, Chenyang Hao<sup>3\*</sup>, Xiongfeng Ma<sup>2\*</sup>, Dongsheng Chen<sup>1\*</sup>

4 <sup>1</sup> Crop Research Institute, Ningxia Academy of Agriculture and Forestry Sciences,  
5 Yinchuan 750002, Ningxia, China

6 <sup>2</sup> State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese  
7 Academy of Agricultural Sciences, Anyang 455000, Henan, China

8 <sup>3</sup> Key Laboratory of Crop Gene Resources and Germplasm Enhancement, Ministry of  
9 Agriculture and Rural Affairs/The National Key Facility for Crop Gene Resources  
10 and Genetic Improvement/Institute of Crop Sciences, Chinese Academy of  
11 Agricultural Sciences, Beijing 100081, China

12 †: Weijun Zhang and Junjie Zhao contributed equally to this work.

13 \*: Corresponding authors: Dongsheng Chen at cdsnky@163.com, Xiongfeng Ma at  
14 maxiongfeng@caas.cn, and Chenyang Hao at haochenyang@caas.cn

15

16 **Abstract**

17 **Background:** The overall genetic distribution and divergences for cloned genes  
18 among wheat varieties that occurred during the breeding process of the past few  
19 decades in Ningxia province of China are poorly understood. Here we report the  
20 genetic diversities of 44 important genes underpinning grain yield, quality, adaptation  
21 and resistance in 121 Ningxia and 86 introduced wheat cultivars and advanced lines.

22 **Results:** Population structure indicated characteristics of genetic components of  
23 Ningxia wheats including landraces of particular genetic resources, introduced  
24 varieties with rich genetic diversities and modern cultivars in different times. Analysis  
25 of allele frequencies showed that dwarfing alleles *Rht-B1b* at *Rht-B1* and *Rht-D1b* at  
26 *Rht-D1*, *1BL/1RS* translocation, *Hap-1* at *GW2-6B* and *Hap-H* at *Sus2-2B* are present  
27 very frequently in Ningxia modern cultivars and introduced varieties from other

28 regions, but absent in landraces, indicating that the introduced wheat germplasm with  
29 numerous beneficial genes are vital for broadening genetic diversities of Ningxia  
30 wheat varieties. Large population differentiation occurred at adaptation genes between  
31 modern cultivars and well-adapted landraces. Founder parents have excellent allele  
32 combinations of important genes with a higher number of favorable alleles compared  
33 with modern cultivars. The gene flows manifested that six founder parents greatly  
34 contributed to breeding improvement in Ningxia province, in particular Zhou 8425B  
35 for yield related genes.

36 **Conclusions:** These results will greatly benefit for wheat breeding in Ningxia  
37 province and other areas with similar ecological environments.

38 **Keywords:** adaptation genes, founder parents, KASP, functional markers, quality  
39 genes, stress resistance genes, yield genes

40

## 41 **Background**

42 Asia is the largest producer and consumer of wheat, with China and India as the two  
43 major countries in wheat production [1]. Wheat (*Triticum aestivum* L.) is widely  
44 cultivated in intricate geographical environments in China, reflecting its features of  
45 wide adaptability and high yield. Ningxia, as a north-western province of China with  
46 complex ecological types, has a long agricultural history of wheat cultivation. Since  
47 1950s, wheat varieties have experienced five times of replacements and become the  
48 first cultivated cereal crops in Ningxia province. Only 14 wheat landraces, introduced  
49 varieties and founder parents play pivotal roles in the five-time update of wheat  
50 varieties. Therefore, it is essential to dissect the substance of wheat improvement in  
51 past several decades for directing future wheat breeding in Ningxia.

52 Bread wheat (AABBDD) has characteristics of large genome size, allopolyploid,  
53 highly complex repetitive genome contents that are shaped by two recent  
54 polyploidization events [2-5], domestication [6], gene flow from frequent intra- and  
55 inter-species introgression [7-8], and post-domestication selection aimed at

56 developing high-yielding locally adapted varieties [9]. Multiple factors drive the  
57 evolution of wheat varieties, particularly many important genetic loci have been  
58 selected during modern wheat breeding. Insight into these genetic loci is important to  
59 understand phenotypic variations in adaptability, resistance to biotic and abiotic  
60 stresses, processing and nutritional quality, and yield stability. The adaptation of  
61 wheat to diverse environments is largely governed by genes for vernalization (*Vrn-A1*,  
62 *Vrn-B1* and *Vrn-D1*) [10], photoperiod (*Ppd-D1* etc.) [11], and plant height (*Rht-B1*  
63 and *Rht-D1*) [12]. Yield related genes include sucrose synthase genes *TaSus1-7A*, *7B*,  
64 *TaSus2-2A*, *2B* for thousand-kernel weight and grain size [13-14], *TaGW2-6A*, *6B* for  
65 grain width [15-17], *TaGS-D1* for grain size [18], *TaCwi-A1* encoding cell wall  
66 invertase [19], *TaCKX6-D1* encoding cytokinin oxidase/dehydrogenase [20] and grain  
67 length-associated gene *TaGASR-A1* [21] in bread wheat. Assessing processing quality  
68 is crucial in wheat quality improvement. Strong-gluten wheat varieties are  
69 characterized by a combination of medium-high kernel hardness, acceptable protein  
70 content, medium-strong dough and good extensibility, and representative varieties  
71 included Yumai 34, Zhengmai 366 for both pan bread and noodles quality [22]. High-  
72 and low-molecular-weight glutenin subunits (HMW-GS and LMW-GS) associated  
73 with dough quality are influenced by *Glu-1* and *Glu-3* loci [23]. Flour color  
74 responsible for noodle quality was influenced by several factors including polyphenol  
75 oxidase (PPO) activity (*Ppo-A1* and *Ppo-D1*) [24-25], phytoene synthase (PSY)  
76 enzymes (*Psy-A1*, *Psy-B1* and *Psy-D1*) [26-27],  $\zeta$  (zeta)-carotene desaturase (ZDS)  
77 enzymes (*Zds-A1*) [28] and peroxidase (*Pod-A1*) [29]. Kernel hardness, which has a  
78 profound effect on milling and end-use quality, is largely determined by the *Pina-D1*  
79 and *Pinb-D1* genes encoding puroindoline a and puroindoline b proteins, respectively  
80 [30]. Increasing biotic- and abiotic-stresses are major challenges with the impacts of  
81 climate and environmental changes for wheat breeding. In the past decades, some  
82 important stress resistance genes were cloned. Dehydration-responsive element  
83 binding (DREB) proteins at *Dreb-B1* locus are induced for improving drought  
84 tolerance [31]. Fusarium head blight (FHB) devastates wheat production worldwide  
85 and its resistance genes *Fhb1* using recombinants [32-33] and *Fhb7* in wheat distant

86 hybridization breeding [34] were cloned recently. The *Lr34/Yr18/Pm38* locus  
87 conferring durable adult-plant resistance to multiple diseases is used in wheat  
88 breeding programs worldwide [30]. The 1BL/1RS translocation (*1BL/1RS*) has been  
89 adopted widely in wheat breeding due to positive impacts on grain yield, adaptation,  
90 and particularly the presence of resistance genes to several diseases and pests  
91 although its translocation is associated with undesirable bread-making quality [35].

92 Modern breeding has imposed selection for improved productivity that largely  
93 influences the frequency of superior alleles for genetic loci underpinning traits of  
94 breeding interest. Therefore, molecular diagnosis for the allelic variations of genes is  
95 important to manipulate beneficial alleles in wheat molecular breeding. Enhanced  
96 capacity in sequencing, along with the availability of high-quality genome sequences  
97 of bread wheat, has allowed researchers to deploy specific favorable alleles using  
98 molecular markers. Currently, 157 functional markers documented for more than 100  
99 loci underpinning adaptability, resistance to biotic and abiotic stresses, quality and  
100 grain yield were converted into high-throughput KASP assays [36]. Such approaches  
101 will promote assessing the distribution of functional genes of wheat germplasms and  
102 applications in wheat breeding.

103 Our objectives for this study were to evaluate genetic structure, diversity,  
104 divergence and allelic variations of wheat germplasm resources in Ningxia province  
105 of China using KASP assays of 44 cloned genes for adaptation, stress resistance,  
106 quality, and grain yield. Genetic characteristics were exhibited in 207 bread wheat  
107 varieties, landraces and advanced lines including founder parents and varieties from  
108 Ningxia and other regions. Gene flow and allelic frequency implicate the distribution  
109 of important functional genes, which may contribute to improvement selection of  
110 future wheat breeding in Ningxia province and provide a robust guiding breeding  
111 foundation for other regions and countries with similar ecological environments.

## 112 **Methods**

### 113 **Plant materials and DNA extraction**

114 A representative sampling of wheat germplasm consisted of 207 wheat varieties,

115 including 121 Ningxia varieties and 86 introduced varieties (Table S1). The latter  
116 were introduced in Ningxia province over the past decades and played a huge role in  
117 local wheat breeding. The Ningxia varieties included 13 landraces and 108 modern  
118 cultivars and advanced lines. In addition, six founder parents among 207 wheat  
119 varieties used in this study include Moba 66, Abbondanza, Beijing 8, Orofen, Xiaoyan  
120 6 and Zhou 8425B. Genomic DNA was extracted from fresh leaves in each accession  
121 using the CTAB method [37].

### 122 **KASP genotyping of functional genes**

123 Conventional functional markers were summarized based on 44 cloned wheat genes  
124 for grain yield, quality, adaptation and stress resistance [30]; these markers were  
125 converted into KASP assays [36] that were widely exploited to characterize wheat  
126 germplasm resources [38-41]. A total of 44 KASP arrays developed from cloned  
127 genes were used for genotyping in this study (Table S2).

128 KASP assays were detected with 5.0  $\mu\text{L}$  mixtures containing 2.2  $\mu\text{L}$  of 40  $\text{ng}/\mu\text{L}$   
129 DNA, 2.5  $\mu\text{L}$  of 1  $\times$  KASP V4.0 2X Master mix (KBS-1016-017), 0.04  $\mu\text{L}$   $\text{Mg}^{2+}$ ,  
130 0.056  $\mu\text{L}$  of primer mixture, and 0.204  $\mu\text{L}$  ddH<sub>2</sub>O and the following amplification  
131 programs: hot start at 95°C for 15 min, followed by ten touchdown cycles (95°C for  
132 20 s; touchdown at 65°C initially and decreasing by 1°C per cycle for 25 s), followed  
133 by 30 additional cycles of annealing (95°C for 10 s; 57°C for 60 s) [41]. KASP  
134 genotyping were performed using QuantStudio™ 7 Flex (Applied Biosystems by Life  
135 Technologies, U.S.), and visualized by QuantStudio™ Real-time PCR software v.1.3  
136 (Applied Biosystems by Life Technologies, U.S.).

### 137 **Population structure and phylogenetic analysis**

138 A neighbor-joining tree was constructed in PowerMarker v3.25 [42] and visualized in  
139 MEGA 7 [43] using genotypic data of 44 genes. The first three eigenvectors of  
140 principal coordinate analysis (PCA) were performed using the R package Adegenet  
141 v2.0.1 [44]. Population structure of 207 accessions based on 44 functional genes was  
142 evaluated using Structure 2.3.4 with a burn-in period at 50,000 iterations and a run of

143 500,000 replications of Markov Chain Monte Carlo (MCMC) [45]. Then the number  
144 of populations was estimated based on the  $\Delta K$  model [46].

145 Allele numbers and frequencies were calculated for all loci. Genetic diversities  
146 were evaluated by PowerMarker v3.25, and the student's *t*-test was further used to  
147 compare the effects of two genotypes at a threshold probability of  $P < 0.05$ . Genetic  
148 flow and *F*-statistics (*F<sub>st</sub>*) were measured of population differentiation with  
149 POPGENE software [47].

## 150 **Results**

### 151 **Genotyping and population structure**

152 Genotyping of 207 wheat varieties using 44 KASP assays identified allelic variations  
153 at 44 loci (Table S1). All selected KASP assays showed clear clustering results of  
154 varieties (Figure S1). In total, these loci underpin grain yield (10), quality (14),  
155 adaptation (6), and stress resistance (14).

156 The neighbor-joining analysis divided 207 varieties into two groups, designated  
157 as Ningxia and Others, respectively (Figure 1A), in agreement with PCA (Figure 1B).  
158 The number of subpopulation (*K*) was plotted against the  $\Delta K$  calculated from the  
159 Structure, and the peak of the broken line graph was observed at  $K = 2$  (Figures 1C,  
160 S2), demonstrating that the population was basically divided into two subgroups. The  
161 first subgroup mainly referred to as landraces and cultivars from Ningxia provinces  
162 (Ningxia), while the second mainly consisted of introduced varieties from foreign  
163 countries and other provinces in China (Others). Moreover, accessions from Ningxia  
164 grouped into two clades for landraces and modern cultivars, respectively (Figure S3).  
165 This indicated characteristics of genetic components of Ningxia wheats, in which  
166 landraces, introduced varieties and modern cultivars in different times together  
167 formed wheat breeding process.

### 168 **Genetic divergence between Ningxia and Others**

169 To further clarify the large genetic differences between germplasms from Ningxia and  
170 Others, genetic diversities and variations were assessed. There was apparent

171 difference in genetic diversity at 44 loci controlling yield, quality, adaptation and  
172 stress resistance between Ningxia wheat germplasms and Others (Figure 2A). Further  
173 exploration indicated a higher genetic diversity at ten grain yield loci in the group  
174 Others compared with Ningxia wheat varieties ( $P < 0.01$ ), whereas Ningxia had a  
175 higher genetic diversity than Others at 14 quality genes ( $P < 0.05$ ) (Figure S4A, B).  
176 Among them, genetic diversities estimated at *Cwi-4A*, *GS-D1*, *Sus2-2B* and *Sus1-7B*  
177 loci for yield were abundant in the Others subgroup, while genes of *Glu-B1*, *Glu-D1*,  
178 *Pina-D1* and *Zds-A1* for quality showed much higher genetic diversities in the  
179 Ningxia subgroup relative to Others (Table S3). In addition, we also found that  
180 genetic divergence was most obviously at quality genes, followed by yield genes  
181 (Figure 2B). An in-depth analysis showed an evident genetic divergence at some loci  
182 such as *Cwi-4A* (0.035), *Sus2-2B* (0.035) for grain yield, *Pinb-D1* (0.057) and *Zds-A1*  
183 (0.064) for quality for Ningxia and Others subgroups (Table S3). To investigate  
184 genetic divergences on gene level, we evaluated allele frequencies at genes for grain  
185 yield, quality, adaptation and stress resistance as indicated below.

#### 186 **Grain yield genes**

187 Among ten yield-related genes, alleles of *Hap-4A-C* (*Cwi-4A*), *GS-D1a* (*GS-D1*),  
188 *Hap-A* (*GW2-6A*) and *Hap-1* (*GW2-6B*) for larger grain size and TKW were  
189 predominant in the subgroup Ningxia compared with Others (Figure 2C), whereas at  
190 *Sus2-2B*, the allele *Hap-H* associated with higher TKW was more frequently in the  
191 subgroup Others. Besides, the desirable alleles at *Sus1-7A* and *TGW6* loci tended to  
192 occur in higher frequencies in both two subgroups. However, the favorable alleles of  
193 *H1c* and *Hap-H* at *GASR-A1* and *MOC-7A*, respectively, were rarely present in two  
194 subgroups, indicating that these were not selected by breeders.

#### 195 **Quality genes**

196 At 14 loci for quality traits, a higher frequency of *Glu-D1d* encoding  
197 high-molecular-weight glutenin subunits (HMW-GS) Dx5+Dy10 occurred more  
198 frequently in the subgroup Ningxia (44%) than Others (27%) (Figure 2D). The  
199 *Ppo-A1b*, *Pod-A1b* and *Zds-A1a* alleles, associated with lower PPO activity, higher  
200 POD activity and lower yellow pigment content, respectively, were more frequent in

201 the subgroup Ningxia than Others. In contrast, the frequencies of HMW-GSs Ax1 or  
202 Ax2\* and *Pinb-D1b* for hard grain texture were significantly higher in the subgroup  
203 Others. Additionally, the majority of two subgroups had high proportion of *Pinb-B2b*  
204 for hard grain texture, *IBL.IBL*, and *Psy-A1b*, *Psy-D1a* and *Psy-B1a or b* for low  
205 yellow pigment content, whereas the minority had alleles of *Pina-D1b*, *Pds-B1b* and  
206 *Wx-B1b* at *Pina-D1*, *Pds-B1* and *Wx-B1* loci, respectively.

#### 207 **Adaptation genes**

208 Concerning adaptation genes, genetic diversities in two subgroups were not  
209 significantly different (Figure S4C). Allelic variations had minor differences at  
210 *Rht-B1*, *Rht-D1*, *Vrn-B1*, *Vrn-D1* and *Ppd-D1* in both subgroups Ningxia and Others  
211 except at *Vrn-A1* locus (Figure S5A), indicating that spring-type alleles of *Vrn-B1b*  
212 and *Vrn-D1a*, photoperiod-insensitive allele *Ppd-D1a*, *Rht-B1b* and *Rht-D1b* reducing  
213 plant height distributed consistently in two subgroups.

#### 214 **Stress resistance genes**

215 Loci for stress resistance included *PHS1*, *Sdr-B1*, *VP-1B* and *MFT-A1* associated with  
216 seed dormancy, *1-fehw3* and *Dreb-B1* related to drought stress, *Lr14a*, *Lr34*, *Lr68*,  
217 *Yr15* and *Fhb1* for disease resistance, *DRO-5A* and *DRO-5B* for root architectures.  
218 The favorable alleles *Sdr-B1a*, *Vp-1Bc* and *Dreb-B1a* were higher in the subgroup  
219 Ningxia than Others (Figure S5B), whereas the favorable alleles of *PHS+* at *PHS1*,  
220 *PHS+* at *MFT-A1* and *Lr14a* occurred in higher frequencies in the subgroup Others  
221 compared with Ningxia. The drought-resistance alleles of *Westonia type* were not  
222 significantly different between two subgroups Ningxia and Others, while alleles  
223 *Lr68+*, *Yr15+* and *Fhb1+* were rarely present in two subgroups.

#### 224 **Genetic divergence between landraces and modern cultivars in Ningxia** 225 **accessions**

226 To evaluate the population differentiation during breeding improvement in Ningxia  
227 province, we investigated the genetic structure based on 44 genes (Figures 3A, S3).  
228 The results showed that wheat accessions from Ningxia clustered in two clades, i.e.  
229 landraces and modern cultivars, respectively. Therefore, we further analyzed genetic

230 relationship between landraces and modern cultivars. We found a higher genetic  
231 diversity in modern cultivars compared with landraces (Figure 3B). Moreover, the  
232 difference of genetic diversities was clearly manifested in adaptation-related genes,  
233 while no significant differences between two groups were observed in other genes  
234 controlling yield, quality and resistance (Figure S6). Population differentiation (*Fst*)  
235 and gene frequency against the four types of genes were also analyzed to reveal  
236 substantial divergences between two subgroups below.

### 237 **Adaptation genes**

238 The *Fst* between modern cultivars and landraces was very high at *Vrn-A1* (0.39),  
239 followed by *Rht-B1* (0.16) (Figure 4A). Similarly, the spring-type allele *Vrn-A1a* at  
240 *Vrn-A1* influencing vernalization occurred frequently in modern cultivars (57%), but  
241 absent in landraces (Figure 4B). In contrast, *Vrn-B1b* associated with spring type was  
242 predominant both in modern cultivars (58%) and landraces (82%); and *Vrn-D1a* for  
243 spring type retained a towering scaling in modern cultivars (57%) and landraces (86%)  
244 (Figure S7). The dwarfing allele *Rht-B1b* (*Rht-B1*) was present in 28% modern  
245 cultivars but absent in landraces (Figure 4B). A similar situation did happen in  
246 another dwarfing gene *Rht-D1b*. Interestingly, the photoperiod-insensitive allele  
247 (*Ppd-D1a*) predominated in modern cultivars and landraces (Figure S7).

### 248 **Stress resistance genes**

249 For genes controlling stress resistance, the *Fst* among two subgroups at *Lr34* and  
250 *1-fehw3* loci were extremely high compared with other resistance genes, which were  
251 0.55 (*Lr34*) and 0.43 (*1-fehw3*), respectively (Figure 4C). Allele frequency analyses  
252 showed that the favorable alleles of *Lr34+* and *Westonia type* at *1-fehw3* were  
253 predominant in landraces (Figure 4D). The genetic differentiation (*Fst*) of *DRO-5B*,  
254 *Lr68*, *PHS1*, *VP-1B* and *Lr14a* was 0.30, 0.22, 0.22, 0.14 and 0.12, respectively  
255 (Table S4), whereas corresponding favorable allele frequencies had distinct  
256 differences between modern cultivars and landraces (5% vs 54%, 5% vs 47%, 50% vs  
257 92%, 67% vs 30%, 22% vs 0%) (Figure S8).

### 258 **Quality genes**

259 For quality genes, the most extreme genetic differentiation occurred in *Pod-A1* (0.23),

260 followed in *Pinb-D1* (0.24) between two subgroups (Figure 4E). The majority  
261 of modern cultivars (55%) had the *Pod-A1b* allele, whereas few landraces had this  
262 allele. The hard grain texture allele (*Pinb-D1b*) was frequently present in modern  
263 cultivars (38%), but it was absent in landraces (Figure 4F), which was verified at  
264 genetic differentiation between two subgroups (Figure 4E). Additionally, we found  
265 that loci including *Glu-D1*, *PSY1-D1* and *Pinb2-B2* had obvious genetic difference  
266 (Table S4), and corresponding allele frequencies were significantly different between  
267 cultivars and landraces (Figure S9).

### 268 **Grain yield genes**

269 For yield-related genes, the most significant difference occurred in *TaGW2-6B* (Figure  
270 4G), at which the favorable allele *Hap-1* was predominant in modern cultivars (76%),  
271 but absent in landraces (Figure 4H). At *TGW6*, *Cwi-4A* and *GS-D1* loci, the favorable  
272 allele frequencies were higher in modern cultivars than in landraces (Figure S10),  
273 whereas a contrary situation was observed at *GASR-A1*, *Sus1-7A* and *GW2-6A* loci.

### 274 **Genetic contribution from founder parents for Ningxia wheat cultivars**

275 Founder parents, as one of important genetic resources, have greatly promoted  
276 improvement of wheat varieties in China since the 1950s. In this study, we analyzed  
277 genetic contributions of six founder parents, including Moba 66, Abbondanza, Beijing  
278 8, Orofen, Xiaoyan 6 and Zhou 8425B, to modern cultivars in Ningxia province. To  
279 clearly understand the importance of founder parents, we counted the number of  
280 favorable alleles of genes for yield, quality, adaptation and stress resistance in these  
281 cultivars (Figure 5). Results showed that the number of favorable alleles for higher  
282 TKW in six founder parents ranged from three to seven at ten yield genes. The  
283 founder parent Xiaoyan 6 carried seven favorable alleles that were *Hap-4A-C* at  
284 *Cwi-4A*, *GS-D1a* at *GS-D1*, *Hap-H* at *Sus1-7A* and *2B*, *Hap-A* at *GW2-6A*, *Hap-1* at  
285 *GW2-6B* and *TGW6-A1a* at *TGW6* (Figure S11). At 14 quality genes, the average  
286 favorable alleles were about six, ranging from five to nine. The founder parent  
287 Xiaoyan 6 carried the most of favorable alleles, which was *Ax1* or *Ax2\** at *Glu-A1*,  
288 *Glu-D1d* at *Glu-D1*, *Zds-A1a* at *Zds-A1*, *Pds-B1b* at *Pds-B1*, *Pod-A1b* at *Pod-A1*,

289 *Psy-A1b* at *Psy-A1*, *Psy-D1a* at *Psy1-D1*, *Psy-B1a* or *b* at *Psy-B1* and *Wx-B1b* at  
290 *Wx-B1*. All six founder parents carried *Psy-A1b*, *Psy-D1a* and *Psy-B1a* or *b*  
291 associated with low YP content except *Psy-D1* for Zhou 8425B (Figure S12). For  
292 stress resistance, the founder parent Abbondanza had eight favorable alleles including  
293 *PHS+* at *PHS1*, *Vp-1Bc* at *VP-1B*, *PHS+* at *MFT-A1*, *Westonia type* at *1-fehw3*,  
294 *Dreb-B1a* at *Dreb-B1*, *Hap-5A-A* at *DRO-5A*, *Hap-5B-II* at *DRO-5B* and *Lr14+* at  
295 *Lr14a* (Figure S13). At six adaptation genes, the photoperiod-insensitive allele  
296 (*Ppd-D1a*) was present in all founder parents except Orofen. Besides, Moba 66,  
297 Xiaoyan 6 and Zhou 8425B had the dwarfing allele *Rht-B1b*. And Zhou 8425B also  
298 had another dwarfing allele *Rht-D1b* (Figure S14). Such evaluation of these founder  
299 parents from different types of functional genes allowed us to aim at contributions to  
300 breeding improvement in Ningxia province.

301 To compare differences at four types of genes between founder parents and  
302 modern cultivars, we investigated relationship number of alleles with the proportion  
303 of accessions. The 17% of founder parents had seven and nine favorable alleles at  
304 yield and quality genes, respectively, while the proportion was 7% on average in  
305 modern cultivars (Figures 5B, C). Most founder parents and modern cultivars had  
306 three to six allelic variations at resistance genes and carried dwarfing, spring-type and  
307 photoperiod-insensitive alleles at adaptation genes (Figures 5D, E).

308 The gene flow value at yield genes was 2.47 between modern cultivars and Zhou  
309 8425B, which is most frequent among all founder parents, indicating that Zhou 8425B  
310 had the largest genetic exchange with modern cultivars and played an important role  
311 in yield potential in Ningxia wheats (Figure 6). At quality and resistance genes, all  
312 founder parents had nearly equal gene flow to modern cultivars, with an average gene  
313 flow of 0.60 at quality loci ranging from 0.48 to 0.76, and 0.53 at resistance loci  
314 ranging from 0.40 to 0.69. For adaptation genes, the gene flow values between  
315 founder parents Abbondanza, Orofen and modern cultivars were 0.82 and 0.63,  
316 respectively. In a word, founder parents with different favorable alleles pushed  
317 together improvement of Ningxia wheat cultivars.

318 **Discussion**

319 **Population structure indicated wheat genetic components in Ningxia province**

320 Ningxia province, located in north-western of China, is one of the major spring wheat  
321 grown regions and has been experienced five times of varietal replacements, i.e.,  
322 direct utilization of ‘Quality Florence’, a wheat germplasm introduced from Australia  
323 in 1950s; efficient use of ‘Abbondanza’ wheat resource from Italy in 1960s; breeding  
324 of milestone variety ‘Doudi 1’ in 1970s; application and improvement of ‘Ningchun 4’  
325 in 1980s, and release of ‘Ningchun 50’ in 2000s [48]. Looking back to wheat breeding  
326 history, it is not difficult to find foreign wheat germplasm that has played an  
327 important role in China, especially from establishment of the People’s Republic of  
328 China in 1949 to present. Of course, landraces with particular genetic resources have  
329 been applied to improve cultivars in breeding program. In this study, all varieties  
330 could be clustered into three subgroups based on the population structure analyses,  
331 namely modern cultivars, landraces and others, which showed their genetic  
332 differences based on 44 important functional genes (Figures 3A, S3). The three  
333 genetic divergent groups were in accordance with wheat breeding history of utilizing  
334 wheat genetic resources from other regions and local landraces in Ningxia province.

335 Ningxia, being diversified ecological types, is among the provinces that could  
336 benefit from the use of introduced germplasm and landraces from genebanks in China.  
337 To further clarify the contribution of introduced germplasm and landraces during the  
338 successive decades of breeding process in Ningxia province, genetic diversities and  
339 the frequency spectra divergence were evaluated based on 44 genes for yield, quality,  
340 adaptation and resistance. In this study, genetic diversities are most enriched for the  
341 subgroup Others comprising introductions outside of Ningxia province, followed by  
342 modern cultivars, and landraces have minimum genetic diversities at 44 important loci  
343 (Figures 2A, 3B). This indicates that conventional artificial hybridizing breeding  
344 using external resources to improve varieties has increased diversities by promoting  
345 gene exchanges and recombination in gene coding regions in particular important  
346 cloned functional genes for self-pollinated wheat [49]. Besides, introduced varieties

347 from other regions with rich genetic diversities facilitate local breeding improvement,  
348 which means introduced varieties made a significant genetic contribution to Chinese  
349 modern cultivars [41]. Allele frequency spectra divergence also supports this point on  
350 gene levels. Dwarfing alleles *Rht-B1b* at *Rht-B1* and *Rht-D1b* at *Rht-D1*, the  
351 well-known green revolution genes that swept through China and occupied  
352 significantly positive influence on wheat breeding, are very widespread in Ningxia  
353 modern cultivars and introductions from other regions but absent in landraces  
354 (Figures 4B, S5A, S7), indicating that introduced wheat germplasm with numerous  
355 beneficial genes are vital for broadening the genetic diversities of Ningxia wheat  
356 cultivars. Similar types of genes also include *IBL/IRS* translocation, *GW2-6B*, and  
357 *Sus2-2B* (Figures 2D, S8, S10). As autochthonous traditional varieties, evaluation of  
358 wheat landraces stored in gene banks with highly beneficial untapped diversity and  
359 sources of stress adaptation should be used for wheat improvement [50]. Due to  
360 colonization of diverse ecological environments in the process of domestication and  
361 selections by ancient farmers in Ningxia province, landraces contain wider specific  
362 genetic loci than most breeding programs and form the basis of early wheat breeding,  
363 especially for China in pre-1950s. Zhou et al [5] highlights environmental stresses and  
364 independent selection efforts that have resulted in considerable genome-wide  
365 divergence at the population level in Chinese wheat landraces. Of course, this  
366 characteristic has been exploited in other countries, where the first improved varieties  
367 consisted of selections of local landraces [50], such as landrace population 'Catalan  
368 de Monte' in Spain [52-53] and 'Turkey Red' in the United States [54]. Overall, our  
369 results indicate the point that landraces with good adaptation and introduced varieties  
370 with wide diversities will co-promote wheat breeding in Ningxia province.

### 371 **Dissecting allele frequency identified selection direction of important genes**

372 The modern wheat breeding practices accompanying intensive selection pressure have  
373 always focused on economically important loci [55-56]. For each of those loci  
374 contributing to agronomic phenotypes, causal polymorphisms were identified with  
375 increased frequencies of favorable alleles consistent with selection during modern

376 breeding [57]. For improvement selection for Ningxia wheat from landraces to  
377 modern cultivars in this study, we found the evidence of convergent increases in allele  
378 frequencies at targeted genes. *VP-1B* is one of important seed dormancy genes for  
379 PHS tolerance during harvest [58] and the favorable allele *Vp-1Bc* was predominant  
380 with a frequency of 68% in Ningxia modern cultivars. Similarly, *Hap-1* at *GW2-6B*  
381 locus influencing strongly kernel width and thousand kernel weight was observed in  
382 76% modern cultivars, while the desirable allele was absent in landraces, indicating  
383 that breeders selected intensively the favorable allele at *GW2-6B* locus due to the  
384 demand for increasing grain yield in breeding. HMW-GS was influenced at *Glu-D1*  
385 [23]. The allele *Glu-D1d* associated with strong gluten contents and superior  
386 bread-making quality attributes had relatively high frequency in modern cultivars  
387 compared with landraces, which is concordant with the reports of most cultivars in  
388 Pakistani and China [22, 55-56]. Besides, the favorable allele *Hap-4A-C* at *Cwi-4A*  
389 encoding CWI enzyme that converts sucrose to glucose associated with grain size was  
390 present in 85% of modern cultivars, showing the effective use of this gene in wheat  
391 molecular breeding.

392 For some important genes, their favorable alleles had maintained high values in  
393 varieties before modern wheat breeding [41]. The photoperiod insensitive allele,  
394 *Ppd-D1a*, was fixed with a frequency of 100% in both landraces and modern cultivars,  
395 showing that this gene is so important that it has been selected completely before  
396 modern breeding in Ningxia. As we know, flowering time is one of the most important  
397 developmental traits for wheat adaptability and yield stability in target environments,  
398 and the photoperiod insensitive allele at photoperiod response (*Ppd1*) is known to be  
399 major determinants of flowering time optimization [59]. Therefore, early flowering  
400 with varieties carrying photoperiod insensitive allele was fixed during long-time  
401 selection process. High yellow pigment content is favored for durum wheat pasta, but  
402 is considered undesirable for Chinese steamed bread and white noodles [60-62] and  
403 the alleles *Psy-A1b* and *Psy-B1a/b* are encouraged. The frequencies of two alleles  
404 were approached almost 100% and were fixed in both landraces and modern cultivars  
405 before breeding selection in Ningxia.

406           However, to breed perfect wheat varieties in Ningxia, these favorable alleles with  
407 minor frequencies in modern cultivars should be regarded. In this study, the root  
408 architecture related gene *DRO-5B* is an IAA response gene which is responsible for  
409 reduced height and increased thousand kernel weight [63-64]. The favorable allele  
410 *Hap-5B-II* was present in 54% of landraces, but only 5% of modern cultivars,  
411 indicating that this allele has not been focused by breeders in the past. Besides, *H1c* at  
412 *GASR-A1* locus which influences grain length were pre-dominant in landraces, but  
413 had a low frequency in modern cultivars. This comes to the point that breeding is a  
414 process of aggregating desirable genes and eliminating undesirable or even  
415 deleterious alleles. Low frequencies of favorable alleles for important genes in  
416 modern cultivars identified the improvement direction in future wheat breeding in  
417 Ningxia and are helpful to further performing breeding by design.

#### 418 **Founder parents contain the combination of important functional genes**

419 Founder parents, which serve as important germplasm resources, play a pivotal role in  
420 update of new varieties [65]. They exhibited not only superior phenotypes, high  
421 combining ability, but also wide adaptation and prominent specific characteristic [66].  
422 Previous studies found that genes controlling important traits were the combination  
423 rather than randomly distributed on chromosomes in founder parents [67-72]. For  
424 example, the pedigree analysis of rice Huanghuazhan showed 61.79% of 50 kb blocks  
425 are HTBs (Huanghuazhan traceable blocks) together with the elite performance of  
426 Huanghuazhan, and large scale of important genes locate on HTBs, supporting that  
427 they are the combination of elite allele of important genes [72]. In this study, counting  
428 number of favorable alleles at ten yield-related genes had successfully clarified 3~7  
429 favorable alleles for higher TKW in six founder parents. Furthermore, *Hap-4A-C* at  
430 *Cwi-4A*, *Hap-H* at *Sus1-7A*, *GS-D1a* at *GS-D1*, *Hap-A* at *GW2-6A*, *Hap-1* at *GW2-6B*  
431 and *TGW6-A1a* at *TGW6* were conserved in founder parents Zhou 8425B, Xiaoyan 6  
432 and Abbondanza except for *GS-D1a* allele. For quality-related genes, the average  
433 number of favorable alleles was about six, and *Psy-A1b*, *Psy-D1a*, *Psy-B1a* or *b*  
434 associated with low YP content were conserved across all founder parents, showing

435 that favorable alleles of these important genes had been conserved in modern wheat  
436 breeding. Interestingly, six founder parents carried different favorable alleles at  
437 resistance and adaptation related genes, probably due to the reason that these genes  
438 are randomly selected to respond to varied environments, so that these founder  
439 parents can maintain high yield and good quality wherever cultivated.

440 Founder parents have excellent allele combinations of important genes for  
441 agronomical desirable traits, and many varieties have been derived from them. In this  
442 study, yield traits improvement of modern cultivars is the main achievement using  
443 founder parents in Ningxia wheat breeding. Gene flow was most frequent (2.47)  
444 comparing the modern cultivars with founder parents Zhou 8425B for yield related  
445 genes, meaning that founder parents especially from Zhou 8425B contributed greatly  
446 in yield improvement of Ningxia wheat. Zhou 8425B is a founder parent fitting  
447 current breeding needs with features of dwarfing, high yielding and disease resistance,  
448 and more than 300 wheat varieties (lines) such as AK58, Zhou 16 were bred from this  
449 parent [65, 72-73]. High yield is an ever-important objective of wheat breeding, and  
450 analysis of the breeding history of many crop species revealed the presence and roles  
451 of founder parents [67]. Li et al [75] found that Beijing 8 serving as a founder parent  
452 has contributed many loci in close proximity to the positions of known yield  
453 component genes conferring important traits in breeding. The pedigree analysis  
454 showed that inherited ancestor genome segments donated to rice variety  
455 Huanghuazhan extremely enrich in grain yield category [72].

#### 456 **New era of functional markers is advanced through progress of cloning genes** 457 **with published reference genome**

458 Functional markers that have strong associations with relevant phenotypes, are ideal  
459 for gene tagging and allelic variants can be diagnosed with functional genes in  
460 breeding [30, 76]. Liu et al. [30] documented 97 functional markers that detect 93  
461 alleles at 30 loci in bread wheat. Rasheed et al [36] converted gel-based functional  
462 markers to high-throughput KASP markers. In this study, we evaluate molecular  
463 characterization and improvement history of Ningxia wheat breeding in terms of

464 important genes related with adaptation, stress resistance, quality and yield utilizing  
465 these KASP functional markers. However, an objective fact is that these in our  
466 research are only a few number of predicted genes relative to wheat whole genome.  
467 Currently, most gene mapping studies (both QTL and GWAS) could identify more  
468 genes controlling agronomic traits and have been converted to KASP markers [76].  
469 More than 150 KASP markers for almost 100 functional genes were developed and 72  
470 were validated in a bread wheat diversity panel [77]. With the innovations of  
471 whole-genome assemblies, revolutionized advances of reference genome sequences  
472 were recently gained for bread wheat ‘Chinese Spring’ [78] and its progenitors, *T.*  
473 *turgidum* spp. *dicoccoides* [79], *Aegilops tauschii* [80-81] and *T. urartu* [82]. These  
474 wheat genome data provide new opportunities to uncover genetic variation in traits of  
475 breeding interest and enable genome-based breeding to deliver wheat cultivars.  
476 High-throughput genotyping technology of KASP is now used in wheat not only for  
477 breeding by design but also for high-density genome-wide genotyping. These  
478 advances will promote wheat research to dig more important genes of targeting traits  
479 and ensure sustainable wheat production through developing widely adapted wheat  
480 cultivars with high yield and nutritional quality.

## 481 **Conclusions**

482 In this study, we report a comprehensive functional gene assessment of modern  
483 improved wheat based on 44 important genes underpinning grain yield, quality,  
484 adaptation and resistance in 207 cultivars and lines in Ningxia province. Introduced  
485 varieties from other regions with rich genetic diversities and landraces with  
486 well-adapted genetic resources have been applying to improve modern cultivars.  
487 Founder parents in particular Zhou 8425B for yield related genes contributed greatly  
488 to breeding improvement of wheat in Ningxia province. This work reports genetic  
489 characteristics on gene levels and advances improvement selection of future wheat  
490 breeding at Ningxia province.

## 491 **Additional files**

492 **Additional file1:** Table S1. Detailed information of materials and their allelic  
493 variations of 44 genes used in this study.

494 **Additional file1:** Table S2. Basic information including allelic variations and primer  
495 sequences for the 44 KASP assays.

496 **Additional file1:** Table S3. Gene diversities and genetic differentiation (*Fst*) of 44  
497 polymorphic genes between Ningxia and Others accessions.

498 **Additional file1:** Table S4. Gene diversities and genetic differentiation (*Fst*) of 44  
499 polymorphic genes between landraces and modern cultivars in Ningxia province.

500 **Additional file2:** Figure S1. KASP genotyping at *Rht-D1*, *VP-1B* and *Glu-A1*. Red  
501 and blue dots show homozygous varieties; Green dots show heterozygous varieties;  
502 Black dots show negative control; X shows missing types.

503 **Additional file3:** Figure S2. Plot of Delta K against putative K ranging from 1 to 8.

504 **Additional file4:** Figure S3. Population structure of all accessions based on Structure  
505 from K=2 to K=3.

506 **Additional file5:** Figure S4. Genetic diversities on four types of genes between  
507 Ningxia and Others subgroups. (A) Genetic diversities on grain yield genes. (B)  
508 Genetic diversities on quality genes. (C) Genetic diversities on adaptation genes. (D)  
509 Genetic diversities on stress resistance genes.

510 **Additional file6:** Figure S5. Allele frequencies between Ningxia and Others  
511 subgroups at adaptation (A) and stress resistance (B) genes.

512 **Additional file7:** Figure S6. Genetic diversities on four types of genes between  
513 modern cultivars and landraces subgroups in Ningxia province. (A) Genetic  
514 diversities on adaptation genes. (B) Genetic diversities on stress resistance genes. (C)  
515 Genetic diversities on quality genes. (D) Genetic diversities on grain yield genes.

516 **Additional file8:** Figure S7. Allele frequencies of adaptation genes between modern  
517 cultivars and landraces in Ningxia province.

518 **Additional file9:** Figure S8. Allele frequencies of stress resistance genes between  
519 modern cultivars and landraces in Ningxia province.

520 **Additional file10:** Figure S9. Allele frequencies of quality genes between modern  
521 cultivars and landraces in Ningxia province.

522 **Additional file11:** Figure S10. Allele frequencies of grain yield genes between  
523 cultivars and landraces in Ningxia province.

524 **Additional file12:** Figure S11. Distribution of allelic variations of grain yield genes in  
525 six founder parents.

526 **Additional file13:** Figure S12. Distribution of allelic variations of quality genes in six  
527 founder parents.

528 **Additional file14:** Figure S13. Distribution of allelic variations of stress resistance  
529 genes in six founder parents.

530 **Additional file15:** Figure S14. Distribution of allelic variations of adaptation genes in  
531 six founder parents.

## 532 **Abbreviations**

533 *Fst*: F-statistics; HWM-GS: High-Molecular-Weight Glutenin Subunits; KASP:  
534 Kompetitive Allele Specific PCR; PCA: Principal Coordinate Analysis; MCMC:  
535 Markov Chain Monte Carlo; TKW: Thousand Kernel Weight; SNP: Single Nucleotide  
536 Polymorphism; LGC: Laboratory of the Government Chemist; HTBs: Huanghuazhan  
537 Traceable Blocks

## 538 **Acknowledgements**

539 We gratefully acknowledge help from Dr. Xianchun Xia, Institute of Crop Sciences,  
540 Chinese Academy of Agricultural Sciences, with valuable suggestions and English  
541 editing.

## 542 **Funding**

543 This work was supported by Introduction of KASP technology and its application in  
544 identification and evaluation of wheat germplasm resources in Ningxia  
545 (DWHZC-2017006), Ningxia Agricultural Breeding Project (2018NYYZ02) and  
546 Ningxia Province central guidance local science and technology development project  
547 (2020YDDF0056).

## 548 **Authors' Contributions**

549 Weijun Zhang carried out the experiments and wrote the manuscript; Junjie Zhao  
550 helped to carry out the experiments and analyzed the data; Jinshang He guided the use  
551 of instruments; Ling Kang prepared a part of figures for the manuscript; Xiaoliang  
552 Wang and Fuguo Zhang gave suggestions for the experiments; Chenyang Hao  
553 participated in the design of experiments and contributed to writing the manuscript;  
554 Xiongfeng Ma contributed to overall design of the experiments; Dongsheng Chen  
555 provided advice for data analysis and assisted in writing the manuscript. All authors  
556 have read and approved the final version.

557 **Availability of data and materials**

558 The datasets generated and analyzed during the current study are available from the  
559 corresponding author on reasonable requests.

560 **Conflicts of interest/Competing interests**

561 The authors declare no conflicts of interest.

562 **References**

- 563 1. Zhang QF, Xu ML, Xia XC, Komatsuda T, Varshney RK, Shi K. Crop genetics  
564 research in Asia: improving food security and nutrition. *Theor Appl Genet.*  
565 2020;133:1339-1344.
- 566 2. Dvorak J, McGuire PE, Cassidy B. Apparent sources of the A genomes of wheats  
567 inferred from polymorphism in abundance and restriction fragment length of  
568 repeated nucleotide sequences. *Genome.* 1988;30:680-689.
- 569 3. Dvorak J, Terlizzi PD, Zhang H, Resta P. The evolution of polyploid wheats:  
570 Identification of the A genome donor species. *Genome.* 1993;36:21-31.
- 571 4. Kilian B, Ozkan H, Deusch O, Effgen S, Brandolini A, Kohl J, et al. Independent  
572 wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes.  
573 *Mol Biol Evol.* 2017;24:217-227.
- 574 5. Feldman M, Levy AA. Genome evolution due to allopolyploidization in wheat.  
575 *Genetics.* 2012;192:763-774.
- 576 6. Tanno K, Willcox G. How fast was wild wheat domesticated? *Science.*  
577 2006;311:1886.
- 578 7. Pont C, Leroy T, Seidel M, Tondelli A, Duchemin W, Armisen D, et al. Tracing  
579 the ancestry of modern bread wheats. *Nat Genet.* 2019;51:905-911.
- 580 8. He F, Pasam RK, Shi F, Kant S, Keeblegagnere G, Kay P, et al. Exome  
581 sequencing highlights the role of wild-relative introgression in shaping the  
582 adaptive landscape of the wheat genome. *Nat Genet.* 2019;51:896-904.
- 583 9. Jordan KW, Wang SC, Lun YN, Gardiner LJ, MacLachlan R, Hucl P, et al. A  
584 haplotype map of allohexaploid wheat reveals distinct patterns of selection on  
585 homoeologous genomes. *Genome Biol.* 2015;16:48.
- 586 10. Zhang XK, Xiao YG, Zhang Y, Xia XC, Dubcovsky J, He ZH. Allelic variation at  
587 the vernalization genes *Vrn-A1*, *Vrn-B1*, *Vrn-D1*, and *Vrn-B3* in Chinese wheat  
588 cultivars and their association with growth habit. *Crop Sci.* 2008;48:458-470.
- 589 11. Yang FP, Zhang XK, Xia XC, Laurie DA, Yang WX, He ZH. Distribution of the  
590 photoperiod insensitive *Ppd-D1a* allele in Chinese wheat cultivars. *Euphytica.*

- 591 2009;165: 445-452.
- 592 12. Yediay FE, Andeden EE, Baloch FS, Borner A, Kilian B, Ozkan H. The allelic  
593 state at the major semi-dwarfing genes in a panel of Turkish bread wheat cultivars  
594 and landraces. *Journal of Plant Genetic Resources*. 2011;9:423-429.
- 595 13. Jiang QY, Hou J, Hao CY, Wang LF, Ge HM, Dong YS, et al. The wheat (*T.*  
596 *aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is  
597 associated with yield traits. *Funct Integ Genomics*. 2011;11:49-61.
- 598 14. Hou J, Jiang QY, Hao CY, Wang YQ, Zhang HN, Zhang XY. Global selection on  
599 sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol*.  
600 2014;164:1918-1929.
- 601 15. Su ZQ, Hao CY, Wang LF, Dong YC, Zhang XY. Identification and development  
602 of a functional marker of *TaGW2* associated with grain weight in bread wheat  
603 (*Triticum aestivum* L.). *Theor Appl Genet*. 2011;122:211-223.
- 604 16. Yang ZB, Bai ZY, Li XL, Wang P, Wu QX, Yang L, et al. SNP identification and  
605 allelic-specific PCR markers development for *TaGW2*, a gene linked to wheat  
606 kernel weight. *Theor Appl Genet*. 2012;125:1057-1068.
- 607 17. Qin L, Hao CY, Hou J, Wang YQ, Li T, Wang LF, et al. Homologous haplotypes,  
608 expression, genetic effects and geographic distribution of the wheat yield gene  
609 *TaGW2*. *BMC Plant Biol*. 2014;14.
- 610 18. Ma L, Li T, Hao CY, Wang YQ, Chen XH, Zhang XY, et al. *TaGS5-3A*, a grain  
611 size gene selected during wheat improvement for larger kernel and yield. *Plant*  
612 *Biotechnol J*. 2016;14:1269-1280.
- 613 19. Ma DY, Yan J, He ZH, Wu L, Xia XC. Characterization of a cell wall invertase  
614 gene *TaCwi-A1* on common wheat chromosome 2A and development of  
615 functional markers. *Mol Breeding*. 2012;29:43-52.
- 616 20. Zhang L, Zhao YL, Gao LF, Zhao GY, Zhou RH, Zhang BS, et al. *TaCKX6 - D1*,  
617 the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat.  
618 *New Phytol*. 2012;195:574-584.
- 619 21. Dong L, Wang F, Liu T, Dong Z, Li A, Jing RL, et al. Natural variation of  
620 *TaGASR7-A1* affects grain length in common wheat under multiple cultivation

- 621 conditions. Mol Breeding. 2014;34:937-947.
- 622 22. He ZH, Xia XC, Zhang Y, Zhang Y, Xiao YG, Chen XM, et al. China-CIMMYT  
623 collaboration enhances wheat improvement in China. Front Agric Sci Eng. 2019;  
624 6:233-239.
- 625 23. Luo C, Griffin WB, Branlard G, Mcneil DL. Comparison of low and high  
626 molecular weight wheat glutenin allele effects on flour quality. Theor Appl Genet.  
627 2019;102:1088-1098.
- 628 24. Sun DJ, He ZH, Xia XC, Zhang LP, Morris CF, Appels R, et al. A novel STS  
629 marker for polyphenol oxidase activity in bread wheat. Mol Breeding.  
630 2005;16:209-218.
- 631 25. He XY, He ZH, Zhang LP, Sun DJ, Morris CF, Fuerst EP, et al. Allelic variation  
632 of polyphenol oxidase (*PPO*) genes located on chromosomes 2A and 2D and  
633 development of functional markers for the *PPO* genes in common wheat. Theor  
634 Appl Genet. 2007;115:47-58.
- 635 26. He XY, He ZH, Ma WJ, Appels R, Xia XC. Allelic variants of phytoene synthase  
636 1 (*Psy1*) genes in Chinese and CIMMYT wheat cultivars and development of  
637 functional markers for flour colour. Mol Breeding. 2009;23:553-563.
- 638 27. He XY, Wang JW, Ammar K, Peña RJ, Xia XC, He ZH. Allelic variants at the  
639 *Psy-A1* and *Psy-B1* loci in durum wheat and their associations with grain  
640 yellowness. Crop Sci. 2009;49:6.
- 641 28. Zhang CY, Dong CH, He XY, Zhang LP, Xia XC, He ZH. Allelic variants at the  
642 *TaZds-D1* locus on wheat chromosome 2DL and their association with yellow  
643 pigment content. Crop Sci. 2011;51:1580-1590.
- 644 29. Wei JX, Geng HW, Zhang Y, Liu JD, Wen WE, Zhang Y, et al. Mapping  
645 quantitative trait loci for peroxidase activity and developing gene-specific  
646 markers for *TaPod-A1* on wheat chromosome 3AL. Theor Appl Genet.  
647 2015;128:2067-2076.
- 648 30. Liu YN, He ZH, Appels R, Xia XC. Functional markers in wheat: Current status  
649 and future prospects. Theor Appl Genet. 2012;125:1-10.
- 650 31. Wei B, Jing RL, Wang CS, Chen JB, Mao XG, Chang XP, et al. *Dreb1* genes in

- 651 wheat (*Triticum aestivum* L.): development of functional markers and gene  
652 mapping based on SNPs. *Mol Breeding*. 2012;23:13-22.
- 653 32. Li GQ, Zhou JY, Jia HY, Gao ZX, Fan M, Luo YJ, et al. Mutation of a  
654 histidine-rich calcium-binding-protein gene in wheat confers resistance to  
655 *Fusarium* head blight. *Nat Genet*. 2019;51:1106-1112.
- 656 33. Su ZQ, Bernardo A, Tian B, Chen H, Wang S, Ma HX, et al. A deletion mutation  
657 in *TaHRC* confers *Fhb1* resistance to *Fusarium* head blight in wheat. *Nat Genet*.  
658 2019;51:1099-1105.
- 659 34. Wang HW, Sun SL, Ge WY, Zhao LF, Hou BQ, Wang K, et al. Horizontal gene  
660 transfer of *Fhb7* from fungus underlies *Fusarium* head blight resistance in wheat.  
661 *Science*. 2020;368:eaba5435.
- 662 35. Howell T, Hale L, Jankuloski L, Bonafede M, Gilbert M, Dubcovsky J. Mapping  
663 a region within the 1RS.1BL translocation in common wheat affecting grain yield  
664 and canopy water status. *Theor Appl Genet*. 2014;127:2695-2709.
- 665 36. Rasheed A, Wen WE, Gao FM, Zhai SN, Jin H, Liu JD, et al. Development and  
666 validation of KASP assays for genes underpinning key economic traits in bread  
667 wheat. *Theor Appl Genet*. 2016;129:1843-1860.
- 668 37. Chen DH, Ronald PC. A rapid DNA minipreparation method suitable for AFLP  
669 and other PCR applications. *Plant Mol Biol Rep*. 1999;17:53-57.
- 670 38. Roncallo PF, Beaufort V, Larsen AO, Dreisigacker S, Echenique V. Genetic  
671 diversity and linkage disequilibrium using SNP (KASP) and AFLP markers in a  
672 worldwide durum wheat (*Triticum turgidum* L. *Var Durum*) collection. *PLoS*  
673 *ONE*. 2019;14:e0218562 .
- 674 39. Zou JW, Jia WL, Li LX, Chen X, Jia D, Yan CS, et al. KASP marker assays for  
675 functional genes of important trait in 120 wheat cultivars (lines). *Mol Plant*  
676 *Breeding*. 2019;17:3945-3959.
- 677 40. Khalid M, Afzal F, Gul A, Amir R, Subhani A, Ahmed Z, et al. Molecular  
678 characterization of 87 functional genes in wheat diversity panel and their  
679 association with phenotypes under well-watered and water-limited conditions.  
680 *Front Plant Sci*. 2019;10:717.

- 681 41. Zhao JJ, Wang ZW, Liu HX, Zhao J, Li T, Hou J, et al. Global status of 47 major  
682 wheat loci controlling yield, quality, adaptation and stress resistance selected over  
683 the last century. *BMC Plant Biol.* 2019;19:1-14.
- 684 42. Liu K, Muse SV. Powermarker: An integrated analysis environment for genetic  
685 marker analysis. *Bioinformatics.* 2005;21:2128-2129.
- 686 43. Kumar S, Stecher G, Tamura K. Mega7: Molecular evolutionary genetics analysis  
687 version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33:1870-1874.
- 688 44. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components:  
689 a new method for the analysis of genetically structured populations. *BMC Genet.*  
690 2010;11:94.
- 691 45. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in  
692 structured populations. *Am J Hum Genet.* 2000;67:170-181.
- 693 46. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals  
694 using the software STRUCTURE: a simulation study. *Mol Ecol.*  
695 2005;14:2611-2620.
- 696 47. Yeh FC, Yang RC, Boyle TB, Ye ZH, Mao JX, Yeh C, et al. Popgene version 1.32:  
697 The user friendly software for population genetic analysis. Molecular Biology  
698 and Biotechnology Centre, University of Alberta, Canada. 1999.
- 699 48. Yuan HM, Zhang FG, Chen DS, Yuan HJ, Wang XL, Kang L, et al. Foreign  
700 wheat germplasm investigation, introduction and utilization in Ningxia. *Journal*  
701 *of Plant Genetic Resources.* 2012;13:308-312.
- 702 49. Hao CY, Wang YQ, Chao SM, Li T, Liu HX, Wang LF, et al. The iselect 9K SNP  
703 analysis revealed polyploidization induced revolutionary changes and intense  
704 human selection causing strong haplotype blocks in wheat. *Sci Rep.*  
705 2017;7:41247.
- 706 50. Lopes MS, Ibrahim EB, Baenziger PS, Singh S, Roto C, Ozbek K, et al.  
707 Exploiting genetic diversity from landraces in wheat breeding for adaptation to  
708 climate change. *J Exp Bot.* 2015;12:3477-3486.
- 709 51. Zhou Y, Chen ZX, Cheng MP, Chen J, Zhu TT, Wang R, et al. Uncovering the  
710 dispersion history, adaptive evolution and selection of wheat in China. *Plant*

- 711 Biotechnol J. 2018;16:280-291.
- 712 52. Gadea M. Trigos cultivados en españa y nuevas variedades recomendadas.  
713 Madrid, Spain: Ministerio de Agricultura. 1958.
- 714 53. Royo C, Briceño-Félix GA. Spanish wheat pool. In: Bojean AP, Angus WJ, van  
715 Ginkel M, eds. The world wheat book. A history of wheat breeding. 2011;Vol 2.  
716 Lavoisier, 121–154.
- 717 54. Olmstead AL, Rhode PW. The Red Queen and the Hard Reds: productivity  
718 growth in American wheat, 1800–1940. Journal of Economic History 62. 2002.
- 719 55. Rasheed A, Xia XC, Mahmood T, Quraishi UM, Aziz A, Bux H, et al.  
720 Comparison of economically important loci in landraces and improved wheat  
721 cultivars from Pakistan. Crop Sci. 2016;56:287-301.
- 722 56. Meng LZ, Xiang C, Liu HW, Yang L, Mai CY, Yu LQ, et al. The impact of  
723 modern plant breeding on dominant chinese wheat cultivars (*Triticum aestivum* L.)  
724 revealed by SSR and functional markers. Genet Resour Crop Evol.  
725 2018;65:55-65.
- 726 57. Wang BB, Lin ZC, Li X, Zhao YP, Zhao BB, Wu GX, et al. Genome-wide  
727 selection and genetic improvement during modern maize breeding. Nat Genet.  
728 2020;52:565-571.
- 729 58. Yang Y, Chen XM, He ZH, Roder M, Xia LQ. Distribution of *Vp-1* alleles in  
730 Chinese white-grained landraces, historical and current wheat cultivars. Cereal  
731 Res Commun. 2009;37:169-177.
- 732 59. Kamran A, Iqbal M, Spaner D. Flowering time in wheat (*Triticum aestivum* L.): a  
733 key factor for global adaptability. Euphytica. 2014;197:1-26.
- 734 60. Huang S, Morrison WR. Aspects of proteins in Chinese and British common  
735 (hexaploid) wheats related to quality of white and yellow Chinese noodles. J  
736 Cereal Sci. 1988;8:177-187.
- 737 61. He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX, et al. Characterization of  
738 phytoene synthase 1 gene (*PSY1*) located on common wheat chromosome 7A and  
739 development of a functional marker. Theor Appl Genet. 2008;116:213-221.
- 740 62. Wang JW, He XY, He ZH, Wang H, Xia XC. Cloning and phylogenetic analysis

- 741 of phytoene synthase 1 (*Psy1*) genes in common wheat and related species.  
742 *Hereditas*. 2009;146:208-256.
- 743 63. Uga Y, Sugimoto K, Ogawa S, Rane J, Ishitani M, Hara N, et al. Control of root  
744 system architecture by DEEPER ROOTING 1 increases rice yield under drought  
745 conditions. *Nat Genet*. 2013;45:1097-1102.
- 746 64. Zhang WJ, Li T, Qin L, Zhao J, Zhao JJ, Liu HX, et al. *TaDRO*, a gene associated  
747 with wheat root architectures, its global distribution and evolution in breeding.  
748 *Scientia Agricultura Sinica*. 2018;51:1813-1829.
- 749 65. Li YX, Wang TY, Li Y. Formation, research and utilization of founder parents in  
750 major crops. *Journal of Plant Genetic Resources*. 2019;20:1093-1102.
- 751 66. Zhuang QS. Chinese Wheat Improvement and Pedigree Analysis. China  
752 Agricultural Press, Beijing. 2003. (in Chinese).
- 753 67. Guo J, Shi W, Zhang Z, Cheng J, Sun D, Yu J, et al. Association of yield-related  
754 traits in founder genotypes and derivatives of common wheat (*Triticum aestivum*  
755 L.). *BMC Plant Biol*. 2018;18:38.
- 756 68. Hulbert SH, Webb CA, Smith SM, Sun Q. Resistance gene complexes: Evolution  
757 and utilization. *Annu Rev Phytopathol*. 2001;39:285-312.
- 758 69. Huang XQ, Kempf H, Ganai MW, Rder MS. Advanced backcross QTL analysis  
759 in progenies derived from a cross between a German elite winter wheat variety  
760 and a synthetic wheat (*Triticum aestivum* L.). *Theor Appl Genet*.  
761 2004;109:933-943.
- 762 70. Li SS, Jia JZ, Wei XY, Zhang XC, Li LZ, Chen HM, et al. A intervarietal genetic  
763 map and QTL analysis for yield traits in wheat. *Mol Breeding*. 2007;20:167-178.
- 764 71. Ji XL, Xie CL, Ni ZF, Yang TM, Nevo E, Fahima T, et al. Identification and  
765 genetic mapping of a powdery mildew resistance gene in wild emmer (*Triticum*  
766 *Dicoccoides*) accession IW72 from Israel. *Euphytica*. 2008;159:385-390.
- 767 72. Zhou DG, Chen W, Lin ZC, Chen HD, Wang CR, Li H, et al. Pedigree-based  
768 analysis of derivation of genome segments of an elite rice reveals key regions  
769 during its breeding. *Plant Biotechnol J*. 2016;14:638-648.
- 770 73. Tang JW, Yin GH, Gao Y, Wang LN, Han YL, Huang F, et al. Comprehensive

- 771 analysis on agronomic traits and processing quality of core parent Zhou8425B  
772 and its derivatives. *Journal of Triticeae Crops*. 205;35:777-784.
- 773 74. Xiao YG, Yin GH, Li HH, Xia XC, Yan J, Zheng TC, et al. Genetic diversity and  
774 genome-wide association analysis of Stripe rust resistance among the core wheat  
775 parent Zhou8425B and its derivatives. *Scientia Agricultura Sinica*.  
776 2011;44:3919-3929.
- 777 75. Li XJ, Xu X, Yang XM, Li XQ, Liu WH, Gao AN, et al. Genetic diversity among  
778 a founder parent and widely grown wheat cultivars derived from the same origin  
779 based on morphological traits and microsatellite markers. *Crop Pasture Sci*.  
780 2012;63:303-310.
- 781 76. Rasheed A, Xia XC, et al. From markers to genome-based breeding in wheat.  
782 *Theor Appl Genet*. 2019;132:767-784.
- 783 77. Rasheed A, Jin H, Xiao YG, Zhang Y, Hao YF, Zhang Y, et al. Allelic effects and  
784 variations for key bread-making quality genes in bread wheat using  
785 high-throughput molecular markers. *J Cereal Sci*. 2019;85:305-309.
- 786 78. IWGSC. Shifting the limits in wheat research and breeding using a fully  
787 annotated reference genome. *Science*. 2018;361:eaar7191.
- 788 79. Avni R, Nave M, Barad O, Baruch K, Twardziok S, Gundlach H, Hale I, et al.  
789 Wild emmer genome architecture and diversity elucidate wheat evolution and  
790 domestication. *Science*. 2017;357:93-97.
- 791 80. Luo MC, Gu YQ, Puiu D, Wang H, Twardziok SO, Deal KR, et al. Genome  
792 sequence of the progenitor of the wheat D genome *Tegilops tauschii*. *Nature*.  
793 2017;551:498-502.
- 794 81. Zhao GY, Zou C, Li K, Wang K, Li TB, Gao LF, et al. The *Aegilops tauschii*  
795 genome reveals multiple impacts of transposons. *Nat plants*. 2017;3:946-955.
- 796 82. Ling HQ, Ma B, Shi XL, Liu H, Dong LL, Sun H, et al. Genome sequence of the  
797 progenitor of wheat a subgenome *Triticum urartu*. *Nature*. 2018;557:424-428.
- 798

799 **Legends of Figures**

800 Figure 1. The population structure of 207 wheat accessions based on 44 genes. (A) A  
801 neighbor-joining tree of all accessions. Different lines are presented in different colors.  
802 (B) Plots of first three principal components of all accessions. (C) Population  
803 structure of all accessions based on Structure.

804 Figure 2. Genetic divergence on all 44 genes between Ningxia and Others subgroups.  
805 (A) Genetic diversities between Ningxia and Others subgroups. (B) Genetic  
806 differentiation (*Fst*) between Genetic diversities on grain yield genes. (C) Allele  
807 frequency at grain yield genes between Ningxia and Others subgroups. (D) Allele  
808 frequencies of quality genes between Ningxia and Others subgroups.

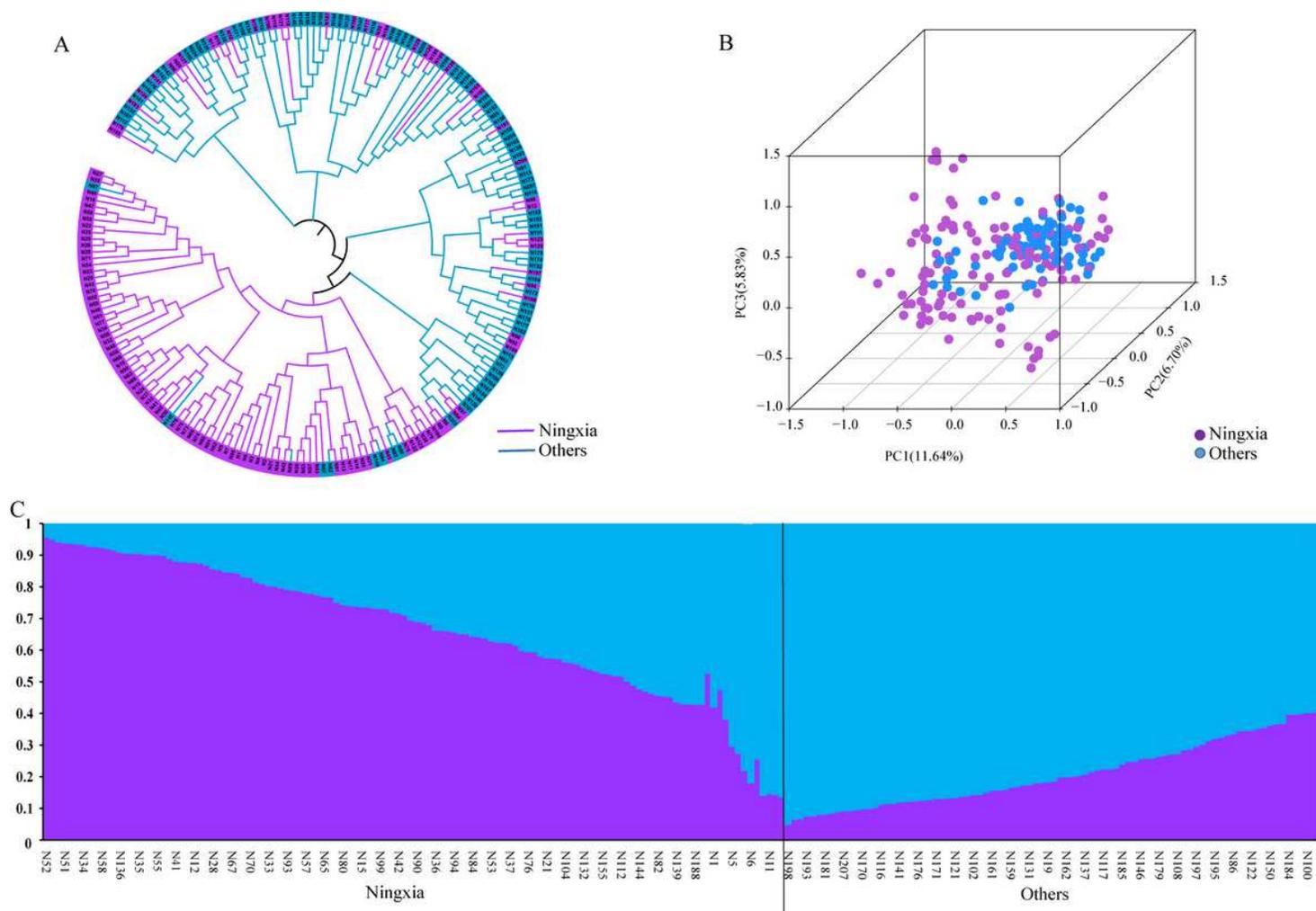
809 Figure 3. A neighbor-joining tree and genetic diversity based on 44 genes between  
810 modern cultivars and landraces in Ningxia province.

811 Figure 4. Genetic differentiation (*Fst*) and allele frequencies of 44 genes between  
812 modern cultivars and landraces in Ningxia province. (A) (B) *Fst* and allele  
813 frequencies at adaptation genes between modern cultivars and landraces. (C) (D) *Fst*  
814 and allele frequencies of stress resistance genes between modern cultivars and  
815 landraces. (E) (F) *Fst* and allele frequencies of quality genes between modern  
816 cultivars and landraces. (G) (H) *Fst* and allele frequency of grain yield-related genes  
817 between modern cultivars and landraces.

818 Figure 5. Distributions of favorable alleles of yield, quality, stress resistance and  
819 adaptation genes in founder parents and modern cultivars. (A) Number of favorable  
820 alleles at four types of genes in six founder parents. (B) Proportion of accessions  
821 carrying different numbers of favorable alleles of yield genes in founder parents and  
822 modern cultivars. (C) Proportion of accessions carrying different numbers of  
823 favorable alleles of quality genes in founder parents and modern cultivars. (D)  
824 Proportion of accessions carrying different numbers of favorable alleles of stress  
825 resistance genes in founder parents and modern cultivars. (E) Proportion of accessions  
826 carrying different numbers of favorable alleles of adaptation genes in founder parents  
827 and modern cultivars.

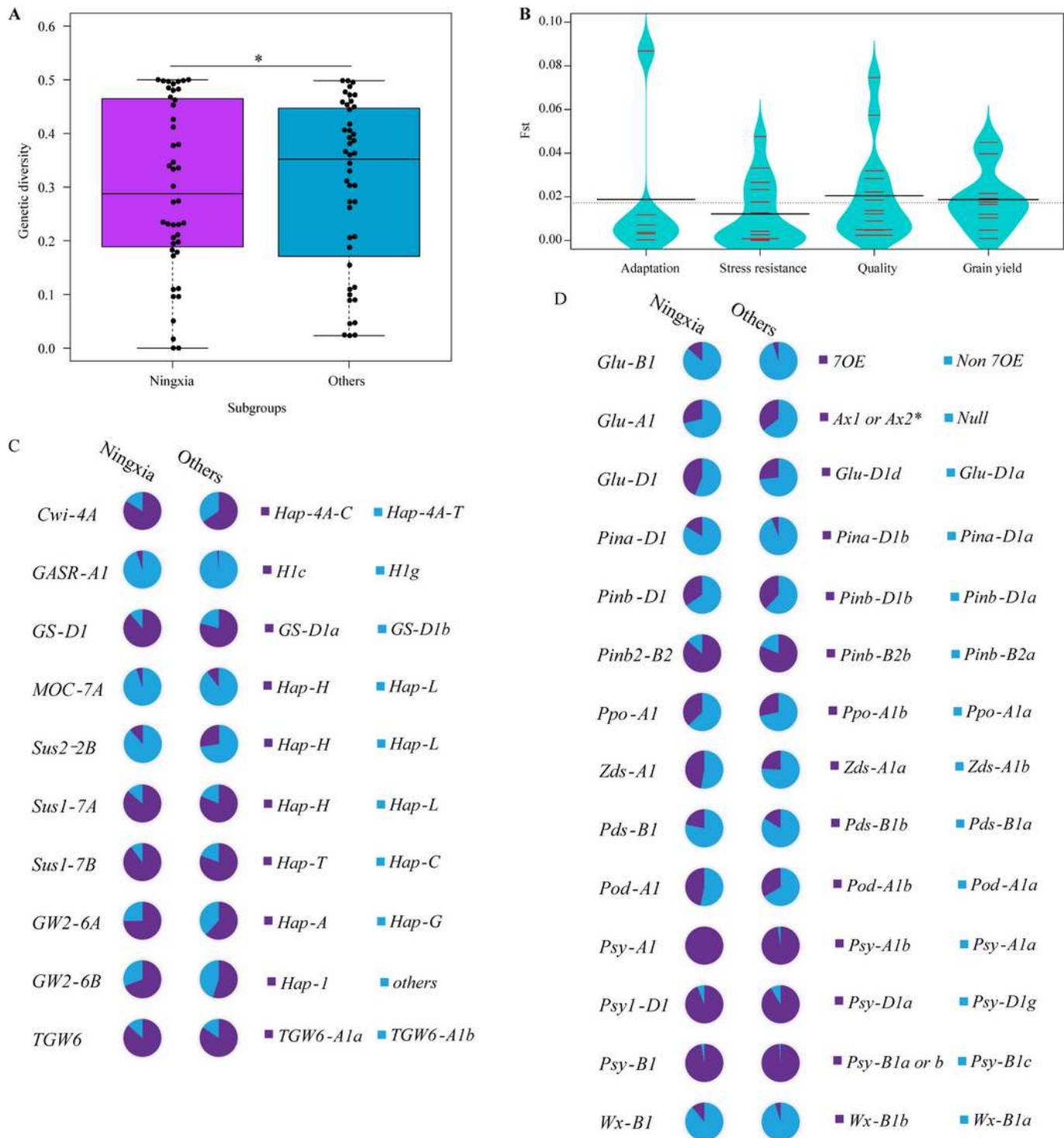
828 Figure 6. Gene flow between six founder parents and modern cultivars at yield,  
829 quality, stress resistance and adaptation genes.

# Figures



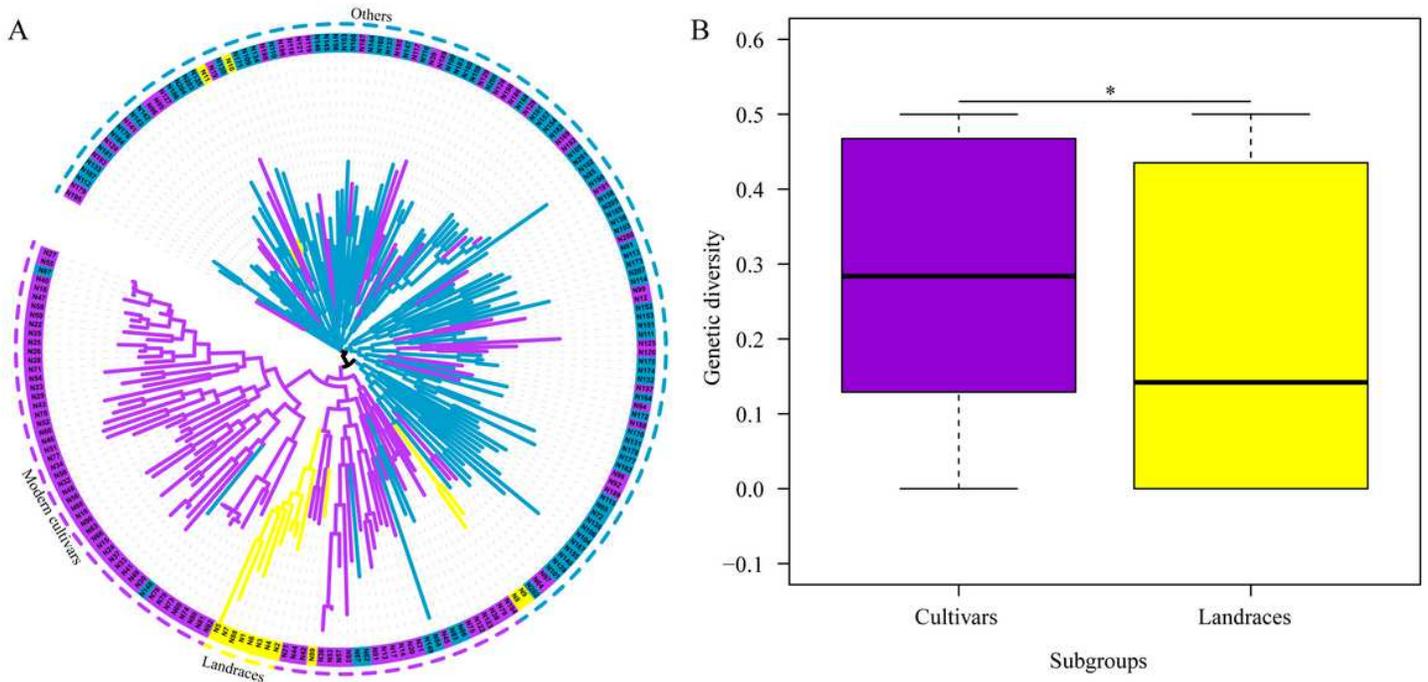
**Figure 1**

The population structure of 207 wheat accessions based on 44 genes. (A) A neighbor-joining tree of all accessions. Different lines are presented in different colors. (B) Plots of first three principal components of all accessions. (C) Population structure of all accessions based on Structure.



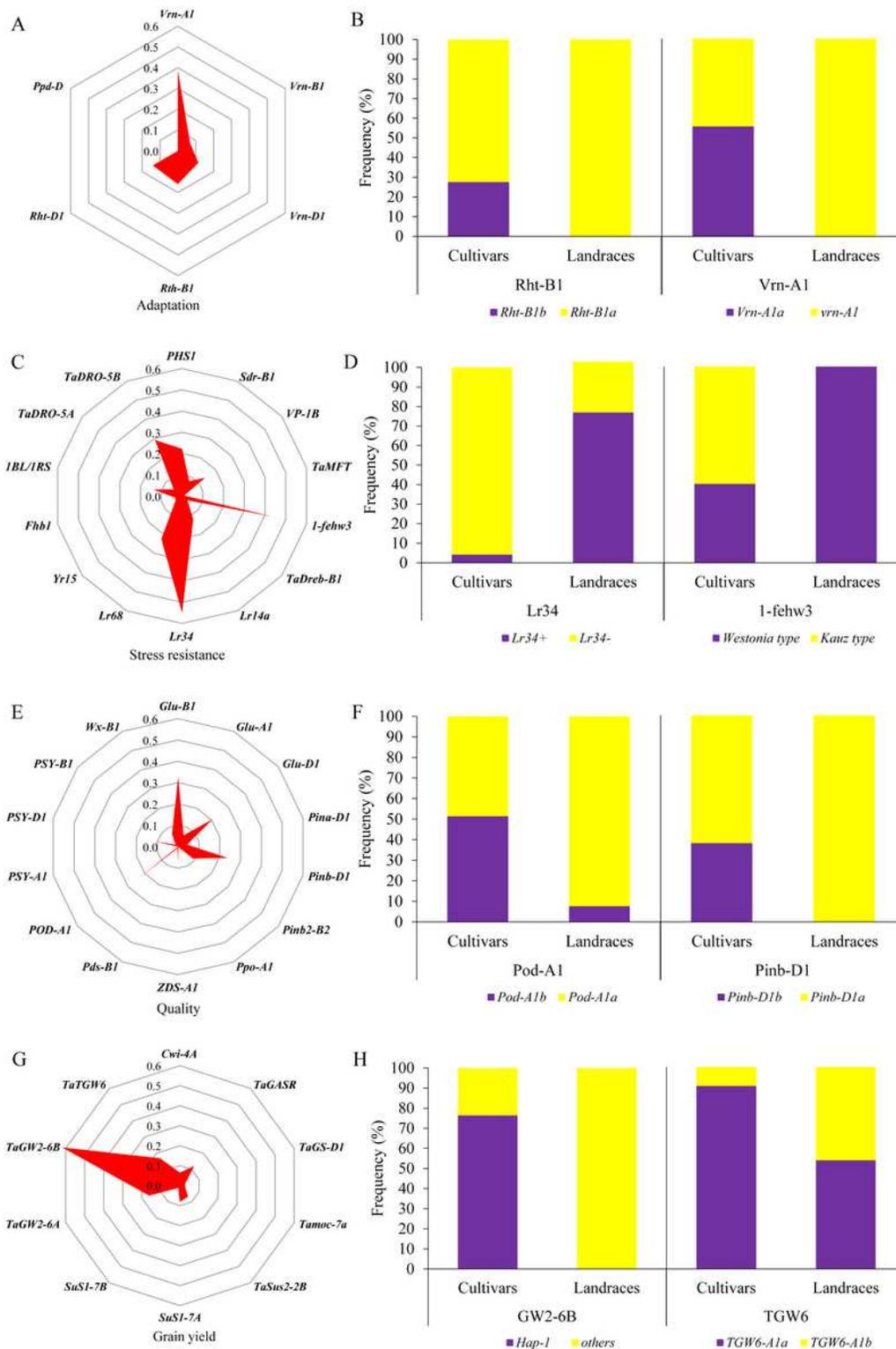
**Figure 2**

Genetic divergence on all 44 genes between Ningxia and Others subgroups. 8 (A) Genetic diversities between Ningxia and Others subgroups. (B) Genetic differentiation ( $F_{st}$ ) between Genetic diversities on grain yield genes. (C) Allele frequency at grain yield genes between Ningxia and Others subgroups. (D) Allele frequencies of quality genes between Ningxia and Others subgroups.



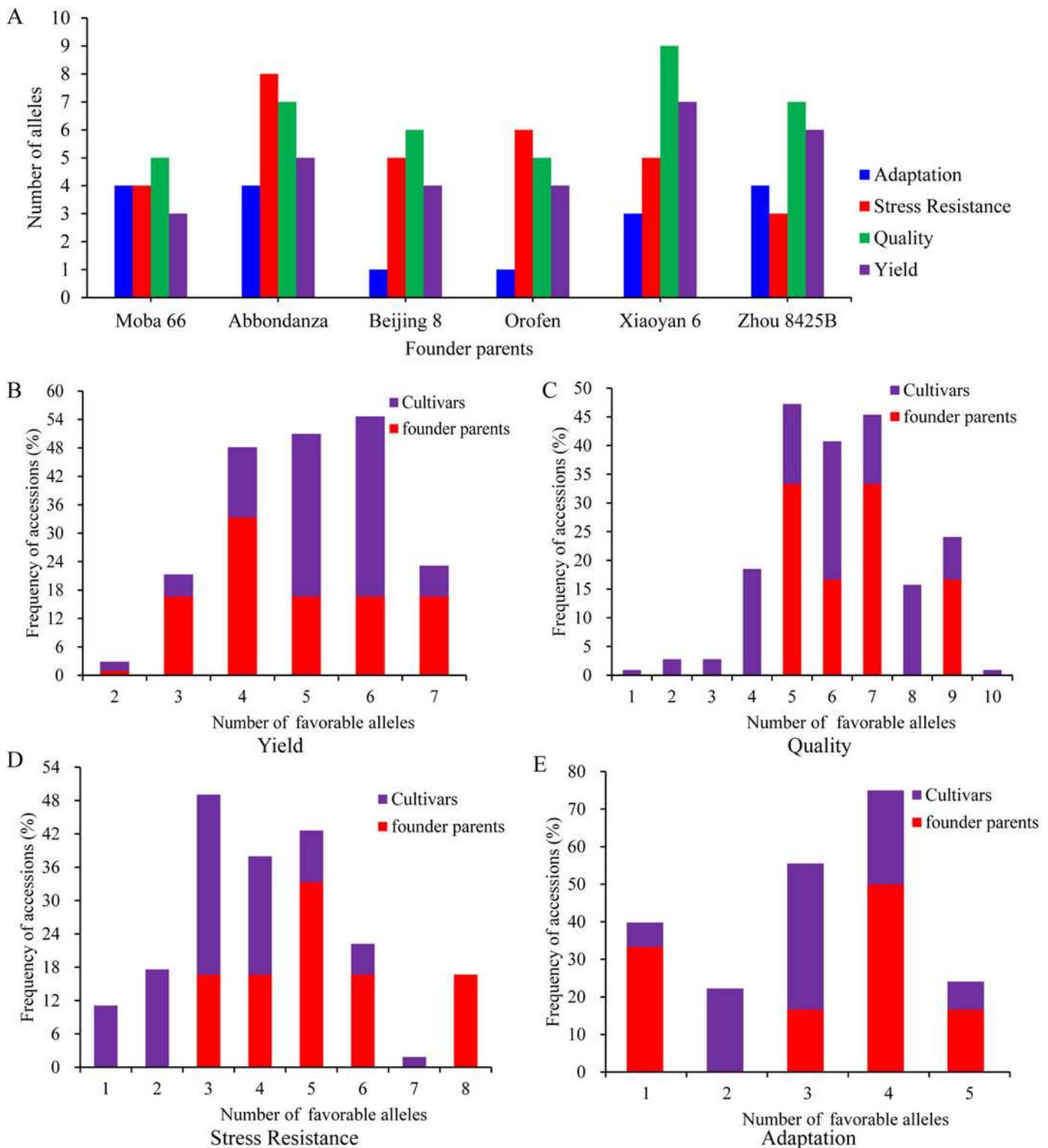
**Figure 3**

A neighbor-joining tree and genetic diversity based on genes between modern cultivars and landraces in Ningxia province.



**Figure 4**

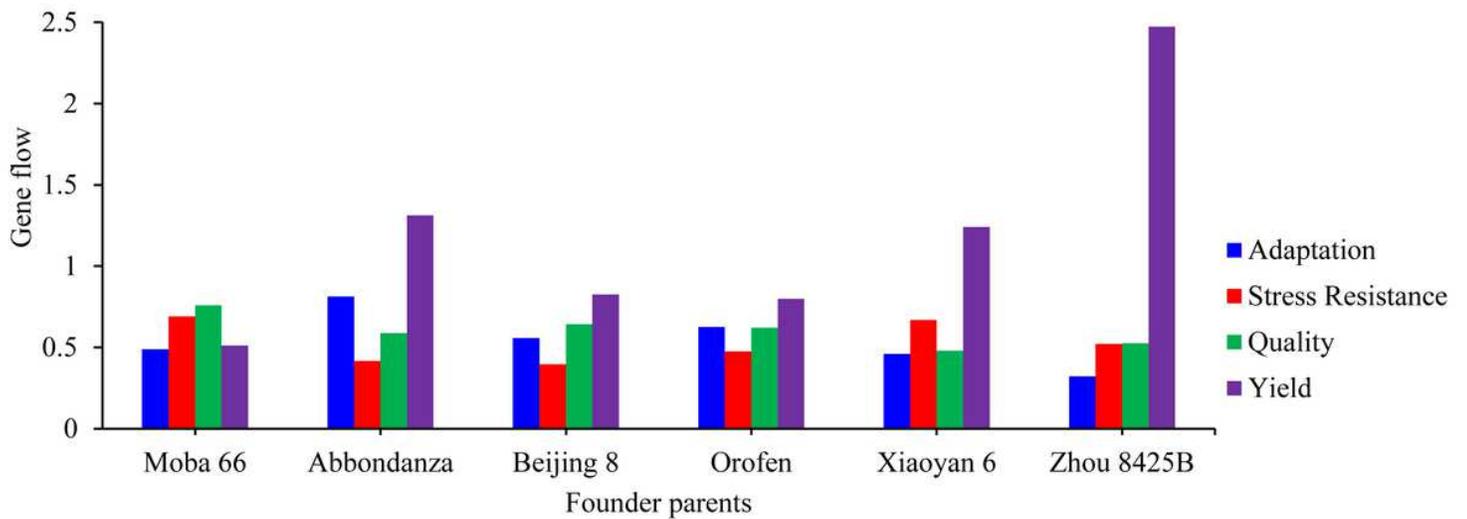
Genetic differentiation ( $F_{st}$ ) and allele frequencies of 44 genes between modern cultivars and landraces in Ningxia province. (A) (B)  $F_{st}$  and allele frequencies at adaptation genes between modern cultivars and landraces. (C) (D)  $F_{st}$  and allele frequencies of stress resistance genes between modern cultivars and landraces. (E) (F)  $F_{st}$  and allele frequencies of quality genes between modern cultivars and landraces. (G) (H)  $F_{st}$  and allele frequency of grain yield-related genes between modern cultivars and landraces.



**Figure 5**

Distributions of favorable alleles of yield, quality, stress resistance and adaptation genes in founder parents and modern cultivars. (A) Number of favorable alleles at four types of genes in six founder parents. (B) Proportion of accessions carrying different numbers of favorable alleles of yield genes in founder parents and modern cultivars. (C) Proportion of accessions carrying different numbers of favorable alleles of quality genes in founder parents and modern cultivars. (D) Proportion of accessions

carrying different numbers of favorable alleles of stress resistance genes in founder parents and modern cultivars. (E) Proportion of accessions carrying different numbers of favorable alleles of adaptation genes in founder parents and modern cultivars.



**Figure 6**

Gene flow between six founder parents and modern cultivars at yield, quality, stress resistance and adaptation genes.

## Supplementary Files

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

- [FigureS1.tif](#)
- [FigureS10.tif](#)
- [FigureS11.tif](#)
- [FigureS12.tif](#)
- [FigureS13.tif](#)
- [FigureS14.tif](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)
- [FigureS4.tif](#)
- [FigureS5.tif](#)
- [FigureS6.tif](#)
- [FigureS7.tif](#)
- [FigureS8.tif](#)
- [FigureS9.tif](#)

- [SupplementalTable.xlsx](#)