Dissection and validation of minor quantitative trait loci (QTLs) conferring grain size and weight in rice

Ping Sun[1], Yuanyuan Zheng[1], Pingbo Li, Hong Ye, Hao Zhou, Guanjun Gao, Yuqing He[2]

National Key Laboratory of Crop Genetic Improvement and National Center of Crop Molecular Breeding, Huazhong Agricultural University, Wuhan 430070, China.

[1]These authors contributed equally to this work.

[2]Corresponding author.

E-mail:

Ping Sun: sunping121605@163.com

Yuanyuan Zheng: yuanyuanzh1005@163.com

Pingbo Li: lipingboxwmr@126.com

Hong Ye: 19881212hong@163.com

Hao Zhou: zhouhao@webmail.hzau.edu.cn

Guanjun Gao: gaojun8199@webmail.hzau.edu.cn

Yuqing He: yqhe@mail.hzau.edu.cn

Abstract

Grain size and weight contribute greatly to the grain yield of rice. In order to identify minor QTLs conferring grain size and weight, an F_2 population derived from a cross between two *indica* rice lines showing small difference on grain size, Guangzhan 63-4S (GZ63-4S) and Dodda, and its derived $F_{2:3}$ population were developed and used for QTL analysis. Totally, 36 QTLs for grain size and weight were detected, and 7 were repeatedly detected, of which the number of beneficial alleles was contributed roughly equally by the two parents. In order to further validate effects of QTLs detected, a BC_1F_2 population derived from a backcross of a mixture of F_2 lines with GZ63-4S was developed and subjected to QTL selection. Heterozygous regions of 3 QTLs, *qGS3*, *qTGW6.2* and *qGT7* were identified, and corresponding near-isogenic lines (NILs) of each QTL were constructed with three rounds of self-crosses. In the background of NILs, *qGS3* was responsible for GL, LWR, GT and TGW, *qTGW6.2* was for GL and TGW, and *qGT7* was for GT and TGW. These results have laid the foundation of further fine mapping and cloning of underlying genes, and could be of great use in breeding and improvement of rice lines with desirable size and yield.

Keywords: grain size, grain weight, minor QTL, validation, NIL, rice

Introduction

Rice is one of the staple crops worldwide, and feeds more than half of the world's population. In the face of continuously increasing population and reduced arable land, how to further improve the grain yield of rice is a major concern of scientists and breeders. Grain size, characterized by four factors viz., grain length (GL), grain width (GW), length-to-width ratio (LWR) and grain thickness (GT), contributes greatly to grain weight, which is a key determinant of grain yield [1]. Therefore, dissection of the genetic basis that underlies grain size and weight would be of great use in developing rice lines with high grain yield.

Considerable efforts have been made to investigate the genetic basis of grain size and weight in the past two decades, and results showed that the four factors of grain size, GL, GW, LWR and GT, and thousand-grain weight (TGW) are quantitative traits, and subjected to control of many genes [2, 3]. Up to now, large numbers of quantitative trait loci (QTLs) have been identified, however, only a small proportion of QTLs displaying large effect have been cloned, such as *GS3* [4, 5], *OsMADS1* [6, 7], *GL3.3*/*TGW3* [8-10], *GW5*/*GSE5* [11, 12], *GS5* [13], *GW8* [14], *GS2*/*GL2* [15-17], *GL7*/*GW7* [18, 19], etc. Although the knowledge of molecular regulation of grain size and weight has greatly increased, the mining and cloning of more QTLs, especially minor QTLs, is still of great importance to have a better understanding of underlying mechanisms and provide breeding programs with valuable gene resources.

Rice lines displaying large difference on grain size and weight were always selected to develop segregating populations for QTL analysis, which resulted in the repeated

detection of several major QTLs/genes. For example, two major genes for grain size, *GW2* and *GL3.1* were identified and cloned from genetic populations derived from FAZ1 and WY3, of which the TGW values differ by 23.12g [20, 21]. The two genes above, together with another two major genes, *GS3* and *GW5*/*GSE5*, contributed to the huge variation of grain size and weight between N411 and N643, of which the TGW values differ by 54.33g [22]. The existence of major genes is likely to interfere the mapping and validation of minor QTLs, exemplified by the fine mapping of *GS5* [13]. Therefore, in order to identify minor QTLs for grain size and weight, rice lines displaying small difference should be preferred.

Quantitative traits are easily affected by environment, which leads to the instability of QTL detection. Therefore, genetic validation of QTLs is of great necessity in further breeding utilization or cloning. The most frequently used method is evaluation the effect of a QTL using near-isogenic lines (NILs), which are lines that carry segregating regions at target QTL but homozygous regions in the rest of genome [23]. NILs for a QTL are always developed by backcrossing lines carrying the QTL region from donor to the receipt several times until the non-target QTL regions were completely from the receipt, which could achieve the simultaneous improvement of target traits of recipient [24, 25]. Another simple method is to select lines carrying segregating target QTL regions from inbred populations that have undertaken several rounds of self-crosses, also known as residual heterozygous lines (RHLs) [26, 27]. This method is sometimes utilized for absence of laborious hybridization work. The NIL of *Ghd8*, a major QTL with pleiotropic effects on grain yield, heading date and

89 plant height, was constructed by screening lines carrying segregation target regions
90 from a RIL population of the F_7 generation [28, 29].
91 In this study, in order to identify minor QTLs for grain size, two *indica* rice lines
92 displaying small difference, Guangzhan 63-4S (GZ63-4S) and Dodda were selected to
93 develop the F_2 and derived $F_{2:3}$ populations, and QTL analysis of grain size and
94 weight were performed. In order to validate QTL detected, lines carrying
95 heterozygous QTL regions were screened from a BC_1F_2 population derived from a
96 backcross of a mixture of F_2 lines with GZ63-4S. NILs of three QTLs were developed
97 by a series of self-crosses of screened BC_1F_2 lines, and further used for evaluation
98 their genetic effect on grain size and TGW.
99

100 **Materials and methods**

101 Population development and cultivation
102 Guangzhan 63-4S (GZ63-4S) is a leading *indica* two-line male sterile line developed
103 by the China North Japonica Hybrid Rice Research Center and Hefei Fengle Seed
104 Company, and has been mated with many restorer lines to produce promising hybrid
105 combinations in recent years [30]. Dodda is an *indica* cultivar with unknown origin,
106 belonging to the core germplasm collections of our lab. The TGW values of GZ63-4S
107 and Dodda differ by less than 10 g (data not shown).
108 As displayed in Fig.1, 1000 F_2 lines were produced from a cross between GZ63-4S
109 and Dodda, and were subjected to selection of the *TMS5* locus conditioning
110 thermo-sensitive genic male sterility with a closely linked marker [31]. 214 lines

carrying homozygous *TMS5* regions were selected to make up the F_2 mapping population, which was further self-crossed to produce the $F_{2:3}$ population. Both the F_2 and $F_{2:3}$ population was exploited to map QTLs for grain size and TGW. In addition, 1200 BC_1F_2 lines were produced by backcrossing a mixture of F_2 lines to GZ63-4S, followed by a self-cross. These lines were subjected to *TMS5* selection, and 250 lines carrying homozygous *TMS5* regions were selected to perform heterozygous QTL regions screening with flanking makers in the mapping process (Table 2). BC_1F_2 lines carrying heterozygous QTL regions were further self-crossed three times to produce the BC_1F_5 populations, which were utilized to validate the effect of QTLs.

The F_2, $F_{2:3}$ and BC_1F_5 populations were planted in year 2014, 2015 and 2018, respectively, during the normal rice growing seasons at the Experimental Farm of Huangzhong Agricultural University in Wuhan, China. Each F_3 line consisted of 12 plants, and each BC_1F_5 population consisted of 100 plants.

Trait evaluation

GL, GW, LWR and TGW were measured with more than 200 grains per line or plant using the yield traits scorer [32]. GT was determined for each grain individually using an electronic digital caliper (Guanglu Measuring Instrument Co. Ltd., China), and thirty grain values were averaged for each line or plant. For the $F_{2:3}$ population, the phenotypic value of each line was the average value of 12 plants.

Genetic map construction

A total of more than 1000 simple sequence repeat markers or insert/deletion markers were employed to screen for polymorphic markers between GZ63-4S and Dodda, and 143 markers were identified. Among that, 111 markers were selected to perform genotyping of the F_2 population with 4% polyacrylamide gels migration and silver staining [33]. A genetic linkage map was constructed using MapMaker/Exp3.0 with the Kosambi mapping function [34].

Data analysis

Correlation analysis was performed using the data analysis module in Microsoft Office Excel 2016. QTL analysis was performed by composite interval mapping using the software package QTLCartographer V2.5 with a logarithm of odds (LOD) threshold of 3.0 [35]. ANNOVA analysis was performed using the IBM SPSS Statistics 22.

Results

Phenotypic variation and correlation of the F_2 and $F_{2:3}$ populations

GZ63-4S is a typical photoperiod- and thermo-sensitive genic male sterile line, and shows male sterility in the normal growing seasons in Wuhan. Therefore, the seeds could not be harvested, which abolished comparison of grain size and TGW between the two parents. All the five traits of the F_2 and $F_{2:3}$ populations showed continuous variation and followed normal distribution in year 2014 and 2015, respectively (Fig. 2).

All the four grain size factors were significantly positively correlated with TWG in both years, except for LWR (Table 1). GL was significantly positively correlated with LWR and GT in both years, while GW was only significantly negatively correlated with LWR in both years. The three highest correlation coefficients were observed between GW and TGW in year 2014, GL in two years, and GW and LWR in year 2014, with values of 0.739, 0.731 and 0.728, respectively.

QTLs detected in the F_2 and $F_{2:3}$ populations

GL

Ten QTLs for GL were detected in two populations and distributed on seven chromosomes, with phenotypic variation explained by each QTL ranging from 2.58% to 25.39% (Table 2, Fig.3). Among those, the beneficial alleles of *qGL3*, *qGL4.1* and *qGL4.2* were from GZ63-4S, while that of others were from Dodda. Two QTLs, *qGL3* and *qGL6* were repeatedly detected, and explained 2.58% and 13.71% of the variation in the F_2 population, and 8.12% and 7.41% of the variation in the $F_{2:3}$ population, respectively. The remaining QTLs were detected only in the $F_{2:3}$ population, excluding *qGL4.1*.

GW

Five QTLs were detected for GW in the $F_{2:3}$ population, while none in the F_2 population (Table 2, Fig.3). Among those, the beneficial alleles of two were from GZ63-4S, while that of the other three were from Dodda.

LWR

177 Five QTLs for LWR were identified in the two populations, and distributed on four
178 chromosomes (Table 2, Fig.3). Among those, two QTLs, *qLWR3* and *qLWR9*, were
179 repeatedly detected, and displayed nearly the same values of additive effect in
180 opposite direction. The remaining were minor QTLs accounting for less than 6% of
181 the variation and were detected only in one population.

182 GT

183 Seven QTLs were identified for GT in the two populations and were distributed on
184 chromosome 2, 3, 4, 7 and 11 (Table 2, Fig.3). The beneficial allele of all eight QTLs
185 were from GZ63-4S, except for that of *qGT11*. The three QTLs, *qGT3*, *qGT7* and
186 *qGT11* were stably detected, and explained 15.16%, 6.60% and 5.83% of the variation
187 in the F_2 population, and 20.98%, 10.25% and 6.24% of the variation in the $F_{2:3}$
188 population, respectively. *qGT2.1* and *qGT2.2* were only detected in the $F_{2:3}$
189 population, while *qGT7* and *qGT11* were only in the F_2 population.

190 TGW

191 Nine QTLs for TGW were detected in the two populations, which were distributed on
192 chromosome 2, 3, 4, 6, 7 and 11 (Table 2, Fig.3). Among those, four QTLs, *qTGW3*,
193 *qTGW4*, *qTGW6.2* and *qTGW11*, were repeatedly detected, which accounted for
194 0.83%, 15.76%, 6.89% and 20.68% of the variation in the F_2 population, and 3.90%,
195 12.03%, 5.11% and 10.98% of the variation in the $F_{2:3}$ population, respectively. The
196 beneficial alleles of *qTGW3* and *qTGW4* were from GZ63-4S, while that of *qTGW6.2*
197 and *qTGW11* were from Dodda. The remaining QTLs were all only detected in the
198 $F_{2:3}$ population.

199 The region flanked by marker LRJ99 and RM232 on chromosome 3 and consisting of
200 four QTLs, was responsible for GL, LWR, GT and TGW in both the F_2 and $F_{2:3}$
201 population, and was term *qGS3*, hereafter.
202
203 Validation of QTL effects using NILs
204 In the BC_1F_2 population derived from backcrossing some F_2 lines with GZ63-4S,
205 heterozygous regions were screened for QTLs repeatedly detected in both the F_2 and
206 $F_{2:3}$ populations or QTLs accounting for more that 10% of variation in one population
207 using flanked markers (Fig.2, Table 1). Lines carrying heterozygous regions of three
208 QTLs, *qGS3*, *qTGW6.2* and *qGT7*, were identified respectively, and were self-crossed
209 three times to produce NIL populations for each QTL.
210 In the NIL population of *qGS3*, significant differences were observed in the average
211 values of GL, GT and TGW among the three genotypes, $qGS3^{Dodda}$, $qGS3^{H}$ and
212 $qGS3^{GZ63\text{-}4S}$ (Table 3). Compared to NIL ($qGS3^{Dodda}$), NIL ($qGS3^{GZ63\text{-}4S}$) showed
213 increased values by 0.21 mm in GL, 0.10 in LWR, 0.07mm in GT and 1.47g in TGW.
214 For *qTGW6.2*, significant differences were observed in the average values of GL and
215 TGW between $qTGW6.2^{Dodda}$ and $qTGW6.2^{GZ63\text{-}4S}$, while no difference between
216 $qTGW6.2^{H}$ and $qTGW6.2^{GZ63\text{-}4S}$ in the NIL population (Table 4). Compared to NIL
217 ($qTGW6.2^{Dodda}$), NIL ($qTGW6.2^{GZ63\text{-}4S}$) showed decreased values by 0.14 mm in GL
218 and 1.24g in TGW.
219 For *qGT7*, significant differences were observed in the average values of GT among
220 the three genotypes, $qGT7^{Dodda}$, $qGT7^{H}$ and $qGT7^{GZ63\text{-}4S}$, and in that of TGW between

221 $qGT7^{Dodda}$ and $qGT7^{GZ63\text{-}4S}$ in the NIL population (Table 5). Compared to NIL
222 ($qGT7^{Dodda}$), NIL ($qGT7^{GZ63\text{-}4S}$) showed increased values by 0.07 mm in GT and 1.27g
223 in TGW.
224

225 **Discussion**

226 Evaluation of grain size and weight

227 Genotyping and phenotyping are two key processes in genetic analysis. With the
228 completion of high-quality genome sequences of several rice cultivars and
229 development of sequencing techniques, genotyping a population is becoming
230 increasingly simple and cheap [36, 37]. Therefore, high-throughput and time-saving
231 methods of phenotyping are in urgent need. In previous studies, grain size was always
232 evaluated using electronic digital-display vernier caliper, and about 30 randomly
233 chosen filled grains was used for each line, which is both pains taking and time
234 consuming [19, 38]. *SmartGrain* is a phenotyping software developed for measuring
235 grain size through image analysis, which improved greatly the throughput, but is still
236 time consuming for the separation of adjacent seeds in the scanning process [39]. In
237 this study, evaluation of GL, GW, LWR, and TGW was performed using the yield
238 traits scorer (YTS) that could fulfil the measurement of a rice line represented by
239 about 500 seeds within one minute [28]. Therefore, the YTS dramatically increases
240 the amount of seeds evaluated and reduces the time of phenotyping, demonstrating its
241 great power in phenotype evaluation.
242

243 Minor QTLs for grain size and weight

244 In this study, a total of 37 QTLs were identified for GL, GW, LWR, GT and TGW in
245 the F_2 and $F_{2:3}$ populations, and 7 QTL regions were repeatedly detected, of which the
246 additive effects were far less than that of cloned major genes for grain size, such as
247 *GS3*, *GL3.1/qGL3*, *GW5*/*GSE5* and *GW2* [4, 5, 11, 12, 20-22], demonstrating minor
248 QTLs for grain size and weight. Moreover, the number of beneficial alleles was
249 contributed roughly equally by the two parents, indicating that novel minor QTLs
250 could be detected from rice lines that differ little in grain size.

251 Among QTLs detected, *qGS3*, the pleiotropic QTL for GL, LWR, GT and TGW on
252 chromosome 3, is co-located with *OsMADS1*, of which a natural allele was reported
253 responsible for GL, GT and TGW in two separate studies [6, 7]. However, the
254 difference of GL between the two NILs in Yu et al. (2018) was almost twice that
255 between our NILs, while the difference of GT was the same [7]. An appropriate
256 explanation is that another gene for GL in the Nipponbare background may interact
257 with *OsMADS1* to amplify the difference in the NILs, as reported by Xia et al. (2018)
258 [9]. Therefore, *qGS3* is likely to be *OsMADS1*. In addition, the region of *qTGW6.2*
259 overlaps with that of two QTLs for TGW in the chromosomal segment substitution
260 lines (CSSLs) population derived from Yamadanishiki or Takanari in the background
261 of Koshihikari, respectively [40, 41]. *qGT7*, the QTL for GT on chromosome 7, is
262 co-located with a region for GL, GW, LWR and GT reported by Liu et al (2015),
263 which contains *GL7*/*GW7*, a major gene influencing GL and GW simultaneously [18,
264 19, 42]. As *qGT7* has no effect on GL and GW in the NIL background, it is a novel

gene different from *GL7*/*GW7*. The region of *qLWR9* was also detected for LWR only by Yin et al. (2015) [43]. The remaining QTLs are seldom reported, or maybe novel.

Validation of minor QTLs using NILs

QTLs detected in primary populations are sometimes unstable, and thus should be further validated, especially for minor QTLs. The best way to validate QTLs is the use of NILs. In this study, lines carrying heterozygous QTL regions were screened from the BC_1F_2 population, in case of the loss of target regions in subsequent self-crosses. Then, selected lines were subjected to three rounds of self-crosses, in order to reduce the heterozygosity of non-target regions. The method we preferred ensures that NILs for QTLs of interest are constructed, and eliminates laborious hybridization work.

In this study, the NILs of three QTLs, *qGS3*, *qTGW6.2* and *qGT7*, were constructed, and effects on grain size and TGW were evaluated. The beneficial alleles of *qGS3* from GZ63-4S could increase the value by 0.21mm in GL, 0.07mm in GT, and 1.47g in TGW in homozygous NILs, which was consistent with the values of additive effect in the F_2 and $F_{2:3}$ population on the whole (Table 2, Table 3), suggesting that *qGS3* is a stable and pleiotropic QTL for GL, GT and TGW. *qTGW6.2* was initially detected as a QTL for TGW, but was validated to have effect on both GL and TGW in the NIL population and act in a dominant manner (Table 4). The failure in detection of *qTGW6.2* on GL in F_2 and $F_{2:3}$ population may be attributed to the complexity of genome background and the low variation explained, which further supported the necessity of validation of QTLs using NILs. *qGT7* was repeatedly confirmed as a

287 QTL for GT, and had no effect on GL and GW in the F_2, $F_{2:3}$, and NIL populations
288 (Table2, Table 5). Being one of the four factors of grain size, GT has received less
289 attention, and several cloned genes conditioning GT are responsible for GL and/or
290 GW at the same time, such as *GS2*, *GW8* [17, 44]. Therefore, *qGT7* is a good
291 candidate for further research of the molecular mechanism underlying GT.
292
293 Improvement of grain size, quality and yield in rice
294 Grain size contributes to not only grain yield, but also grain quality, especially
295 appearance quality and milling quality [45 46]. Abundant variation of grain size is
296 observed in rice germplasm cultivated worldwide, thus providing valuable resources
297 for breeding or improvement of rice grain with desirable size and yield. QTLs or
298 genes for grain size mined from germplasm resources have been or are being
299 exploited in rice breeding programs. The *gs3* allele and *GW5* allele from 93-11,
300 together with beneficial alleles controlling fine eating and cooking quality from
301 Nipponbare, were introduced into the Teqing background, and the resulting lines
302 displayed dramatic improvement in grain size and quality [47]. The *gs3* allele and the
303 *GW7* allele from TFA were pyramided into the background of HJX74, resulting in
304 simultaneously improvement of grain yield and quality [18]. In this study, seven QTL
305 regions were repeated detected in both the F_2 and $F_{2:3}$ populations, and three of them
306 were further validated in the NIL background, demonstrating the reality and stability
307 these QTLs. *qGL6* and *qTGW6.2*, the two QTLs for GL located on chromosome 6,
308 could be used in improvement of GL and KGW of GZ63-4S with the beneficial

alleles from Dodda, and further in improvement of the yield and quality performance of hybrid combinations using GZ63-4S as the maternal parent. In addition, the QTLs detected in this study, together with QTLs or genes in other studies, could be combined in the breeding and improvement of rice grains with desirable size, quality and yield.

Acknowledgement

This work was supported by grants from the National Program on R&D of Transgenic Plants (2016ZX08009003-004, 2016ZX08001002-002), and the earmarked fund for the China Agriculture Research System (CARS-01-03) of China.

Competing interests

The authors declare that they have no competing financial interests.

Author contributions

Ping Sun and Yuanyuan Zheng performed most of experiments and analyzed the date. Pingbo Li wrote the paper. Hong Ye and Hao Zhou constructed the F_2 mapping population. Guanjun Gao participated in the field management. Yuqing He designed and supervised the study.

Reference

331 **1.** Xing YZ, Zhang QF (2010) Genetic and molecular bases of rice yield. Annu Rev
332 Plant Biol 61:421-442
333 **2.** Huang RY, Jiang LR, Zheng JS et al (2013) Genetic bases of rice grain shape: so
334 many genes, so little known. Trends Plant Sci 18:218-226
335 **3.** Zuo JR, Li JY (2014) Molecular genetic dissection of quantitative trait loci
336 regulating rice grain size. Annu Rev Genet 48:99-118
337 **4.** Fan CH, Xing YZ, Mao HL et al (2006) *GS3*, a major QTL for grain length and
338 weight and minor QTL for grain width and thickness in rice, encodes a putative
339 transmembrane protein. Theor Appl Genet 112:1164-1171
340 **5.** Takano-Kai N, Jiang H, Kubo Y et al (2009) Evolutionary history of *GS3*, a gene
341 conferring grain length in rice. Genetics 182:1323-1334
342 **6.** Liu Q, Han RX, Wu K et al. (2018) G-protein βγ subunits determine grain size
343 through interaction with MADS-domain transcription factors in rice. Nat Commun
344 9: 852
345 **7.** Yu JP, Miao JL, Zhang ZY et al (2018) Alternative splicing of *OsLG3b* controls
346 grain length and yield in *japonica* rice. Plant Biotechnol J 16:1667-1678
347 **8.** Hu ZJ, Lu SJ, Wang MJ et al (2018) A novel QTL *qTGW3* encodes the
348 GSK3/SHAGGY-like kinase OsGSK5/OsSK41 that interacts with OsARF4 to
349 negatively regulate grain size and weight in rice. Mol Plant 11:736-749
350 **9.** Xia D, Zhou H, Liu RJ et al (2018) *GL3.3*, a novel QTL encoding a
351 GSK3/SHAGGY-like kinase, epistatically interacts with *GS3* to produce
352 extra-long grains in rice. Mol Plant 11:754-756

353 **10.** Ying JZ, Ma M, Bai C et al (2018) *TGW3*, a major QTL that negatively modulates
354 grain length and weight in rice. Mol Plant 11:750-753
355 **11.** Duan PG, Xu JS, Zeng DL et al (2017) Natural variation in the promoter of *GSE5*
356 contributes to grain size diversity in rice. Mol Plant 10:685-694
357 **12.** Liu JF, Chen J, Zheng XM et al (2017) *GW5* acts in the brassinosteroid signalling
358 pathway to regulate grain width and weight in rice. Nat Plants 3:17043
359 **13.** Li YB, Fan CC, Xing YZ et al (2011) Natural variation in *GS5* plays an important
360 role in regulating grain size and yield in rice. Nat Genet 43:1266-1269
361 **14.** Wang SK, Wu K, Yuan QB et al (2012) Control of grain size, shape and quality
362 by *OsSPL16* in rice. Nat Genet 44:950-954
363 **15.** Che RH, Tong HN, Shi BH et al (2015) Control of grain size and rice yield by
364 GL2-mediated brassinosteroid responses. Nat Plants 2:15195
365 **16.** Duan PG, Ni S, Wang JM et al (2015) Regulation of *OsGRF4* by OsmiR396
366 controls grain size and yield in rice. Nat Plants 2:15203
367 **17.** Hu J, Wang YX, Fang YX et al (2015) A rare allele of *GS2* enhances grain size
368 and grain yield in rice. Mol Plant 8:1455-1465
369 **18.** Wang SK, Li S, Liu Q et al (2015) The *OsSPL16-GW7* regulatory module
370 determines grain shape and simultaneously improves rice yield and grain quality.
371 Nat Genet 47:949-954
372 **19.** Wang YX, Xiong GS, Hu J et al (2015) Copy number variation at the *GL7* locus
373 contributes to grain size diversity in rice. Nat Genet 47:944-948
374 **20.** Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width

375 and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat
376 Genet 39:623-630
377 **21.** Qi P, Lin YS, Song XJ et al (2012) The novel quantitative trait locus *GL3.1*
378 controls rice grain size and yield by regulating Cyclin-T1;3. Cell Res
379 22:1666-1680
380 **22.** Zhang XJ, Wang JF, Huang J et al (2012) Rare allele of *OsPPKL1* associated with
381 grain length causes extra-large grain and a significant yield increase in rice. P Natl
382 Acad Sci USA 109:21534-21539
383 **23.** Yano M, Sasaki T (1997) Genetic and molecular dissection of quantitative traits in
384 rice. Plant Mol Biol 35:145-153
385 **24.** Xiao C, Hu J, Ao YT et al (2016) Development and evaluation of near-isogenic
386 lines for brown planthopper resistance in rice cv. 9311. Sci Rep 6: 38159
387 **25.** Zhou RH, Zhu ZD, Kong XY et al (2005) Development of wheat near-isogenic
388 lines for powdery mildew resistance. Theor Appl Genet 110:640-648
389 **26.** Darvasi A, Soller M (1995) Advanced intercross lines, an experimental population
390 for fine genetic mapping. Genetics 141:1199-1207
391 **27.** Tuinstra MR, Ejeta G, Goldsbrough PB (1997) Heterogeneous inbred family
392 (HIF) analysis: a method for developing near-isogenic lines that differ at
393 quantitative trait loci. Theor Appl Genet 95:1005-1011
394 **28.** Yan WH, Wang P, Chen HX et al (2011) A major QTL, *Ghd8*, plays pleiotropic
395 roles in regulating grain oroductivity, plant height, and heading date in rice. Mol
396 Plant 4:319-330

397 **29.** Zhang YS, Luo LJ, Xu CG, Zhang QF, Xing YZ (2006) Quantitative trait loci for
398 panicle size, heading date and plant height co-segregating in trait-performance
399 derived near-isogenic lines of rice (*Oryza sativa*). Theor Appl Genet 113:361-368
400 **30.** Shen QW, Jiang H (2014) Performance of male sterile line Guangzhan 63-4S on
401 breeding and its enlightenment of Hubei middle rice. Hubei Agricultural Sciences
402 53: 2490-2492 (Chinese with English abstract)
403 **31.** Zhou H, Zhou M, Yang YZ et al (2014) RNase Z^{S1} processes Ub_{L40} mRNAs and
404 controls thermosensitive genic male sterility in rice. Nat Commun 5: 4884
405 **32.** Yang WN, Guo ZL, Huang CL et al (2014) Combining high-throughput
406 phenotyping and genome-wide association studies to reveal natural genetic
407 variation in rice. Nat Commun 5: 5087
408 **33.** Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers
409 and characterization of simple sequence length polymorphism (SSLP) in rice
410 (*Oryza sativa* L.). Mol Gen Genet 252:597-607
411 **34.** Lincoln SE, Daly MJ and Lander ES (1992) Constructing genetic maps with
412 MAPMAKER/EXP3.0. Whitehead Institute Technical Report, Whitehead
413 Institute, Cambridge
414 **35.** Wang SC, Christopher JB and Zeng ZB (2012) Windows QTL Cartographer 2.5.
415 Department of Statistics, North Carolina State University, Raleigh, NC
416 **36.** Huang XH, Lu TT, Han B (2013) Resequencing rice genomes: an emerging new
417 era of rice genomics. Trends Genet 29:225-232
418 **37.** Xie WB, Feng Q, Yu HH et al (2010) Parent-independent genotyping for

419 constructing an ultrahigh-density linkage map based on population sequencing. P
420 Natl Acad Sci USA 107:10578-10583

421 **38.** Sun L, Ma DP, Yu HH et al (2013) Identification of quantitative trait loci for grain
422 size and the contributions of major grain-size QTLs to grain weight in rice. Mol
423 Breeding 31:451-461

424 **39.** Tanabata T, Shibaya T, Hori K, Ebana K, Yano M (2012) SmartGrain:
425 High-throughput phenotyping software for measuring seed shape through image
426 analysis. Plant Physiol 160:1871-1880

427 **40.** Okadao S, Onogi A, Iijima K et al (2018) Identification of QTLs for rice grain
428 size using a novel set of chromosomal segment substitution lines derived from
429 Yamadanishiki in the genetic background of Koshihikari. Breeding Sci
430 68:210-218

431 **41.** Takai T, Ikka T, Kondo K et al (2014) Genetic mechanisms underlying yield
432 potential in the rice high-yielding cultivar Takanari, based on reciprocal
433 chromosome segment substitution lines. BMC Plant Biol 14:295

434 **42.** Liu DL, Kang MH, Wang F et al (2015) Mapping of the genetic determinant for
435 grain size in rice using a recombinant inbred line (RIL) population generated from
436 two elite *indica* parents. Euphytica 206:159-173

437 **43.** Yin CB, Li HH, Li SS, Xu LD, Zhao ZG, Wang JK (2015) Genetic dissection on
438 rice grain shape by the two-dimensional image analysis in one *japonica* x *indica*
439 population consisting of recombinant inbred lines. Theor Appl Genet
440 128:1969-1986

441 **44.** Gao FY, Luo ZL, Ren JS et al (2015) Fine mapping and candidate gene analysis
442 of *qGT8*, a najor QTL for grain thickness in rice. Sci Agric Sin 48:4859-4871
443 **45.** Fitzgerald MA, McCouch SR, Hall RD (2009) Not just a grain of rice: the quest
444 for quality. Trends Plant Sci 14:133-139
445 **46.** Zheng TQ, Xu JL, Li ZK, Zhai HQ, Wan JM (2007) Genomic regions associated
446 with milling quality and grain shape identified in a set of random introgression
447 lines of rice (*Oryza sativa* L.). Plant Breeding 126:158-163
448 **47.** Zeng DL, Tian ZX, Rao YC et al (2017) Rational design of high-yield and
449 superior-quality rice. Nat Plants 3: 17031

450

451 Table 1 Correlations of GL, GW, LWR, GT and TGW of the F_2 and $F_{2:3}$ populations
452 in year 2014 and 2015

	GL14	GL15	GW14	GW15	LWR14	LWR15	GT14	GT15	TGW14	TGW15
GL14										
GL15	0.731**									
GW14	0.157	0.055								
GW15	0.046	0.139	0.489**							
LWR14	0.525**	0.453**	-0.728**	-0.424**						
LWR15	0.477**	0.585**	-0.348**	-0.692**	0.657**					
GT14	0.387**	0.381**	0.126	0.269**	0.144	0.053				
GT15	0.225*	0.452**	0.109	0.235*	0.057	0.118	0.677**			
TGW14	0.656**	0.558**	0.378**	0.415**	0.129	0.046	0.739**	0.495**		
TGW15	0.484**	0.712**	0.261**	0.450**	0.093	0.132	0.584**	0.707**	0.700**	

453 GL14, GL in 2014; GL15, GL in 2015; GW14, GW in 2014; GW15, GW in 2015; LWR14, LWR
454 in 2014; LWR15, LWR in 2015; GT14, GT in 2014; GT15, GT in 2015; TGW14, TGW in 2014;
455 TGW15, TGW in 2015.
456 *, ** Significantly at $P<0.05$ and $P<0.01$, respectively.
457

458 Table 2 QTLs detected for GL, GW, LWR, GT and TGW in the F_2 and $F_{2:3}$

459 populations in year 2014 and 2015, respectively

Trait	QTL	Chr[a)]	Interval	2014 F_2				2015 $F_{2:3}$			
				LOD	Add[b)]	Dom[c)]	R^2(%)[d)]	LOD	Add	Dom	R^2(%)
GL	*qGL1.1*	1	RM283-RM259					3.68	0.09	-0.02	6.14
	qGL1.2	1	RM582-RM562					3.04	0.07	-0.04	5.80
	qGL3	3	LRJ99-RM232	4.52	-0.09	-0.07	2.58	8.26	-0.12	-0.03	8.12
	qGL4.1	4	RM252-RM470	3.51	-0.10	-0.01	5.36				
	qGL4.2	4	RM470-RM349					16.71	-0.20	-0.02	25.39
	qGL5.1	5	RM516-RM3381					3.34	0.12	0.03	7.15
	qGL5.2	5	RM249-RM163					4.28	0.15	0.09	6.54
	qGL6	6	RM402-RM5963	5.29	0.10	-0.08	13.71	4.04	0.07	-0.06	7.41
	qGL9	9	RM566-YH16.8					5.95	0.13	-0.03	12.69
	qGL11	11	RM286-RM26085					4.86	0.09	-0.02	7.86
GW	*qGW4*	4	RM349-RM567					8.21	-0.05	-0.01	13.97
	qGW8	8	RM126-RM515					3.35	-0.03	0.00	6.96
	qGW11	11	RM27181-RM224					4.19	0.03	0.01	4.58
	qGW12.1	12	RM511-RM313					4.15	0.03	0.01	5.11
	qGW12.2	12	RM309-RM3726					4.67	0.03	0.01	5.09
LWR	*qLWR2.1*	2	RM7252-RM233					3.95	0.06	0.02	4.07
	qLWR2.2	2	RM279-RM555	3.23	0.08	0.02	5.96				
	qLWR3	3	LRJ99-RM232	7.03	-0.08	-0.06	3.96	9.57	-0.09	-0.02	12.09
	qLWR9	9	RM566-YH16.8	5.13	0.08	0.03	8.02	4.89	0.07	0.03	10.00
	qLWR11	11	RM27181-RM224					5.23	-0.06	-0.05	1.46
GT	*qGT2.1*	2	RM7252-RM233					6.84	-0.02	0.02	11.73
	qGT2.2	2	RM327-RM263					6.00	-0.03	-0.01	11.16
	qGT3	3	LRJ99-RM232	16.94	-0.04	-0.02	15.16	16.07	-0.04	-0.01	20.98
	qGT4.1	4	RM6659-RM16616	4.31	-0.02	0.01	7.99				
	qGT4.2	4	RM16653-RM16820	4.52	-0.02	0.01	10.05				
	qGT7	7	RM560-RM234	4.95	-0.02	-0.00	6.60	6.58	-0.02	0.00	10.25
	qGT11	11	RM286-RM26085	3.00	0.01	-0.01	5.83	3.26	0.01	-0.01	6.24
TGW	*qTGW2.1*	2	RM7252-RM233					3.01	0.00	0.63	3.01
	qTGW2.2	2	RM327-RM263					5.08	-0.68	-0.13	8.28
	qTGW3	3	LRJ99-RM232	3.54	-0.49	-0.54	0.83	7.60	-0.65	-0.43	3.90
	qTGW4	4	RM470-RM349	5.08	-0.83	0.46	15.76	7.63	-0.83	-0.12	12.03
	qTGW6.1	6	RM402-RM5963					4.47	0.60	-0.01	7.25
	qTGW6.2	6	RM3183-RM20048	3.09	0.55	-0.25	6.89	6.26	0.67	0.29	5.11
	qTGW7.1	7	RM21242-RM542					3.16	-0.40	0.29	5.76
	qTGW7.2	7	RM455-RM234					5.75	-0.62	0.50	12.83

qTGW11	11	RM286-RM26085	7.70	0.85	-0.58	20.68	6.22	0.68	-0.13	10.98

a) Chr, chromosome.

b) Add, additive effect of QTL: positive values indicate that alleles from Dodda increase the trait values, and negative values indicate that alleles from GZ63-4S increase the trait values.

c) Dom, dominant effect of QTL.

d) R^2, variation explained by QTL.

Table 3 Genetic effect of *qGS3* in the NIL population

	Number of lines	GL (mm) Mean±SE	GW (mm) Mean±SE	LWR Mean±SE	GT (mm) Mean±SE	TGW (g) Mean±SE
NIL($qGS3^{Dodda}$)	21	9.06±0.14 a	2.77±0.05	3.31±0.08 a	1.97±0.02 a	24.10±0.96 a
NIL($qGS3^{H}$)	40	9.15±0.14 b	2.77±0.06	3.33±0.07 a	2.00±0.02 b	24.88±1.10 b
NIL($qGS3^{GZ63-4S}$)	19	9.27±0.11 c	2.75±0.01	3.41±0.05 b	2.04±0.03 c	25.57±1.08 c

Lower-case letters indicate statistically significant (P<0.05) differences between the mean values within each column (Duncan test).

NIL($qGS3^{Dodda}$) and NIL($qGS3^{GZ63-4S}$) are the lines carrying homozygous *qGS3* regions from Dodda and GZ63-4S in the NIL population, respectively, while NIL($qGS3^{H}$) is the line carrying heterozygous *qGS3* regions.

Table 4 Genetic effect of *qTGW6.2* in the NIL population

	Number of lines	GL (mm) Mean±SE	GW (mm) Mean±SE	LWR Mean±SE	GT (mm) Mean±SE	TGW (g) Mean±SE
NIL($qTGW6.2^{Dodda}$)	19	9.64±0.11 b	2.64±0.05 b	3.70±0.06	2.15±0.03	26.86±1.25 b
NIL($qTGW6.2^{H)}$	34	9.55±0.11 a	2.61±0.06 a	3.71±0.08	2.14±0.03	25.96±1.10 a
NIL($qTGW6.2^{GZ63-4S}$)	18	9.50±0.11 a	2.61±0.04 ab	3.69±0.07	2.14±0.03	25.62±1.08 a

Lower-case letters indicate statistically significant (P<0.05) differences between the mean values within each column (Duncan test).

NIL($qTGW6.2^{Dodda}$) and NIL($qTGW6.2^{GZ63-4S}$) are the lines carrying homozygous *qTGW6.2* regions from Dodda and GZ63-4S in the NIL population, respectively, while NIL($qTGW6.2^{H}$) is the line carrying heterozygous *qTGW6.2* regions.

Table 5 Genetic effect of *qGT7* in the NIL population

	Number of lines	GL (mm) Mean±SE	GW (mm) Mean±SE	LWR Mean±SE	GT (mm) Mean±SE	TGW (g) Mean±SE
NIL($qGT7^{Dodda}$)	14	9.66±0.12	2.57±0.06	3.81±0.10 b	1.96±0.03 a	21.84±1.20 a
NIL($qGT7^{H}$)	21	9.62±0.14	2.61±0.08	3.74±0.08 a	1.99±0.03 b	22.81±1.19 b
NIL($qGT7^{GZ63\text{-}4S}$)	13	9.60±0.13	2.60±0.06	3.76±0.07 ab	2.03±0.03 c	23.17±1.05 b

Lower-case letters indicate statistically significant (P<0.05) differences between the mean values within each column (Duncan test).

NIL($qGT7^{Dodda}$) and NIL($qGT7^{GZ63\text{-}4S}$) are the lines carrying homozygous *qGT7* regions from Dodda and GZ63-4S in the NIL population, respectively, while NIL($qGT7^{H}$) is the line carrying heterozygous *qGT7* regions.

Figure legends

Fig. 1 Schematic representation of the experimental design.

Fig. 2 Frequency distribution of the F_2 and $F_{2:3}$ populations for GL, GW, LWR, GT and TGW in year 2014 and 2015.

Arrow indicates the value of Dodda.

Fig. 3 Distribution of putative QTLs for GL, GW, LWR, GT and TGW identified in the F_2 and $F_{2:3}$ populations on the linkage map.

GL14, QTLs for GL detected in the F_2 population in year 2014; GL15, QTLs for GL detected in the $F_{2:3}$ population in year 2014. The QTLs for GW, LWR, GT and TGW are represented as the same manner as that for GL.

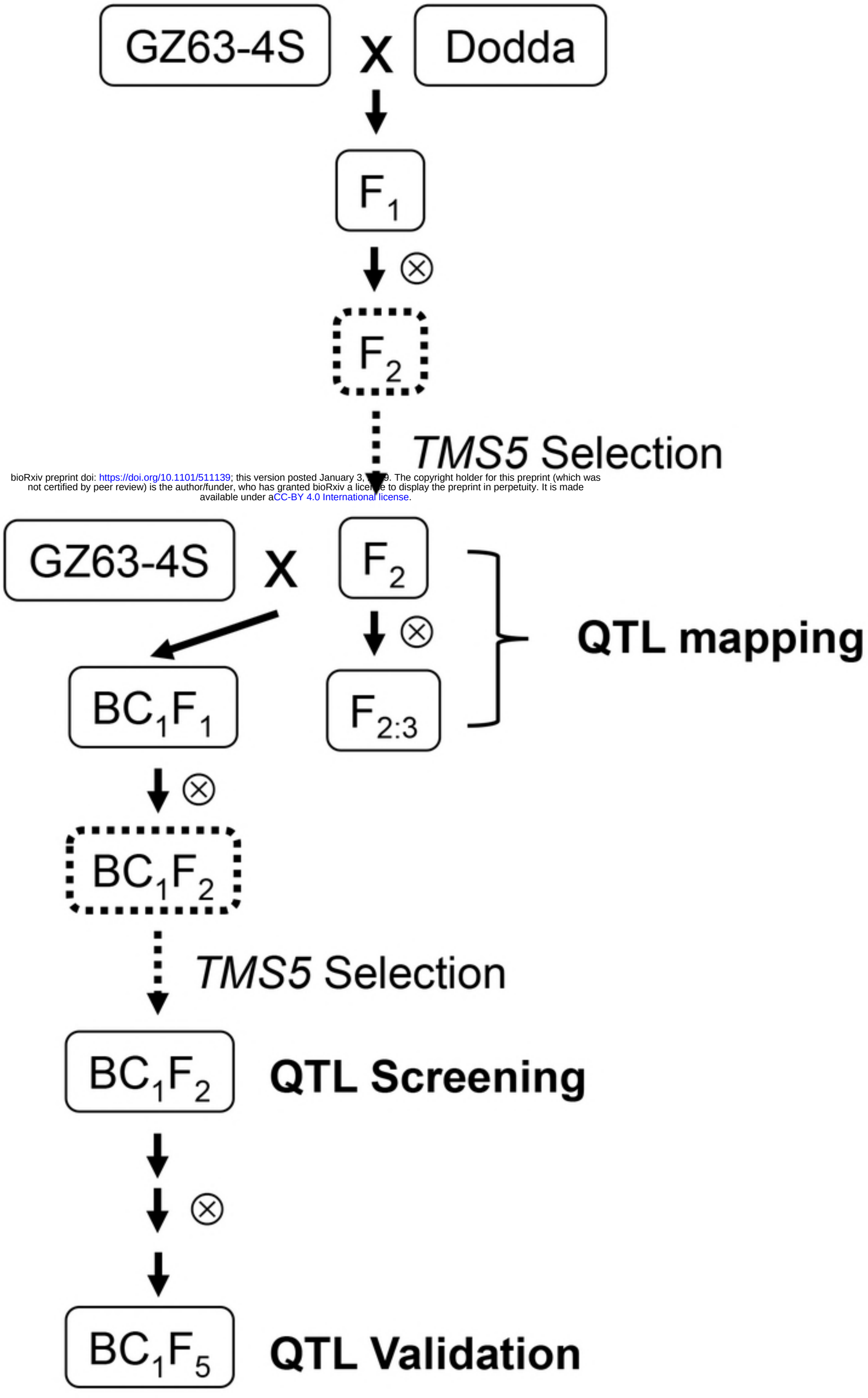
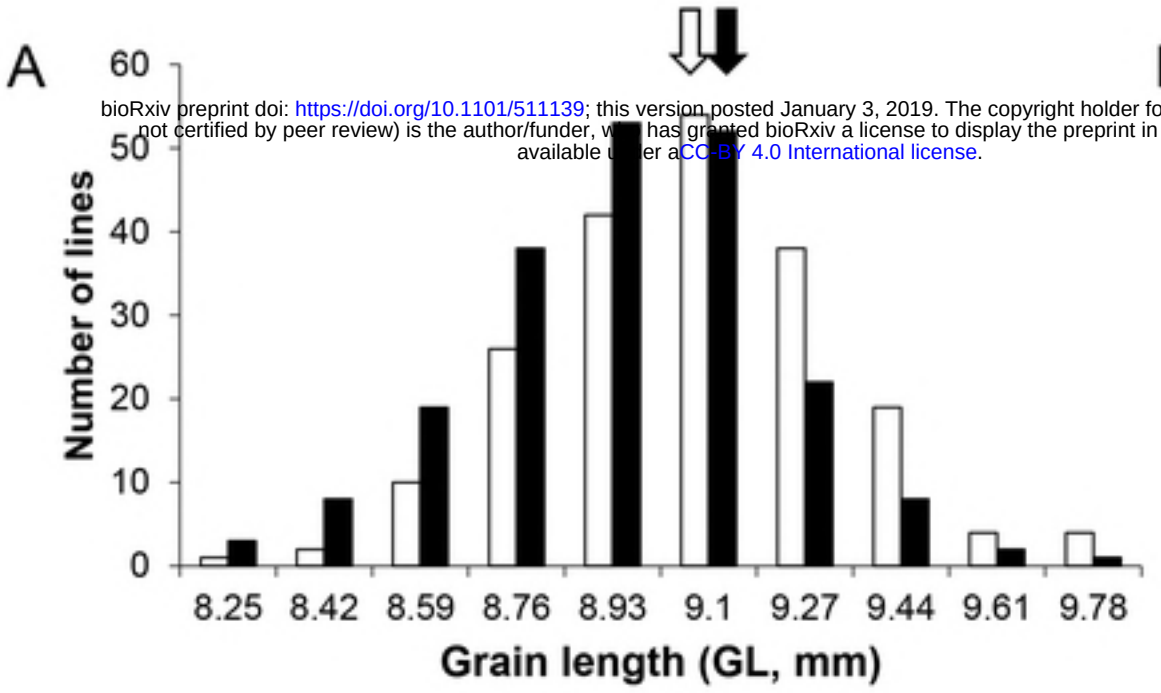
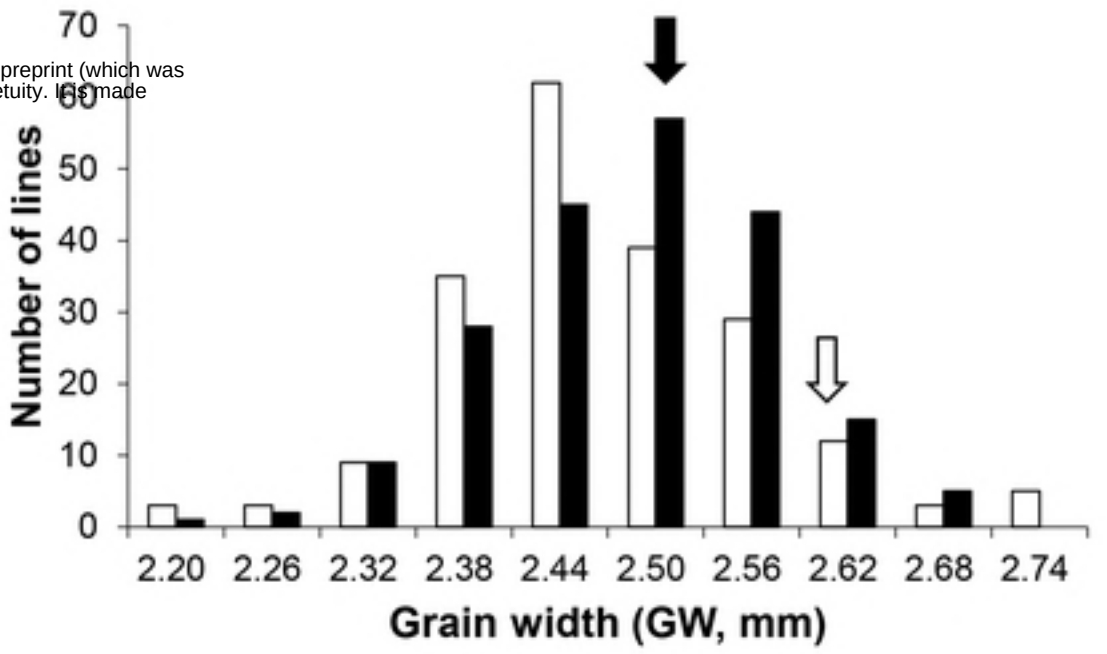
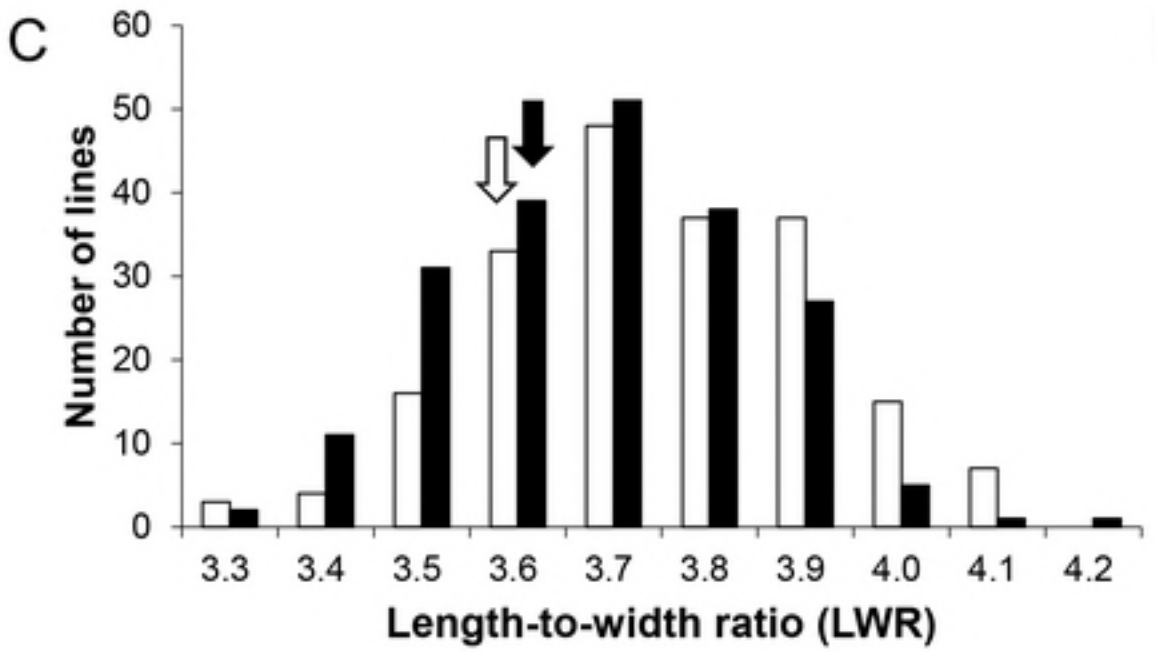
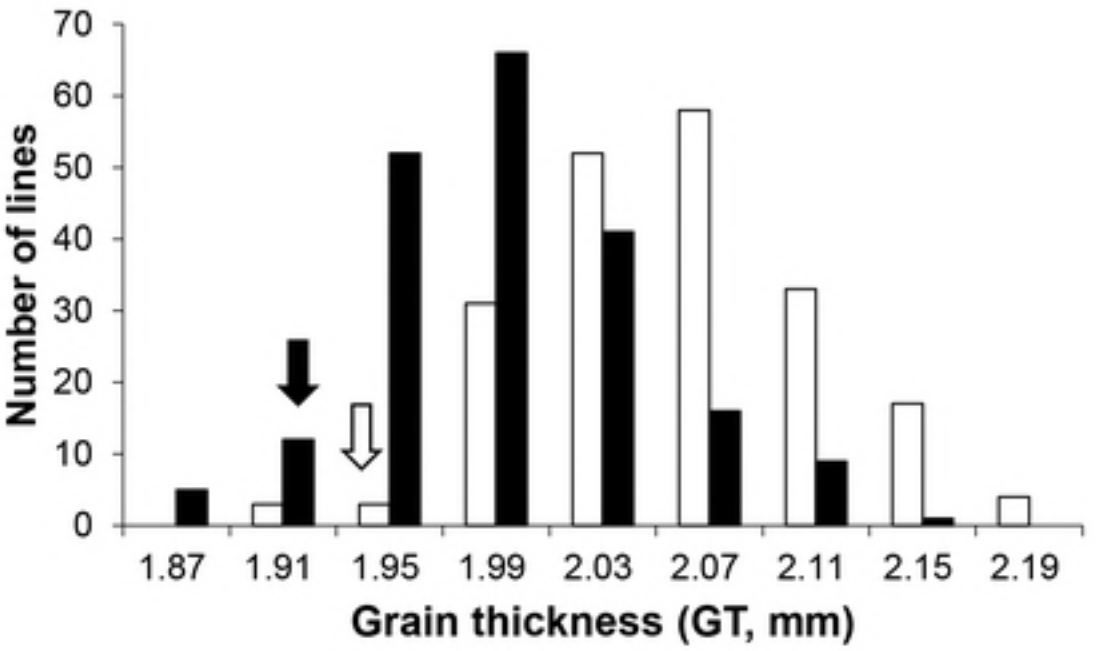


Figure 1









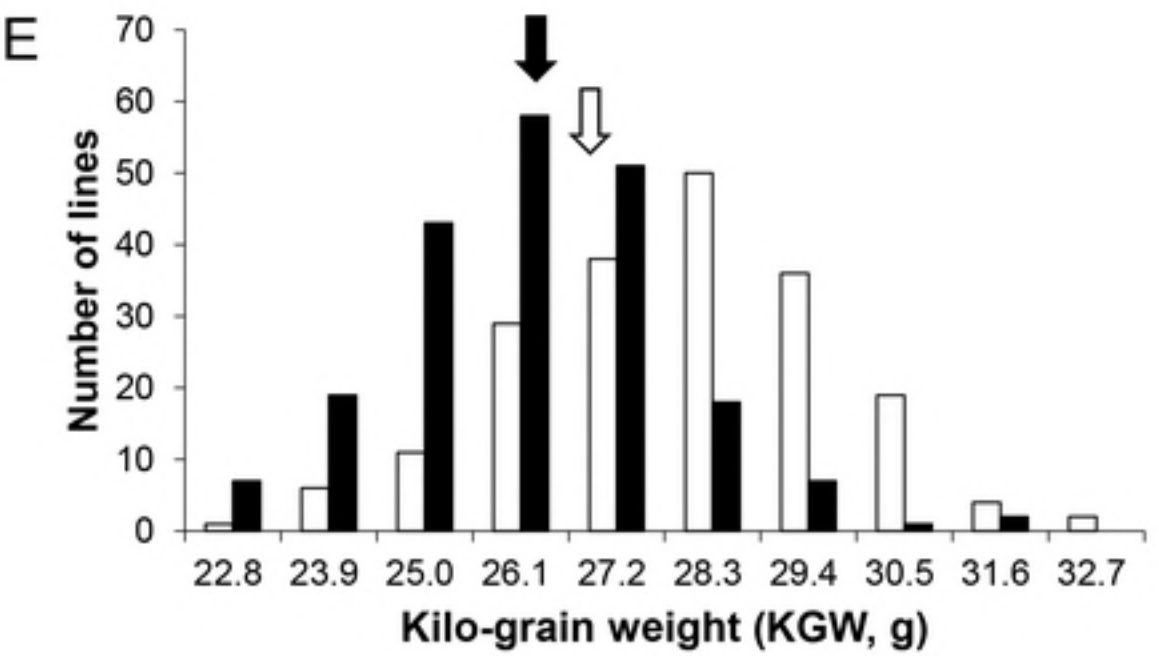


Figure 2

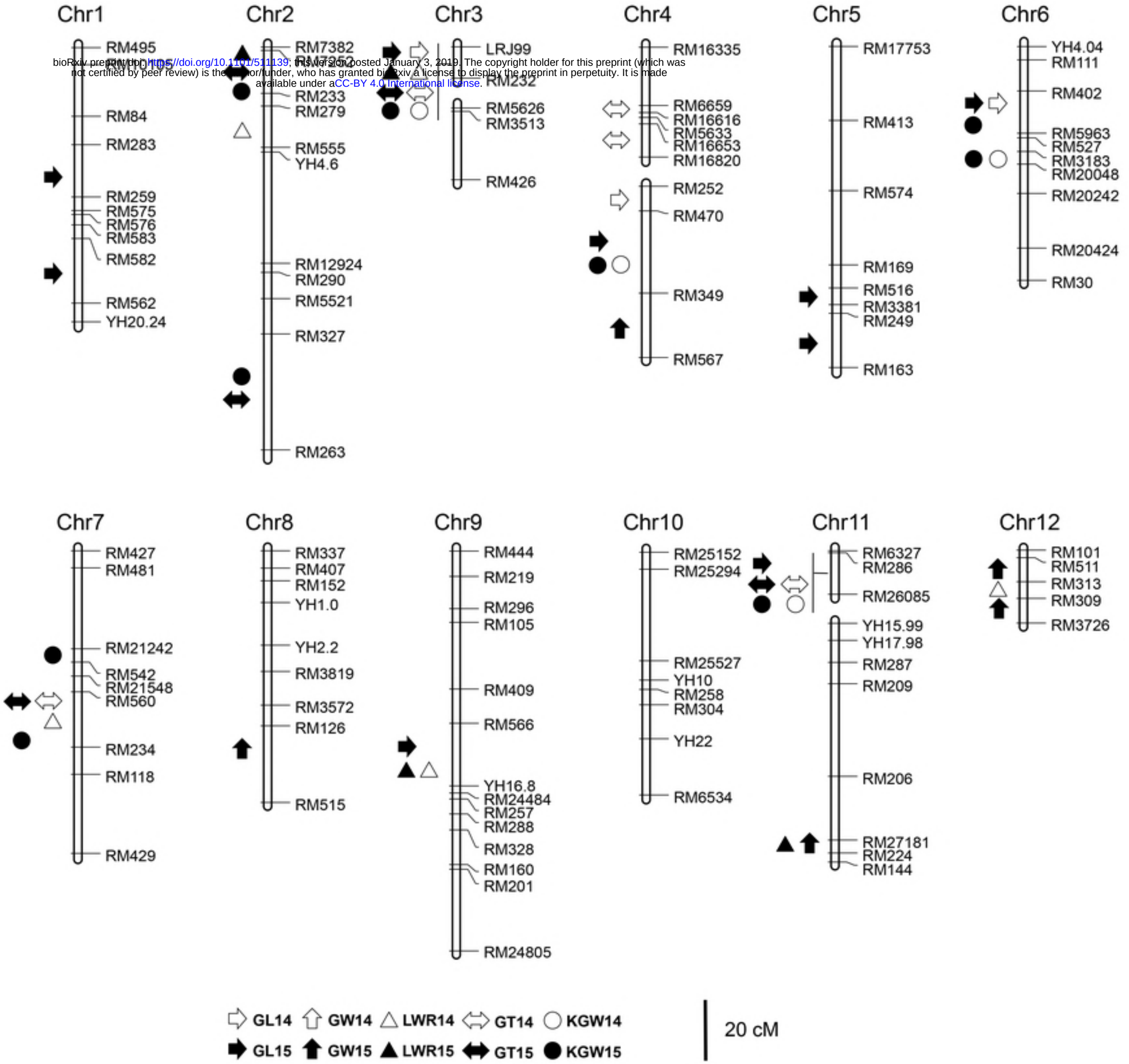


Figure 3