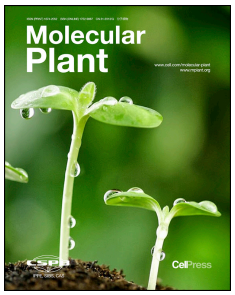
Accepted Manuscript



TGW3, a Major QTL that Negatively Modulates Grain Length and Weight in Rice

Jie-Zheng Ying, Ming Ma, Chen Bai, Xuehui Huang, Jian-Li Liu, Ye-Yang Fan, Xian-Jun Song

PII: S1674-2052(18)30093-5
DOI: 10.1016/j.molp.2018.03.007
Reference: MOLP 604

To appear in: *MOLECULAR PLANT*
Accepted Date: 7 March 2018

Please cite this article as: **Ying J.-Z., Ma M., Bai C., Huang X., Liu J.-L., Fan Y.-Y., and Song X.-J.** (2018). *TGW3*, a Major QTL that Negatively Modulates Grain Length and Weight in Rice. Mol. Plant. doi: 10.1016/j.molp.2018.03.007.

TGW3, a major QTL that negatively modulates grain length and weight in rice

Jie-Zheng Ying[1,4], Ming Ma[2,4,5], Chen Bai[2,4,5], Xuehui Huang[3], Jian-Li Liu[2], Ye-Yang Fan[1], Xian-Jun Song[2,6]

1. National Key Laboratory of Rice Biology and Chinese Center of Rice Improvement, China National Rice Research Institute, Hangzhou 310006, China.

2. Key Laboratory of Plant Molecular Physiology, Institute of Botany, the Chinese Academy of Sciences, No.20 Nanxincun, Xiangshan, Beijing 100093, China.

3. College of Life and Environment Sciences, Shanghai Normal University, Shanghai 200234, China

4. These authors contributed equally to this work.

5. Graduate University of Chinese Academy of Sciences

6. Correspondence should be addressed to X.-J.S. (songxj@ibcas.ac.cn).

Dear Editor,

Grain length (size) and weight is an essential component trait of crop yield. To date, several QTLs for the trait have been identified. *GS3* encodes a putative transmembrane protein and functions as a negative regulator, and its larger-grain allele contains a nonsense mutation causing a 178-aa truncation (Fan et al., 2006). *GL3.1*/*qGL3* acts also as a negative regulator encoding a putative protein phosphatase (Qi et al., 2012; Zhang et al., 2012). Another negative regulator is TGW6 that hydrolyzes indole-3-acetic acid (IAA)-glucose into IAA and glucose (Ishimaru et al., 2013). In contrast, *GW6a* is a positive regulator that encodes a novel histone H4 acetyltransferase (Song et al., 2015). Copy number variations at the *GL7*/*GW7* locus cause elevated expression of *GL7* and thus an increase in grain length (Wang et al., 2015a; Wang et al., 2015b). *GL2*/*GS2* encodes a plant-specific transcription factor OsGRF4, and its larger-grain allele harbors a mutation perturbing the cleavage by miR396c that has elevated *GL2*/*GS2* expression (Hu et al., 2015; Che et al., 2016). *GLW7* encodes the plant-specific transcription factor OsSPL13, and its higher expression is associated with larger grains (Si et al., 2016). These findings greatly enhanced our knowledge of grain length and weight regulation; however, there still exists gaps to integrate these factors into genetic network(s). Here we report a thorough dissection of QTL compositions of a super large grain, and characterization of a novel QTL *qTGW3* that regulates grain length and weight in rice.

In a previous study, to understand how a super-large rice grain was formed, we performed QTL analyses, and identified a single grain-length QTL on chromosome 6 (Ying et al., 2012). We assumed that there existed undetected major QTL(s) shaping the giant grain. We then produced an F7 RIL population derived from JZ1560 and the small-grain cultivar Huanghuazhan (HHZ) (Figure 1A). As expected, phenotypic variations for almost investigated traits approximate a normal distribution, and we observed transgressive segregation for grain length (GL) and grain yield per plant (GY) (Supplemental Figure 1). We thus considered the population to be suitable for

QTL mapping, and inferred that novel QTL(s) or a combination of different loci contributed to the transgressive GL and GY.

We utilized a SLAF-seq technology for genotyping the RIL population, and ultimately achieved a total of 18,194 efficient SLAFs to construct a genetic map (Supplemental Figure 2). Using the map and phenotypes of GL, grain width (GW), thousand grain weight (TGW), and GY in either year 2015 or 2016, we mapped more than 40 QTLs including the known loci *GW2* and *qSW5*/*GW5* (Ying et al., 2012). It was worth noting that most of the mapped GL, GW, and TGW QTLs (91%, i.e. 29 in 32) could be repeatedly detected; however, for GY QTLs, only approximately 42% of them could be detected again (Supplemental Figure 3 and Supplemental Table 1).

To confirm these results, we resorted to PCR genotyping and QTL analysis; subsequently, we observed that most of the mapped QTLs, especially of those with relative larger effects, could be reproduced (Supplemental Figure 4). Nevertheless, we have obtained a deeper understanding of the genetic architecture of the super-large grain JZ1560 and achieved novel major QTLs, such as *qGL1-1*, *qGL3*/*qTGW3-2* (we here named *qTGW3*) (see below) for grain size, weight and yield.

We next scrutinized the *qTGW3* locus (Supplemental Figure 5A and 5B). Also, we screened residual heterozygous lines (RHLs) that are heterozygous for *qTGW3*; GL segregation in the RHL populations revealed characteristic single-gene regulation (Supplemental Figure 5C). We then bred a nearly isogenic line, NIL(*TGW3*), containing a introgression segment from JZ1560 between markers JD3014 and JD3015. Compared with the HHZ isogenic control, we observed a considerable increase in GL (7.6%) but a relative less increase in GT (4.2%) and GW (1.1%) in NIL(*TGW3*) (Supplemental Figure 6A－6C); We also observed a significant increase (8.5%) in TGW (Supplemental Figure 6D). We thus concluded that the enhanced TGW in NIL(*TGW3*) was primarily due to an increase in GL.

We further compared the grain yields of NIL(*TGW3*) with the control plants; We

subsequently observed over 10% higher GY in NIL(*TGW3*) (Supplemental Figure 6E). We also examined whether *TGW3* has pleiotropic effects. The plants of NIL(*TGW3*) did not differ in appearance (Supplementary Figure 6F), whereas they were slightly taller ($P < 0.05$) than, and have an indistinguishable panicle number from, the control plants; both total grain number per plant and empty grain number per plant in NIL(*TGW3*) were reduced, whereas filled grain number per plant did not differ, and more interestingly, seed setting rate in NIL(*TGW3*) was significantly boosted ($P < 0.01$) (Supplementary Figure 7). These results suggested that *TGW3* affects other traits, beyond its profound impact on grain size.

The detailed analysis of RHLs for *qTGW3* enabled us to map the target gene to an interval between JD3014 and JD3015, where a total of 54 recombinants were obtained (Figure 1B and 1C). Finally, we localized *qTGW3* to a region of 18.7 kb (Figure 1D), in which we identified three predicted ORFs as viable candidates for *qTGW3*. Of these ORFs, *LOC_Os03g62500* encodes a GSK3/SHAGGY-like kinase. Prior study revealed that the rice GSK3/SHAGGY-like kinase GSK2 plays an important role in grain size regulation (Tong et al., 2012). Especially, sequence analysis showed that, compared with the HHZ (small grain) allele of *LOC_Os03g62500*, the correspondingly entire exons 3 and 4 of its JZ1560 allele have been missing where an ATP Binding Site (ABS) domain localized (Figure 1E), suggesting that the large-grain allele could be, at least partially, functionally incapacitated. To test the hypothesis, we produced transgenic rice plants causing altered gene expression; enhanced expression of the *LOC_Os03g62500* HHZ allele driven by the 35S promoter reduced grain length (Figure 1F, Supplemental Figure 8A and 8B), while down-regulated endogenous expression of the gene enlarged grain (Figure 1G, Supplemental Figure 8C and 8D). In addition, gene edition of the gene by a CRISPR/Cas9 strategy caused enlarged grain as well (Figure 1H, Supplemental Figure 8E). Collectively, we concluded that *LOC_Os03g62500* represented QTL *TGW3* for grain length.

We sought to reveal the molecular basis for the *TGW3* regulation of grain size, and examined its temporal and spatial expression profile using real-time PCR, and the results suggested that the gene was preferentially expressed in young panicles, which was consistent with its biological function (Supplemental Figure 9A). Meanwhile, we analyzed the subcellular localization of a green fluorescent protein (GFP)-tagged TGW3 in transiently transformed tobacco (*N. benthamiana*) leaf epidermal cells, which showed that the fused protein localized to the nucleus and the cytoplasm (Supplemental Figure 9B).

We found two nucleotide substitutions in cDNA sequence of the JZ1560 allele ($TGW3^{JZ}$) that did not cause any amino acid changes and a 333-bp fragment deletion (equivalent to 111 amino acid residues) that corresponds to the entire exons 3 and 4 in that of the *TGW3* HHZ allele ($TGW3^{HHZ}$) (Figure 1E, Supplemental Figure 10 and Supplemental Figure 11A). Subsequent SWISS-MODEL database analysis indicated that the deletion resulted in entire loss of ABS domain and severe damage to dimer interface domain (Supplemental Figure 11B), suggested that the $TGW3^{HHZ}$ protein could form dimerization structure, yet $TGW3^{JZ}$ could not (Supplemental Figure 11C). We tested these data using yeast two-hybrid (Y2H) experiments, and the results indicated that $TGW3^{HHZ}$ could interact with itself in yeast cells, yet $TGW3^{JZ}$ could not (Supplemental Figure 11D). We also tested our prediction that altered nucleotides in the 3rd intron of $TGW3^{JZ}$ allele was the reason underlying its distinct RNA splicing pattern; indeed, our *Agrobacterium*-mediated transient assays in *N. benthamiana* revealed that a single fragment containing the first four exons and three introns of $TGW3^{JZ}$ allele caused its 3rd intron retention (Supplemental Figure 12), and then the retention, we inferred that, would result in the entire loss of exon 3 and 4, although the detailed molecular mechanism remains to be determined in future studies.

We next sought to uncover cytological basis underlying *TGW3*'s regulation of grain size, and compared the center part of the spikelet hull (lemma) at mature stage in NIL(*TGW3*) and the corresponding control by scanning electron microscope analysis

(Figure 1J). The cell of spikelet hull of NIL(*TGW3*) was larger in size (by 10.1%) and significantly longer (10.6%, in the grain-length orientation) than that of the control, although cell width did not differ; however, the estimated total cell number was found to be a significant decrease in NIL(*TGW3*) (Figure 1K and 1L). We also inspected the cytological features in transgenic rice grains of *TGW3*. Consistent with above results, the cell sizes and cell lengths of spikelet hulls in down-regulated and/or CRIPR-Cas9-edited *TGW3* were significantly enlarged, while that of overexpressed *TGW3* was clearly reduced (Supplemental Figure 13A and 13B). Furthermore, the estimated total cell number of grain containing reduced expression or gene edition of *TGW3* became significantly less than the control, yet those of overexpressed *TGW3* did not differ (Supplemental Figure 13C). Collectively, the *TGW3* regulation of spikelet hull length presumably results from its contrast effect on cell size and cell number, suggesting that it regulates grain size by coordinated alteration of cell expansion and cell division.

To investigate whether *TGW3* was a domestication-associated gene, we analyzed the genomic sequences of 1,083 *O. sativa* and 446 *O. rufipogon* accessions. However, we did not detect any domestication signals in the 100 kb region at the *TGW3* locus (Supplemental Figure 14A). We also analyzed the ORF cDNA sequence of *TGW3* from Sancunli (SCL) and Changxiangdao (CXD), two other rice varieties that have longer grains, and we found a sequence almost identical to its JZ1560 allele (Supplemental Figure 14B). These observations support a conclusion that *TGW3* might not be a domestication gene and its large-grain allele was a rare event.

In summary, we reported the in-depth dissection of QTL composition of a super-large grain，in which we identified more than 40 QTLs for grain size and yield, and importantly, most of which have not yet been cloned. We also map-based cloned a major QTL *qTGW3* that altered grain length. Impressively, *TGW3* modulates the trait also as a negative factor，and the modulation coordinately alters cell size and cell number in the spikelet hull. Our findings thus inform on the genetic architecture of a

super-large rice grain and uncover a novel mechanism of grain size regulation.

FUNDING

This work was supported by grants from the Chinese Academy of Sciences (XDA08020108-3), the National Natural Science Foundation of China (91435113, 31371605), National Program on Key Basic Research Project (973 Project, 2014CB943404) and Chinese High-yielding Transgenic Program (2016ZX08001-004).

Author Contributions

J.-Z.Y., M.M. and C.B. conducted most of the experiments; X.-H.H. performed population sequence comparison; J.-L.L., and Y.-Y.F. performed some of the experiments and experimental field management; X.-J.S. designed the experiments and wrote the paper.

REFERENCE

Che R.H., Tong H.N., Shi B.H., Liu Y.Q., Fang S.R., Liu D.P., Xiao Y.H., Hu B., Liu L.C., Wang H.R., Zhao M.F., and Chu C.C. (2016). Control of grain size and rice yield by GL2-mediated brassinosteroid responses. Nat. Plants **2**:15195-15201.

Fan C.C., Xing Y.Z., Mao H.L., Lu T.T., Han B., Xu C.G, Li X.H., Zhang Q.F. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor. Appl. Genet. **112**:1164-1171.

Hu J., Wang Y.X., Fang Y.X., Zeng L.J., Xu J., Yu H.P., Shi Z.Y., Pan J.J., Zhang D., Kang S.J., Zhu L., Dong G.J., Guo L.B., Zeng D.L., Zhang G.H., Xie L.H., Xiong G.S., Li J.Y., and Qian Q. (2015). A rare allele of GS2 enhances grain size and grain yield in rice. Mol. Plant **8**: 1455-1465.

Ishimaru K., Hirotsu N., Madoka Y., Murakami N., Hara N., Onodera H.,

Kashiwagi T., Ujiie K., Shimizu B., Onishi A., Miyagawa H., and Katoh E. (2013). Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. Nat. Genet. **45**:707-711.

Qi P., Lin Y.S., Song X.J., Shen J.B., Huang W., Shan J.X., Zhu M.Z., Jiang L.W., Gao J.P., and Lin H.X. (2012). The novel quantitative trait locus GL3.1 controls rice grain size and yield by regulating cyclin-T1;3. Cell Res. **22**: 1666-1680.

Si L.Z., Chen J.Y., Huang X.H., Gong H., Luo J.H., Hou Q.Q., Zhou T.Y., Lu T.T., Zhu J.J., Shangguan Y.Y., Chen E.W., Gong C.X., Zhao Q., Jing Y.F., Zhao Y., Li Y., Cui L.L., Fan D.L., Lu Y.Q., Weng Q.J., Wang Y.C., Zhan Q.L., Liu K.Y., Wei X.H., An K., and Han B. (2016). OsSPL13 controls grain size in cultivated rice. Nat. Genet. **48**:447-456.

Song X.J., Kuroha T., Ayano M., Furuta T., Nagai K., Komeda N., Segami S., Miura K., Ogawa D., Kamura T., Suzuki T., Higashiyama T., Yamasaki M., Mori H., Inukai Y., Wu J.Z., Kitano H., Sakakibara H., Jacobsen S.E., and Ashikari M. (2015). Rare allele of a previously unidentified histone H4 acetyltransferase enhances grain weight, yield, and plant biomass in rice. Proc. Natl. Acad. Sci. USA **112**:76-81.

Tong H.N., Liu L.C., Jin Y., Du L., Yin Y.H., Qian Q., Zhu L.H., and Chu C.C. (2012). DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate Brassinosteroid responses in rice. The Plant Cell **24**:2562-2577.

Wang S.K., Li S., Liu Q., Wu K., Zhang J.Q., Wang S.S., Wang Y., Chen X.B., Zhang Y., Gao C.X., Wang F., Huang H.X., and Fu X.D. (2015a). The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nat. Genet. **47**:949-954.

Wang Y.X., Xiong G.S., Hu J., Jiang L., Yu H., Xu J., Fang Y.X., Zeng L.J., Xu E.B., Xu J., Ye W.J., Meng X.B., Liu R.F., Chen H.Q., Jing Y.H., Wang Y.H., Zhu X.D., Li J.Y., and Qian Q. (2015b). Copy number variation at the GL7 locus contributes to grain size diversity in rice. Nat. Genet. **47**:944-948.

Ying J.Z., Gao J.P., Shan J.X., Zhu M.Z., Shi M., and Lin H.X. (2012). Dissecting

the genetic basis of extremely large grain shape in rice cultivar 'JZ1560'. Journal of Genetics and Genomics **39**:325-333.

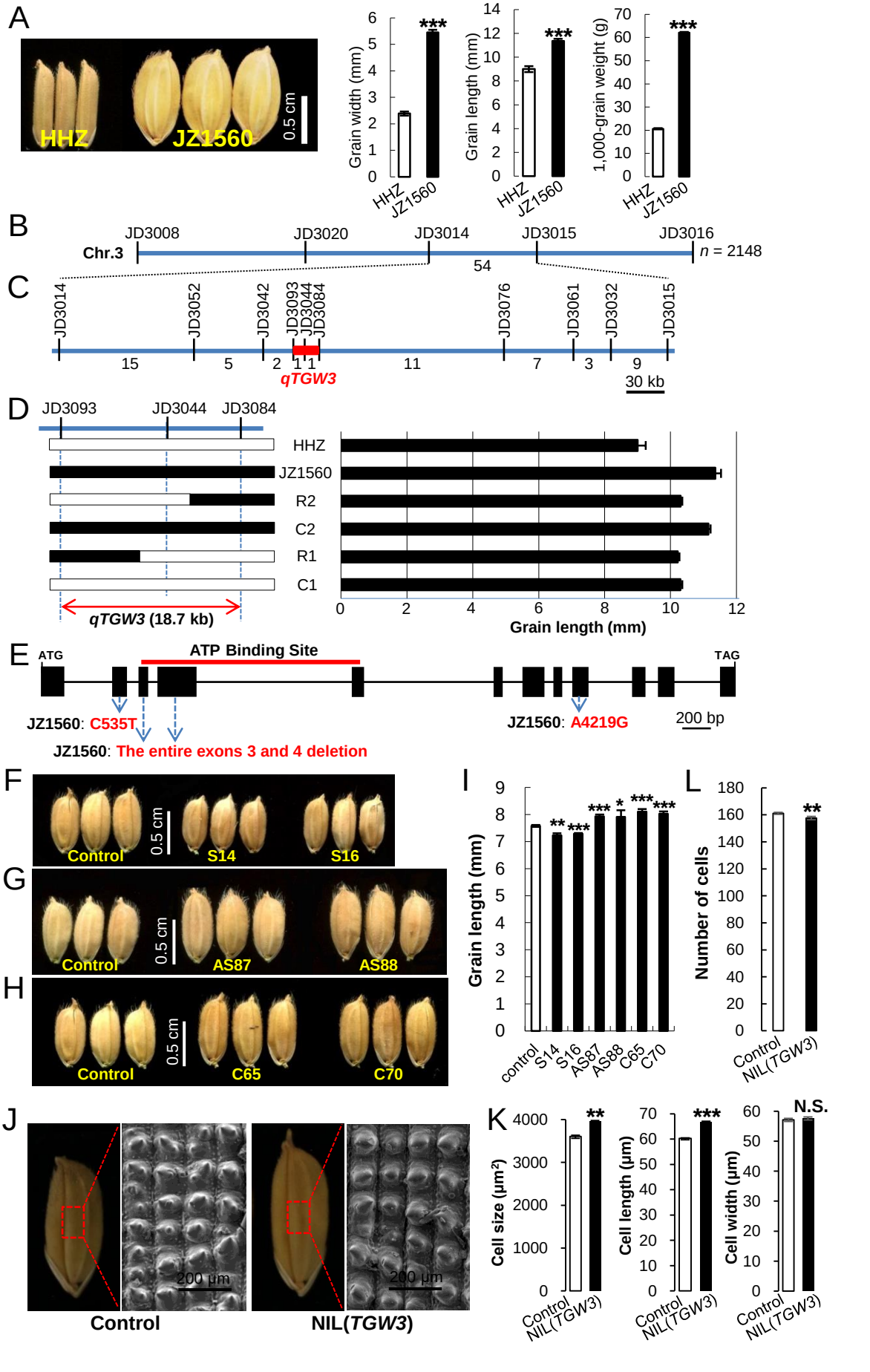
Zhang X.J., Wang J.F., Huang J., Lan H.X., Wang C.L., Yin C.F., Wu Y.Y., Tang H.J., Qian Q., Li J.Y., and Zhang H.S. (2012). Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. Proc. Natl. Acad. Sci. USA **109**: 21534-21539.

FIGURE LEGENDS:

Figure 1. QTL cloning at *qTGW3*, and its controlling of grain size might through coordinated alteration of cell size and cell number.

(A) Grain phenotypes of small-seed HHZ and super large-seed JZ1560. **(B)** The *qTGW3* locus was mapped initially on chromosome 3 and defined by the molecular markers JD3014 and JD3015. **(C)** High-revolution genetic mapping was conducted by analyzing a chromosome segment substitution segregating population of 2148 individuals of the *qTGW3* target region. **(D)** Progeny testing of fixed recombinant plants (F4) narrowed *qTGW3* to an 18.7-kb interval between molecular markers JD3084 and JD3093. Grain lengths (mean ± s.d.) of the recombinant line (R1) and control 1 (C1; homozygous for HHZ in the target region) did not differ, and another recombinant line (R2) was significantly higher than that of control 2 (C2; homozygous for JZ1560 in the target region). Filled and open bars represent homozygous chromosomal segments, respectively, for JZ1560 and HHZ. **(E)** *TGW3* gene structure and mutation sites, including nucleotide substitutions and deletions in JZ1560. The location of the predicted ATP Binding Site domain is indicated by a red line. Black boxes represent exons; thin lines between exons represent introns. Boxes and lines are drawn to scale as indicated. **(F–I)** Grain phenotypes of transgenic plants with overexpression of the HHZ allele **(F),** down-regulated endogenous expression **(G),** and gene edition by a CRISPR/Cas9 strategy **(H),** of the LOC_Os03g62500 gene, and quantification of the grain length traits in **F**, **G**, and **H (I).** S14 and S16 are transgenic lines carrying HHZ sense *TGW3* cDNA; AS87 and AS88 are transgenes carrying antisense *TGW3* cDNA from HHZ, and C65 and C70 are independent

CRISPR-mediated gene edition lines of *TGW3*. **(J, K,** and **L)** The spikelet hull of NIL(*TGW3*) is longer and contains much larger, but less cells than that of the HHZ isogenic control. **(J)** Images of spikelet hulls and cells by scanning electron microscopic observations of NIL(*TGW3*) and the corresponding control. **(K)** Comparisons of cell sizes, length, and width, and **(L)** total cell number, of the spikelet hulls in NIL(*TGW3*) and the HHZ isogenic control. Data in **A, I** (n = 20)**, K** and **L** (n = 30) are the means ± SD; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; N.S., not significant. Student's *t*-test was used to generate the *P* values.


A
HHZ
JZ1560
0.5 cm
Grain width (mm)
Grain length (mm)
1,000-grain weight (g)
***
HHZ
JZ1560
B
Chr.3
JD3008
JD3020
JD3014
JD3015
JD3016
n = 2148
54
C
JD3014
JD3052
JD3042
JD3093
JD3044
JD3084
JD3076
JD3061
JD3032
JD3015
15
5
2
1
1
11
7
3
9
qTGW3
30 kb
D
JD3093
JD3044
JD3084
HHZ
JZ1560
R2
C2
R1
C1
qTGW3 (18.7 kb)
Grain length (mm)
E
ATG
ATP Binding Site
TAG
JZ1560: C535T
JZ1560: A4219G
200 bp
JZ1560: The entire exons 3 and 4 deletion
F
Control
S14
S16
0.5 cm
G
Control
AS87
AS88
0.5 cm
H
Control
C65
C70
0.5 cm
I
Grain length (mm)
control
S14
S16
AS87
AS88
C65
C70
L
Number of cells
Control
NIL(TGW3)
J
200 μm
Control
NIL(TGW3)
K
Cell size (μm²)
Cell length (μm)
Cell width (μm)
N.S.
Control
NIL(TGW3)