The Plant Cell, Vol. 32: 923–934, April 2020, www.plantcell.org © 2020 ASPB.

A Single Amino Acid Substitution in STKc_GSK3 Kinase Conferring Semispherical Grains and Its Implications for the Origin of *Triticum sphaerococcum*[OPEN]

Xuejiao Cheng,[1] Mingming Xin,[1] Ruibin Xu, Zhaoyan Chen,[1] Wenlong Cai, Lingling Chai, Huanwen Xu, Lin Jia, Zhiyu Feng, Zihao Wang, Huiru Peng, Yingyin Yao, Zhaorong Hu, Weilong Guo, Zhongfu Ni,[2] and Qixin Sun[2]

State Key Laboratory for Agrobiotechnology/Key Laboratory of Crop Heterosis and Utilization, the Ministry of Education/Key Laboratory of Crop Genetic Improvement, National Plant Gene Research Centre, Joint Laboratory for International Cooperation in Crop Molecular Breeding, Ministry of Education, China Agricultural University, Beijing, People's Republic of China

ORCID IDs: 0000-0002-5665-8931 (X.C.); 0000-0003-4010-4165 (M.X.); 0000-0002-6976-5410 (R.X.); 0000-0003-3955-4134 (Z.C.); 0000-0002-9448-5736 (W.C.); 0000-0002-7905-8875 (L.C.); 0000-0001-7414-7142 (H.X.); 0000-0002-7801-4942 (L.J.); 0000-0001-5335-5823 (Z.F.); 0000-0002-5242-1910 (Z.W.); 0000-0002-8329-9428 (H.P.); 0000-0003-4122-4118 (Y.Y.); 0000-0002-1815-648X (Z.H.); 0000-0001-5199-1359 (W.G.); 0000-0003-4524-7720 (Z.N.); 0000-0002-3819-6892 (Q.S.)

Six subspecies of hexaploid wheat (*Triticum aestivum*) have been identified, but the origin of Indian dwarf wheat (*Triticum sphaerococcum*), the only subspecies with round grains, is currently unknown. Here, we isolated the grain-shape gene *Tasg-D1* in *T. sphaerococcum* via positional cloning. *Tasg-D1* encodes a Ser/Thr protein kinase glycogen synthase kinase3 (STKc_GSK3) that negatively regulates brassinosteroid signaling. Expression of *TaSG-D1* and the mutant form *Tasg-D1* in Arabidopsis (*Arabidopsis thaliana*) suggested that a single amino acid substitution in the Thr-283-Arg-284-Glu-285-Glu-286 domain of TaSG-D1 enhances protein stability in response to brassinosteroids, likely leading to formation of round grains in wheat. This gain-of-function mutation has pleiotropic effects on plant architecture and exhibits incomplete dominance. Haplotype analysis of 898 wheat accessions indicated that the origin of *T. sphaerococcum* in ancient India involved at least two independent mutations of *TaSG-D1*. Our results demonstrate that modest genetic changes in a single gene can induce dramatic phenotypic changes.

INTRODUCTION

Hexaploid wheat (*Triticum aestivum*) is a typical allopolyploid species containing three distinct genomes (A, B and D; AABBDD) resulting from two sequential allopolyploidization events. The first episode, which occurred 0.36 to 0.50 million years ago, involved a cross between *Triticum urartu* (AA genome) and an unidentified species (BB genome; Dvořák, 1976; Huang et al., 2002; Dvorak and Akhunov, 2005; Pont and Salse, 2017). The second polyploidization event between *Triticum turgidum* (AABB) and *Aegilops tauschii* (DD genome) ultimately led to the origin of hexaploid wheat ~10,000 years ago (Kihara, 1944; McFadden and Sears, 1946; Dvorak et al., 1998; Huang et al., 2002).

To date, six subspecies of hexaploid wheat (AABBDD) have been reported: *Triticum aestivum*, *Triticum spelta*, *Triticum vavilovi*, *Triticum macha*, *Triticum compactum*, and *Triticum sphaerococcum*. *Triticum zhukovskyi* was excluded from the list since it has a different genome constitution (AABBGG). *T. sphaerococcum*, also known as Indian dwarf wheat, is a landrace endemic to India and Pakistan (Ellerton, 1939; Hosono, 1954). Archaeological evidence supports the contribution of this wheat to ancient civilization in the Indus valley dating back to 9000 to 10,000 years ago (Josekutty, 2008; Mori et al., 2013). *T. sphaerococcum* exhibits semispherical grains, a semidwarf stature, short rigid culms, straight flag leaves, and compact and dense spikes (Percival, 1921; Sears, 1947). In addition, the grain has a shallow crease and high protein content compared to other wheat subspecies, making it a valuable genetic resource in wheat breeding programs (Singh, 1946). Thus, *T. sphaerococcum* was cultivated in the Indian subcontinent for thousands of years before the Green Revolution.

The spike-related morphological traits of hexaploid wheat, including seed threshability, spike compactness, and grain shape, are strongly influenced by three major genes: *Q* (free-threshing seeds; Mac Key, 1954; Faris et al., 2005), *C* (compact ears; Rao, 1972), and *s* (sphaerococcum grains; Rao, 1977; Salina et al., 2000). However, the genetic and biological basis of round grain development in *T. sphaerococcum* wheat is largely unknown. Miczynski (1930) crossed *T. aestivum* wheat with *T. sphaerococcum* and observed complete dominance of the *T. aestivum* trait over the *T. sphaerococcum* trait, indicating that the inheritance of the *T. sphaerococcum* trait is likely controlled by a single pleiotropic gene. Sears (1947) located the corresponding gene (denoted as *s*) on chromosome 3D by crossing *T. sphaerococcum* with nullisomics of wheat cv Chinese Spring and concluded that the *T. sphaerococcum* gene is hemizygous ineffective,

[1] These authors contributed equally to this work as co-first authors.
[2] Address correspondence to qxsun@cau.edu.cn and nizf@cau.edu.cn. The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the instructions for Authors (www.plantcell.org) are: Qixin Sun (qxsun@cau.edu.cn) and Zhongfu Ni (nizf@cau.edu.cn).
[OPEN]Articles can be viewed without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.19.00580

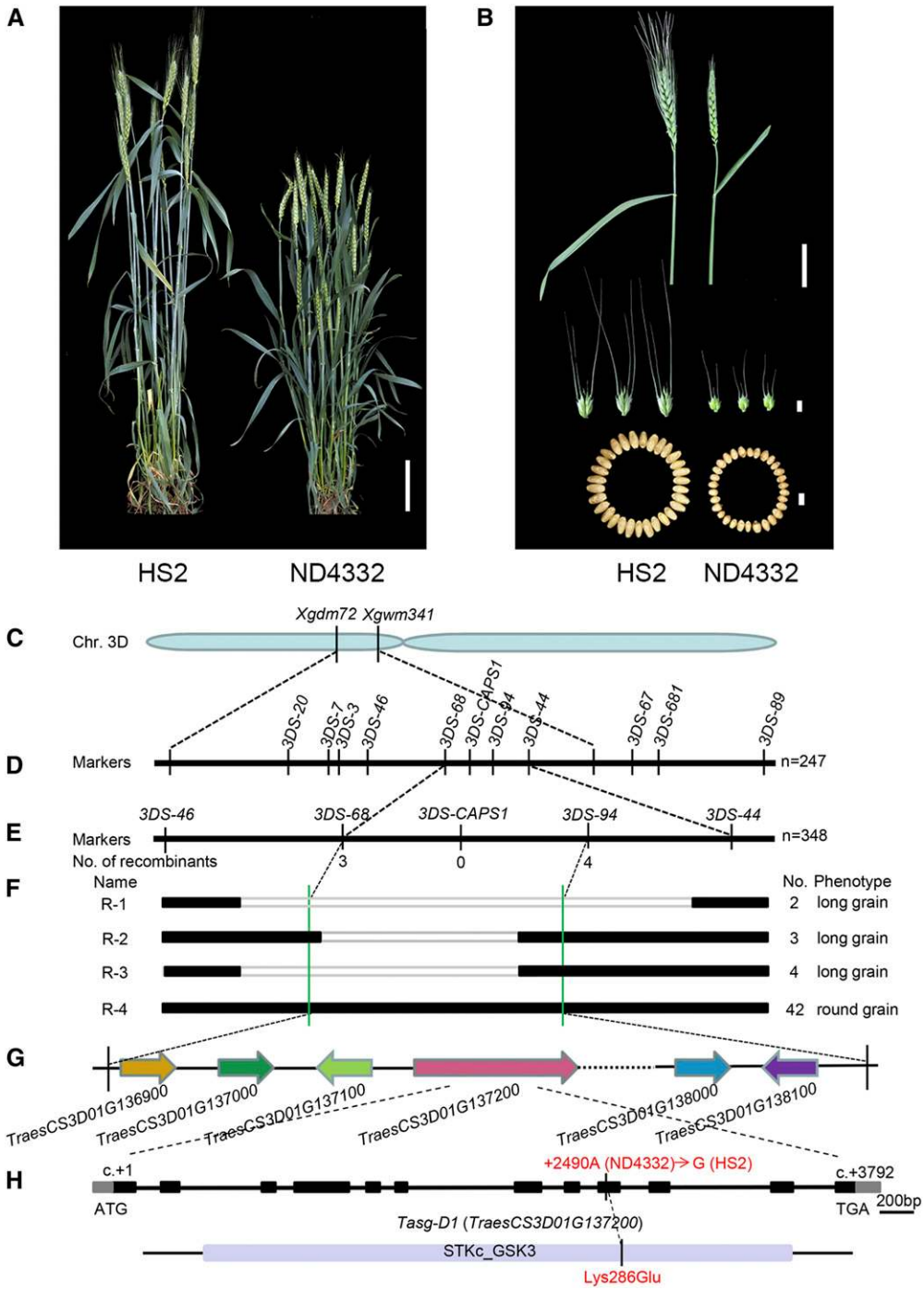

Figure 1. Characterization and Map-Based Cloning of the Semispherical Grain Gene *Tasg-D1*.

(A) Plant architecture of HS2 and ND4332. Bar = 10 cm.
(B) Spike, spikelet, and grain morphology of HS2 and ND4332; bar = 5 cm, 5 mm, and 5 mm, respectively.
(C) The *Tasg-D1* locus was mapped between markers *Xgdm72* and *Xgwm341* on chromosome (Chr.) 3DS near the centromeric region.
(D) High-resolution linkage map of *Tasg-D1*.
(E) The *Tasg-D1* gene was fine-mapped to a 1.01-Mb region between markers *3DS-68* and *3DS-94* using recombinants. The number of recombinants between the molecular marker and *Tasg-D1* is indicated.
(F) Analysis of genomic architecture and grain shape analysis for each recombinant type. The number of recombinants used for phenotypic analysis was indicated on the right. The black and white rectangles represent the homozygous ND4332 and HS2 regions, respectively.

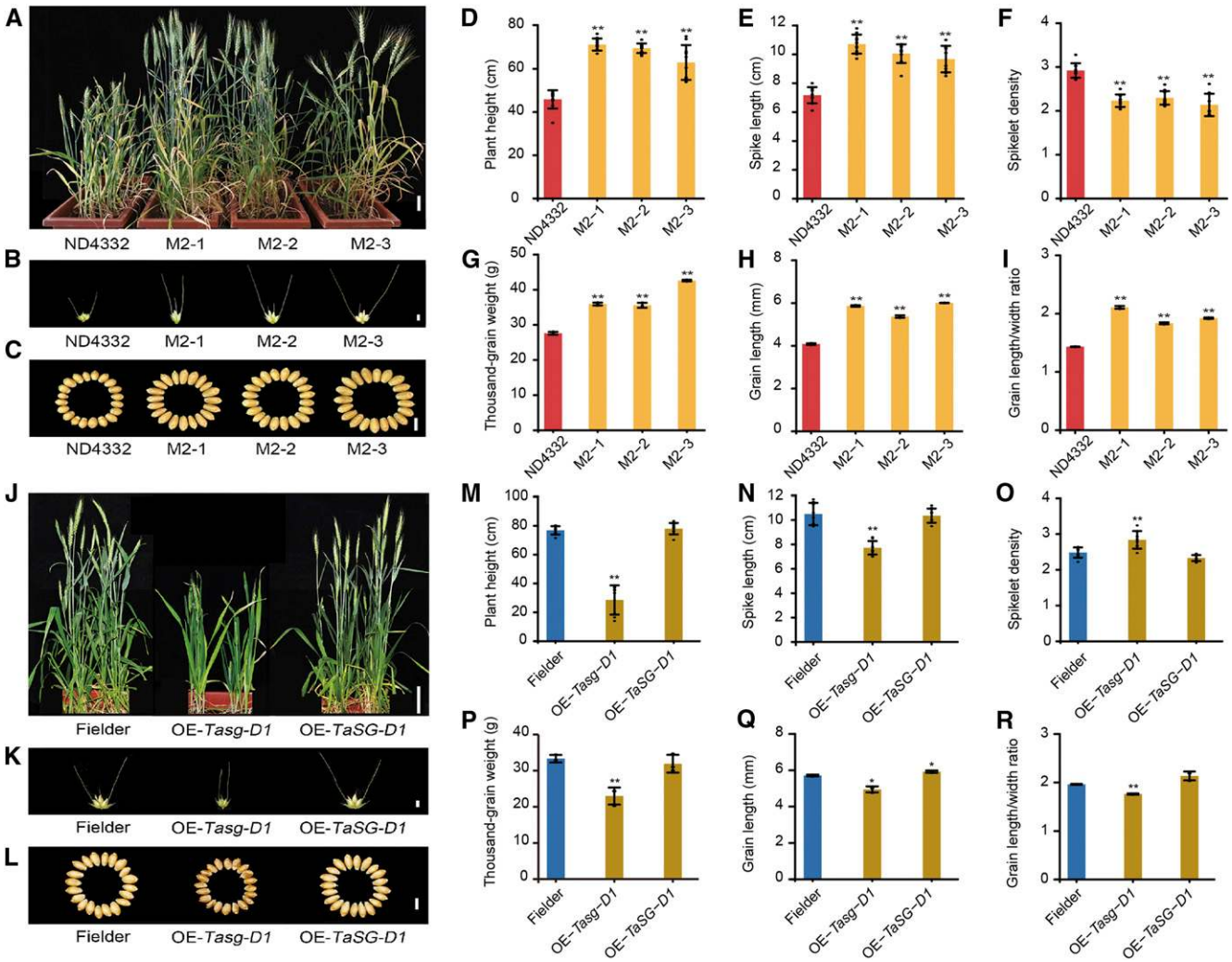

Figure 2. Functional Analysis of *Tasg-D1*.

(A) Gross morphology of ND4332 and its three independent EMS mutants. Bar = 10 cm.
(B) and **(C)** Spikelet **(B)** and grain morphology **(C)** of ND4332 and the mutants; bar = 5 mm and 5 mm, respectively. Twenty grains were aligned for photograph.
(D) to **(I)** ANOVA of plant height **(D)**, spike length **(E)**, spikelet density **(F)**, thousand-grain weight **(G)**, grain length **(H)**, and grain length/width ratio **(I)** between ND4332 and the mutants. Ten plants were selected for trait analysis.
(J) Gross morphology of Fielder (control) and OE-*Tasg-D1* and OE-*TaSG-D1* transgenic plants. Bar = 10 cm. **(K)** and **(L)** Spikelet **(K)** and grain morphology **(L)** of Fielder (control), OE-*Tasg-D1*, and OE-*TaSG-D1*; bar = 5 mm and 5 mm, respectively. Twenty grains were aligned for photograph.
(M) to **(R)** ANOVA of plant height **(M)**, spike length **(N)**, spikelet density **(O)**, thousand-grain weight **(P)**, grain length **(Q)**, and grain length/width ratio **(R)** among Fielder, OE-*Tasg-D1*, and OE-*TaSG-D1*. Eight plants were selected for data analysis. Values are means, with bars showing the SD. ** and * indicate significant differences at the 1 and 5% levels, respectively.

in contrast to the observation that the phenotypes segregate at a ratio of ~1(SS):2(Ss):1(ss; Ellerton, 1939). Subsequent studies confirmed this observation and mapped the *s* locus close to the centromere region of chromosome 3D (Rao, 1977; Koba and Tsunewaki, 1978; Singh, 1987). The *sphaerococcum* trait is not only restricted to the responsible gene in the D genome but also can be attributed to its homoeologs on the A and B genomes (Schmidt et al., 1963; Schmidt and Johnson, 1963, 1966). Indeed,

Figure 1. (continued).

(G) Predicted high-confidence genes in the mapping region according to the International Wheat Genome Sequencing Consortium wheat genome. Colored arrows indicate the orientation and order of the annotated genes.
(H) Gene structure and sequence analysis of *Tasg-D1* between HS2 and ND4332. The SNP information is indicated in red.

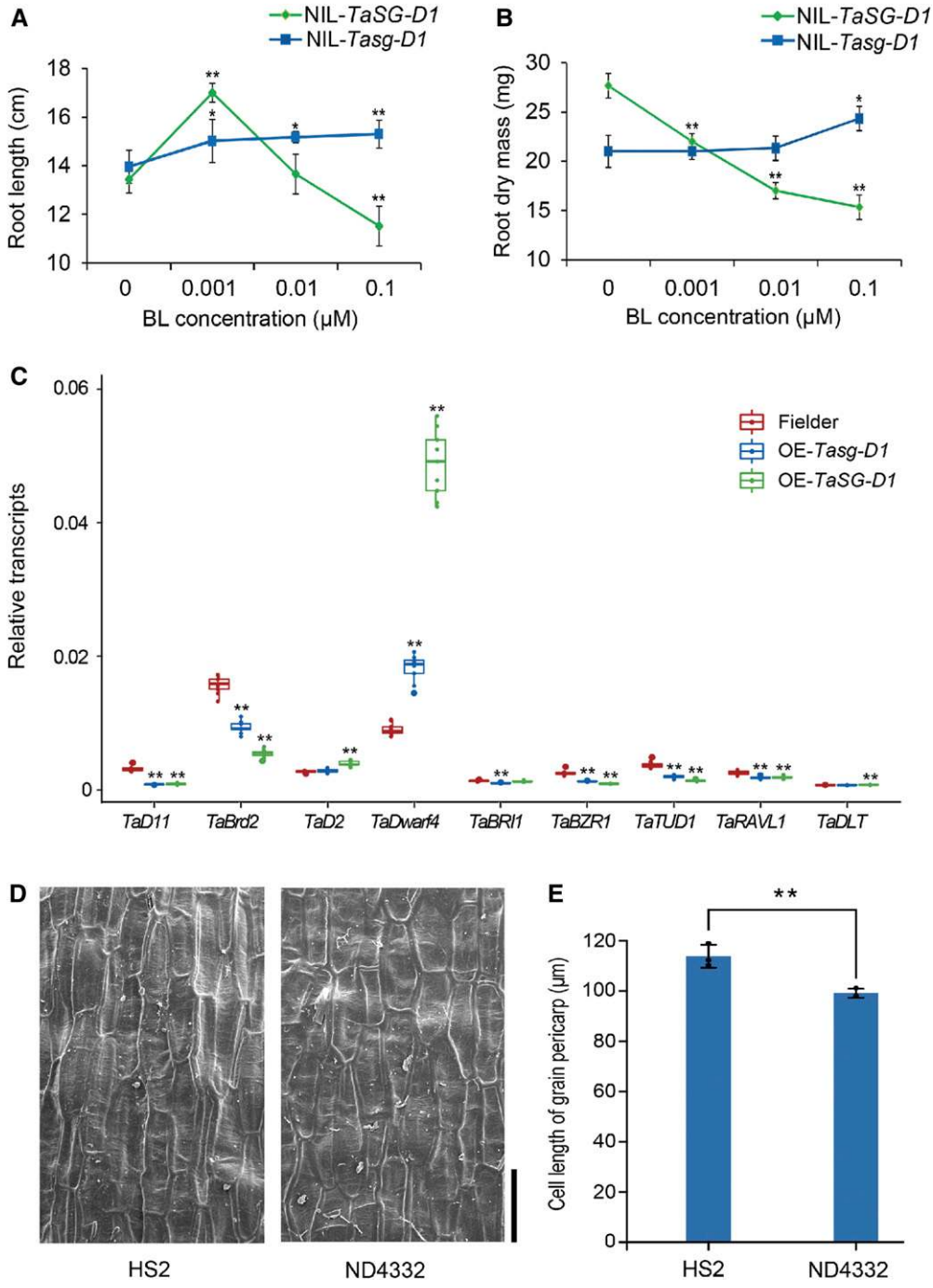

Figure 3. Effects of Point Mutations in *TaSG-D1* on Wheat.

(A) and **(B)** Comparison of root length **(A)** and dry mass **(B)** between NIL-*Tasg-D1* and NIL-*TaSG-D1* with different concentrations of BL at the seedling stage. NIL-*Tasg-D1* and NIL-*TaSG-D1* are recombinants with 90% background similarity, identified from the fine-mapping population. The BL concentrations were 0, 0.001, 0.01, and 0.1 μM, respectively. Three independent biological replicates were used (five plants per treatment).

(C) Expression analysis of BR biosynthesis-related and signaling-related genes in Fielder (control), OE-*Tasg-D1*, and OE-*TaSG-D1*. *TaD11*, *TaBrd2*, *TaD2*, and *TaDwarf4* are BR biosynthesis-related genes. *TaBRI1*, *TaBZR1*, *TaTUD1*, *TaRAVL1*, and *TaDLT* are BR signaling-related genes. Values are means, with bars showing the SD.

(D) and **(E)** Scanning electron microscopy of grain pericarp in HS2 and ND4332 **(D)** and statistical analysis **(E)**. Bar = 100 μm. A minimum of 100 cells per sample were measured using the instrument's software. Each parent was performed with three biological replicates. ANOVA method was conducted for statistical analysis. Values are means, with bars showing the SD. Asterisks indicate significance determined by ANOVA: *P < 0.05 and **P < 0.01.

Salina et al. (2000) mapped the *T. sphaerococcum* genes *s1*, *s2*, and *s3* on chromosomes 3D, 3B, and 3A, respectively.

Recent studies have highlighted the importance of brassinosteroids (BRs) in controlling grain shape through BR biosynthetic/homeostatic or signaling pathways (Huang et al., 2013; Jiang et al., 2013; Che et al., 2015; Zhu et al., 2015; Tian et al., 2019). BR-deficient mutants often exhibit smaller and less elongated seeds than the wild-type plants due to reduced cell length (Jiang et al., 2013; Che et al., 2015). The glycogen synthase kinase 3-like kinase (GSK) BRASSINOSTEROID INSENSITIVE2 (BIN2), a critical repressor of BR signaling, plays a crucial role in regulating grain shape. For example, loss of function of *OsGSK3*, a rice (*Oryza sativa*) ortholog of *BIN2*, alters the phosphorylation status of BRASSINAZOLE RESISTANT1 and induces its nuclear localization, leading to enhanced BR sensitivity and ultimately resulting in increased seed length (Gao et al., 2019). In addition, GSK2, another ortholog of BIN2, interacts with and represses the transcriptional activation activity of GROWTH REGULATION FACTOR4 in rice (Che et al., 2015), leading to reduced BR responses and small seeds (Che et al., 2015; Hu et al., 2015). However, the role of BR responses in determining grain shape in *T. sphaerococcum* remains unknown. Identifying the mechanism underlying this trait would elucidate the evolution of this subspecies.

Here, by performing positional cloning based on the semispherical seed phenotype, we isolated the subspecies-forming gene underlying grain shape in *T. sphaerococcum*. This gene, *Tasg-D1*, encodes a Ser/Thr protein kinase glycogen synthase kinase 3 (STKc_GSK3), the wheat ortholog of *BIN2*. In *T. sphaerococcum*, a single amino acid substitution of STKc_GSK3 enhances protein stability in response to BR, leading to round grain formation. Evolutionary analysis provided evidence that the origin of *T. sphaerococcum* wheat involved at least two independent mutations of *TaSG-D1*.

RESULTS AND DISCUSSION

Map-Based Cloning of *Tasg-D1*

To explore the genetic basis of semispherical grain trait and its origin in *T. sphaerococcum*, we generated an F2 segregating population using wheat lines HeSheng 2 (HS2, common wheat) and Nongda 4332 (ND4332, derived from a cross between *T. aestivum* and *T. sphaerococcum*; Figure 1A), which exhibit significant differences in grain shape (Figure 1B; Supplemental Data Set 1). The corresponding segregation ratio fits a Mendelian model of 3:1 (208 long grains/62 semispherical grains; $\chi^2 < \chi^2_{0.05,\,1} = 3.84$), indicating that the semispherical grain trait is controlled by a single nuclear gene, which is consistent with previous findings (Sears, 1947; Salina et al., 2000).

The gene that determines semispherical grain (hereafter referred to as *Tasg-D1* according to the official nomenclature rules in wheat; https://wheat.pw.usda.gov/GG2/Triticum/wgc/2013/; *Ta*, *Triticum aestivum*; *sg*, semispherical grain; *D1*, the first *sg* gene identified in genome D) was mapped between markers *Xgwm341* and *Xgdm72* on the short arm of chromosome 3D near the centromeric region (Figure 1C). To perform fine mapping of *Tasg-D1*, we generated an F7 recombinant inbred line (RIL) population of 247 lines derived from a cross between HS2 and ND4332 and confirmed that the locus is located between markers *3DS-68* and *3DS-44* (Figure 1D; Supplemental Data Set 2). We self-pollinated a residual heterozygous line (RHL) from the F7 RIL population to produce a segregating population and screened for recombinants. Using recombinant-derived progeny, we narrowed the candidate region to a 1.01-Mb region between markers *3DS-68* and *3DS-94* (Figures 1E and 1F; Supplemental Data Set 2).

This region contains 13 predicted high-confidence genes, as annotated based on the Ensembl Plants database (http://plants.ensembl.org/index.html; Figure 1G; Supplemental Data Sets 3 and 4; Appels et al., 2018). Resequencing of the 13 predicted high-confidence genes in the 1.01-Mb region revealed only one single-nucleotide polymorphism (SNP) in the coding sequences (CDSs). This SNP (A/G) is located in the ninth exon of *TraesCS3D01G137200* between ND4332 (*Tasg-D1*) and HS2 (*TaSG-D1*; Figure 1H). Although 253 SNPs and 32 insertions/deletions were detected in the intergenic regions, we focused on the SNP in the coding region, which led to an amino acid substitution from Lys (286K) to Glu (286E; Figure 1H).

TraesCS3D01G137200 is predicted to encode an STKc_GSK3 kinase containing a Thr-283-Arg-284-Glu-285-Glu-286 (TREE) domain (Figure 1H). *TraesCS3D01G137200* is an ortholog of *OsGSK1* in rice (Supplemental Figure 1A; Supplemental Data Set 3; Koh et al., 2007). Interestingly, the orthologous *OsGSK1* gene in rice is involved in regulating stress responses (Koh et al., 2007), but not grain shape, indicating that the roles of these orthologs have diverged. By contrast, *OsGSK2*, a homolog of *OsGSK1*, helps regulate seed development (Tong et al., 2012). Phylogenetic analysis revealed that the TREE domain is highly conserved among different species, including monocots and dicots (Supplemental Figure 1B).

By using the publicly available expression data for wheat (eFP browser; http://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi), we predicted the expression profiles of *TaSG-D1* in different tissues and found that it was highly expressed in shoot meristem, root, spike, grain, shoot axis, and ovary tissue. To verify these predictions, we performed reverse transcription quantitative PCR (RT-qPCR) analysis of *TaSG-D1* in 13 tissue/time points, including spikes and grains at different developmental stages (since we were focusing on seed shape formation). As expected, the expression patterns of the candidate gene were consistent with the publicly available data, with high levels of expression in roots, immature spikes, and seeds (Supplemental Figure 2A).

TaSG1 has three homoeologs since wheat is a typical hexaploid containing three genomes. We investigated their expression patterns according to the publicly available data and found that the three homoeologs shared similar expression patterns (Supplemental Figure 2B).

Genetic Analysis of *Tasg-D1*

Previous studies have revealed a discrepancy in the inheritance pattern of the *sphaerococcum* gene, which either has a hemizygous-ineffective recessive effect or an incompletely dominant effect (Sears, 1947; Schmidt et al., 1963; Salina et al., 2000). Since we determined that *Tasg-D1* is a gain-of-function allele in wheat

line ND4332, we evaluated its genetic effects on grain shape and other traits using a segregating population (F8-RHL-II; see Methods). The grain length showed a 1:2:1 segregation ratio in the examined populations ($\chi^2 = 0.46 < \chi^2_{0.05,2} = 5.99$; Supplemental Figure 3), supporting the notion that *Tasg-D1* shows incomplete dominance. Moreover, the phenotypes of heterozygous RHL-*TaSG-D1*/*Tasg-D1* individuals were significantly different from those of homozygous RHL-*TaSG-D1* and RHL-*Tasg-D1* individuals, including grain length, plant height, spike length, spikelet density, and thousand-grain weight (Supplemental Data Set 5), suggesting that the effects of *Tasg-D1* on plant architecture associated with the *s* locus are indeed pleiotropic.

Functional Analysis of *Tasg-D1*

The TREE domain likely plays a crucial role in GSK activity in plants (Li and Nam 2002; Koh et al., 2007; Tong et al., 2012). The overexpression of a mutant *OsGSK2* allele with a missense mutation in the TREE domain led to the production of small round seeds in rice (Tong et al., 2012). To determine whether the STKc_GSK3 kinase is associated with the *s* locus responsible for grain shape, we obtained three independent mutants (M2 generation) of *TaSG-D1* via ethyl methanesulfonate (EMS) mutagenesis in wheat line ND4332 (Supplemental Figure 4). Homozygous premature termination mutations resulted in plants with longer grains, larger spikes, and greater plant height than the wild-type plants (Figures 2A to 2I).

To investigate the potential functions of the *TaSG-D1* homoeologs, we obtained a plant with a mutation in the TREE domain of the *TaSG-A1* homoeolog (namely, line IIA236) in the Ningchun 4 (NC4) background by EMS mutagenesis (Supplemental Figure 5A). NC4 contained the same *TaSG-A1* allele as HS2, and the NC4 plants exhibited tall stature, large leaf angles, and long grain shape. Interestingly, the IIA236 mutant showed contrasting phenotypic variation, including semispherical grains (Supplemental Figures 5B to 5G). This result is consistent with the finding that the *Tasg-D1* gene homoeolog in the A genome contributes to the semispherical grain development (Schmidt et al., 1963; Schmidt and Johnson, 1963, 1966; Salina et al., 2000). In addition, BLASTP analysis (match length, >70%; sequence identity, >70%; *e* value, $<1 \times 10^{-5}$) revealed 22 homologs of *TaSG-D1* in the wheat genome. Only six of the homologs might be involved in negatively regulating the BR-mediated signaling pathway (according to the annotation in Ensembl Plant), indicating that neo- or subfunctionalization of the gene duplicates occurred independently during evolution, a topic that merits further investigation.

To further confirm our observations, we overexpressed the full CDS of this gene from HS2 (OE-*TaSG-D1*) and ND4332 (OE-*Tasg-D1*) driven by the *Ubiquitin* promoter in wheat cv Fielder. The overexpression of *Tasg-D1* appeared to cause more severe phenotypic variation compared to *TaSG-D1* (Figures 2J to 2R). Specifically, the OE-*Tasg-D1* lines had reduced grain length, whereas the OE-*TaSG-D1* lines had increased grain length compared to the control (Figures 2K, 2L, and 2Q). In addition, the OE-*Tasg-D1* lines exhibited delayed flowering, decreased plant height and spike length, increased spikelet density, decreased thousand-grain weight, and the failure to produce ears in tillers with unelongated internodes compared to control and OE-*TaSG-D1* plants (Figures 2J and 2M to 2P). Together, these findings confirm the notion that the STKc_GSK3 kinase is the *s* gene and that a single amino acid substitution is sufficient to confer pleiotropic effects on phenotype, including grain shape, between *T. sphaerococcum* and other wheat subspecies.

Tasg-D1 Functions as a Negative Regulator of the BR Signaling Pathway

The homologs of *TaSG-D1* in Arabidopsis and rice are negative regulators of BR signaling (Li and Nam, 2002; Koh et al., 2007; Tong et al., 2012). Therefore, to explore how *Tasg-D1* regulates grain shape, we treated the roots of plants of the near-isogenic lines (NILs) NIL-*TaSG-D1* and NIL-*Tasg-D1* with *epi*-brassinolide (BL). Root growth was significantly promoted in plants treated with a low BL concentration (0.001 μM, $P = 1.3 \times 10^{-6}$) and inhibited with a high BL concentration (0.1 μM, $P = 9.7 \times 10^{-4}$), compared to untreated NIL-*TaSG-D1* plants (Figure 3A). In addition, the root dry mass significantly decreased with increasing BL concentration (Figure 3B). By contrast, NIL-*Tasg-D1* plants showed less sensitivity to BL application than NIL-*TaSG-D1* plants, with only a gradual increase in root length in response to increasing BL concentrations from 0.001 to 0.1 μM. Moreover, the expression of BR biosynthesis and BR signaling-related genes decreased in both *Tasg-D1* and *TaSG-D1* overexpression plants compared to the wild type, including *TaD11*, *TaBrd2*, *TaBZR1*, *TaTUD1*, and *TaRAVL1* (Figure 3C). Consistent with the effects of BR on cell growth (Gasperini et al., 2012; Jiang et al., 2013; Che et al., 2015), cell length was significantly lower in ND4332 (*Tasg-D1*) than in HS2 (*TaSG-D1*), with a decrease of 12.97% ($P = 0.006$; Figures 3D and 3E), as revealed by scanning electron microscopy of the grain

	Nucleotide sequence (5'→3')	Allele type	Amino acid	No. of accessions
Ae. tauschii	A C C C G T G A G G A A	I	TREE	107
T. aestivum	A C C C G T G A G G A A	I	TREE	656
T. spelta	A C C C G T G A G G A A	I	TREE	121
T. sphaerococcum	A C C C G T G A G **A** A A	II	TREK	10
T. sphaerococcum	A C C **G** G T G A G G A A	III	TGEE	4

Figure 4. Haplotype Analysis of *TaSG-D1*.

Three types of TREE domains were detected in 898 wheat accessions with varying ploidy levels. I indicates the HS2-type allele; II indicates the ND4332-type allele; and III indicates the third type in which the TREE domain was changed to TGEE.

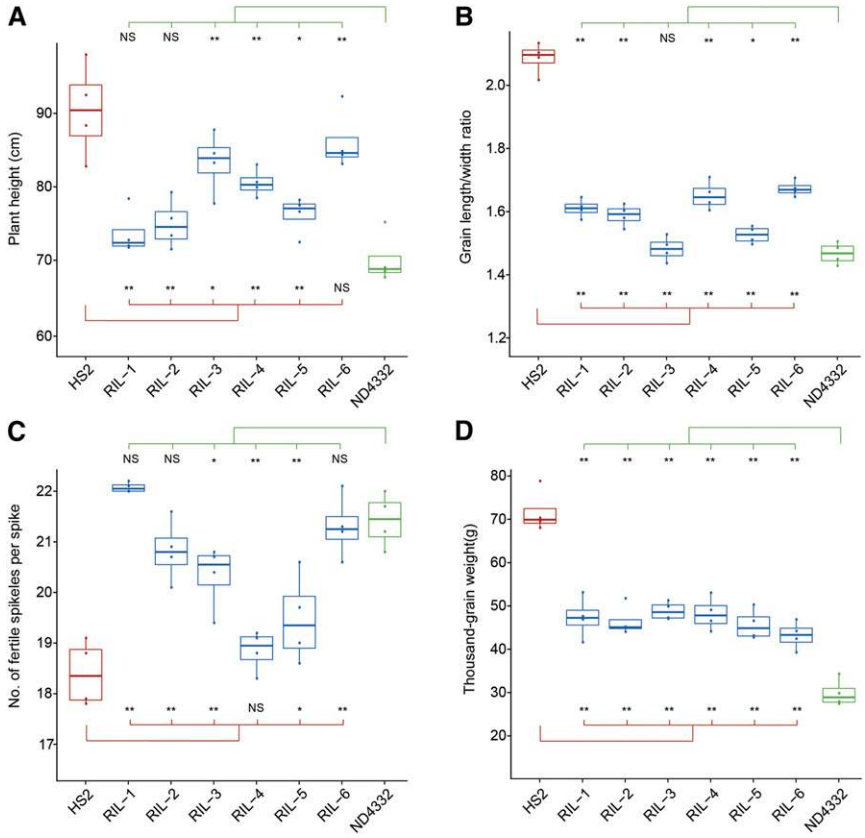

Figure 5. Phenotypic Comparison of the Parents and Six RILs.

(A) to **(D)** ANOVA of plant height **(A)**, grain length/width ratio **(B)**, number of fertile spikelets per spike **(C)**, and thousand-grain weight **(D)** among HS2, ND4332, and six RILs. Asterisks indicate significance determined by ANOVA: *P < 0.05 and **P < 0.01.

pericarp. These results indicate that *Tasg-D1* has a negative effect on cell elongation.

To further investigate the effect of *Tasg-D1* at the molecular level, we fused TaSG-D1 and Tasg-D1 with GFP (*TaSG-D1-GFP* and *Tasg-D1-GFP* respectively), and introduced them into Arabidopsis to test whether BR affects TaSG-D1 protein abundance. Similar to the results in wheat, *Tasg-D1-GFP* transgenic Arabidopsis plants showed more severe phenotypes than *TaSG-D1-GFP* plants (Supplemental Figure 6A). Furthermore, treatment with brassinazole, a specific inhibitor of BR biosynthesis, enhanced the dwarf phenotype of *TaSG-D1-GFP* plants but had little effect on *Tasg-D1-GFP* plants (Supplemental Figure 6B). BL treatment resulted in a rapid decrease in TaSG-D1-GFP abundance but no change in Tasg-D1-GFP abundance after 3, 6, 9, and 12 h of treatment (Supplemental Figure 6C). These results indicate that Tasg-D1 is more stable than TaSG-D1 in response to BR. Therefore, our results indicate that a single amino acid substitution in the TREE domain influences the accumulation of TaSG-D1.

The Origin of *T. sphaerococcum*

To explore the origin of *T. sphaerococcum*, we analyzed the sequences of the TREE domain of TaSG-D1 in 898 wheat accessions with varying ploidy levels. Of these, 14 *T. sphaerococcum* accessions (Supplemental Data Set 6) harbored two haplotypes, one identical to that of ND4332 and the other exhibiting an amino acid substitution from R to G in the TREE domain (Figure 4). Interestingly, all of the examined *Ae. tauschii* accessions contained the same allele to HS2 (Supplemental Data Set 7). Therefore, we propose that the *Tasg-D1* gene or related traits of *T. sphaerococcum* might have originated from *T. aestivum* due to a spontaneous mutation in the TREE domain of TaSG-D1 after hexaploidization. The mutations most likely occurred on two separate occasions. All of the other hexaploid wheat accessions with long grain shape contained the allele identical to that in HS2, including all *T. spelta* accessions (Supplemental Data Sets 8 and 9), indicating that the mutation likely occurred after domestication.

Potential Use of *Tasg-D1* for Wheat Yield Improvement

Although *Tasg-D1* itself had a negative effect on grain weight, as the NIL with the ND4332 allele (~28.9 g) exhibited significantly lower thousand-grain weight than the NIL containing the HS2 allele (~45.8 g; Supplemental Data Set 5), we evaluated the potential use of *Tasg1-D1* combined with other yield-related genes in wheat breeding programs by comprehensively examining yield-related traits in various RILs. Compared with the parent HS2, six selected lines had compact plant architecture, with erect leaves,

decreased plant height (~11.2 cm, 12.4%; Figure 5A), reduced grain length/width ratio (~0.5, 23.8%; Figure 5B; Supplemental Figure 7), and more fertile spikelets per spike (~2.1, 11.3%; Figure 5C), but the thousand-grain weight was reduced in these lines by an average of 34.9%. However, the thousand-grain weight of these RILs was 56.1% (16.8 g) larger compared to that of the other parental line, ND4332 (~29.9 g; Figure 5D; Supplemental Figure 7). Collectively, these results indicate that combining *Tasg-D1* with other favorable yield-related genes could improve wheat yield potential substantially, even though *Tasg-D1* itself has a negative effect on grain weight.

Finally, elliptical grains produced by plants harboring *Tasg-D1* could be beneficial for enhancing the flour extraction rate during wheat processing (Evers et al., 1990; Gegas et al., 2010). To test this hypothesis, we performed a flour extraction experiment using ND4332 (*Tasg-D1*) and HS2 (*TaSG-D1*) plants. The flour extraction rate was significantly lower in ND4332 (54.8%) versus HS2 (63.4%; Supplemental Figure 8), suggesting that elliptical grains might be one factor affecting the flour extraction rate. Other factors such as seed size and the depth of the crease might also contribute to the flour extraction rate. Thus, we conducted the experiment again using RILs with the same thousand-grain weight (~31.0 g). The flour extraction rate was indeed higher for semispherical grains (62.1%) than for long grains (57.4%; Supplemental Figure 8).

In conclusion, positional cloning and haplotype analysis of *Tasg-D1*, a gene affecting grain shape in wheat, broadened our view of the origin and evolution of *T. sphaerococcum* species. Introducing *Tasg-D1* into modern wheat cultivars could facilitate the genetic engineering of grain shape and potentially improve the flour extraction rate.

METHODS

Plant Materials

Line ND4332 was derived from a cross between common wheat and natural *Triticum sphaerococcum* wheat. HS2 is a common wheat (*Triticum aestivum*) cultivar with large grains. An F2 population of 270 individuals was generated by self-pollinating the F1 progeny of a cross between HS2 and ND4332 and used for genetic analysis and candidate gene mapping. A population comprising 247 RILs was developed by crossing HS2 and ND4332 and was advanced to the F7 generation by single-seed descent. We identified two residual heterozygous lines in the F7 RIL population, including one line (RHL-I) carrying the heterozygous segment in the genetic region from simple sequence repeat (SSR) markers *3DS-68* to *3DS-94* and the other (RHL-II) carrying the heterozygous segment from SSR markers *3DS-20* to *3DS-94*. These two lines were self-pollinated to produce the F8 generations (F8-RHL-I and F8-RHL-II, respectively). The recombinants of F8-RHL-I were self-pollinated to generate a segregation population for fine-mapping of candidate genes. The F8-RHL-II population was used to analyze the effect of *Tasg-D1* on grain shape and other agronomic traits.

Seeds from self-pollinated ND4332 plants were used to construct the EMS mutant population. More than 10,000 seeds were soaked in water for 12 h at room temperature and imbibed in a 0.4% (w/v) aqueous solution of EMS (M0880, Sigma-Aldrich) in the dark for 16 h at room temperature. The seeds were washed in water for 12 h and air-dried at room temperature. M1 plants with variant phenotypes were selected and self-pollinated to produce the M2 generation. In addition, IIA236, a *sphaerococcum* mutant, was identified from the EMS-mutagenized population of the spring wheat cv NC4 kindly provided by Ligeng Ma (Capital Normal University, China) and used for analysis of *TaSG-A1*.

Using the marker *3D-CASP1*, 898 wheat accessions with varying ploidy levels were used to test the allele frequency of *TaSG-D1*; these accessions included 791 allohexaploid wheat accessions (14 *T. sphaerococcum* wheat accessions, kindly provided by Yuming Wei, Sichuan Agricultural University, China), 296 Chinese common wheat accessions, and 481 accessions from other countries (Supplemental Data Sets 6, 8, and 9, respectively) and 107 diploid accessions (*Ae. tauschii*; Supplemental Data Set 7).

The Fielder accession (United States spring hexaploid wheat) and Arabidopsis (*Arabidopsis thaliana*) Columbia (Col-0) ecotype were used for transformation. Transgenic wheat plants were grown in a greenhouse (400W HPS lamps, light intensity of 3000 lux) under 16 h of light at 24°C and 8 h of dark at 16°C. The transgenic Arabidopsis plants were grown in a greenhouse (white light, ~100 μmol m^{-2} s^{-1}) at 22°C/18°C under a 16-h-light/8-h-dark cycle.

Scanning Electron Microscopy

HS2 and ND4332 seeds were cleaned, placed crease down onto aluminum specimen stubs, and sputter coated with gold. The grains were observed and photographed under an S-3400 scanning electron microscope (Hitachi; Wu et al., 2016). Cell length was measured in a minimum of 100 cells per sample using Photoshop 7.0 (Adobe). Each experiment was performed with three biological replicates using seeds from at least three plants per biological replicate. Analysis of variance (ANOVA) for cell size was conducted using SPSS 22.0 software with default parameters. Asterisks indicate significant differences determined by ANOVA (Supplemental File 1): *$P < 0.05$ and **$P < 0.01$.

Trait Evaluation

Seeds from the F2, F8-RHL-I, and F8-RHL-II segregating populations, the parent, and M0 seeds of ND4332 were planted in a seasonal growing location in Shangzhuang Experimental Station in Beijing, China, with 20 seeds per row (1.5-m-long rows). The F7 RIL population was grown in Beijing, Shandong, and Hebei using a randomized complete block design with three biological replicates. Each line was hand sown in two-row plots of 1.5-m-long rows with 30 seeds per row. The seeds of each M1 plant were planted in one row in Beijing. The three independent M2 lines and parent (ND4332) were grown in the greenhouse with 10 seeds per plot with three biological replicates. The Fielder and two independent transgenic plants of T1 generation were also grown in the greenhouse with six plants per row and two rows per pot. Agronomic traits, including plant height, spike length, and spikelet density, were examined before harvest. Thousand-grain weight, grain length, and grain width were determined using a camera-assisted phenotyping system (Wanshen Detection Technology). Statistical analysis was performed via one-way ANOVA using SPSS 22.0 software with default parameters. ** and * indicate significant differences at the 1 and 5% levels, respectively. A chi-square (χ^2) test was used to test whether the distribution of the test statistic for each trait fit a chi-square distribution.

Marker Development

Microsatellite markers were designed and used to construct a genetic map of chromosome 3DS. The chromosome 3DS reference sequence of the Chinese Spring wheat (http://www.wheatgenome.org/) and the *Aegilops tauschii* (accession AL8/78) genome sequence (http://aegilops.wheat.ucdavis.edu/ATGSP/) were used to design new SSR markers as described by Cheng et al. (2015). A parental polymorphism survey and identification of polymorphic SSR markers were conducted using the PCR conditions described by Cheng et al. (2015).

A cleaved amplified polymorphic sequence (CAPS) marker, *3DS-CAPS1*, was developed as follows. Based on the sequence of rice gene *Os01g0205700* (*OsGSK1*), the cDNA of the wheat orthologous gene was cloned and sequenced to detect polymorphisms between the two parents. Sequence analysis revealed an SNP in the cDNA sequences between the parents, that is, a *MnlI* site (CCTC) mutation in ND4332 (TCTC). This SNP was used to develop a chromosome 3DS specific CAPS marker (*3DS-CAPS1*). The PCR product (10 μL) was digested with restriction enzyme *MnlI* restriction enzymes for 4 h, and the digestion products were separated in 2% (w/v) agarose gels and visualized under UV light after ethidium bromide staining. As expected, 107-bp PCR products were completely cleaved by *MnlI* in HS2, but not in ND4332. Detailed information about the polymorphic markers used for genetic map construction is provided in Supplemental Data Set 10.

Linkage analysis of the markers was performed using JoinMap 4.0 software (http://www.kyazma.nl/index.php/mc.JoinMap/sc.General).

BL Treatment of Roots

Two individuals with 90% background similarity were selected from the F9 population, which were detected using 210 SSR primer pairs on 21 chromosomes. The two individuals had homozygous segments from markers *3DS-68* to *3DS-94* of ND4332 and HS2; the individuals were named NIL-*Tasg-D1* and NIL-*TaSG-D1*, respectively. The seeds were surface sterilized with 75% (v/v) ethanol for 30 s and 1% (v/v) NaClO for 10 min and germinated for 1 d in Petri dishes. The germinated seeds were incubated at 4°C for 3 d in the dark and exposed to white light for 1 d. Seedlings with uniform growth were sown in boxes containing the appropriate BL (E1641, Sigma-Aldrich) concentration (0, 0.001, 0.01, and 0.1 μM) and grown in a greenhouse for 4 d (16 h of light at 24°C and 8 h of dark at 16°C; 75% humidity). Primary roots were stretched out with forceps and measured with a ruler. The experiment was repeated three times, with at least five plants per treatment. The roots of all five seedlings were removed and placed together in a single kraft-paper bags. Root dry mass was measured after the roots were dried for 3 d at 60°C.

Genome Resequencing

Genomic DNA from HS2 and ND4332 was used to build paired-end sequencing libraries with insert sizes of ~500 bp as described previously (Chai et al., 2018). We performed sequencing with an average 5× coverage of the assembled genome using the Illumina NovaSeq 6000 platform. High-quality reads were aligned to the reference genome of Chinese Spring (RefSeq v1.0; Appels et al., 2018) using the Burrows-Wheeler Aligner 0.7.15 program with default parameters (Li and Durbin, 2009). SNP and insertions/deletions calling was performed using the HaplotypeCaller module. Illumina reads of all samples were deposited in the Bioproject at the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA533588.

Gene Cloning and Sequence Analysis

Homoeolog-specific primer pairs were designed to amplify the genomic and full-length CDSs of the three *TaSG1* homoeologs. The PCR assays were performed with high-fidelity Tks Gflex polymerase (TaKaRa) for 35 cycles (94°C for 1 min; cycles of 98°C for 10 s, 58°C for 15 s, and 68°C for 30 s; followed by 68°C for 10 min). The PCR products were confirmed using 1% (w/v) agarose gels and visualized under UV light after ethidium bromide staining. The purified PCR products were cloned into the pEASY-Blunt vector (TaKaRa) and sequenced. DNAMAN software and the online tool Conserved Domain Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) were used to analyze the nucleotide and amino acid sequences.

Phylogenetic Analysis

We downloaded the sequences of TaSG-D1 homologous proteins from various species, including both monocots and dicots, by performing a BLASTP (nr) search in the NCBI database (Supplemental File 2). The amino acid sequences were aligned with ClustalW, and a phylogenetic tree was constructed using the neighbor-joining method with MEGA 6.06 software with default parameters (Tamura et al., 2013). The evolutionary distances were calculated using Poisson model. The phylogeny test was computed using bootstrap method with 1000 replications (Supplemental File 3).

RT-qPCR Analysis

For expression analysis, 13 different tissues from HS2 and leaves from transgenic wheat lines and Arabidopsis lines were collected from at least three plants per biological replicate for RNA extraction. Total RNA was extracted from the samples using the standard TRIzol RNA isolation protocol (Invitrogen). The detailed protocols and procedures were as described by Han et al. (2016). RT-qPCR was performed using SYBR Green PCR Master Mix (TaKaRa) with a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.). The RT-qPCR conditions and analytical methods were the same as those used in Wang et al. (2018). The wheat *β-ACTIN* gene (Paolacci et al., 2009) was used as an endogenous control. The specific primer pairs used for amplification are listed in Supplemental Data Set 10. Each experiment was performed with three biological replicates.

Vector Construction and Plant Transformation

For wheat, the full-length CDS of *TaSG-D1* and *Tasg-D1*of HS2 and ND4332 were amplified using specific primer pairs and inserted into the pWMB110 vector via homologous recombination technology with a ClonExpress II One Step Cloning Kit (Vazyme). The sequencing-confirmed vectors were transformed into 15-d-old immature wheat embryos via *Agrobacterium*-mediated transformation.

For Arabidopsis, the *35S*:*TaSG-D1-GFP* and *35S*:*Tasg-D1-GFP* constructs were generated by inserting the full-length CDS of HS2 and ND4332 into a modified pCAMBIAsuper1300 vector containing the green fluorescent protein (GFP) gene. The resulting vectors were introduced into *Agrobacterium tumefaciens* GV3101 and transformed into the wild-type Arabidopsis plants using the floral dip method (Clough and Bent, 1998). All primers used for PCR amplification are listed in Supplemental Data Set 10.

Chemical Treatment of Arabidopsis Seedlings

As described by Peng et al. (2008), seeds from two independent transgenic Arabidopsis lines of T2 generation were germinated on one-half-strength Murashige and Skoog (MS) medium supplemented with 2 μM brassinazole (SML1406, Sigma-Aldrich), a specific inhibitor of BR biosynthesis, and grown in a greenhouse at 22°C under a 16-h-light/8-h-dark cycle for 2 weeks. The seedlings were transferred to half strength MS liquid medium containing 1 mM BL and incubated for 0, 3, 6, 9, and 12 h prior to immunoblot analysis.

Immunoblot Analysis

Transgenic Arabidopsis seedlings grown on half strength MS medium with or without chemical treatment were collected, transferred to liquid nitrogen, and ground into a fine powder. Total protein was extracted from the samples via incubation at 100°C for 10 min in 5× SDS-PAGE sample buffer (Zhang, et al., 2017), followed by centrifugation at 12,000*g* for 10 min at room temperature. The protein samples were separated in an 8% (v/v) SDS-PAGE gel for 1 h and 10 min, transferred to a polyvinylidene difluoride

membrane (Millipore), and blocked with 5% (w/v) fat-free milk in Tris-buffered saline with 0.1% (v/v) Tween-20 (TBST; Guo et al., 2015) for 1 h. The membrane was incubated for 1 h in antibodies [anti-ACTIN antibody [A0480, Sigma] and anti-GFP antibody [ab290, Abcam] diluted 1:5000), washed three times (10 min each) with TBST, incubated with secondary antibody diluted 1:5000 for 1 h, and washed three times (5 min each) with TBST. Specific protein bands were visualized with Immobilon Western Chemiluminescent horseradish peroxidase substrate (http://www.millipore.com). ACTIN served as a loading control.

Flour Extraction Experiment

The flour extraction rate was examined using 300 g of seeds from HS2, ND4332, and two RILs harvested in Beijing in 2017. The thousand-grain weight of HS and ND4332 was 67.4 and 25.0 g, respectively, whereas the two RILs had the same thousand-grain weight (31.0 g). Seed moisture content was measured with a Grain analyzer machine (Infratec 1241, FOSS), and water was added until the moisture content rose to 14%. The seeds were stored at room temperature for 24 h before starting the flour extraction experiment using a flour-milling machine (CD1, Chopin Moulin). Each sample was performed with three biological replicates. Finally, ANOVA was performed to calculate differences in flour extraction rates using SPSS 22.0 software with default parameters. Asterisks indicate significant differences determined by ANOVA: *$P < 0.05$ and **$P < 0.01$.

Accession Numbers

Sequence data from this article for RT-qPCR can be found in the Ensembl Plants data library or NCBI data library under the following accession numbers: TraesCS2B02G350400 (*TaD11*); TraesCS7A02G559400 (*TaBrd2*); TraesCS3D02G106100 (*TaD2*); TraesCS4B02G234100 (*TaDwarf4*); TraesCS3D02G246500 (*TaBRI1*); TraesCS2A02G187800 (*TaBZR1*); TraesCS4A02G060800 (*TaTUD1*); TraesCS2D02G392400 (*TaRAVL1*); TraesCS4A02G430600 (*TaDLT*); NP_175350.1 (*AtActin*); TraesCS5B02G124100 (*TaActin*). Sequences used for phylogenetic analysis can be found in NCBI under the following accession numbers: *OsGSK1*, Q9LWN0.1; *OsGSK2*, XP_015637571.1; *Aegilops tauschii* (*AeGSK*), XP_020190077.1; *Hordeum vulgare* (*HvGSK2.1*), BAJ91341.1; *Brachypodium distachyon* (*BdGSK1*), XP_003565257.1; *Oryza sativa*, XP_015617780.1; *Sorghum bicolor* (SbGSK1), XP_002454973.1; *Zea mays*, NP_001131812.1; *Arabidopsis thaliana* (AtBIN2), NP_193606.1; *Brassica napus*, XP_013702388.1; *Ananas comosus*, XP_020097081.1; *Panicum miliaceum,* RLM91534.1; *Chlamydomonas reinhardtii*, PNW74746.1; *Vitis vinifera*, XP_010657061.1; *Nelumbo nucifera*, XP_010274430.1; *Elaeis guineensis*, XP_010915474.1; *Citrus sinensis*, XP_006468992.1; *Ipomoea nil*, XP_019194482.1; *Carica papaya*, XP_021898006.1; *Musa acuminate*, XP_009394828.1; *Daucus carota*, XP_017248488.1; *Helianthus annuus*, XP_022002937.1; *Solanum tuberosum*, XP_006345509.1; *Solanum lycopersicum*, XP_004240041.1; *Nicotiana tabacum*, NP_001312816; *Populus trichocarpa*, XP_002320754.2; *Glycine max*, XP_003528879.1; *Ziziphus jujube*, XP_015874161.1; *Hevea brasiliensis*, XP_021678477.1; *Camellia sinensis*, XP_028055688.1; *Sesamum indicum*, XP_011082988.1; *Malus domestica*, XP_008383613.1; *Medicago truncatula*, XP_003596712.2; *Prunus persica*, XP_007211435.1; *Morus notabilis*, EXC00262.1.

Genome resequencing data of all samples were deposited in the Bioproject at the NCBI (https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA533588.

Supplemental Data

Supplemental Figure 1. Sequence analysis of TaSG-D1 in different species.

Supplemental Figure 2. Expression analysis of *TaSG-D1* and its homoeologs.

Supplemental Figure 3. Distribution of grain length in the segregating population.

Supplemental Figure 4. Sequence analysis of *TaSG-D1* in various EMS mutants.

Supplemental Figure 5. Sequence analysis and phenotypic characterization of NC4 and the EMS mutant IIA236.

Supplemental Figure 6. Functional analysis of *TaSG-D1* and *Tasg-D1* in Arabidopsis.

Supplemental Figure 7. Grain morphology of the parents and six RILs.

Supplemental Figure 8. Flour extraction rate comparison analysis of RILs and their parents.

Supplemental Data Set 1. Phenotypic performance of HS2 and ND4332.

Supplemental Data Set 2. Genotype data of RIL population and recombinants.

Supplemental Data Set 3. Comparative analysis of orthologous genes between wheat and rice at *Tasg-D1* locus.

Supplemental Data Set 4. Detailed information about the 13 annotated high-confidence genes in the 1.01-Mb region.

Supplemental Data Set 5. Statistical analysis of the effect of *Tasg-D1* on agronomic traits in a segregating population differing in grain shape.

Supplemental Data Set 6. Origin and identity of the *T. sphaerococcum* wheat accessions (AABBDD) used for the allele frequency analysis.

Supplemental Data Set 7. Origin and identity of the *Ae. tauschii* accessions (DD) used for the allele frequency analysis.

Supplemental Data Set 8. Origin and identity of the Chinese wheat accessions (AABBDD) used for the allele frequency analysis.

Supplemental Data Set 9. Origin and identity of the wheat accessions from other countries (AABBDD) used for the allele frequency analysis.

Supplemental Data Set 10. Primers used in this study.

Supplemental File 1. Statistical analysis.

Supplemental File 2. Text file of the alignment used for the phylogenetic analysis shown in Supplemental Figure 1.

Supplemental File 3. Machine-readable tree file of the phylogenetic analysis.

ACKNOWLEDGMENTS

We thank Ligeng Ma (Capital Normal University, China) for providing the seeds and for information about NC4 and the IIA236 mutant, Yuming Wei (Sichuan Agricultural University, China) for providing the seeds and for information about *T. sphaerococcum* wheat accessions, and Na Song (China Agricultural University) for help with wheat transformation. We are also grateful to Zhenqi Su (China Agricultural University) for editing the article. This work was supported by the Major Program of the National Natural Science Foundation of China (grants 91935304 and 31991210) and the National Key Research and Development Program of China (grant 2017YFD0101004).

AUTHOR CONTRIBUTIONS

Q.S. and Z.N. conceived the project; X.C. and R.X. performed the experiments; X.C., M.X., R.X., Z.C., W.C., L.C., H.X., L.J., and Z.F. collected the plant materials; X.C., M.X., Z.W., H.P., Y.Y., Z.H., and W.G. analyzed the data; M.X., X.C., Z.N., and Q.S. wrote the article.

Received August 1, 2019; revised January 13, 2020; accepted February 11, 2020; published February 14, 2020.

REFERENCES

Appels, R., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science **361:** eaar7191.

Chai, L., et al. (2018). Dissection of two quantitative trait loci with pleiotropic effects on plant height and spike length linked in coupling phase on the short arm of chromosome 2D of common wheat (*Triticum aestivum* L.). Theor. Appl. Genet. **131:** 2621–2637.

Che, R., Tong, H., Shi, B., Liu, Y., Fang, S., Liu, D., Xiao, Y., Hu, B., Liu, L., Wang, H., Zhao, M., and Chu, C. (2015). Control of grain size and rice yield by *GL2*-mediated brassinosteroid responses. Nat. Plants **2:** 15195.

Cheng, X., Chai, L., Chen, Z., Xu, L., Zhai, H., Zhao, A., Peng, H., Yao, Y., You, M., Sun, Q., and Ni, Z. (2015). Identification and characterization of a high kernel weight mutant induced by gamma radiation in wheat (*Triticum aestivum* L.). BMC Genet. **16:** 127.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Ellerton, S. (1939). The origin and geographical distribution of *Triticum sphaerococcum* Perc. and its cytogenetical behaviour in crosses with *T. vulgare* Vill. J. Genet. **38:** 307–324.

Evers, A., Cox, R., Shaheedullah, M., and Withey, R. (1990). Predicting milling extraction rate by image analysis of wheat grains. Asp. Appl. Biol. **25:** 417–426.

Gasperini, D., Greenland, A., Hedden, P., Dreos, R., Harwood, W., and Griffiths, S. (2012). Genetic and physiological analysis of Rht8 in bread wheat: An alternative source of semi-dwarfism with a reduced sensitivity to brassinosteroids. J. Exp. Bot. **63:** 4419–4436.

Gegas, V.C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J.H., and Snape, J.W. (2010). A genetic framework for grain size and shape variation in wheat. Plant Cell **22:** 1046–1056.

Dvořák, J. (1976). The relationship between the genome of *Triticum urartu* and the A and B genomes of *Triticum aestivum*. Can. J. Genet. Cytol. **18:** 371–377.

Dvorak, J., and Akhunov, E.D. (2005). Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the *Aegilops-Triticum* alliance. Genetics **171:** 323–332.

Dvorak, J., Luo, M.C., Yang, Z.L., and Zhang, H.B. (1998). The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor. Appl. Genet. **97:** 657–670.

Faris, J.D., Simons, K.J., Zhang, Z., and Gill, B.S. (2005). The wheat super domestication gene *Q*. Wheat Inf. Serv. **100:** 129–148.

Gao, X., et al. (2019). Rice qGL3/OsPPKL1 functions with the GSK3/SHAGGY-like kinase OsGSK3 to modulate brassinosteroid signaling. Plant Cell **31:** 1077–1093.

Guo, W., Yang, H., Liu, Y., Gao, Y., Ni, Z., Peng, H., Xin, M., Hu, Z., Sun, Q., and Yao, Y. (2015). The wheat transcription factor *TaGAMyb* recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. Plant J. **84:** 347–359.

Han, Y., Xin, M., Huang, K., Xu, Y., Liu, Z., Hu, Z., Yao, Y., Peng, H., Ni, Z., and Sun, Q. (2016). Altered expression of *TaRSL4* gene by genome interplay shapes root hair length in allopolyploid wheat. New Phytol. **209:** 721–732.

Hosono, S. (1954). Classification and distribution of wheat. In Studies of Wheat, H. Kihara, ed (Tokyo: Yokendo).

Hu, J., Wang, Y., Fang, Y., Zeng, L., Xu, J., Yu, H., Shi, Z., Pan, J., Zhang, D., Kang, S., Zhu, L., and Dong, G., et al. (2015). A rare allele of *GS2* enhances grain size and grain yield in rice. Mol. Plant **8:** 1455–1465.

Huang, H.Y., Jiang, W.B., Hu, Y.W., Wu, P., Zhu, J.Y., Liang, W.Q., Wang, Z.Y., and Lin, W.H. (2013). BR signal influences Arabidopsis ovule and seed number through regulating related genes expression by *BZR1*. Mol. Plant **6:** 456–469.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., and Gornicki, P. (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA **99:** 8133–8138.

Jiang, W.B., Huang, H.Y., Hu, Y.W., Zhu, S.W., Wang, Z.Y., and Lin, W.H. (2013). Brassinosteroid regulates seed size and shape in Arabidopsis. Plant Physiol. **162:** 1965–1977.

Josekutty, P.C. (2008). Defining the Genetic and Physiological Basis of *Triticum sphaerococcum* Perc.. (New Zealand: University of Canterbury).

Kihara, H. (1944). Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric. Hortic. **19:** 889–890.

Koba, T., and Tsunewaki, K. (1978). Mapping of the *s* and *Ch2* genes on chromosome 3D of common wheat. Wheat Inf. Serv. **45:** 18–20.

Koh, S., Lee, S.C., Kim, M.K., Koh, J.H., Lee, S., An, G., Choe, S., and Kim, S.R. (2007). T-DNA tagged knockout mutation of rice *OsGSK1*, an orthologue of Arabidopsis *BIN2*, with enhanced tolerance to various abiotic stresses. Plant Mol. Biol. **65:** 453–466.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:** 1754–1760.

Li, J., and Nam, K.H. (2002). Regulation of brassinosteroid signaling by a GSK3/SHAGGY-like kinase. Science **295:** 1299–1301.

Mac Key, J. (1954). Neutron and X-ray experiments in wheat and a revision of the speltoid problem. Hereditas **40:** 65–180.

McFadden, E.S., and Sears, E.R. (1946). The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. **37:** 81–107, 107.

Miczynski, K., Jr. (1930). O dziedziczeniu sie niektorych cech u pszenioy w krzyzowkach Triticum pyramidale × T. durum i T. vulgare × T. sphaerococoum. Roczn. Nauk. Rol. **23:** 27–62.

Mori, N., Ohta, S., Chiba, H., Takagi, T., Niimi, Y., Shinde, V., Kajale, M.D., and Osada, T. (2013). Rediscovery of Indian dwarf wheat (*Triticum aestivum* L. ssp *sphaerococcum* (Perc.) MK.) an ancient crop of the Indian subcontinent. Genet. Resour. Crop Evol. **60:** 1771–1775.

Paolacci, A.R., Tanzarella, O.A., Porceddu, E., and Ciaffi, M. (2009). Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol. Biol. **10:** 11.

Peng, P., Yan, Z., Zhu, Y., and Li, J. (2008). Regulation of the Arabidopsis GSK3-like kinase BRASSINOSTEROID-INSENSITIVE 2 through proteasome-mediated protein degradation. Mol. Plant **1:** 338–346.

Percival, J. (1921). The Wheat Plant, a Monograph. (London: Duckworth and Co).

Pont, C., and Salse, J. (2017). Wheat paleohistory created asymmetrical genomic evolution. Curr. Opin. Plant Biol. **36:** 29–37.

Rao, M.P. (1972). Mapping of the compactum gene *C* on chromosome 2D of wheat. Wheat Inf. Serv. **35:** 9.

Rao, M.P. (1977). Mapping of the sphaerococcum gene '*S*' on chromosome 3D of wheat. Cereal Res. Commun. **5:** 15–17.

Salina, E., Borner, A., Leonova, I., Korzun, V., Laikova, L., Maystrenko, O., and Roder, M.S. (2000). Microsatellite mapping of the induced sphaerococcoid mutation genes in *Triticum aestivum*. Theor. Appl. Genet. **100:** 686–689.

Schmidt, J.W., and Johnson, V.A. (1963). A sphaerococcum-like tetraploid wheat. Crop Sci. **3:** 98–99.

Schmidt, J.W., and Johnson, V.A. (1966). Inheritance of the sphaerococcum effect in tetraploid wheat. Wheat Inf. Serv. **22:** 5–6.

Schmidt, J.W., Weibel, D.E., and Johnson, V.A. (1963). Inheritance of an incompletely dominant character in common wheat simulating *Triticum sphaerococcum* L. Crop Sci. **3:** 261–264.

Sears, E.R. (1947). The sphaerococcum gene in wheat. Genetics **32:** 102–103.

Singh, D. (1987). Mapping of the complex *T. sphaerococcum* locus. Wheat Inf. Serv. **64:** 17–20.

Singh, R. (1946). *Triticum sphaerococcum* Perc. (Indian dwarf wheat). Indian J. Genet. **6:** 34–37.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. **30:** 2725–2729.

Tian, P., et al. (2019). *GW5-Like*, a homolog of *GW5*, negatively regulates grain width, weight and salt resistance in rice. J. Integr. Plant Biol. **61:** 1171–1185.

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

Wang, H., et al. (2018). Three genomes differentially contribute to the seedling lateral root number in allohexaploid wheat: Evidence from phenotype evolution and gene expression. Plant J. **95:** 976–987.

Wu, Y., Fu, Y., Zhao, S., Gu, P., Zhu, Z., Sun, C., and Tan, L. (2016). *CLUSTERED PRIMARY BRANCH 1*, a new allele of *DWARF11*, controls panicle architecture and seed size in rice. Plant Biotechnol. J. **14:** 377–386.

Zhang, N., et al. (2017). The E3 ligase TaSAP5 alters drought stress responses by promoting the degradation of DRIP proteins. Plant Physiol. **175:** 1878–1892.

Zhu, X., Liang, W., Cui, X., Chen, M., Yin, C., Luo, Z., Zhu, J., Lucas, W.J., Wang, Z., and Zhang, D. (2015). Brassinosteroids promote development of rice pollen grains and seeds by triggering expression of Carbon Starved Anther, a MYB domain protein. Plant J. **82:** 570–581.