Supporting Information for

Rice Big Grain 1 promotes cell division to enhance organ development, stress tolerance and grain yield

Shuen-Fang Lo[1,2], Ming-Lung Cheng[1,3], Yue-Ie Hsing[4], Yi-Shih Chen[1], Kuo-Wei Lee[1],Ya-Fang Hong[4], Yu Hsiao[1], An-Shan Hsiao[4], Pei-Jing Chen[1,2], Lai-In Wong[1,2], Nan-Chen Chen[1,4], Christophe Reuzeau[5]*, Tuan-Hua David Ho[2,3,4,6]*, Su-May Yu[1,2,3,6]*

[1] Institute of Molecular Biology, Academia Sinica, Nankang, Taipei 115, Taiwan, ROC

[2] Agricultural Biotechnology Center, National Chung Hsing University, Taichung 402, Taiwan, ROC

[3] Department of Life Sciences, National Cheng Kung University, Tainan 701, Taiwan, ROC

[4] Institute of Plant and Microbial Biology, Academia Sinica, Nankang, Taipei 115, Taiwan, ROC

[5] CropDesign N.V., BASF Plant Science, Kerrebroek 66b, B-9850 Nevele, Belgium

[6] Department of Life Sciences, National Chung Hsing University, Taichung 402, Taiwan, ROC

Running title: RBG1 enhances organ size and stress recovery

Key words: rice, grain size, cell division, auxin, microtubule, root, abiotic stress

* Co-corresponding authors:

Su-May Yu
Phone: 886-2-2788-2695
FAX: 886-2-2782-6085
e-mail : sumay@imb.sinica.edu.tw

Christophe Reuzeau
Phone : +32-92423-463
FAX :+ 32-9-388 9600
e-mail : christophe.reuzeau@cropdesign.com

Tuan-Hua David Ho
Phone: 886-2-2787-1188
FAX: 886-2-2782-1605
E-mail: tho@gate.sinica.edu.tw; ho@biology2.wustl.edu

Supporting Text

RBG1 encodes a novel protein with six conserved motifs

Amino acid alignment revealed that six highly-conserved motifs are shared among most members of the RBG1 and RBG1-L clusters (Figure S7). RBG1-L contains five conserved motifs (motifs 1 to 5) that are similar to those in RBG1 even though the overall sequence similarity to RBG1 is only 30%, mostly due to the aforementioned divergence in inter-motif sequences (Figure S7).

To understand the function of the six conserved motifs in regulating grain size, we individually deleted each of the six motifs from RBG1 and the mutants were overexpressed in transgenic rice. We compared RBG1 variant mRNA levels and seed lengths (Figure S8b) and found that the motif 3-deleted RBG1-overexpressing line (referred to as RBG1-ΔD3) exhibited a similar seed length to the full-length WT RBG1-overexpressing line (RBG1-Ox-5). The seed lengths of the remaining truncated RBG1-overexpressing lines, i.e., RBG1-△D1, RBG1-△D2, RBG1-△D4, RBG1-△D5, and RBG1-△D6, were all shorter than those of full-length RBG1-Ox-5. These results indicate that all of these domains (i.e., with the exception of domain 3) are crucial for RBG1 regulation of seed size. Notably, RBG1 transcript levels appeared to be lower than that of WT in the RBG1-ΔD4 and RBG1-ΔD5 overexpressing lines, suggesting that motifs 4 and 5 could be essential for the stability of RBG1 transcripts.

RBG1 is a unique protein with six conserved motifs **separated by stretches of amino acid homopolymers**

RBG1 homologs have been found in both monocots and dicots, but not in other plant groups, indicating that these genes are specific to angiosperms and that they probably arose in a recent evolutionary event. The protein structure of RBG1 is unusual. It has six conserved domains that are shared among its homologs (Figure S7), but these motifs reveal no significant similarity with sequences of known functions. Apart from motif 3, all these motifs appear to be essential for the function of RBG1, as deletions of any of them failed to generate the big grain phenotype (Figure S8b). The six conserved domains are also present in dicot RBG1 homologs, but overexpression of *RBG1* in Arabidopsis did not lead to the big grain phenotype (Figure S9d-f). Although we have revealed substantial information about the role of *RBG1* in rice, the function of *RBG1* homologs in dicots is virtually unknown. More work in dicots or other species would be needed to support the results obtained in rice.

QTL associates with RBG1

We examined the QTL Annotation Rice Online (Q-TARO) database (http://qtaro.abr.affrc.go.jp/) to assess the QTL information associated with the chromosome region where *RBG1* is mapped to (Yonemaru et al., 2010). Intriguingly, *RBG1* is located within QTLs for regulating seed length, soil stress tolerance, and drought stress tolerance (Figure S12 and Table S5) (Moncada et al., 2001; Qi et al., 2008; Yoshida et al., 2002). The well-known QTLs related to grain shape are *GW2*, *GS5* and *GW5*/*qSW5* located on chromosomes 2, 3, and 5, respectively (Lu et al., 2013). A QTL on chromosome 11 has been suggested to have the largest impact on grain length (Yoshida et al., 2002), but the identity of the target gene is unknown. We suggest that *RBG1* is likely the candidate gene in this QTL on chromosome 11. Moreover, co-localization with stress QTLs revealed that *RBG1* may regulate not only grain length but also abiotic stress tolerance (Table S5).

Supporting methods

Construction of *RBG1*-truncated mutants

To generate *RBG1*-truncated mutants, each comprising a deletion of one of the six conserved motifs, full-length *RBG1* cDNA in the pENTR/MCS vector was employed as template for inverse PCR, using primers flanking conserved regions. KOD DNA polymerase (TOYOBO) was used to generate *RBG1* mutants lacking either motif 1, 2, 3, 4 or 5, whereas VioTaq DNA polymerase was employed to generate the motif 6-lacking *RBG1* mutant. PCR products were heated at 70 ºC for 5 minutes, 5'end-phosphorylated by T4 polynucleotide kinase, and then self-ligated. Resulting plasmids were transferred to XL1-blue *E. coli* competent cells. Correct DNA sequences of *RBG1*-truncated mutants were confirmed by nucleotide sequencing of genomic DNA isolated from *E. coli* and transgenic rice.

RT-PCR and qRT-PCR analysis

Total RNA was extracted from rice tissues, and RT-PCR and qRT-PCR analyses were conducted as described (Lo et al., 2008). For the RT-PCR analysis of *RBG1*, 5% DMSO was included in the RT-PCR reaction solution.

Plasmid constructions

We generated the various plasmids for gene overexpression by constructing a Gateway (Invitrogen)-based destination vector, including *Ubi*:attR1-Cm-ccdB-attR2-*Nos*, in pZP200 vector containing the chloramphenicol (Cm)-resistance gene (bacteria selection marker), a recombination selection marker (ccdB) and an LR Clonase recombination site (attR1 and attR2) between the *Ubi* promoter and the transcription terminator (*Nos*). *RBG1* cDNA was amplified from mRNA isolated from developing rice panicles and cloned to the Gateway entry vector pENTR/MCS. After sequence confirmation, *RBG1* cDNA was subcloned into the destination binary vector by LR Clonase to obtain pZP200 (*Ubi:RBG1-Nos*). For RNAi constructs, a destination vector pZP200 (*Ubi*:attR12-RNAi-*Nos*), which has a partial *GUS* cDNA flanked by attR in both the sense and antisense orientations, was constructed. The pZP200 (*Ubi:RBG1*-RNAi-*Nos*) plasmid was obtained by recombination with pENTR (*RBG1*) using LR Clonase (Invitrogen).

For studying temporal and spatial activity of the *RBG1* promoter, a 1.6 kb DNA fragment containing sequences upstream of the translation initiation codon ATG and the promoter region of *RBG1* was inserted into the pENTR/MCS vector to generate pENTR(Pro*RBG1*). LR Clonase was used to transfer the *RBG1* promoter into the pHGWFS7.0 binary vector that contains the coding

region of *GUS* protein (Karimi et al., 2005), generating the construct *RBG1:GUS*.

For subcellular localization studies, *RBG1* cDNA was first inserted into the pENTR (*mOrange2*-gene) or pENTR (*eGFP*-gene) vector, generating the pENTR (*mOrange2-RBG1*) or pENTR(*eGFP-RBG1*) plasmids. *mOrange2-RBG1* or *eGFP-RBG1* was subcloned into pBS (*CaMV35S*:attR12-Nos) to generate the pBS (*CaMV35S:mOrange2-RBG1-Nos*) or pBS (*CaMV35S:eGFP-RBG1-Nos*) plasmids using LR Clonase.

Rice transformation

All rice cultivars were transformed using the Agrobacterium method as described (Hiei et al., 1994).

GUS activity staining

GUS activity staining was conducted as described (Chen et al., 2015).

Scanning electron microscopy

The abaxial epidermis cells of hull lemmae were captured by scanning electron microscopy. The numbers of protuberances in the longitudinal direction were counted in 700 μm length intervals, which were then used to calculate the deduced cell number per seed length. The means ± SD represent an average deduced cell number per mm or per seed length.

Immunofluorescence staining of microtubules

Four-day-old seedlings were used for microtubule assays following the method described by Deng et al. (2015) (Deng et al., 2015). Approximately 1-cm segments of root tips were cut and fixed in 4% paraformaldehyde in PME buffer 1 (50 mM PIPES, 2 mM $MgSO_4$, 2 mM EGTA, pH 6.9) containing 0.05% Triton X-100 for 30 min. After washing thoroughly with PME buffer 1, samples were digested with 2% cellulase R-10 (Yakult Pharmaceutical Industry) and 1% pectolyase Y-23 (Yakult Pharmaceutical Industry) in PME buffer 1 at 37 °C for 30 min. The softened root tips were washed gently with PME buffer 1 and frozen and thawed twice. Samples were blocked with blocking buffer [3% BSA in PBST (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, 0.05% Triton X-100)] for 1 h at room temperature. Samples were incubated with 1/50 diluted primary antibody anti-α-tubulin (T9026, Sigma) in blocking buffer at 4 °C overnight. After washing with PBST five times, samples were incubated with 1/800 diluted secondary antibody AlexaFluor488 goat anti-mouse IgG (A-11001, Invitrogen) in PBST at 37 °C for 3 h. After washing four times with PBST and once with PBS, samples were mounted with PBS containing 50% glycerol and 0.1% *o*-phenylenediamine.

Images were obtained with a 63 X water objective using a Zeiss LSM 880 equipped with argon and HeNe lasers as excitation sources. Fluorescence was excited at 488 nm and collected with a 492-560 nm filter. Z-series sections were obtained at 0.45-μm intervals and the images of maximum intensity projection were from 10 optical slices.

Rice cell culture

De-hulled rice seeds were surface sterilized and placed on an MS medium plate containing 2 mg/liter of 2,4-D for callus induction. Calli were subcultured in 50 ml of MS medium containing 3% sucrose and 2 mg/liter of 2,4-D with shaking at 90 rpm, and subcultured weekly.

Microarray analysis

Total RNA was purified from coleoptiles and embryos of seedlings at 2 DAI, using Trizol® reagent (Invitrogen), and further purified with an RNeasy Mini Kit (QIAGEN). RNA quality assessment and array experiments were conducted in the Affymetrix Gene Expression Service Lab of Academia Sinica, using the GeneChip® Rice Genome Array (Affymetrix). The Affymetrix CEL files were imported into GeneSpring 12.6 software (Agilent Technologies) for data normalization, and genes with a signal ratio greater than a 3-fold change were used for function classification as described (Lo et al., 2017).

Abiotic stress treatment

Seeds were germinated and seedlings were cultivated in 0.5 X Kimura solution for 21 days. Seedlings were incubated in a 42 ℃ incubator for 4 days, in 250 mM NaCl solution for 4 days, or in 30% PEG 6000 at 28 ˚C for 18 h. Salt- and PEG-treated plants were washed thoroughly with water, and all stressed plants were recovered in water for 7 days in a 28 ℃ incubator, and survival rates were determined.

Table S1. List of RBG1 and RBG1-L homologous proteins.

Plants	Accession number
RBG1 homologous proteins	
Arabidopsis lyrata	XP_002867785
Arabidopsis thaliana	NP_567650.2
Brachypodium distachyon	XP_003576041
Cicer arietinum	XP_004486194
Capsella rubella	EOA16953
Cucumis sativus	XP_004134782
Fragaria vesca	XP_004295297
Glycine max(1)	*XP_006584450*
Glycine max(2)	*KRH11496*
Medicago truncatula	XP_003594150
Ricinus communis	XP_002518417
Solanum lycopersicum	XP_004232680
Sorghum bicolor	XP_002449526
Setaria italica	XP_004979307
Theobroma cacao	EOY28817
Vitis vinifera	XP_002270044
Zea mays(1)	XP_008678123.1
Zea mays(2)	XP_008670732.1
RBG1-L homologous proteins	
*Brachypodium distachyon*_L	XP_014754150.1
*Nicotiana tomentosiformis*_L	XP_009628473.1
*Phaseolus vulgaris*_L	XP_007153991.1
*Setaria italica*_L	XP_004971090.1
*Solanum tuberosum*_L	XP_006348166.1
*Sorghum bicolor*_L	EES04120.1
*Triticum aestivum*_L	CDM85783.1
*Vigna angularis*_L	BAT77345.1
*Zea mays*_L	NP_001145328.1

Table S2. Overexpression of *RBG1* elevates the expression of stress tolerance-related genes.

Classification	MSU locus	log_2 FC RBG1ox / WT	Description
Abiotic stress			
Response to heat	LOC_Os03g53340	5.836	heat shock factor protein HSF8, putative, expressed
	LOC_Os11g13980	5.695	22.0 kDa class IV heat shock protein precursor, putative, expressed
	LOC_Os06g09560	4.004	dnaJ protein, putative, expressed
	LOC_Os02g54140	3.698	17.5 kDa class II heat shock protein, putative, expressed
	LOC_Os05g44340	3.693	heat shock protein 101, putative, expressed
	LOC_Os04g36750	3.670	22.0 kDa class IV heat shock protein precursor, putative, expressed
	LOC_Os02g52150	3.569	heat shock 22 kDa protein, mitochondrial precursor, putative, expressed
	LOC_Os12g38180	3.526	heat shock cognate 70 kDa protein 2, putative, expressed
	LOC_Os03g53340	3.485	heat shock factor protein HSF8, putative, expressed
	LOC_Os03g53340	3.480	heat shock factor protein HSF8, putative, expressed
	LOC_Os03g16030	3.274	17.4 kDa class I heat shock protein 3, putative, expressed
	LOC_Os03g16020	3.117	17.4 kDa class I heat shock protein 2, putative, expressed
	LOC_Os04g01740	3.083	heat shock protein 82, putative, expressed
	LOC_Os04g01740	2.958	heat shock protein 82, putative, expressed
	LOC_Os04g45480	2.871	heat shock protein STI, putative, expressed
	LOC_Os03g16030	2.840	17.4 kDa class I heat shock protein 3, putative, expressed
	LOC_Os04g01740	2.823	heat shock protein 82, putative, expressed
	LOC_Os09g35790	2.733	heat shock factor protein 7, putative, expressed
	LOC_Os01g04360	2.597	16.9 kDa class I heat shock protein 3, putative, expressed
	LOC_Os02g54140	2.560	17.5 kDa class II heat shock protein, putative, expressed
	LOC_Os10g28340	2.518	heat shock factor protein HSF30, putative, expressed
	LOC_Os04g01740	2.466	heat shock protein 82, putative, expressed
	LOC_Os04g01740	2.464	heat shock protein 82, putative, expressed
	LOC_Os04g01740	2.429	heat shock protein 82, putative, expressed
	LOC_Os06g08154	2.342	receptor-like protein kinase 5 precursor, putative, expressed
	LOC_Os06g14240	2.224	17.4 kDa class I heat shock protein 3, putative, expressed
	LOC_Os02g54140	2.215	17.5 kDa class II heat shock protein, putative, expressed
	LOC_Os01g42190	2.149	dnaJ-like protein, putative, expressed
	LOC_Os01g04380	2.113	16.9 kDa class I heat shock protein 2, putative, expressed
	LOC_Os05g35400	2.083	luminal-binding protein 5 precursor, putative, expressed
	LOC_Os03g16920	2.025	heat shock cognate 70 kDa protein, putative, expressed
	LOC_Os03g15960	1.993	17.4 kDa class I heat shock protein 3, putative, expressed
	LOC_Os05g38530	1.804	heat shock cognate 70 kDa protein, putative, expressed
	LOC_Os06g46900	1.727	phosphosulfolactate synthase-related protein, putative, expressed
	LOC_Os02g08490	1.717	chaperone clpB 1, putative, expressed
	LOC_Os06g35960	-1.716	heat shock factor protein 3, putative, expressed
Receptor protein kinases	LOC_Os05g06240	2.351	receptor protein kinase TMK1 precursor, putative
	LOC_Os02g34490	1.832	Leucine Rich Repeat family protein, expressed
Others	LOC_Os05g28740	3.834	universal stress protein, putative, expressed
	LOC_Os05g03130	1.757	expressed protein
	LOC_Os10g30150	1.734	ethylene-responsive protein, putative, expressed

Table S3. Effect of *RBG1* overexpression driven by the *GOS2* promoter on agronomic traits in *japonica* rice (J) and in *indica* rice (I). TraitMill greenhouse experiments were performed in 2012 (I) and in 2014 (J). Percentage change over WT is shown for each trait together with p-values from statistical analysis.

Constructs	**Biomass**	**Height**	**Thousand Kernel Weight**	**Fill rate**	**Seed yield per plant**	**Harvest Index**
pGOS2 ::RBG1-016A (J)	28.1	16.1	13.8	20.6	76.4	40.7
p-value	0.0158	0.0002	0	0.0003	0.0001	0
pGOS2 ::RBG1-041A (J)	3.2	7	6.6	14.2	29.6	25.1
p-value	0.6301	0.0207	0.0012	0.0018	0.0815	0.0218
pGOS2 ::RBG1-049A (J)	7.7	3.3	10	14.3	45.3	39.2
p-value	0.3831	0.2323	0	0.0004	0.0333	0.0169
pGOS2 ::RBG1-51A (J)	14.2	9.8	8.3	21	50.2	25.4
p-value	0.2671	0.0189	0.0017	0.0006	0.0294	0.0243
pGOS2 ::RBG1-205A (I)	16.8	8.1	11.2	26.1	74.8	36
p-value	0.072	0.2643	0.0888	0.3654	0.2536	0.4663
pGOS2 ::RBG1-002A (I)	8.8	8.1	0.7	65.6	193.6	161.2
p-value	0.5328	0.3584	0.829	0.2439	0.1101	0.128
pGOS2 ::RBG1-123A (I)	12.6	5.9	25.4	8.9	177.7	160.8
p-value	0.0836	0.3498	0.0001	0.8494	0.2195	0.114

Table S4. List of primers.

Primer	Sequence	Gene and purpose	Reference
Genotyping of T-DNA insertion mutants			
M35973-GT-F1	TGGGCATAGACCCAGAGTAACACA	*RBG1*$_{Act}$ genotyping (830 bp)	
M35973-GT-R1	CAGCCCGCGTTGTTTTC		
RB	AACTCATGGCGATCTCTTACC	TRIM mutant genotyping	
***RBG1* RT-PCR and qRT-PCR**			
RBG1-RT -Ac-F1	CCGTCGCGGATGATGATGAT	*RBG1* (555 bp)	
RBG1-RT -Ac-R1	TTGGCTTGTCCTTGTGGTGC		
RBG1-q-PCR-F2	AACCGGCGGCGCCAAGAA	*RBG1* qRT-PCR (96 bp)	
RBG1-q-PCR-R2	CGGTACCCGATGCAGGAGTTGAT		
30430-RT -Ac-F	GAAGCAGTCTCCACGCCGA	*RBG1* cDNA (139 bp)	
30430-RT -Ac-R	CAGCCATCATCATCATCA		
Actin 1-5'	CTGATGGACAGGTTATCACC	*Actin 1* (567 bp)	
Actin 1-3'	CAGGTAGCAATAGGTATTACAG		
18S rRNA-5'	CAACTTTCGATGGTAGGATAGGG	*18s rRNA* (197 bp)	
18S rRNA-3'	CCAATTACCAGACACTAAAGCGC		
Ubi-5'	CAATGGAGCTATGGTTGTCTGGT	Rice *Ubiquitin* (204 bp)	
Ubi-3'	CAGCACAAAAGGGTATAGCAGAA		
RBG1 domain truncation			
RBG1-dom1F	CTCGACGCCGACCCCGCC	RBG1 dom1 truncated	
RBG1-dom1R	GGCGCTGCTGCTCCGCCG		
RBG1-dom2F	CTCCAACCCCTCCCCTACC	RBG1 dom2 truncated	
RBG1-dom2R	GGAGGGGGATGCGGCGGG		
RBG1-dom3F	TCCCGGAGCCGCTCCGCC	RBG1 dom3 truncated	
RBG1-dom3R	CGATCCGCCTCCGCCTCC		
RBG1-dom4F5	CCGGGCGAGGCGGCGGCGGTGGCGT	RBG1 dom4 truncated	
RBG1-dom4R5	TCGCCCGGCGGCGGGCGGGGCTTC		
RBG1-dom5F	CCGCCATCGCCGGCGGCG	RBG1 dom5 truncated	
RBG1-dom5R	CGGCACGCCGCCGCCTGC		
RBG1-dom6F2	GGGTACCGCCACCAGGTGA	RBG1 dom6 truncated	
RBG1-dom6R2	CAGGTTGATGCCGCGTATCC		
Cloning of native *RBG1* promoter			
30430-PG-F3	GCACGGCTCGGCAGTTTATT	*RBG1* promoter (1623 bp)	
30430-PG-R	CGGATCGGCGTGGAGACT		

RT-PCR of grain size-related genes

BG1-RT-F	CATCCACGCTGCTCGACGC	*BG1* RT-PCR (576 bp)	(Liu et al., 2015)
BG1-RT-R	CGCTGTCCAAGAACCGCACG		
DEP2-F	TGCGTGATAGCCTAGAACGAAG	*DEP2* RT-PCR (23 bp)	(Abe et al., 2010; Li et al., 2010)
DEP2-R	CTGGAATCAGCACTCCTGGATG		
GIF1-F	CATCGCGCAACCCGAACATG	*GIF1* RT-PCR (143 bp)	(Wang et al., 2008)
GIF1-R	CAGGATGCACGCCTTTCCTC		
GL3-F	TCACAACTCCCAGGATAGG	*GL3* RT-PCR (126 bp)	(Qi et al., 2012)
GL3-R	TTTGTCTCGCTCGCTCAT		
GS3-F	CATCGGAGAAGCGAAGTCA	*GS3* RT-PCR (109 bp)	(Fan et al., 2006)
GS3-R	CAGCAGCAGATCCAGGAGA		
GS5-F	AGTGGACTGCTTCCAGGGAAG	*GS5* RT-PCR (180 bp)	(Li et al., 2011)
GS5-R	CACGCAGTACCGAGAACTGA		
GW2-F	CTGCAGCAGGGAAGTGGTAA	*GW2* RT-PCR (142 bp)	(Song et al., 2007)
GS2-R	GTTCTCGGCGAAAGGCAGC		
GW8-F	AGGAGTTTGATGAGGCCAAG	*GW8*-RT-PCR (204 bp)	(Wang et al., 2012)
GW8-R	GCGTGTAGTATGGGCTCTC		
RGA1-F	GCAGCAAGCCTGACCGTGTG	*RGA1* RT-PCR (120 bp)	(Seo et al., 1995)
RGA1-R	CTTCCCTGGAGCGTCTCATGC		

Table S5. Detailed information on the QTLs on chromosome 11 that regulate seed length and abiotic stresses in rice.

QTL entry				
Q-TARO ID.	154	940	904	905
QTL		qDLR11	gpl11.1	gw11.1
Major category	Morphological trait	Resistance or Tolerance	Resistance or Tolerance	Resistance or Tolerance
Category of object character	**Seed**	**Other soil stress tolerance**	**Drought tolerance**	**Drought tolerance**
Character	**grain length**	**Dead leaf rate at 20 days**	**Grains per plant**	**1000-grain weight**
Marker	SSR	SSR	RFLP	RFLP
Genome Start (bp)	2827521	383711	17246592	17246592
Genome End (bp)	20702000	21835946	23651853	23651853
Genome length (Mb)	17.87	21.45	6.41	6.41
Parent A	*Reiho*	*Gaochan 106*	*Caiapo*	*Caiapo*
Parent B	*Yamadanishiki*	*Changbai 9*	*O. rufipogon (IRGC 105491)*	*O. rufipogon (IRGC 105491)*
References	Yoshida et al. (2002)(Yoshida et al., 2002)	Qi et al. (2008)(Qi et al., 2008)	Moncada et al. (2001)(Moncada et al., 2001)	Moncada et al. (2001)(Moncada et al., 2001)

Supporting references

Abe, Y., Mieda, K., Ando, T., Kono, I., Yano, M., Kitano, H. and Iwasaki, Y. (2010) The SMALL AND ROUND SEED1 (SRS1/DEP2) gene is involved in the regulation of seed size in rice. *Genes Genet. Syst.* **85**, 327-339.

Chen, Y.S., Lo, S.F., Sun, P.K., Lu, C.A., Ho, T.H. and Yu, S.M. (2015) A late embryogenesis abundant protein HVA1 regulated by an inducible promoter enhances root growth and abiotic stress tolerance in rice without yield penalty. *Plant Biotechnol. J.* **13**, 105-116.

Deng, Z.Y., Liu, L.T., Li, T., Yan, S., Kuang, B.J., Huang, S.J., Yan, C.J. and Wang, T. (2015) OsKinesin-13A Is an Active Microtubule Depolymerase Involved in Glume Length Regulation via Affecting Cell Elongation. *Scientific reports* **5**, 9457.

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X. and Zhang, Q. (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.

Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (Oryza sativa L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**, 271-282.

Karimi, M., De Meyer, B. and Hilson, P. (2005) Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103-105.

Li, F., Liu, W., Tang, J., Chen, J., Tong, H., Hu, B., Li, C., Fang, J., Chen, M. and Chu, C. (2010) Rice DENSE AND ERECT PANICLE 2 is essential for determining panicle outgrowth and elongation. *Cell Res.* **20**, 838-849.

Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Sun, L., Shao, D., Xu, C., Li, X., Xiao, J., He, Y. and Zhang, Q. (2011) Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nat. Genet.* **43**, 1266-1269.

Liu, L., Tong, H., Xiao, Y., Che, R., Xu, F., Hu, B., Liang, C., Chu, J., Li, J. and Chu, C. (2015) Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. *Proc. Natl. Acad. Sci. USA* **112**, 11102-11107.

Lo, S.F., Ho, T.D., Liu, Y.L., Jiang, M.J., Hsieh, K.T., Chen, K.T., Yu, L.C., Lee, M.H., Chen, C.Y., Huang, T.P., Kojima, M., Sakakibara, H., Chen, L.J. and Yu, S.M. (2017) Ectopic expression of specific GA2 oxidase mutants promotes yield and stress tolerance in rice. *Plant Biotechnol J* **15**, 850-864.

Lo, S.F., Yang, S.Y., Chen, K.T., Hsing, Y.I., Zeevaart, J.A., Chen, L.J. and Yu, S.M. (2008) A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell* **20**, 2603-2618.

Lu, L., Shao, D., Qiu, X., Sun, L., Yan, W., Zhou, X., Yang, L., He, Y., Yu, S. and Xing, Y. (2013) Natural variation and artificial selection in four genes determine grain shape in rice. *New Phytol.* **200**, 1269-1280.

Moncada, P., Martínez, C.P., Borrero, J., Chatel, M., Gauch Jr, H., Guimaraes, E., Tohme, J. and McCouch, S.R. (2001) Quantitative trait loci for yield and yield components in an Oryza sativa×Oryza rufipogon BC2F2 population evaluated in an upland environment. *Theor. Appl. Genet.* **102**, 41-52.

Qi, D., Guo, G., Lee, M.-c., Zhang, J., Cao, G., Zhang, S., Suh, S.-c., Zhou, Q. and Han, L. (2008) Identification of quantitative trait loci for the dead leaf rate and the seedling dead rate under alkaline stress in rice. *J. of Genet. and Genomics* **35**, 299-305.

Qi, P., Lin, Y.S., Song, X.J., Shen, J.B., Huang, W., Shan, J.X., Zhu, M.Z., Jiang, L., 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.

Seo, H.S., Kim, H.Y., Jeong, J.Y., Lee, S.Y., Cho, M.J. and Bahk, J.D. (1995) Molecular cloning and characterization of RGA1 encoding a G protein alpha subunit from rice (Oryza sativa L. IR-36). *Plant Mol. Biol.* **27**, 1119-1131.

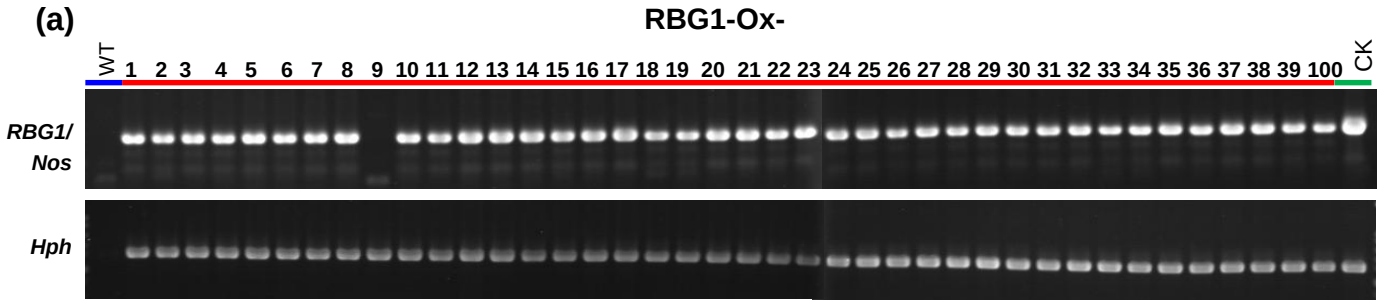
Song, X.J., Huang, W., Shi, M., Zhu, M.Z. and Lin, H.X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet* **39**, 623-630.

Wang, E., Wang, J., Zhu, X., Hao, W., Wang, L., Li, Q., Zhang, L., He, W., Lu, B., Lin, H., Ma, H., Zhang, G. and He, Z. (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat. Genet.* **40**, 1370-1374.

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., Zeng, R., Zhu, H., Dong, G., Qian, Q., Zhang, G. and Fu, X. (2012) Control of grain size, shape and quality by OsSPL16 in rice. *Nat. Genet.* **44**, 950-954.

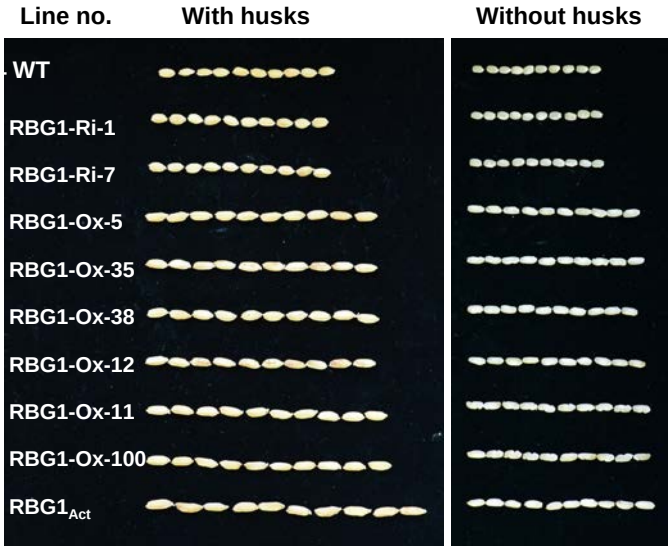
Yonemaru, J.-i., Yamamoto, T., Fukuoka, S., Uga, Y., Hori, K. and Yano, M. (2010) Q-TARO: QTL Annotation Rice Online Database. *Rice* **3**, 194-203.

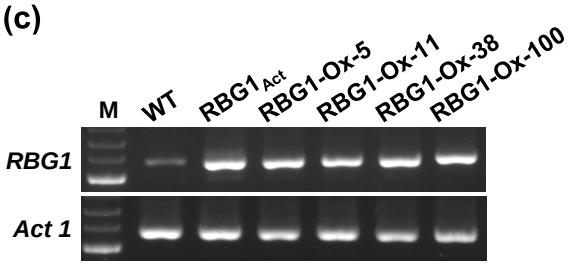
Yoshida, S., Ikegami, M., Kuze, J., Sawada, K., Hashimoto, Z., Ishii, T., Nakamura, C. and Kamijima, O. (2002) QTL Analysis for Plant and Grain Characters of Sake-brewing Rice Using a Doubled Haploid Population. *Breeding Sci.* **52**, 309-317.



(b)

Line no.	Grain length
WT	6.87 ± 0.04
$RBG1_{Act}$	10.05 ± 0.13
RBG1-Ox-5	9.17 ± 0.13
RBG1-Ox-10	7.56 ± 0.09
RBG1-Ox-11	9.63 ± 0.04
RBG1-Ox-12	8.73 ± 0.18
RBG1-Ox-23	7.59 ± 0.07
RBG1-Ox-34	8.13 ± 0.09
RBG1-Ox-35	9.27 ± 0.14
RBG1-Ox-38	8.63 ± 0.04
RBG1-Ox-100	9.75 ± 0.05





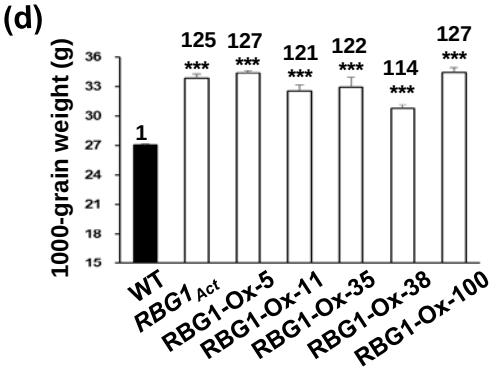


Figure S1. Large grain morphology was recapitulated by overexpression of *RBG1* gene. Transgenic rice overexpressing *RBG1* driven by the *Ubi* promoter. **(a)** Genomic DNA PCR analysis for the intact *RBG1* cDNA (confirmed by *RBG1/Nos* primers) and hygromycin resistance gene (*Hph*) in 39 transgenic rice lines (*RBG1-Ox-*). **(b)** Characterization of 9 independent lines showed overexpression of *RBG1* recapitulated the long grain morphology. **(c)** *RBG1* is highly expressed in transgenic plants relative to WT. (d) Overexpression of *RBG1* increases the 1000-grain weight.

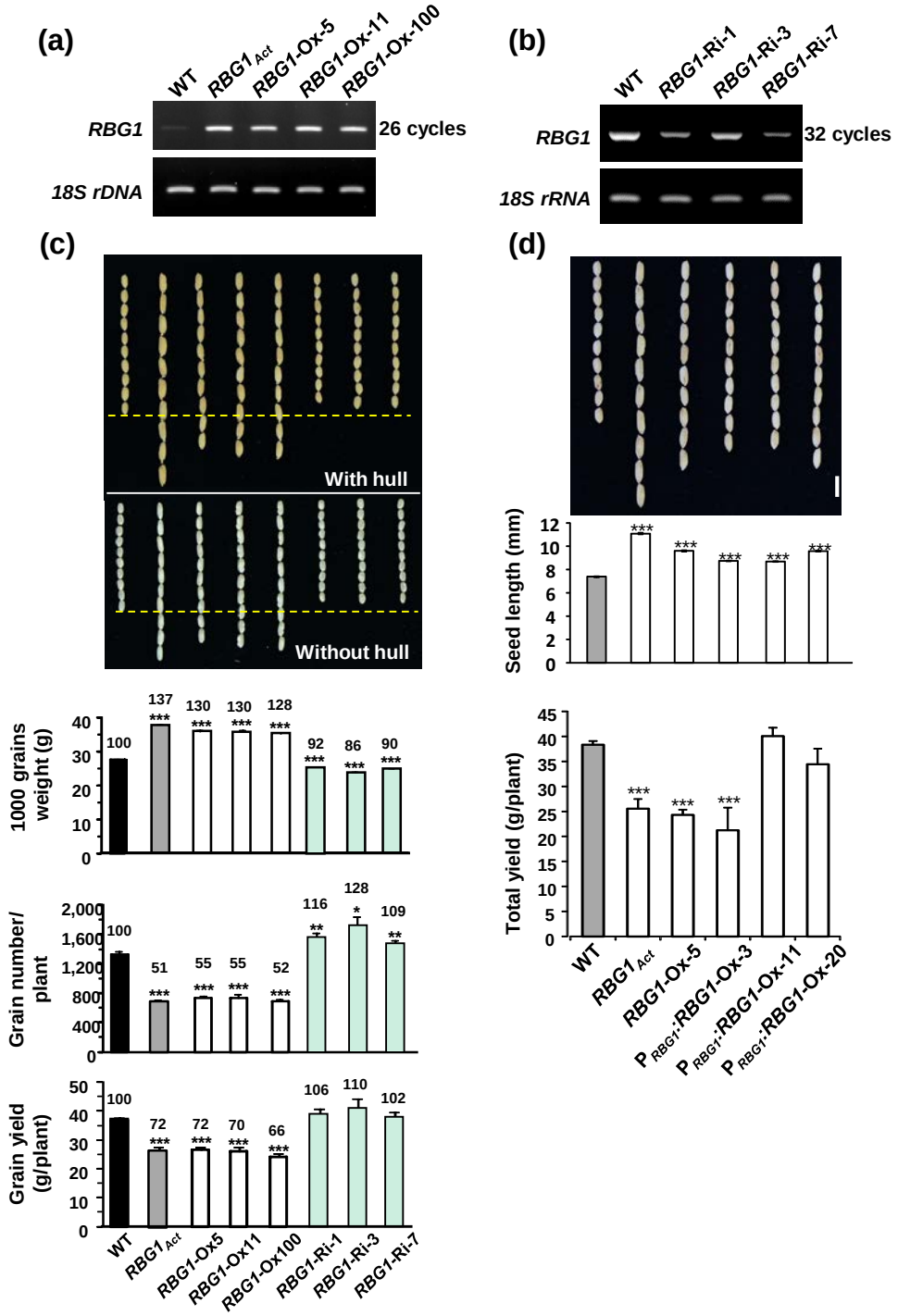


Figure S2. Ectopic expression of *RBG1* driven by the *Ubi* promoter enhances grain size but reduces grain number and yield in transgenic rice.

Total RNA was extracted from 3-5 cm long immature rice panicles and subjected to RT-PCR analyses using gene-specific primers (see Table S4). **(a)** Analysis of *RBG1* overexpression transgenic lines. RT-PCR of the *RBG1* gene was conducted with 26 cycles. **(b)** Analysis of *RBG1* RNAi knockdown transgenic lines. RT-PCR of the *RBG1* gene was conducted with 32 cycles. **(c)** *RBG1* expressed under the control of the *Ubi* promoter enhances grain size. The grain size, weight, number and yield in the $RBG1_{Act}$, *RBG1*-Ox and *RBG1*-Ri lines were compared with those of WT. Numbers above bars are % relative to the value in WT. n = 18 for each line. **(d)** *RBG1* expressed under the control of its native promoter also enhances grain size. Lengths of seeds from three transgenic rice lines carrying *RBG1:RBG1* were compared with those from $RBG1_{Act}$ or *RBG1*-Ox lines. n = 30. Scale bar = 1 cm.

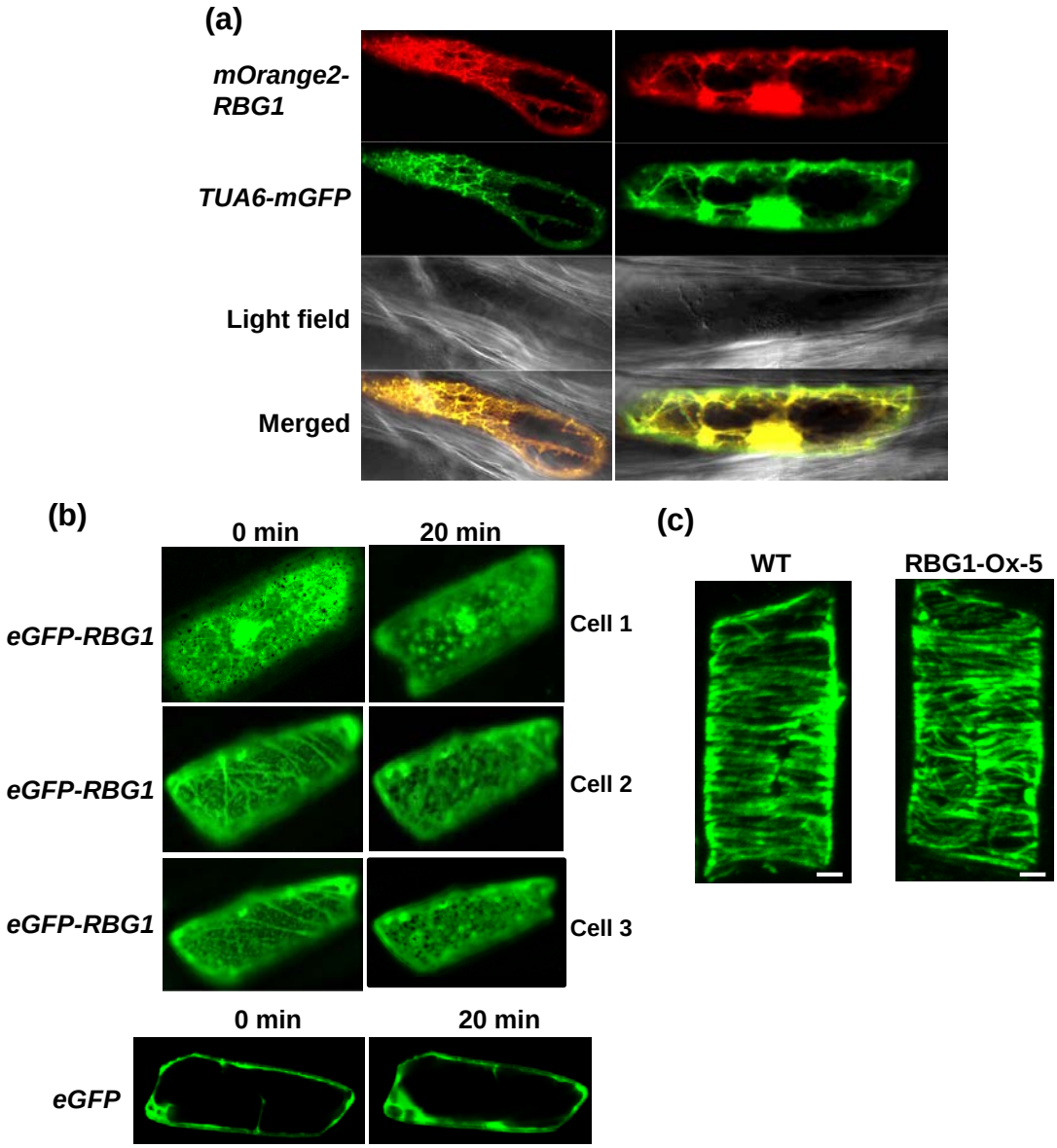


Figure S3. RBG1 is co-localized with microtubules.

Cells were transfected with indicated plasmid constructs by particle bombardment and examined under confocal microscopy. **(a)** Onion epidermal cells co-transfected with the constructs *CaMV35S:mOrange-RBG1* and *Ubi:TUA6-eGFP*. **(b)** Onion epidermal cells transfected with the constructs *CaMV35S:eGFP-RBG1* and *CaMV35S:eGFP* were treated with 2 μM oryzalin for 20 minutes. The filamentous structure of eGFP-RBG1, but not eGFP alone, was abolished after treatment. The images taken from the surface of three representative cells expressing eGFP-RBG1 are shown here. **(c)** Microtubules in root elongation zone cells of WT and *RBG1*-Ox plants at 4 DAI were subjected to immunofluorescence staining using the anti-α-tubulin mouse primary antibody and AlexaFluor488 goat anti-mouse secondary antibody, and visualized by confocal microscopy. The micrographs of microtubule staining are a maximum projection of the fluorescence signals. Scale bars = 2 μm.

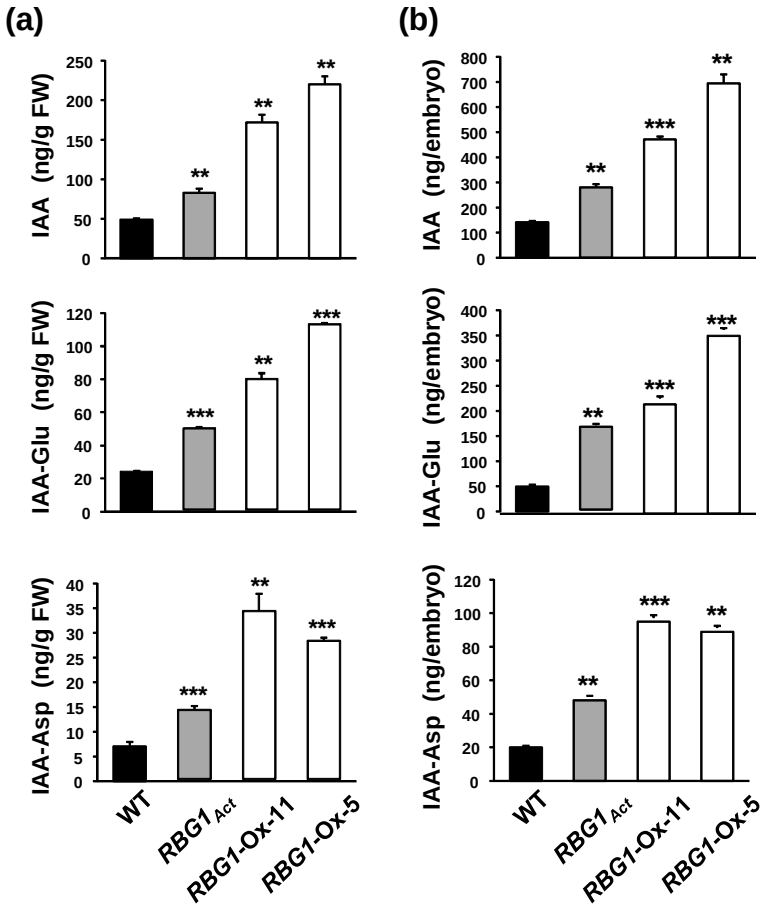


Figure S4. *RBG1* enhances the accumulation of endogenous auxin levels.
Coleoptiles and embryos of seedlings at 2 DAI were used for measurement of IAA, IAA-Glu and IAA-Asp levels. Concentrations are based on **(a)** fresh weight or **(b)** individual coleoptiles/embryos.

(a)

	Classification	Probe id	BG1-Ox5 vs WT	MSU locus	description
IAA	IAA-induced-regulated-responsive-activated	os.15798.1.s1_at	4.109	LOC_Os06g07040	OsIAA20 - Auxin-responsive Aux/IAA gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.23256.1.s1_at	3.904	LOC_Os02g56120	OsIAA9 - Auxin-responsive Aux/IAA gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.9945.1.s1_at	-3.837	LOC_Os03g43410	OsIAA12 - Auxin-responsive Aux/IAA gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.49290.1.s1_at	-4.319	LOC_Os06g39590	OsIAA23 - Auxin-responsive Aux/IAA gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.26512.1.s1_at	-5.636	LOC_Os03g58350	OsIAA14 - Auxin-responsive Aux/IAA gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.39652.1.s1_at	10.232	LOC_Os09g37330	OsSAUR39 - Auxin-responsive SAUR gene family member, expressed
	IAA-induced-regulated-responsive-activated	osaffx.14257.1.s1_x_at	3.611	LOC_Os04g43740	OsSAUR18 - Auxin-responsive SAUR gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.8639.1.s1_at	3.313	LOC_Os02g24700	OsSAUR8 - Auxin-responsive SAUR gene family member, expressed
	IAA-induced-regulated-responsive-activated	osaffx.30167.1.s1_at	-3.693	LOC_Os09g37380	OsSAUR42 - Auxin-responsive SAUR gene family member
	IAA-induced-regulated-responsive-activated	osaffx.2511.1.s1_at	-15.825	LOC_Os02g05060	OsSAUR5 - Auxin-responsive SAUR gene family member, expressed
	IAA-induced-regulated-responsive-activated	os.49607.1.s1_at	4.909	LOC_Os03g09880	AIR12, putative, expressed
	IAA-induced-regulated-responsive-activated	os.11266.1.s1_at	4.208	LOC_Os04g27060	auxin-induced protein PCNT115, putative, expressed
	IAA-induced-regulated-responsive-activated	os.37834.1.s1_a_at	3.737	LOC_Os04g58280	stem-specific protein TSJT1, putative, expressed
	IAA-induced-regulated-responsive-activated	os.11421.1.s1_at	3.605	LOC_Os03g58170	stem-specific protein TSJT1, putative, expressed
	IAA-signal transduction	os.23143.1.s1_at	3.118	LOC_Os09g31478	auxin hydrogen symporter, putative, expressed
	IAA-signal transduction	os.10149.1.s1_at	-3.127	LOC_Os04g32460	transport inhibitor response 1 protein, putative, expressed
	IAA-induced-regulated-responsive-activated	os.11585.1.s1_at	-3.726	LOC_Os09g27080	axi 1 like protein, putative, expressed
	IAA-induced-regulated-responsive-activated	os.25562.1.s1_at	-4.044	LOC_Os05g48270	dopamine beta-monooxygenase, putative, expressed
	IAA-induced-regulated-responsive-activated	osaffx.14314.1.s1_x_at	-6.273	LOC_Os04g47520	auxin-independent growth promoter, putative, expressed
	IAA-signal transduction	os.47814.1.a1_s_at	-7.265	LOC_Os09g38130	auxin Efflux Carrier family protein, expressed
	IAA-induced-regulated-responsive-activated	os.11469.1.s1_at	-8.548	LOC_Os04g26870	auxin-induced protein PCNT115, putative, expressed
	IAA-synthesis-degradation	os.10435.1.s1_at	-8.757	LOC_Os03g62060	IAA-amino acid hydrolase ILR1 precursor, putative, expressed
	IAA-induced-regulated-responsive-activated	osaffx.23807.1.s1_x_at	-16.143	LOC_Os01g56240	auxin responsive protein, expressed
	IAA-induced-regulated-responsive-activated	os.48933.1.s1_at	-74.971	LOC_Os05g37880	axi 1 like protein, putative, expressed

(b)

Classification	Probe id	BG1-Ox5 vs WT	MSU locus	description
Heat	os.50671.1.s1_x_at	34.515	LOC_Os12g38180	heat shock cognate 70 kDa protein 2, putative, expressed
	os.36328.1.s1_at	33.340	LOC_Os01g04340	16.9 kDa class I heat shock protein 3, putative, expressed
	os.55330.1.s1_at	28.489	LOC_Os11g13980	22.0 kDa class IV heat shock protein precursor, putative, expressed
	osaffx.12887.2.s1_at	26.273	LOC_Os03g16920	heat shock cognate 70 kDa protein, putative, expressed
	os.7537.1.s1_at	24.579	LOC_Os04g36750	22.0 kDa class IV heat shock protein precursor, putative, expressed
	os.57495.1.a1_x_at	18.978	LOC_Os11g47760	heat shock cognate 70 kDa protein 2, putative, expressed
	os.4775.1.s1_at	8.772	LOC_Os01g04380	16.9 kDa class I heat shock protein 2, putative, expressed
	os.22731.1.s1_at	8.409	LOC_Os03g16020	17.4 kDa class I heat shock protein 2, putative, expressed
	os.8926.1.s1_at	7.630	LOC_Os03g14180	small heat shock protein, chloroplast precursor, putative, expressed
	os.5817.1.s1_at	7.422	LOC_Os06g09560	dnaJ protein, putative, expressed
	os.12244.1.s1_at	7.082	LOC_Os03g15960	17.4 kDa class I heat shock protein 3, putative, expressed
	osaffx.10905.3.s1_at	6.605	LOC_Os01g04360	16.9 kDa class I heat shock protein 3, putative, expressed
	os.2292.2.s1_at	5.450	LOC_Os03g53340	heat shock factor protein HSF8, putative, expressed
	os.37000.2.s1_x_at	5.104	LOC_Os07g29794	calmodulin binding protein, putative, expressed
	os.5581.1.s1_at	5.039	LOC_Os03g56540	mitochondrial import inner membrane translocase subunit TIM14, putative, expressed
	os.37773.1.s1_x_at	4.959	LOC_Os03g16030	17.4 kDa class I heat shock protein 3, putative, expressed
	os.47474.1.s1_x_at	4.130	LOC_Os06g44160	dnaJ protein, putative, expressed
	os.8971.1.s1_at	3.727	LOC_Os05g44340	heat shock protein 101, putative, expressed
	os.3419.1.s1_a_at	3.697	LOC_Os01g50700	dehydrin family protein, expressed
	os.11039.3.s1_at	3.646	LOC_Os04g01740	heat shock protein 82, putative, expressed
	os.18397.1.s1_at	3.617	LOC_Os02g46640	chaperone protein dnaJ, putative, expressed
	os.50643.2.s1_x_at	3.516	LOC_Os07g44690	AT-HSFB4, putative, expressed
	os.10884.1.s1_at	3.437	LOC_Os03g58790	ATPase 3, putative, expressed
	osaffx.17944.1.s1_at	3.381	LOC_Os09g28200	AT-HSFB4, putative, expressed
	os.51749.1.s1_at	3.356	LOC_Os02g54140	17.5 kDa class II heat shock protein, putative, expressed
	os.12257.1.s1_at	3.348	LOC_Os02g52150	heat shock 22 kDa protein, mitochondrial precursor, putative, expressed
	os.27716.2.s1_a_at	3.180	LOC_Os08g43334	heat shock factor protein 7, putative, expressed
	os.55371.1.s1_at	3.070	LOC_Os02g48140	17.4 kDa class I heat shock protein 3, putative, expressed
	os.50191.1.s1_at	-3.038	LOC_Os05g51360	luminal-binding protein precursor, putative, expressed
	os.30962.1.s1_at	-3.078	LOC_Os02g35000	chaperone protein dnaJ 10, putative, expressed
	osaffx.31338.3.s1_at	-3.157	LOC_Os11g36960	dnaJ domain containing protein, expressed
	os.11822.1.s1_at	-3.931	LOC_Os03g60620	heat shock cognate 70 kDa protein 2, putative, expressed
	os.22676.1.s1_at	-4.356	LOC_Os02g53750	protein kinase APK1A, chloroplast precursor, putative, expressed

Figure S5. Heat maps indicate that the expression of auxin- and HSP-related genes are significantly up-regulated or down-regulated by *RBG*1.

Total RNAs were purified from coleoptiles and embryos of seedlings at 2 DAI and then subjected to microarray analysis.

Figure S6. Phylogenetic analysis of RBG1 homologous proteins in plants.

RBG1 homologs were identified by the BLASTP program from National Center for Biotechnology Information website (NCBI, http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple alignment of RBG1 and homologous proteins from different species was carried out by the ClustalW program of the MEGA5.2 software, and a phylogenetic tree was inferred using the neighbor-joining method. The scale value of 0.1 indicates 0.1 amino acid substitutions per site. The accession numbers of RBG1 and RBG1-like homologous proteins are listed in Table S1. No RBG1 homologs have been found in algae, bryophytes, ferns or gymnosperms.

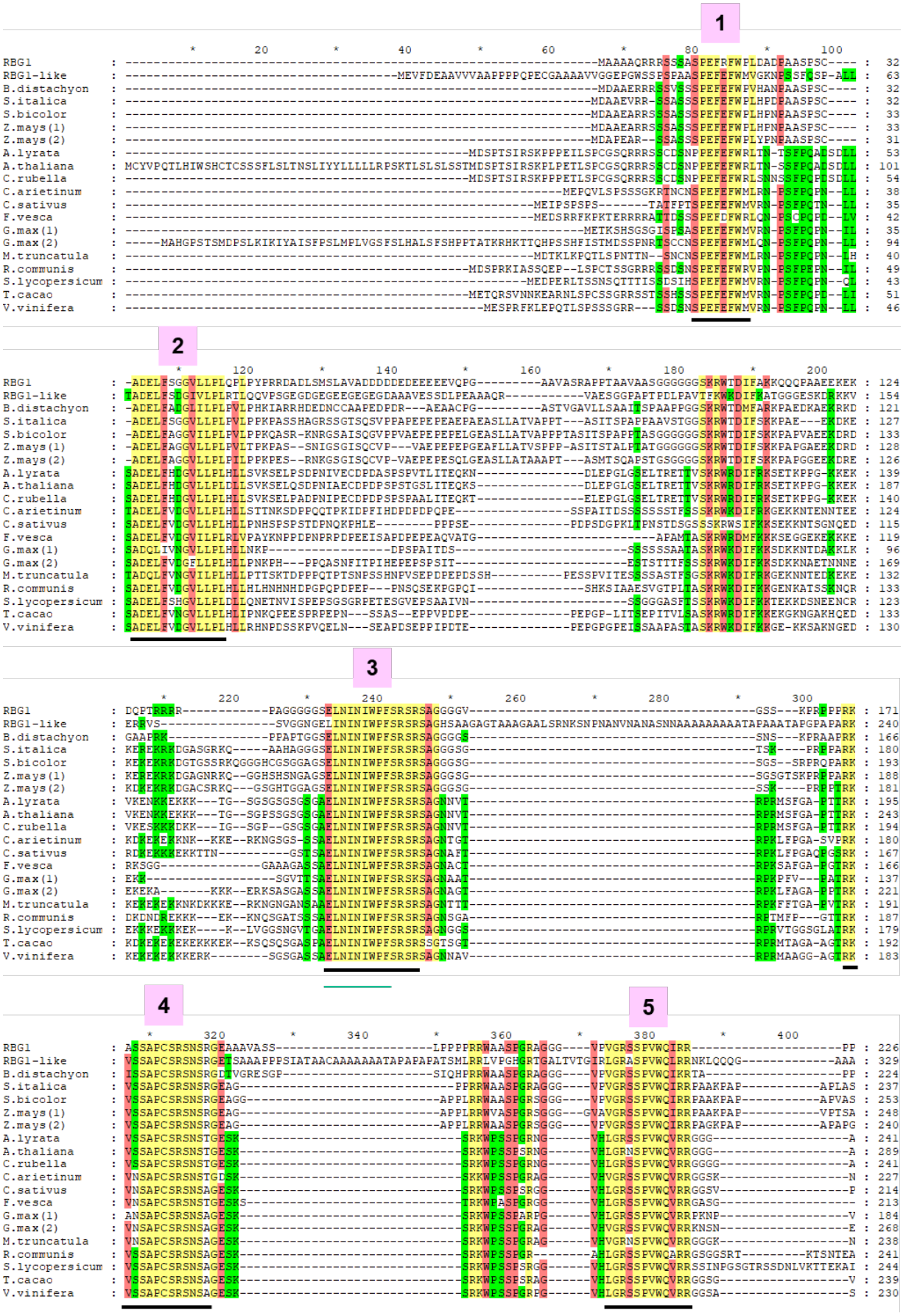

1
2
3
4
5

Figure S7. Six conserved sequences are present in RBG1 and its homologous proteins from different plant species.

The multiple alignment of RBG1 and homologous proteins was carried out using the ClustalW program of the MEGA 5.2 software. Numbers in pink-shaded areas above amino acid sequences denote the six conserved sequences. The accession numbers of the RBG1 homologous proteins are listed in Table S1. Underlined sections indicate the range of amino acids that were deleted from RBG1 to generate RBG1 conserved sequence-deleted mutants used for rice transformations represented in Figure S8.

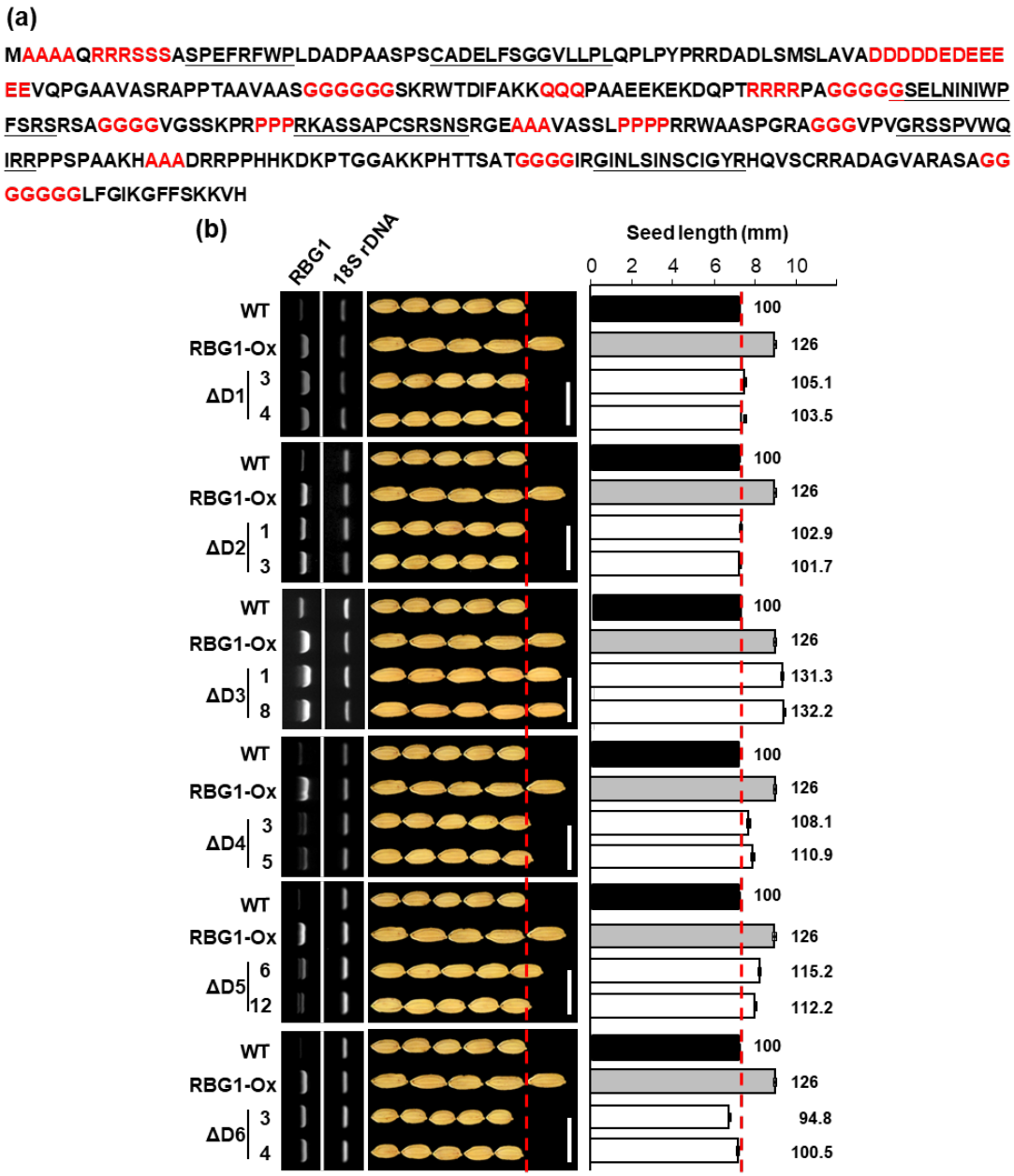


Figure S8. Conserved sequences 3 of RBG1 is not essential for its function. (a) RBG1 amino acid sequence. The inter-conserved sequence regions of RBG1 contain several stretches of amino acid homopolymers (red letters). Underlined regions indicate the six conserved sequences in RBG1 and its homologs. (b) RBG1 truncated at various conserved sequences (ΔD) were expressed under the control of the Ubi promoter in transgenic rice. Total RNAs were purified from 3-5 cm long panicles and subjected to RT-PCR analysis for detection of transgene expression levels. Morphology and seed length measurements (n = 30) of two representative transgenic lines expressing each motif-deleted RBG1 are shown here. The numbers above bars are % relative to the value (100) in WT. Scale bar = 3 cm.

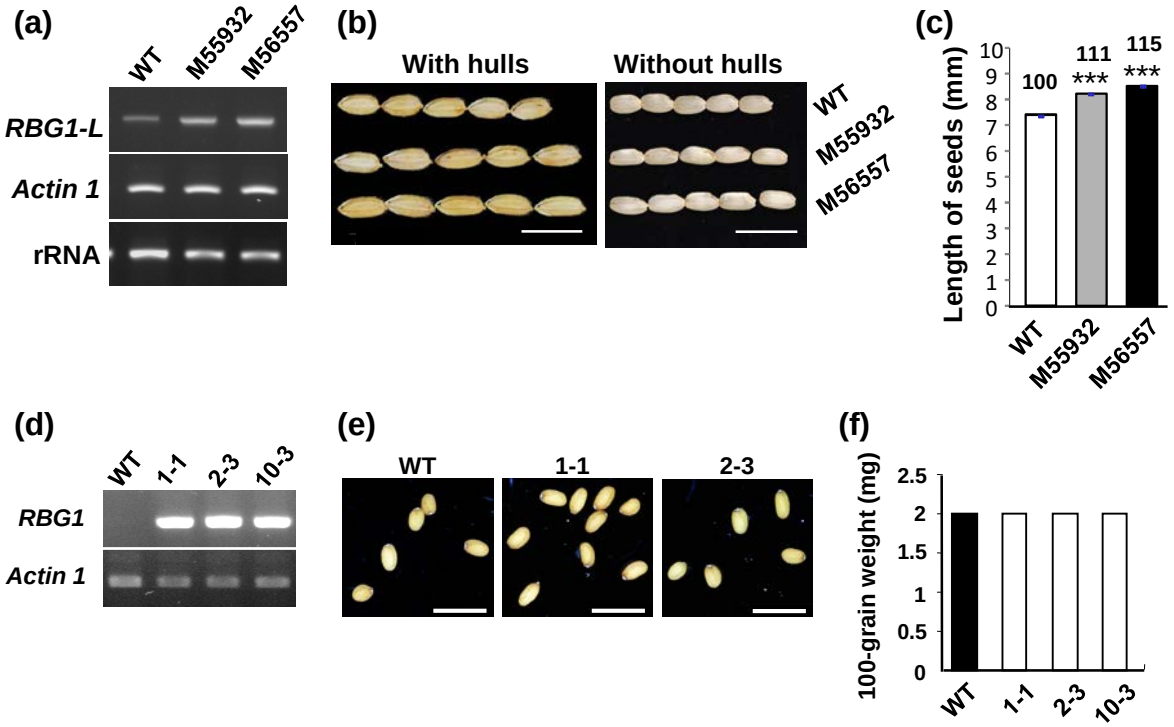


Figure S9. Overexpression of *RBG1-L* induces the big grain phenotype, but overexpression of *RBG1* in Arabidopsis does not induce a bigger grain phenotype.

(a) *RBG1-L* mRNA levels in two *RBG1-L* activation-tagged mutants, M55932 and M56557. Total RNAs were extracted from 11-15 cm long immature panicles and subjected to RT-PCR analyses using gene-specific primers (Table S4). **(b)** The two *RBG1-L* activation mutants display the big grain phenotype: mature seeds with hulls (left panel) and without hulls (right panel). Scale bar = 1 cm. **(c)** Statistical analysis of seed length (including seed coats) of M55932 and M56557. The numbers above bars are % relative to the value in WT, n = 90. **(d)** *RBG1* mRNA levels in T2 transgenic Arabidopsis overexpressing *RBG1*. **(e)** Seed size is not increased in two transgenic Arabidopsis lines overexpressing *RBG1*. **(f)** The 100-grain weight is not increased in three transgenic Arabidopsis lines overexpressing *RBG1*.

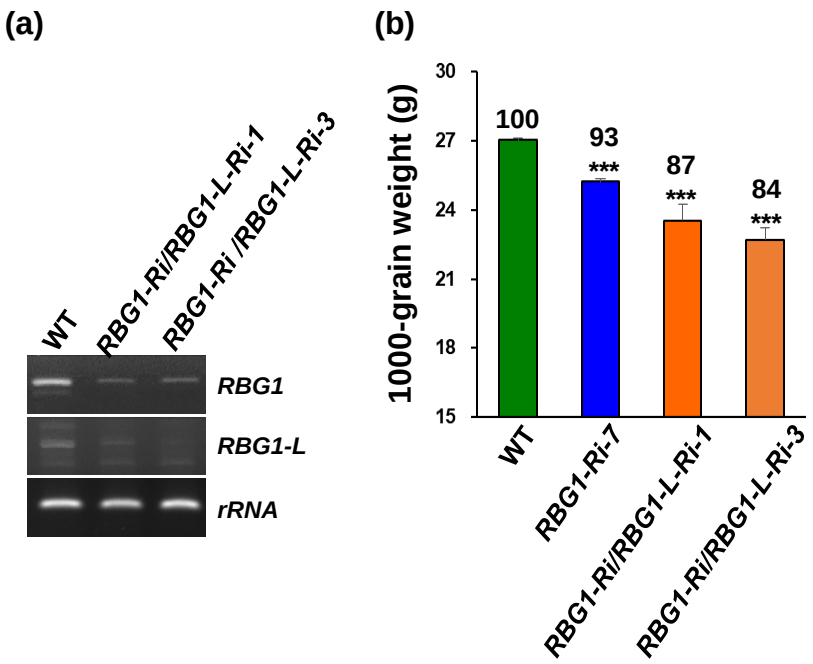


Figure S10. Correlations between level of *RBG1* and *RBG1-L* expression and the grain size. **(a)** Analysis of *RBG1* and *RBG1-L* mRNA levels in the *RBG1* and *RBG1-L Ri* double knockdown transgenic lines.

(a) Knockdown expression of both *RBG1* and *RBG1-L* resulted in lower grain weight than the WT and RBG1-Ri single knockdown. Numbers above bars are % relative to the value in WT. n = 20, 10, 10, 8 for WT, *RBG1-Ri-7*, *RBG1-Ri/RBG1-L-Ri-1* and *RBG1-Ri/RBG1-L-Ri-3* lines, respectively.

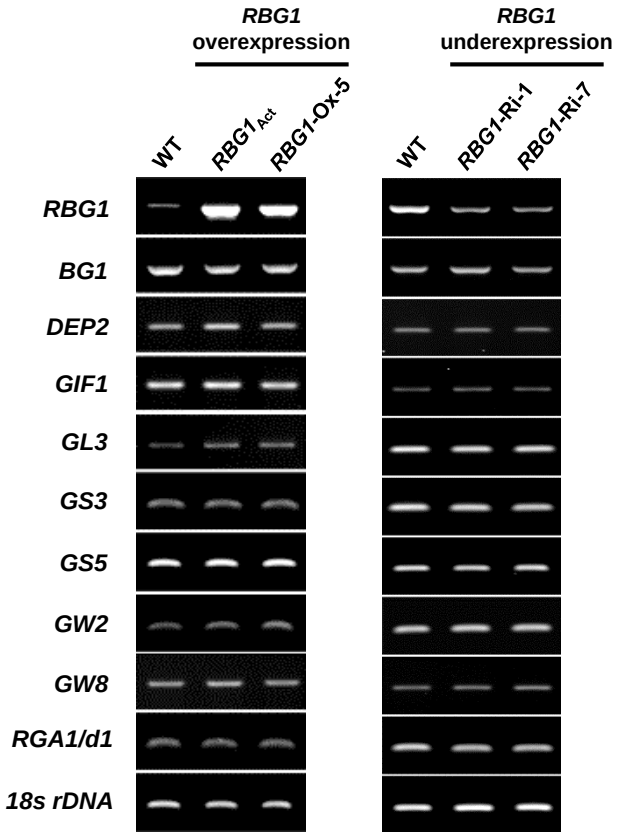


Figure S11. *RBG1* overexpression or underexpression does not affect the expression of genes known to regulate grain size in rice.

Total RNAs were extracted from 3-5 cm long immature panicles and subjected to RT-PCR analyses using the gene-specific primers listed in Table S4. References for these genes are also provided in Table S4.

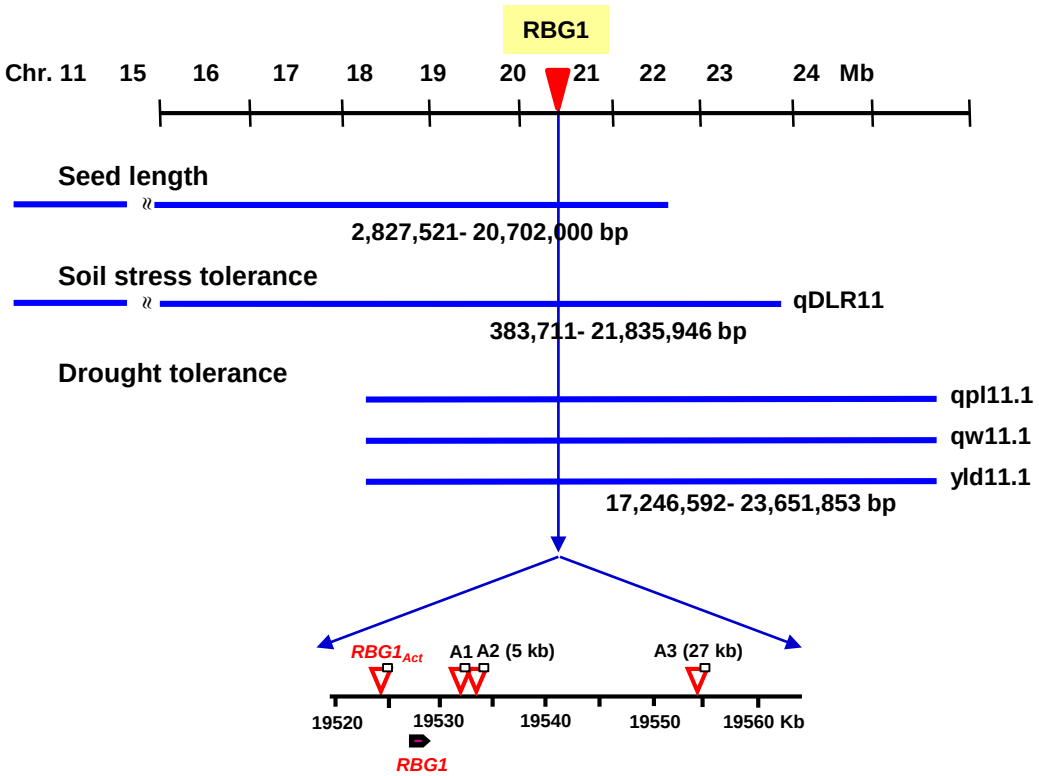


Figure S12. RBG1 is located within QTLs controlling seed length, soil stress tolerance and drought stress tolerance in rice.

Relative genome browser locations of QTLs and *RBG1*. The scale in Mb marks the location on rice chromosome 11 in the Q-TARO database. The positions of $RBG1_{Act}$ and its allelic mutants A1, A2, A3 are denoted.