

Supporting Information

Yoshida et al. 10.1073/pnas.0907896106

SI Text

Cover Caption. The spikelet of rice (*O. sativa*) forms a unique organ, sterile lemma, between a floret and two rudimentary glumes. A mutation in *G1* causes homeotic transformation of the sterile lemma into the lemma. *G1* encodes a land plant-specific

protein containing an uncharacterized domain, named ALOG. Yoshida et al. suggest that *G1* may have been involved in morphological modification of the spikelet during the evolution of rice.

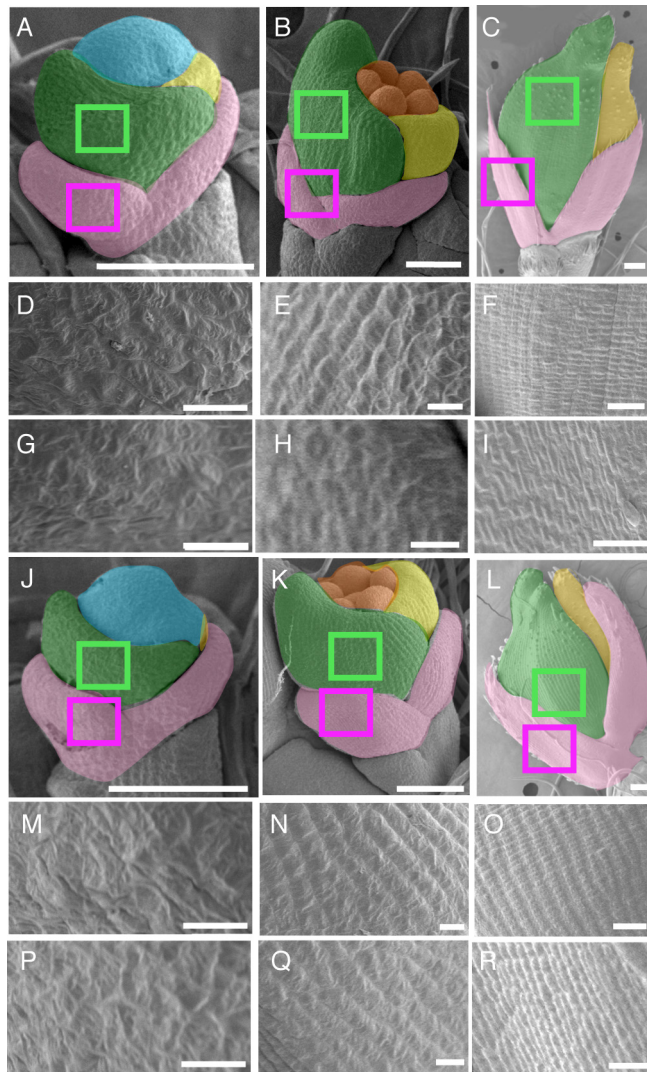


Fig. S1. Spikelet development in WT (A–I) and *g1* (J–R). (A–C, J–L) Developmental patterns of spikelets in WT and *g1*. Each organ is indicated by color: blue, floral meristem; orange stamen; yellow, palea; green lemma; pink, sterile lemma in WT and LSL in *g1*. (D–F, M–O) Epidermal surface of the lemma in WT (D–F) and *g1* (M–O). (G–I) Epidermal surface of the sterile lemma in WT. (P–R) Epidermal surface of the LSL in *g1*. (Scale bars: 100 μm in A–C and J–L; 10 μm in D–I and M–R.)

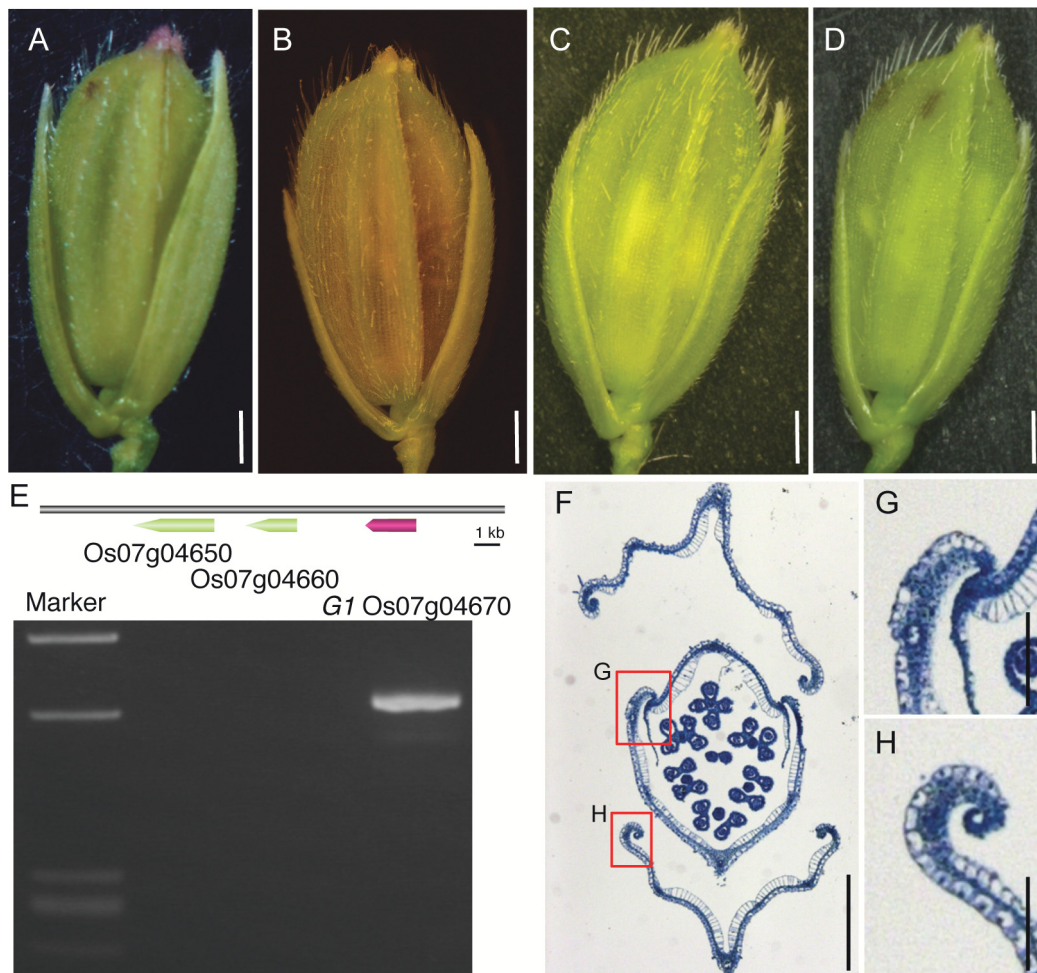


Fig. S2. Phenotypes of spikelets in *g1* mutants and *O. grandiglumis*. (A–D) Spikelet phenotypes of *g1-1*, *g1-2*, *g1-3*, and *g1-4*. (E) RT-PCR analysis of *G1* mRNA in young inflorescence. Among the three genes in the 12-kb deletion (Fig. 2A), only *Os07g04670* is expressed in developing spikelets. (F–H) Cross-section of the spikelets of *O. grandiglumis*. A spikelet (F) and close-up views of marginal region of the lemma/plea (G) and LSL (H) are shown. (Scale bars: 1 mm in A–D; 500 μ m in F; 100 μ m in G and H.)

AGACGCATTCTGTACGTCTCTCTAGTGTTACTGCGTAACGTAGGTGGCCGCTGCCGGCC

ATGTCGTCGTCGTCGCCGCTGCCGCGCTGGGCTCCGACGACGGCTGCTCGCCGGCGGAGCTG 20
M S S S S A A A L G S D D G C S P A E L

CGGCCGAGCCGGTACGAGTCGAGAAGCGCCGGGACTGGCAGACCTTCACGCAGTACCTC 40
R P S R Y E S Q K R R D W Q T F T Q Y L

GCCGCGCACCGCCCGCCTCGAGCTCCGCGCTGCAGCGGCCACGTCCTCGAGTTC 60
A A H R P P L E L R R C S G A H V L E F

CTCCGCTACCTCGACCGTTCGGCAAGACGCGCTCCACGAGCCCGCTGCCGTCGTAC 80
L R Y L D R F G K T R V H E P P C P S Y

GGCGGTCGCTCGCCGTCGCCGCGCCGGTCCGCGCCGCCGCGCCGATGCCAGTGC 100
G G R S P S A A G P V A A A A A A C Q C
T (g1-2, g1-3)

CCGCTGCGCCAGGCGTGGGGCAGCCTCGACGCGCTCGTCGGCCGCTCCGCGCCGCTAC 120
P L R Q A W G S L D A L V G R L R A A Y
A (g1-2)
T (g1-3)

GACGAGCGTCACGGCCGCGCCGGGAGCCCGACGCGCTCGCGGGCGCCGGCGCGTCCGCC 140
D E R H G R A G E P D A V A G A G A V A

ACCGACAGTACCTCCTCCTCCTCCGCCCGCCGCCAACCCCTTCGCCGCGCGCCGCTG 160
T D S T S S S S A A A A N P F A A R A V

AGGCTGTACCTGCGCGACGTCCGCGACGCGCAGGCCATGGCGCGCGGATCTCCTATCAC 180
R L Y L R D V R D A Q A M A R G I S Y H

AAGAAGAAGAAGCGCAGGGGCGGAACATGAACGGCGCCCGCGGGCGGGCGGGCGGGCGGG 200
K K K K R R G G N M N G A R G G G G G G

GCCGCGCGGGCGTGAACGACGGCGATGCGACGGCGCCCGGTGGCGGTGACCCCGGGG 220
A R A G V N D G D A T A P P V A V T P G
deletion (g1-4)

CTACCTCTGCCCGCGTCCACCGTGCCTCAACGGTGTGCCGTTTCGAGTACTGCGACTTC 240
L P L P P L P P C L N G V P F E Y C D F

GGGAGCGTCTCGGGGAGCACATGGCGCCATGGCGGCCATGGCGGGCGGGCGGGCGGG 260
G S V L G G A H G A H G G H G G G G G G

TTCTACGGCGCCGGCTACTTGCATTTCTGTACAACACCTTCAGTTAGTTAGCTAGC 276
F Y G A G V Y L P F L Y N T F S

TAGCTAGTTCGTCGTGATTTGCTGTGCTTCTCACTGTGGTTGCTTCAGTGTACTAGCT

AGCTACATGTGTGATGTGCTGCATCTTGATCATCTTGATCTGTGCTTTGATGCTGAT

Fig. S3. DNA and protein sequences of rice G1. Blue and red boxes indicate the ALOG domain and nuclear localization signal, respectively. Thick lines indicate amino acid insertions in the ALOG domain. Underline indicates the region used for in situ probe. Mutations in the *g1* alleles are shown in orange.

OsG1	1	SRYESQKRRDWTFTQYLAHRPPLLELRCSGAHVLEFLRYLDRFGKTRVHEPPCPSYGGRSPSAAGPVAANAACQCP	80
ZmG1	1	SRYESQKRRDWTFTQYLAHRPPLLELRCSGAHVLEFLRYLDRFGKTRVHAPSAAAYGGGGGRPVE---AAVACQCP	77
SbG1	1	SRYESQKRRDWTFTQYLAHRPPLLELRCSGAHVLEFLRYLDRFGKTRVHAPLCAAYGGGGGPPALV---AAPCQCP	77
BdG1	1	SRYESQKRRDWTFTQYLAHRPPLLELRCSGAHVLEFLRYLDRFGKTRVHHPHCPYGGAAALSSSTANT-HFHPCCQCP	79
OsG1L1	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHAHGCPFFGHPS-----PPAPCP	71
OsG1L2	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHAAGCPFFGHPS-----PPAPCP	71
OsG1L3	1	SRYEAQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHTAACPFPHGN-----PPAPCP	71
OsG1L4	1	SRYEAQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHTAACPFPHGN-----PPAPCP	71
OsG1L5	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHAPACPFPHGPA-----PPAPCP	71
OsG1L6	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHTVPCPFYGHPN-----PPAPCP	71
OsG1L7	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHASCAYFGQPS-----PPGPCP	71
OsG1L8	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHASCAYYQPS-----PPAPCP	71
OsG1L9	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHADGCAFYGEPN-----PPAPCP	71
AtLHS1	1	SRYENQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHQNCAFFGLPN-----PPAPCP	71
AtLHS2	1	SRYENQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHQNCAFFGLPN-----PPAPCP	71
AtLHS3	1	SRYENQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHTNICHFYGHPN-----PPAPCP	71
AtLHS4	1	SRYENQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHTLCPFFGHPN-----PPAPCP	71
AtLHS5	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHATACPFPHGPN-----PPSQC	71
AtLHS6	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHVAACPFGHQ-----PPS	71
AtLHS7	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHIHGCVFFGQVE-----PAG	71
AtLHS8	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHIQGCVFFGQKE-----PPG	71
AtLHS9	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHHQACVFFGQPD-----PPG	71
AtLHS10	1	SRYESQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHVPFCMFGQPE-----PPAP	71
Sm G1-like1	1	SRYEAQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHAPACPFPHGLAH-----PPAP	71
Pp G1-like1	1	SRYEAQKRRDWTFTQYLRNHRPPLLELRCSGAHVLEFLRYLDQFGKTKVHVPFCMFGQPE-----PPHP	71

OsG1	81	QOAWGSLDALVGRRAAYDERHGRAGEPDVAVAGAVATDSTSSSSA-AAANPFAARAVRLYLDRVDRDAQAMARGISY	157
ZmG1	78	QOAWGSLDALVGRRAAYFEERHGARG---IWTSSQPDGGVGVGGGGGANPFAARAVRLYLDRVDRDAQSRARGISY	152
SbG1	78	QOAWGSLDALVGRRAAYFEERHGARGSGTIWTSSQSQSQPQAVVDGDAANPFAARAVRLYLDRVDRDAQSRARGISY	155
BdG1	80	QOAWGSLDALVGRRAAYFADNDNDNNK-----ASINPFAARAVRLYLDRVDRDQSRARGISY	136
OsG1L1	72	QOAWGSLDALVGRRAAYFEEHGGR-----PESNPFGARAVRLYLDRVDRDQSRARGISY	125
OsG1L2	72	QOAWGSLDALVGRRAAYFEEHGGR-----PEANPFGARAVRLYLREVDRDQSRARGISY	125
OsG1L3	72	QOAWGSLDALVGRRAAYFEENGGR-----PESNPFARAVRLYLREVREHQARARGVSY	125
OsG1L4	72	QOAWGSLDALVGRRAAYFEENGGR-----PESNPFARAVRLYLREVREHQARARGVSY	125
OsG1L5	72	QOAWGSLDALVGRRAAYFEENGGR-----PENPFGARAVRLYLREVREHQARARGVSY	125
OsG1L6	72	QOAWGSLDALVGRRAAYFEENGGR-----PEMNPFGARAVRLYLREVREHQARARGVSY	125
OsG1L7	72	QOAWGSLDALVGRRAAYFEENGGR-----PESNPFARAVRLYLREVREHQARARGVSY	125
OsG1L8	72	QOAWGSLDALVGRRAAYFEESGHA-----PESNPFARAVRLYLREVREHQARARGVSY	125
OsG1L9	72	QOAWGSLDALVGRRAAYFEESGGR-----PESNPFARAVRLYLREVREHQARARGVSY	125
AtLHS1	72	QOAWGSLDALVGRRAAYFEENGPP-----PEANPFGARAVRLYLREVDRDQSRARGVSY	125
AtLHS2	72	QOAWGSLDALVGRRAAYFEENGGA-----PETSPPGSRVRIFLREVDRDQSRARGVSY	125
AtLHS3	72	QOAWGSLDALVGRRAAYFEENGK-----PETNPFGARAVRLYLREVDRDQSRARGVSY	125
AtLHS4	72	QOAWGSLDALVGRRAAYFEENGGS-----PETNPFGARAVRLYLREVDRDQSRARGVSY	125
AtLHS5	72	QOAWGSLDALVGRRAAYFEIIGGL-----PESNPFARAVRLYLREVDRDQSRARGVSY	126
AtLHS6	72	QOAWGSLDALVGRRAAYFEENGGR-----PDSNPFARAVRLYLREVDRDQSRARGVSY	125
AtLHS7	72	QOAWGSLDALVGRRAAYFEENGGL-----PERNPFAGGGIRVFLREVDRDQSRARGVSY	125
AtLHS8	72	QOAWGSLDALVGRRAAYFEENGGL-----TEKNPFAGGGIRVFLREVDRDQSRARGVSY	125
AtLHS9	72	QOAWGSLDALVGRRAAYFEHGGGS-----PDTNPFANGSIRVHLREVDRDQSRARGVSY	126
AtLHS10	72	QOAWGSLDALVGRRAAYFEENGPP-----PETNPFASGATRVYLREVDRDQSRARGVSY	125
Sm G1-like1	72	QOAWGSLDALVGRRAAYFEEHGGR-----PESNPFGARAVRLYLREVREHQARARGISY	125
Pp G1-like1	72	QOAWGSLDALVGRRAAYFEENGK-----PESNPFGARAVRLYLREVREHQARARGISY	125

Fig. S4. Amino acid alignment in the ALOG domain. Thick red lines indicate the insertions found in protein in the G1 clade.

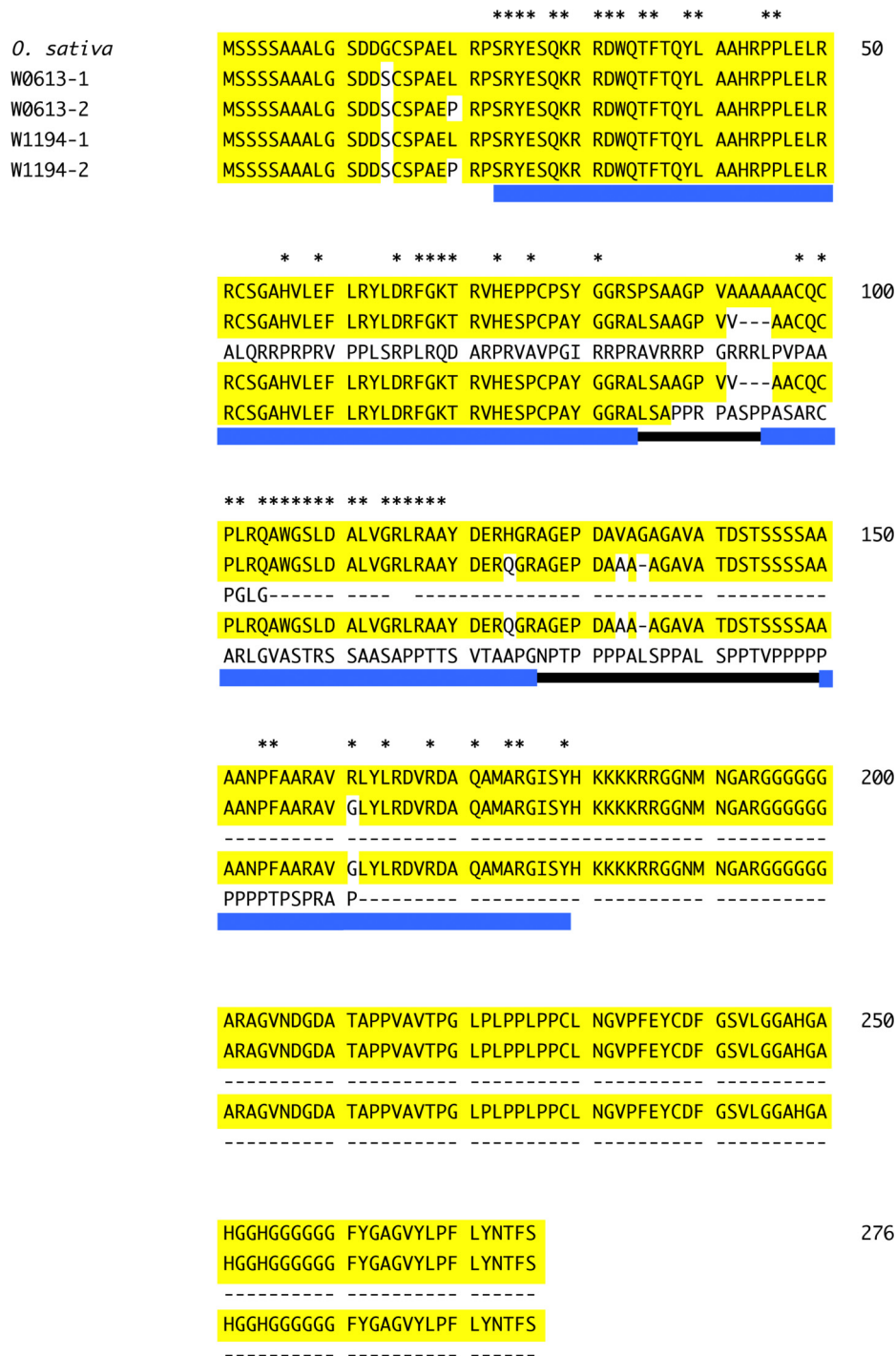


Fig. S5. Amino acid alignment of G1 proteins from *O. sativa* and *O. grandiglumis*. Thick blue lines indicate the ALOG domain and black lines indicate the insertions in G1 proteins. The two genomes of the tetraploid species *O. grandiglumis* (accessions, W0613 and W1194) are tentatively numbered 1 and 2.

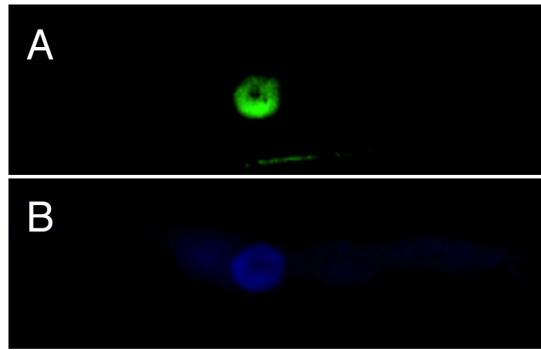


Fig. S6. Subcellular localization of G1. (A) GFP-G1 fusion protein. (B) DAPI staining of the nucleus.

Table S1. Accession numbers

Species	Accession	Gene	Accession no.
<i>O. sativa</i>	—	<i>G1</i>	AB512480
	—	<i>G1L1</i>	AB512490
	—	<i>G1L2</i>	AB512491
	—	<i>G1L3</i>	AB512492
	—	<i>G1L4</i>	AB512493
	—	<i>G1L5</i>	AB512494
	—	<i>G1L6</i>	AB512495
	—	<i>G1L7</i>	AB512496
	—	<i>G1L8</i>	AB512497
<i>O. grandiglumis</i>	—	<i>G1L9</i>	AB512498
	W0613	<i>OgG1-1</i>	AB512486
	W0613	<i>OgG1-2</i>	AB512487
	W1194	<i>OgG1-1</i>	AB512488
<i>B. distachyon</i>	—	<i>OgG1-2</i>	AB512489
	—	<i>BdG1</i>	AB512481
<i>Z. mays</i>	—	<i>ZmG1</i>	AB512482
<i>S. bicolor</i>	—	<i>SbG1</i>	AB512483
<i>P. patens</i>	—	<i>PpG1-like1</i>	AB512484
<i>S. moellendorffii</i>	—	<i>SmG1-like1</i>	AB512485

Table S2. Primers used in this study

	Primer	Marker
Mapping		
g1M01 d1	5'-TATCTACGTCAAAGGCAGTC-3'	CAPS (AfaI)
g1M01 u2	5'-TTGAATGTTACGACATGCA-3'	
g1M02 d1	5'-CATTCAATTTATTTGTACAGATCAC-3'	STS
g1M02 u2	5'-GCATGCTCCCGGTGCATTTTC-3'	
g1M05 d1	5'-CCTGAAACAAAATGGATGATAA-3'	CAPS (DraI)
g1M05 u2	5'-ATGTGAATGAAAAACCTAATCC-3'	
g1M11 d1	5'-CCTGAAACAAAATGGATGATAA-3'	CAPS (PstI)
g1M11 u2	5'-ATTTATTTTATTATGACTTGATTTGT-3'	
g1M19 d1	5'-TGGAACAAGCCGACTGTCATACAC-3'	dCAPS (NdeI)
g1M19 u2	5'-GAGGTTCAATTTCTTTTCTTTTTCAT-3'	
g1M28 d1	5'-TGTCACATTAATATACGGACGTA-3'	dCAPS (AfaI)
g1M28 u2	5'-GATGGATTAATTAGGCTGGC-3'	
g1M32 d1	5'-ATCGTTACGATCGGAAATTTTAA-3'	STS
g1M32 u2	5'-TGATGGTTTAGTAGCGTTTGTAGA-3'	
G1 cDNA amplification		
G1 d1	5'-ACTACCACCTTGGTTGGCTCTATTAAGGC-3'	
G1 u2	5'-CACAAAGATGCAGACACATCACACATGTAGC-3'	
In situ probe		
G1i d1	5'-CCGGGGAGCCCGACGCCGTCGGGGCGCCG-3'	
G1i u2	5'-CACAAAGATGCAGACACATCACACATGTAGC-3'	
Construction of G1-GFP fusion protein		
G1 gfp d1	5'-GGAATTCATGTCGTCGTCGTCGCGCTGCCG-3'	EcoRI*
G1 gfp u2	5'-GGAATTCACTGAAGGTGTTGTACAGAAAT-3'	EcoRI*
Construction of G1-GAL4BD fusion protein		
G1 gal d1	5'-TCCGATATCGATGTCGTCGTCGTCGCGCTGCCGCTGGGC-3'	EcoRV*
G1 gal u2	5'-ACGCGTCGACCTAACTGAAGGTGTTGTACAGAAATGGCAAG-3'	Sall*

*Restriction enzyme used for plasmid construction.