

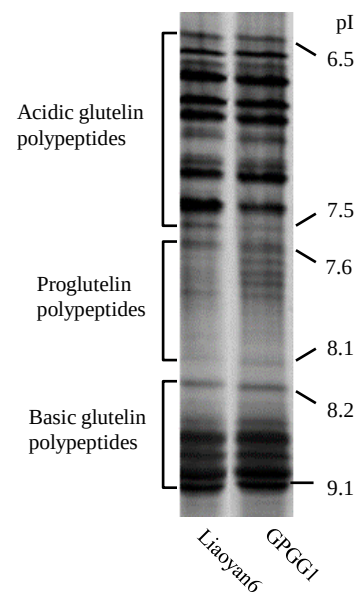


Fig. S1 IEF image of glutelins from *GPGG1* mutant and wild type cultivar Liaoyan6

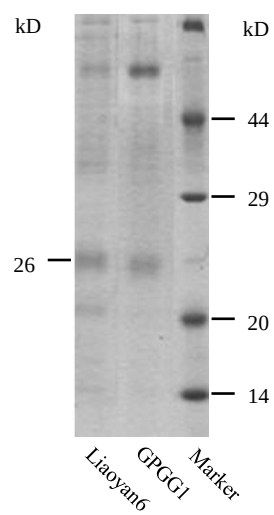


Fig. S2 SDS-PAGE image of globulins from *GPGG1* mutant and wild type cultivar Liaoyan6

Table S1 Markers linked to *GPGG1* mutant gene in short arm of chromosome 12

Marker names	Primer sequences	Fragment size (bp)	Contigs
Indel 12-1	TCTGGAGAGCTGCAGAAACA (forward) GCAATCTCTGCACTTTGATACC (reverse)	127	OJ1769_D07
OJ 1	CGACACCTGCTTCACCACCTT (forward) TTCCTTCCACTAATCACCCATCA (reverse)	122	OJ1311_G04
LS 6	TGCCATCACCACCTCTGT (forward) TCCCACGGTCCTGAAGTA (reverse)	128	OSJNBa0009F13
RM 7315	CACAAAGGCGTGTGGGTTAG (forward) GAGTCACGGGATGTTGCC (reverse)	152	OSJNBb0077O07
RM 3747	AGCAATGCACTCCCTTGATC (forward) TGTCTTCCTCCTTGTTTGG (reverse)	104	OSJNBa0023M16
Indel 12-2	TCTCAAGCATGTCAAGGCTTA (forward) ATGAACATGCAGAGCACCAA (reverse)	158	OSJNBa0095K18
RM27468	CCTCTCTTCTCCTCCCTTCTCC(forward) CACCACCTTGTTGTTCTCCTTCC (reverse)	186	OSJNBb0119L20
RM27469	CTTGCAAGCGGCGTACAGAGC (forward) TTCCTGGATCCTGAAGGTACGTAGG (reverse)	364	OJ1085_G07
RM27470	CTTAATAAACTTCGGCTGGCTTCC (forward) TAACTTCGTCGTGGAGAGTACGG (reverse)	379	OJ1085_G07
RM27471	CAAACGGACACGCAAGAGAGC (forward) TTGCTTGGGTCTAAACGTGAAGG (reverse)	315	OJ1003_C01
RM27473	GGCATGGTATGAGTCCTTATCG (forward) CTATGTTCTGGAAGTTGGAACG (reverse)	397	OJ1003_C01
RM27477	CCATCAAAGAGGAGGTCTTAGTTTGG (forward) CTGGATGTTGCGATCAATGTTCC (reverse)	315	OJ1227_F03

Table S2 ORF-related genes predicted in *GPGG1*-located region

Locus ID	ORF No.	Description of proteins and functions	Forward and reverse primers for PCR
OS12g0126800	3	Protein containing DUF618 domain; Intracellular-membrane transport	TCTCCTATTTACGTGCAAGTTAC TAGTAGCTTGTAGCTTAATCAGC
OS12g0127100	1	Similar to SMC5-6 complex protein; DNA repair/metabolism	CCATGTGCAGATCCTAGCGCACCCCTC AAAAGAGAAAAATGAAAAGAACAAG
OS12g0127200	2	Protein containing Harpin-induced 1 domain; Abiotic stress	TATGTACCAATGCATAAAAAATTCATGC CGATTTTCGTTTTTGTCTCAGCACAAAT
OS12g0127400	2	Protein containing Cyclin-like F-box domain; Cell cycle regulation	GGGTATCGGCGACCTTGTAATAAT ACGAGCGGCGTGGATAAGCGACG
OS12g0128000	4	Protein containing cyclin-like F-box domain; Cell cycle regulation	GACGGGCCGTAGAGGAGTATAATGG CTGGTAATTCGGTTCCTGCAACAAG
OS12g0127600	3	Similar to Ulp1 protease; Proteolysis	TGCAATCATGCAAATGTATCATT GATGAATCTTCGGTGTAGCATCG
OS12g0127650	1	Similar to Ulp1 protease; Proteolysis	GCCTCTTGTGTGGTCATAGGTGG CAAACAATCCATTTTAGCAAAGT
OS12g0127800	1	Similar to VQ motif protein; Transcriptional regulation	TGTTGTGGTTCATTAAACCAAGATG TGTTGTATCTCACAATCAATCTTTT
OS12g0127900	3	Similar to Mpv17/PMP22 protein; Peroxisomal membrane relation	AATATTGCTTATTTTAAATTTTGAAT CGACAAGCCACTACGTACACTTCTCC
OS12g0128200	5	Similar to pentatricopeptide repeat- containing protein; DNA expression, regulation and translation of RNA	TCTCTGCTCCTTGGCTCGTCT GAGACTAAAACCACCGCGGAT
OS12g0128450	5	Similar to protein containing YDG/ SRA domain; Histone methylation	ATCTTCGTCATTCTAGTGCCTGCATT CTACCACAGCAACTAGCAAAAGATAT
OS12g0128600	2	Zinc finger, Tim10/DDP-type family proteins; Mitochondrial membrane transport	GACAGATATATCTAGCTTGAGTTT ACTTGCTACAATCAAAAAGATGAT
OS12g0127500	2	Conserved hypothetical protein	AGAGGGTATTTTGCATTATTATTAG AATCAAGCGATGCTTTCATAAATTC
OS12g0126750	1	-	CTATGCTAGCCGATCTGAATCTG ATCCTGAAAAAAAAAATCTCCT
OS12g0126900	1	-	CGTTCACCGGACAAAAGCAGAGCAAT ATCTTTACACTTTTTACAAATATTAT

Table S3 Whole sequence of wild type gene presumed to encode protein containing PPR motifs

1	TCTCTGCTCCTTGGCTCGTCTCCACGGCGTCATTCCGTGCGGGGAGGAGCTCCAATGTGATGGGTGCT	69
70	GGTGTGTGGGGACGACAGGTGGAGTCTATTTGTGCATAGAGTCTATTTCAAGAAGAAGAACCGGAG	137
138	GAGGAAGAGGTAGAAGCATGAGAGGGCCGACTTGTTCACGACGTTATCGCGCCCACGCTGCAGCCGA	205
206	TCACGCACGTGGAGAGGACCTGGGCGGCTACGGCGTGTGCGGAGAAGAGAACCAACCGGCTGCGC	272
273	GTGAGGGAAGGAGCCGCCTGCCGTGGCGCGCTAGACAAAGGAGCTCGCCGACCACTACTCCGCCGG	339
340	CAAATAGAGAACGGGAGAGGTGAGGAAAAAAGGATGAGATGTTTTTTTTTCATTTTGCTCTTACT	406
407	GATATGTGGGTCCCATAAGAGCCACGTAGGATAAAACCGCTTGCAAAACCATTGAAGGAGTTTATTT	473
474	GCACCGGTTTTCAAAGATAAGGGAGATGTTATACCCGGCTTTACGATTAAGGGAGACGATTCAACAAA	541
542	GGGTAAGAGATGAGGGAGGCAAAATAGGCTTATTCCATGACGAATGCTTCTCTCCCGTAGCAGCCCA	609
610	<u>GGTGCACATGGGCCCAGCAACCGCAGCCGCGGAGGCAGCACTCGCCCGCCGCGCCGCGACCCGCTG</u>	676
677	<u>CCCGCGCTCCGCGCCGCGACGCGCTCCCGCTCCCGGCGCGCCTCTTCGCGCAGCTCCACGCTCTCG</u>	743
744	<u>TCCTCACCGCCGGCCTCGCGCGCCATTCCCCCAACTTCTCCCTCCTCCTCCGCTCGCCTCCCTCT</u>	810
811	<u>CCTCCCCGTCTCCACCGCCTCCGCCTCCTCCTCCTCCCGCTCCCGCCCACCACCTTCTCGCC</u>	877
878	<u>AACTCCCTCCTCCTCGCCTCCTCCTCCTCCCGCTGCCTCCCTCCGCGCTCTCCCTCTACGCCCTCC</u>	944
945	<u>TCTTCCTCTCCTCCTCGCCGCCGCTCCTCCGCCCAACGCCTTACCTACCCGCCGCTGTTCCGCGC</u>	1011
1012	<u>CGCGCCTCCGGCCCTCGCCCTCGCGCTCGCCACCCATTCCGTCAAATTCCTCGGCGCCCATGCTGCCTC</u>	1080
1081	<u>CTGCGACCGCTCCTCGGTGCCGCGCTGCTCGGCGTCTTCGCGCGCTGCGGCAGGATCGCCTCATGCCG</u>	1149
1150	<u>CCGGGTGTTTCGACCGGATCGCGCACCCCGACCTGCCGGCCTGGAACGCGCTGCTCTCCGCTACGCGC</u>	1217
1218	<u>GTCTTCGTGCCAGGGACGTGGCCTGCGCCACCTCTGCGGCGGATGCTATCCTCGAGATGTTTGTGCGGA</u>	1286
1287	<u>TGCTATCCTTAGCCATCGAACCCAACGAGATCACGCTCGTCGCCGTCATCGGCGCGTGTGGCGAGCTTG</u>	1355
1356	<u>GCGCTGTCTAGCCATGGGGTGTGGGCGCACACCTATGCCGTGAAGCGCCGGCTCGCCGTCAACTGCATC</u>	1423
1424	<u>GTGGCCACGGCGCTTGTGGAGATGTACGCCGGGTGCGGGAGGCTGGACCTGGCCGAGCAGGTCTTTGC</u>	1491
1492	<u>AGCCGCTTCGGACAGGGACACACGCTGCTACAATGCGATGCTCCATGGCTTGGCAGTCCATGGCCATG</u>	1559
1560	<u>GCCGTGCTGCCCTCTCCCTGTTGGACAGAATGCACGGTGAGGGCGTGCCGGTGGATGGCGTCACTGTG</u>	1627
1628	<u>CTGTCCGTGATGTGTGCTTGTGCCCATGCTGGATTGGTGGATGAAGGTCTGGACTACTTTGACAGAATG</u>	1696
1697	<u>GAGATCGAATTCGGGATTGAGCCAAGGATTGAGCATTACGGATGCATGGTTGACATGCTGAGCCGGGC</u>	1764
1765	<u>TGGTCGACTTAACGATGCCGAGAAGCTGATCCACGGAATGCCATCGTGCCAAATGCTGCGATATACA</u>	1833
1833	<u>GGTCTCTGATCCGGGCTGTGGCATTTCATGGCAAGCTGGAGCTCGGCGAGAAGATGATCGCAGAATTG</u>	1900
1901	<u>ATGCGACTTGAACCGGATGACAGTGGAACCATGTCCTGATTTCAAACCTTCTACGCGACAACGAACCG</u>	1968
1969	<u>ATGGGAGGACGCCAAGAAGGCGAGGAAGGAGATGAAGTCCATGGGCATTGACAAGAGCCCTGGATCA</u>	2035
2036	<u>AGCTTTGTAGACATCAATGGAGTTCTCCATGAGTTCTTGGTGGGTGACAAGACGCATCCTGCCTCAAA</u>	2103
2104	<u>GGAGATATACGCCATGGTTGAGGATATAGAAACCAGATTAAGCGAGTGCGGACACCGGTCGAGCACCT</u>	2171
2172	<u>CATCTGCGTTGTTTGATGTGAGGAGGAGGACAAGGCGGATGCACTGTCTTACCACAGCGAGAGGCTC</u>	2239
2240	<u>GCCATTGCCTTTGCTCTGATCGCTTCCAATCCAGGAGCTCCCATTAGGATCATCAAGAACCTCCGGGT</u>	2307
2308	<u>GTGTGCCGATTGCCATGAGAGCGCAAAGCTTGATCTCGGGTGTATGGTAGAGAGATTGTCATGAGGG</u>	2375
2376	<u>ACCGTACCCGGTTCATCACTTCAGAGACGGGGTGTGTTCTTGTGGAGATTTCTGGTAAATACTCGTT</u>	2443
2444	<u>GTGTTGTATTGTACATTATACAGTGTAATGTGCTGTAGATTCTCTCATAATTCTGTGGTAATGGGAT</u>	2511
2512	<u>GTGCTGTACTGGTCTCACACATTTGCTCGAGATCACAGGCTGTGTTTAGTTCCACGCTAAAAATTGAAA</u>	2579
2580	<u>GTTTGAAGAAATTAGAACGATGTGACAGAAAAGTTGGAAGTTTGAAGAAAAAAGTTGGAATCTAAATA</u>	2647
2648	<u>GGGCCACAGTTCATTCCGTTGTAGTTTGCATGGTTCAATATCAGTTTTTTTTTACAATCATTTTATTGA</u>	2716
2717	<u>TGACGGATCAAGTGTACATGGCAAGTGGTCAAAGAACTGAAATGAAATTCCGAAACTTGGCCGTCGT</u>	2784
2785	<u>AGTTCAATATCAGTTGTCAGAATTCAAACAGTTGTAGTTTGTATAAATTGCGCAATTCGTGCATATG</u>	2852
2853	<u>TGGCTAGATAGTTCATATGAAAAAGAATTATTTGGGGCATATGTGGTCTATTCTTTCTCTAGATAGTT</u>	2921
2922	<u>CGTATGAAAAGGAAGGAAAAAGTATAAATTACCCCTCCTCCCCCCCCCCCCCAACTATGACGGTCG</u>	2989
2990	<u>TCCGAATTACCCCCGAACCAACAAAACCGGACATTCTTACCATCAACTATGCAACGGGGCGGAATTA</u>	3057
3058	<u>CCCCCTCGACCAATCCGCGGTGGTTTTAGTCTC</u>	3092

* PPR, pentatricopeptide repeat. Grey and underlined font displays the coding region containing two exons

Table S4 Regional composition of exons presumed to encode protein containing PPR motifs

S1-like region	S1-like region	P2-like region	L2-like region	S2-like region	E1-like region	E2-like region	DYW-like region
Exon 1	Exon 1, 2	Exon 2	Exon 2	Exon 2	Exon 2	Exon 2	Exon 2
1088-1180	1364-1377 1432-1510	1511-1615	1616-1723	1724-1819	1820-1921	1922-2023	2024-2431

* PPR, pentatricopeptide repeat; S, short; P, pentatricopeptide; L, long; E, extended; DYW, Asp(D)-Tyr(Y)-Trp(W)