Supplementary materials

Genome-wide profiling of genetic variation in *Agrobacterium*-transformed rice plants

Wen-xu LI, San-ling WU, Yan-hua LIU, Gu-lei JIN,
Hai-jun ZHAO, Long-jiang FAN, Qing-yao SHU

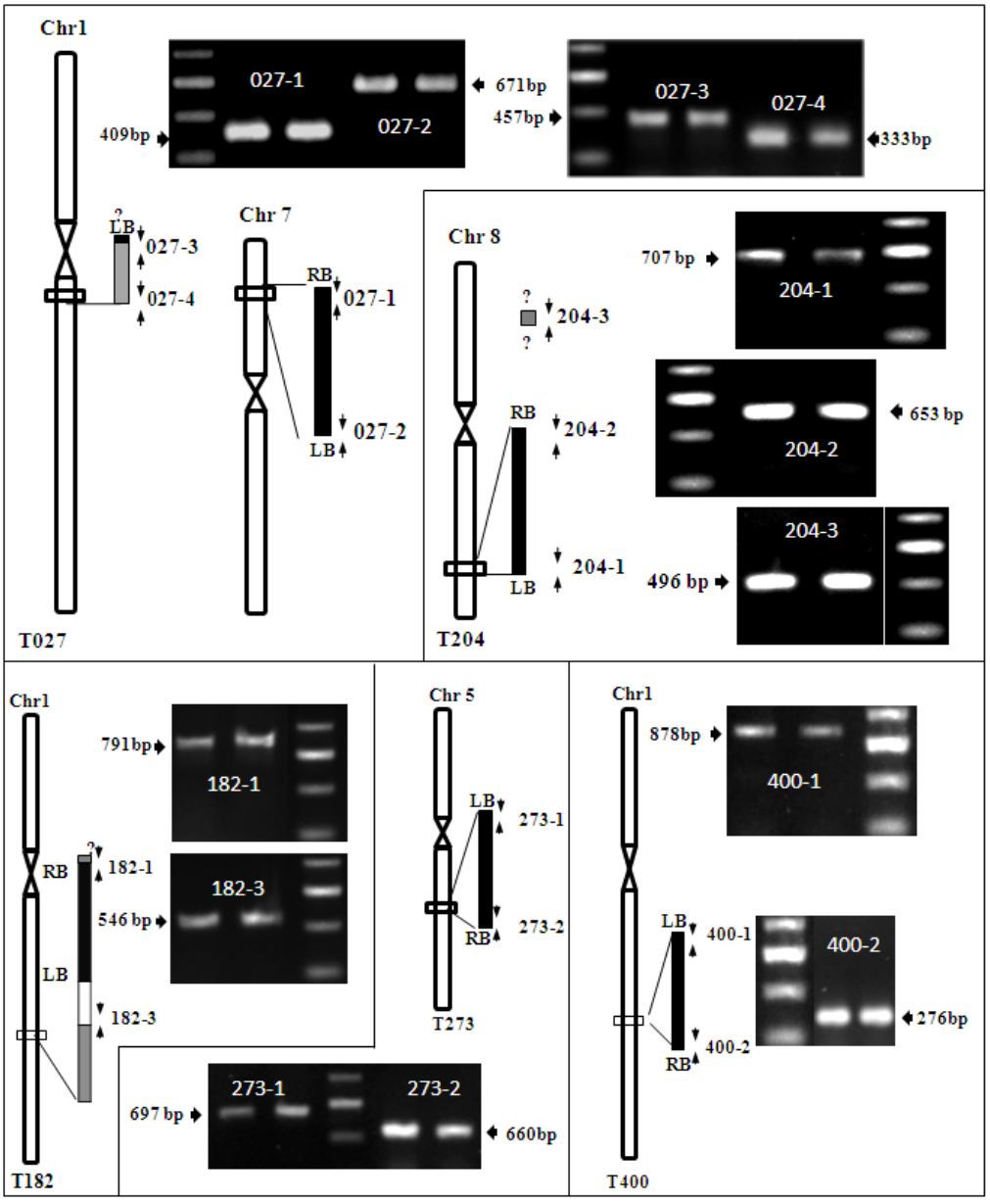

Fig. S1 PCR validation of plasmid DNA in transgenic rice lines

The position of PCR primers are shown alongside the insertions, e.g. 027-1 and 027-2 are primer pairs used for validation of T-DNA insertion in transgenic line T027. Sequence information of these primers are available in Table S1. Expected fragment size is given according to bioinformatics analysis; Different DNA ladders were used to estimate the size of amplicons

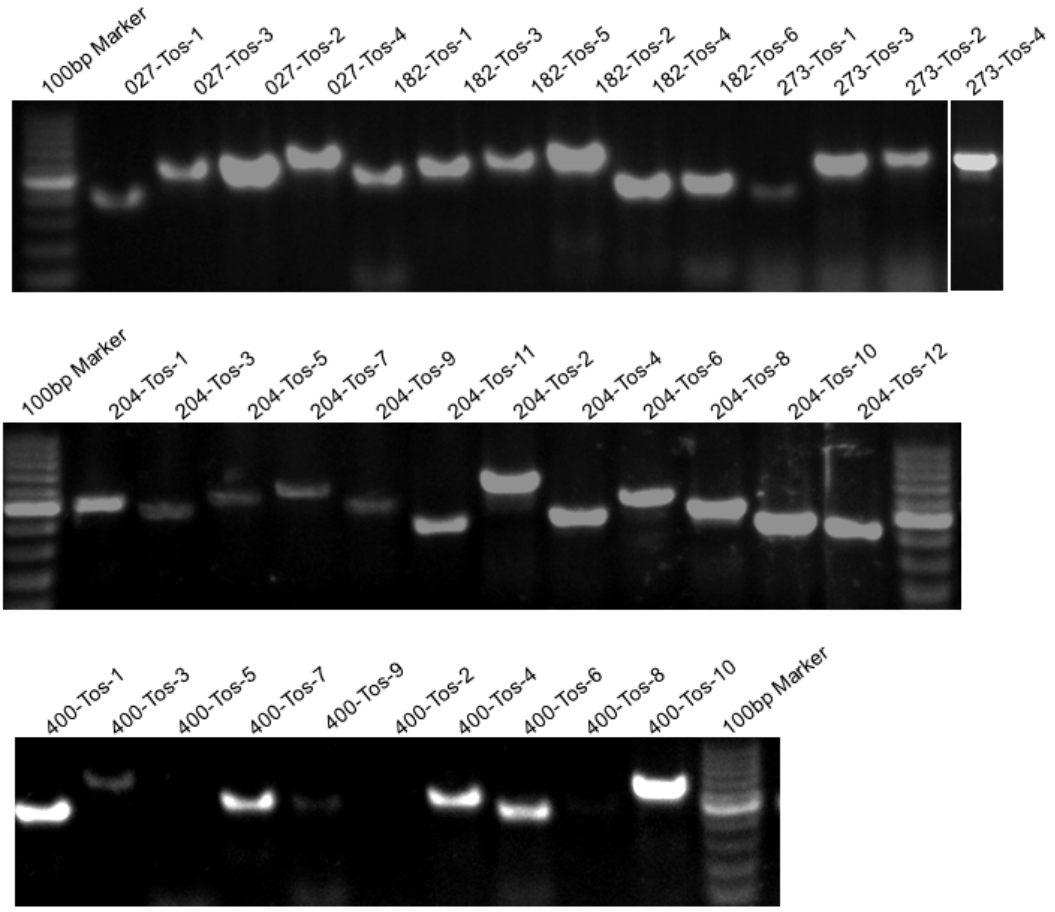

Fig. S2 PCR validation of newly inserted *Tos17* with site specific primers

For primer sequences and expected fragment size, see Table S1 and Table 2

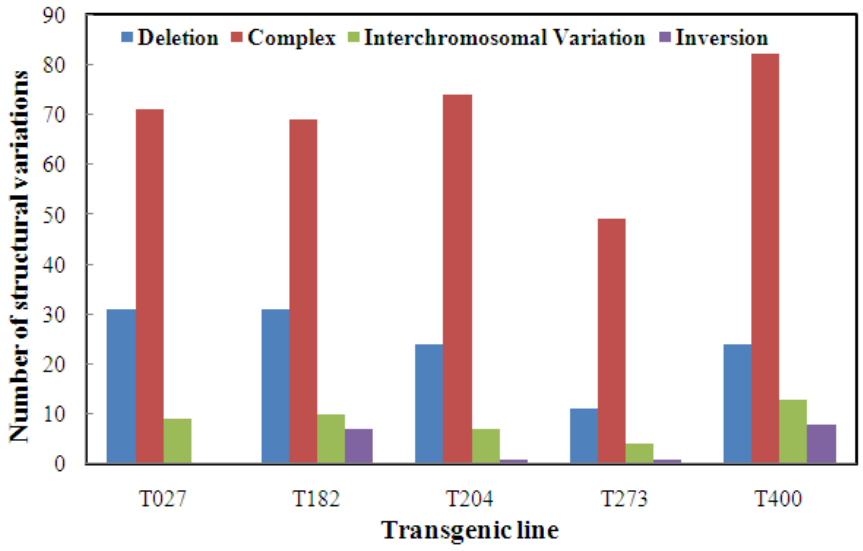

Fig. S3 Chromosome structural variations detected uniquely in transgenic lines

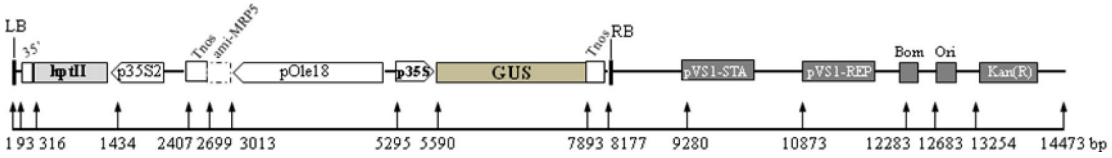

Fig. S4 Schematic diagram of artificial micro-RNA of *OsMRP5* expression plasmid p1301-amiMRP5-OleN

LB and RB are the left and right border of transfer DNA; 35S', the CaMV 35S poly A; hptII, hygromycin phosphotransferase II gene; p35S2, CaMV 35S promoter with a double enhancer sequence; p35S, the 35S promoter from CaMV; pOle18, the promoter of rice 18KDa oleosin gene; Tnos, the nopaline synthase terminator; ami-MRP5, the amiRNA sequence of *OsMRP5* based on osa-miR528 backbone; GUS, *beta*-glucuronidase gene; pVS1-STA, STA region from pVS1 plasmid; pVS1-REP, replication origin from pVS1; Bom, pBR322 bom site; Ori, pBR322 origin of replication. The numbers under the line are the starting nucleotide of functional cassete from the left border

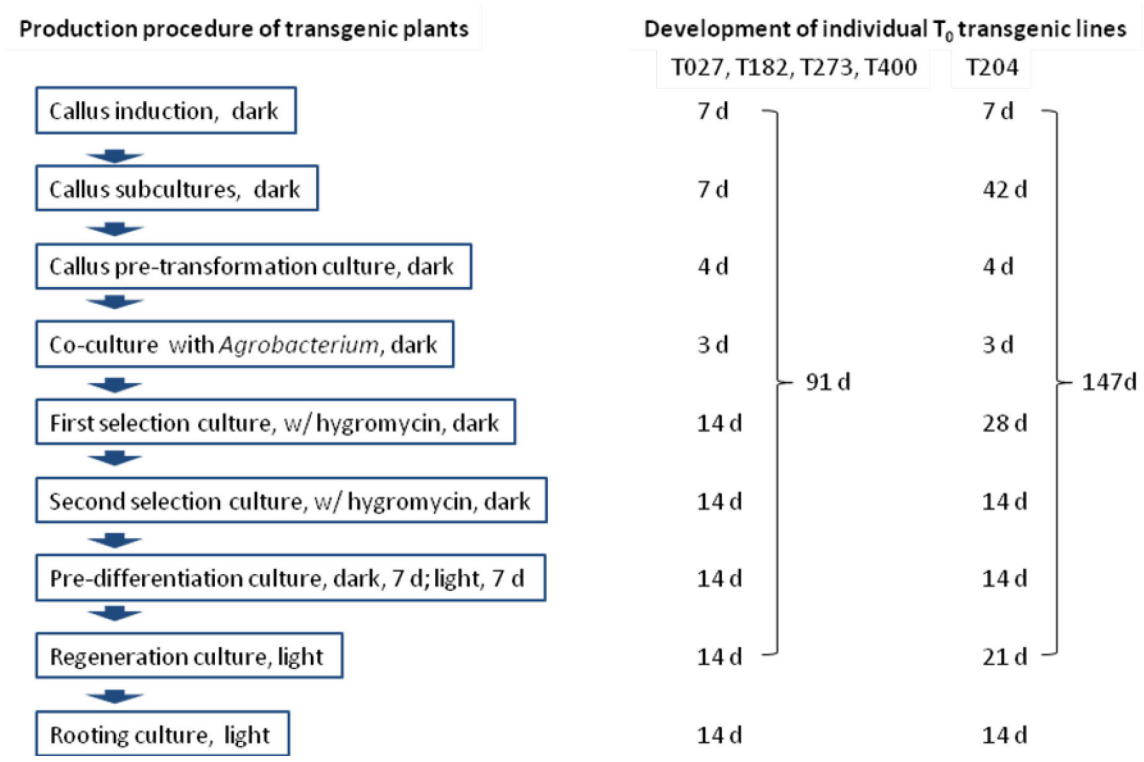

Fig. S5 Procedure used to produce transgenic plants from seed derived calli, and the particulars of 5 individual transgenic T_0 lines used for genome sequencing

Table S1 Summary of the genome sequencing of five transgenic rice lines

Transgenic line	Clean data (bp)	Mapped reads (%)*
T027	11,543,669,100	98.5
T182	11,603,645,460	98.0
T204	11,455,876,620	98.6
T273	11,537,551,800	98.8
T400	11,584,157,220	98.6

'Nipponbare' (RGAP 7) as reference genome

Table S2 Primers for validation of plasmid transfer and backbone DNA insertions

Name[1]	Forward (F) and reverse (R) sequences (5'→3')
027-1	F: TCCTGGCGATGAGCCGAAGATAT
	TD-R: ACCGCAGCAGGGAGGCAAACAAG
027-2	**TD-L**:TATCAGAGCTTGGTTGACGGCAATT
	R: GAACAATGCCAGGTACGTCCACC
027-3	**TD-L2**: CGTCCGAGGGCAAAGAAATAGAG
	BB-L: GCAAAGTCTGCCGCCTTACAACG
027-4	F: CTCTACGTGCCCGTCTGGAAGCT
	R: CGTTGGAGTGACTACATGGGAACTA
182-1	F: GTGCAGCCGGCCGCGTGCAGTTC
	TD-R: ACCGCAGCAGGGAGGCAAACAAG
182-2	**TD-L**:TATCAGAGCTTGGTTGACGGCAATT
	R: GGCACCACCATCGAACAAAGAAT
182-2'	F: TACCTCGTGGAGGCGTGGAGCTGGC
	R: CGAGGGCAAAGAAATAGAGTAGATG
182-3	F: AGAGGTTTACCGAGGAGCTGGTG
	BB-L: GCAAAGTCTGCCGCCTTACAACG
182-4	F: ACGGTCCTGGTCTCCTCCACTCTG
	R: GGGCTAGATGGGTTGCGATGGTC
182-4'	F: ACGGTCCTGGTCTCCTCCACTCTG
	R: ACGGCGGCGCTCGGTTTCTTCA
204-1	**TD-L**:TATCAGAGCTTGGTTGACGGCAATT
	R: CGTAGCGCAAAGGAGATAAGAACG
204-2	F: AAGCCAAGCAACCAATCAAAGCA
	TD-R: ACCGCAGCAGGGAGGCAAACAAG
204-3	F: GGGATAACGCAGGAAAGAACATG
	R: CACCACTTCAAGAACTCTGTAGCACC
273-1	**TD-L**: TATCAGAGCTTGGTTGACGGCAATT
	R: GCCTACTCATCGTCCTTCCTCTT
273-2	F: GATTCGGTGTTTTGGTCTTAGTTT
	TD-R: ACCGCAGCAGGGAGGCAAACAAG
400-1	F: CATCATATCCCTCCGTCTCCAGC
	TD-L: TATCAGAGCTTGGTTGACGGCAATT
400-2	F: ATCCTGTTGCCGGTCTTGCGATGA
	R: CCTACTATTAGCCGCTGACATAAGC

[1] Positions of the primers are illustrated in Supplemental Figure 2. TD-R, TD-L and TD-L2 are the universal primers close to the end of right or left bored of T-DNA which paired with location spesific primers to validat the plasmid transfer and backbone DNA insertions; ***BB-L*** are the universal primer close to the end of left bored of vetor backbone

Table S3 Primers used for validation of newly inserted *Tos17* sequences

Primer[1]	Forward (F) and reverse (R) sequences (5'→3')	Product size (bp)
027-Tos-1	F: CGGGAGATCGACCCGGAAAG	529
027-Tos-2	R: TCAAACAACCCATCAGATAACA	573
027-Tos-3	F: CATCAAAATAACTACTCGAAGGA	555
027-Tos-4	R: GGATTACCAGCGGGGTCTCA	632
182-Tos-1	F: CAGTTGCTCACTACTATGCTCCTT	500
182-Tos-2	R: AATTCGTATTGCGGCATCCT	644
182- Tos-3	F: CCAAATGGGGCTTCACTATG	548
182- Tos-4	R: TCAAACTAAACAGCCAGGAGAT	442
182- Tos-5	F: TGCTCTGTTCAAACCCAATCA	587
182- Tos-6	R: TCTTCTCCCCTTTTGATGGC	451
204-Tos-1	R: TGTTGTGCCTGCATCCACTG	510
204-Tos-2	F: CTTTTGATTTCATTCTCACGCA	688
204-Tos-3	R: TTTTATTCAACCACGCAGGA	474
204-Tos-4	F: TGGAAAGGAATTGTGGCACTG	453
204-Tos-5	R: GGAAGGGATACAAAGCAGGAG	560
204-Tos-6	F: AAAGGCCAGGGCAGAACAGC	573
204-Tos-7	R: CTCCCTCGGTCCCACAAAGA	625
204-Tos-8	F: TATGCCAAGTTTGCACAATGAAG	497
204-Tos-9	R: TCAATGGCAATGGAATACCT	518
204-Tos-10	F: CAGTTCTCCCACAAACATAGCATC	423
204-Tos-11	R: GCTGTGACATCGCATCTATC	418
204-Tos-12	F: ATATGCTGCTGTCCTGTGGC	403
273-Tos-1	F: AGTTTATTGCAGTAGCAGTCAAA	393
273-Tos-2	R: GGATTTGCTAAGCCGTCAAG	612
273-Tos-3	F: GAGAAAATGGTCGGAGCACA	575
273-Tos-4	R: ACGGCGGCTCAAAGGGGACT	617
400-Tos-1	F: CCCCAGTTTGATAGTGTCGC	458
400-Tos-2	R: TGCCCTCCGAGCTACAAGTC	587
400-Tos-3	F: ATCCCCTGTTCCGTGAAACT	656
400-Tos-4	R: GACAGGAAAACTCCGCCAGG	519
400-Tos-5	F: TTTATCCACTGCTCCACTGTCC	427
400-Tos-6	R: ATGGGTTCCAGGATTTATAGGG	423
400-Tos-7	F: CAAATGCCCCAAAACAGTAA	486
400-Tos-8	R: ACGAGCCACGAGCCGATGAA	465
400-Tos-9	F: TTTTGACAAGGTTCCGATGC	482
400-Tos-10	R: CCCACATTCCTGTCAAACCT	611
Tos	F: TGCCCTCCGAGCTACAAGTC	
	R: GGACAGTGGAGCAGTGGATAAA	

[1] A rice flanking forward (e.g. 027-Tos-1) or reverse (e.g. 027-Tos-2) primer is used together with a Tos reverse or forward primer, respectively, for amplification of a particular Tos17 insertion.

Table S4 Information about new *Tos17* insertions in transgenic rice

Transgenic line	Insertion site	Annotation of disrupted gene	Physical position of insertion site (size of chromosome) in million base pairs
T027	LOC_Os04g35210	Leucine rich repeat family protein, expressed	21.4 (35.5)
	LOC_Os12g43840	Ankyrin repeat domain-containing protein, putative, expressed	27.2 (27.5)
T182	LOC_Os01g39770	Calcineurin B, putative, expressed	22.4 (43.3)
	LOC_Os02g58170	Transposon protein, putative, unclassified, expressed	35.6 (35.9)
	LOC_Os08g43760	Carrier, putative, expressed	27.6 (28.4)
T204	LOC_Os01g02860	Transposon protein, putative, unclassified, expressed	1.0 (43.3)
	LOC_Os03g03610	1,3-beta-glucan synthase component domain containing protein, expressed	1.6 (36.4)
	LOC_Os03g07350	CSLA4 - cellulose synthase-like family A; mannan synthase, expressed	3.7 (36.4)
	Chromosome 8, intergenic		27.2 (28.4)
	LOC_Os11g47210	Receptor-like protein kinase 5 precursor, putative, expressed	28.4 (29.0)
	LOC_Os12g01955	Expressed protein	0.6 (27.5)
T273	Chromosome 5, intergenic		28.4 (30.0)
	LOC_Os09g33550	CCT/B-box zinc finger protein, putative, expressed	19.8 (23.0)
T400	LOC_Os02g02660	Retrotransposon protein, putative, unclassified, expressed	1.0 (35.9)
	LOC_Os08g09840	WRKY117, expressed	5.7 (28.4)
	LOC_Os08g44020	Rhamnogalacturonate lyase, putative, expressed	27.7 (28.4)
	LOC_Os09g31502	Dehydrogenase, putative, expressed	19.0(23.0)
	LOC_Os11g36090	Receptor kinase, putative, expressed	21.0 (29.0)

Table S5 Primers for PCR validation of structural variation in transgenic line T182

Name	Forward (F) and reverse (R) sequences (5'→3')	Product size (bp)
182-SV-1	F: AAATAAAGAAACAAATCCCCAACTG R: TTCGGTGCCACTCCAAAACATA	268
182-SV-2	F: ATGTCATTGTCCTATGTGGAAC R: TGACTACGTTGAAACCGCAAGT	434
182-SV-3	F: AGTCCCTCCAGACCCGTCCTGT R: TACTTCTCGCCGAGCGTCACAA	151
182-SV-4	F: TTTCGCCGACTTGAGTCCACCC R: CCTGACAGGCCCAGACAGCATT	376
182-SV-5	F: TCCATAAGAGGGATGATGAAACCA R: GATGACACCGGAGCAGTACCAG	349
182-SV-6	F: GAGGGCAAATTAGTGAGTTCAT R: ACAAGCTGCTACTGGATAGTGTT	382
182-SV-7	F: GCTGTGCGAATCACGAATTGAG R: TATTGGCACGCTCGTGGTAAAA	346
182-SV-8	F: AGTTATTCTCCAGTAGGAAAGGT R: AACCGTGGCGAAGCAAATCTAA	430
182-SV-9	F: ATGGATACTCCCTTTAGCAGGTC R: TGAGGTTTGTTGGCAATGTTCT	434
182-SV-10	F: ACGCTCCTGCTGCTGCAAGACG R: TGATCCCAGTGAATTACCCAGTC	464