



Comparative genomic hybridization and transcriptome sequencing reveal that two genes, *OsI_14279* (*LOC_Os03g62620*) and *OsI_10794* (*LOC_Os03g14950*) regulate the mutation in the γ -*rl* rice mutant

Xulong Wang¹ · Fanhua Wang¹ · Huiqiong Chen¹ · Xiaoyu Liang¹ · Yingmei Huang¹ · Jicai Yi¹

Received: 7 December 2016 / Revised: 29 June 2017 / Accepted: 4 July 2017 / Published online: 22 August 2017
© Prof. H.S. Srivastava Foundation for Science and Society 2017

Abstract We previously established the genetic locus of the rolled-leaf mutant, γ -*rl*, to chromosome 3. In this study, we performed a comparative genomic hybridization (CGH) analysis to identify the genes responsible for the γ -*rl* mutant phenotype. This was combined with RNA transcriptome sequencing (RNA-seq) to analyze differences in the mRNA expression in seeds 12 h after germination. Using the reference genome of the “indica type” rice from GenBank, we created a chip with 386,000 high density DNA probes designed to target chromosome 3. The genomic DNA from γ -*rl* and Qinghuazhan (the wild-type) was used for hybridization against the chip to compare signal differences. We uncovered 49 regions with significant differences in hybridization signals including deletions and insertions. RNA-seq analysis between γ -*rl* and QHZ identified 1060 differentially expressed genes, which potentially regulate numerous biological activities. Moreover, we identified 72 annotated genes in the 49 regions discovered in CGH. Among these, 44 genes showed differential expression in RNA-seq. qRT-PCR validation of the candidate genes confirmed that seven of the 44 genes showed a significant change in their expression levels. Among these, four genes [*OsI_10125* (*LOC_Os03g06654*), *OsI_14045* (*LOC_Os03g62490*), *OsI_14279*

(*LOC_Os03g62620*) and *OsI_14326* (*LOC_Os03g63250*)] were down regulated and three genes [*OsI_10794* (*LOC_Os03g14950*), *OsI_11412* (*LOC_Os03g21250*) and *OsI_14152* (*LOC_Os03g61360*)] were up regulated with a fold change ≥ 2.0 and a *P* value ≤ 0.01 . Finally, we constructed transgenic plants to study the in vivo functions of these genes. RNAi knock down of *LOC_Os03g62620* resulted in rolled-leaf, lower seed-setting and decreased seed growth phenotypes. Transgenic plants with *LOC_Os03g14950* over-expression showed dwarf plants with a shortened leaf phenotype. Our results, *LOC_Os03g62620* and *LOC_Os03g14950* as the essential genes responsible for creating the γ -*rl* mutant phenotypes suggested that these genes may play crucial roles in regulating rice leaf development and seed growth.

Keywords Comparative genomic hybridization (CGH) · RNA transcriptome sequencing (RNA-seq) · Rice rolled-leaf mutant · RNAi · Over-expression

Introduction

Moderate leaf rolling is an ideal architectural feature of the rice plant (Ma and Yuan 2015). However, the molecular mechanism that regulates this phenotype remains unclear. Understanding the genetic mechanism that controls rice leaf rolling may impact rice yield and breeding. In general, molecular studies of the rolled-leaf mutants have mapped a genetic locus via traditional cross breeding. Further mapping was also performed using molecular markers (Shao et al. 2005; Yan et al. 2006; Hu et al. 2010). Such studies identified various genes that regulated the leaf rolling phenotype including genes that encode transcription factors or miRNAs that regulate leaf development (Shi et al. 2007;

Electronic supplementary material The online version of this article (doi:10.1007/s12298-017-0460-9) contains supplementary material, which is available to authorized users.

✉ Jicai Yi
jicai@scau.edu.cn

¹ Key Laboratory of Plant Functional Genomics and Biotechnology of Guangdong Province, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China

Wu et al. 2009; Cho et al. 2013), modulate bulliform cells and sclerenchyma tissue formation (Xiang et al. 2012), and indole-acetic acid (IAA)-mediated regulation of leaf development (Fujino et al. 2008).

In our previous studies, we obtained a rolled-leaf rice mutant, γ -*rl*, with mutation localized on the third chromosome through ^{60}Co - γ irradiation of the native indica rice stable strain, Qinghuazhan (QHZ, *Oryza sativa* L. subsp. *indica*) (Yi et al. 2007, 2013, 2014). Flavin-containing monooxygenase (FMO) plays an essential role in IAA synthesis and growth. Although we identified the potential location of the gene responsible for the mutant phenotype, the molecular mechanism that bridges FMO mutation and rolled-leaf phenotype was not clear. Our failure to amplify the genetic locus of the γ -*rl* mutant by PCR suggested that a complex mutagenesis pattern may exist within that locus. Recently, we uncovered additional phenotypes in the γ -*rl* mutant including increased germination rate and decreased seed growth indicating that γ irradiation may cause extensive genome-wide mutagenesis.

Comparative genomic hybridization (CGH) (Kallioniemi et al. 1992; Freeman et al. 2006; Naseer et al. 2015) has been extensively applied in clinical diagnosis including the detection of genome mutagenesis in pathological tissues (Brady et al. 2013; Byeon et al. 2014; Khoury et al. 2014). CGH is also used for the detection of genome-wide DNA mutations including copy number variations (CNV), genetic deletions, duplications and amplifications (Hegde et al. 2008; Cahan et al. 2009; Marques-Bonet et al. 2009; Naseer et al. 2015), which lead to gene cloning and understanding the mechanism of mutagenesis. RNA transcriptome sequencing (RNA-seq) is able to quantify the changing expression levels of different transcriptomes and plays a crucial role in identifying mutation related genes (Wang et al. 2009; Ozsolak and Milos 2011; Liu et al. 2013; Kalra et al. 2013; Lambert et al. 2014). Recently, some studies have integrated the analyses of genome-wide DNA mutagenesis (e.g. CGH) and mRNA gene expression from transcriptome data (e.g. RNA-seq). These studies have provided novel insights into the molecular mechanisms of the mutants, and effectively accelerated the identification of potential target genes associated with disease pathology (Glaesener et al. 2008; Myllykangas et al. 2008; Jung et al. 2009).

As mentioned above, CGH arrays, widely used to identify genomic variations at a high resolution, can be custom-designed with reference genome sequences to target regions of interest. Thus, CGH has the advantages of simple operation, easy detection of genomic rearrangements and complex regions, and no sample amplification requirement, which is routine in other methods such as resequencing. Therefore, to understand the mechanisms

regulating the γ -*rl*'s phenotypes and to identify the genes responsible for creating the γ -*rl* mutant phenotype, in this study we used a synergistic approach by combining CGH and RNA-seq to analyze the genetic variations in the γ -*rl* locus present in the third chromosome. We identified seven candidate genes and verified their functions by qRT-PCR, RNAi and over-expression analysis in transgenic rice (Zhonghua 11). Our results facilitate a basic understanding of the rolled-leaf phenotype in rice and pave the way for future studies.

Materials and methods

Plant material and growth conditions

The rolled-leaf rice mutant, γ -*rl*, was isolated from a previous ^{60}Co - γ irradiation mutagenesis screen we performed on the dry seeds of Qinghuazhan (QHZ, *O. sativa*, L. subsp. *indica*), which is a common rice strain in Guangdong, China (Yi et al. 2007). Mature seeds of the same size were shelled and soaked in distilled water for 24 h at 25 °C, and grown in petri dishes covered with two layers of wet filter paper. The root and shoot lengths of 30 individual seeds per experiment were measured. This experiment had three independent replicates.

CGH analysis

A custom-designed 385 K microarray containing 386,000 probes were designed using the reference DNA sequence from the third chromosome of 93-11, a typical indica variety (<http://www.ncbi.nlm.nih.gov/nuccore/CM000128.1>). Roche NimbleGen's CGH probe design criteria were used. The average length of a probe was 50–70 nt and the average distance between probes was 109 nt. Probe design, CGH array scanning and data analysis were carried out using the NimbleGen platform (<http://bioservices.capitalbio.com/fwpt/RocheNimbleGenpt/3965.shtml>) and performed at CapitalBio Corporation (Beijing, China, <http://cn.capitalbio.com>). In brief, the genomic DNA from mutants and wild type controls were isolated and labeled with Cy3 and Cy5, respectively. After denaturation and hybridization, the array was washed and probed on NimbleScan (Roche NimbleGen, Inc., Madison, WI, USA) according to the manufacturer's instructions. The CGH analysis was visualized with SignalMAP software (Roche NimbleGen) to identify genomic regions or annotated genes that varied significantly. The gene selection criteria for quality control metrics were determined as follows: Log₂ signal values of ≥ 0.25 or values of ≤ -0.25 , and with at least five consecutive probes.

RNA-seq

RNA was extracted and sequenced from germinating rice seeds (12 h) by CapitalBio Corporation (Beijing, China). Three independent biological replicates were prepared for RNA-sampling. Briefly, total RNA sample was extracted using Trizol (Invitrogen, Carlsbad, USA) according to the manufacturer's protocol, and purified with RNeasy Plant Mini Kit (QIAGEN, Stockach, German). cDNA library was then prepared (equal amounts of three duplicate RNA samples were pooled to make a library) and sequenced on a HiSeq 2500 platform (Illumina, San Diego, USA). The sequence data were filtered for low-quality (unknown nucleotides >5%, or Q20 <20%) bases and adaptor sequences, followed by assembly and annotation. The RNA-seq reads were aligned to the indica rice reference genome sequences (<http://www.ncbi.nlm.nih.gov/nuccore/57015217>) using Bowtie software (Langmead et al. 2009) allowing up to two mismatches. The transcript abundance was designated by the number of reads per kilobase of exon per million mapped reads (RPKM) (Mortazavi et al. 2008). EBSeq (Leng et al. 2013) was employed to evaluate differential gene expression (DEG) between γ -*rl* and QHZ. DEG was considered significant when the \log_2 signal ratio was ≥ 1.0 or ≤ -1.0 with a *P* value ≤ 0.01 between comparisons. Finally, Gene Ontology (GO) functional enrichment analysis was carried out using MAS3.0 system developed by CapitalBio Corporation (<http://bioinfo.capitalbio.com/mas3/>).

RT-PCR

Total RNA was isolated with TRIzol reagent, and treated with RNase-free DNase I (Invitrogen) to remove residual DNA. Using a reverse-transcription kit (Invitrogen), single-stranded cDNA was first synthesized, and used as a template to amplify cDNA. PCR was performed in a 50 μ L volume using 200 ng template containing 5 μ L of 10 \times Buffer, 3 μ L of 25 mM MgSO₄, 0.25 μ L of each of the Forward and Reverse primers (50 μ M), 5 μ L of the 2 mM dNTP mixture and 1.0 U KOD plus DNA polymerase. During cDNA amplification, actin primers (0.25 μ L each) were also added as controls. PCR amplification parameters were as follows: 94 °C for 5 min; 32 cycles of 94 °C for 30 s, 55 °C for 30 s and 68 °C for an appropriate time at the rate of 1.0 kb/min. All cycles had a final extension of 10 min at 68 °C. The experiment was repeated three times.

Quantitative real-time PCR (qRT-PCR)

cDNA was synthesized using SuperScript[®] III Reverse Transcriptase (Invitrogen, Carlsbad, USA) according to the

manufacturer's instruction. qRT-PCR was performed in a 20 μ L volume using 200 ng cDNA, 10 μ L of 2 \times Sybr-green Master Mix (Bio-Rad, USA) and 0.5 μ L of each of the Forward and Reverse primers (10 μ M). PCR cycle conditions were as follows: 95 °C for 2 min; 40 cycles of 95 °C for 10 s and 60 °C for 10 s. The experiment was repeated three times.

RNAi and over expression analysis

RNAi and over-expression vectors were constructed as previously described (Yu et al. 2010). For the RNAi construct, cDNA fragments were cloned by RT-PCR with the corresponding primer pairs listed in Table 1. Each sequence was then cloned into the pYLRNAi.5 vector (kindly provided by Dr. Yao-Guang Liu, South China Agricultural University) between the *Bam*HI and *Hind*III restriction sites. A second round of PCR was then performed using Reverse primers to amplify the DNA inverted repeat and was subsequently cloned into pYLRNAi.5 to generate the final RNAi construct. To construct the over-expression vector, the full-length cDNA was amplified by RT-PCR and digested with *Hind*III and *Bam*HI, and then cloned into pYLox.5 (provided by Dr. Yao-Guang Liu). All cloning were verified by sequencing and transformed into a callus initiated from Zhonghua 11 (*O. sativa* L. subsp. *Japonica*) seeds using *Agrobacterium* (strain EHA105)—mediated transformation. Transgenic plants were confirmed by the presence of the hygromycin-resistant gene (*Hpt*) and transplanted into fields and planted under normal conditions. Verification of the gene insertion in transgenic plants was performed as previously described (Murray and Thompson 1980). Homozygous plants for each independent transgenic line (T₂) were used for further molecular and phenotypical analyses. The primers used for verification PCR are listed in Table 1. Leaf rolled index (LRI) was examined as previously described (Shao et al. 2005). In brief, the flag-leaf width of rice plants was surveyed in both a natural state (*Ln*) and an unfolding state (*Lw*), respectively, at the grain filling stage, and then LRI was calculated using the formula: $LRI = (Lw - Ln)/Lw \times 100\%$.

Primer design, sequence editing and alignment

PCR primers were designed using the online Primer-BLAST tool from the NCBI web site (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). DNASTar software (Burland 2000) was used for DNA sequence editing and processing. NCBI Blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used for sequence alignment and gene annotation.

Table 1 PCR primers used in the present study

Primer name	Primer sequence forward (5′–3′)	Primer sequence reverse (5′–3′)	Primer purpose
iA18	aaaaaa <u>ggatcc</u> gcaagtgaagtggagggtatc	aaaaaa <u>aagctt</u> cagtcgaagtctagccatcg	A18 cDNA synthesis, for RNAi vector construction
iA31	aaaaaa <u>ggatcc</u> ctggttgaggcgcaatg	aaaaaa <u>aagctt</u> gaagacggtagccagcaatg	A31 cDNA synthesis, for RNAi vector construction
iA38	aaaaaa <u>ggatcc</u> gccttagtgcctcaatc	aaaaaa <u>aagctt</u> ctttcaacagaatcggtgtg	A38 cDNA synthesis, for RNAi vector construction
Rev	cacctg <u>acgctg</u> tggttacttctgaagagg	actaga <u>ctgcag</u> cctcagatctaccatggtcg	Reverted cDNA amplification, for RNAi vector construction
oA20	aaaaaa <u>aagctt</u> gcaaacacacgctgatcgat	aaaaaa <u>ggatcc</u> ctgtgcaggtagctcctggat	A20 gene full-length cDNA synthesis, for A20 gene over-expression vector construction
oA46	aaaaaa <u>aagctt</u> tagcaaacagcaatgaactct	aaaaaa <u>ggatcc</u> cagctatctctcagctactt	A46 gene full-length cDNA synthesis, for A46 gene over-expression vector construction
oA50	aaaaaa <u>aagctt</u> ggacgacacgccattctcac	aaaaaa <u>ggatcc</u> gccatctcgcgttcttgctag	A50 gene full-length cDNA synthesis, for A50 gene over-expression vector construction
riA18	attccaaggacctcttgctt	gtggccatggcaactaagat	For A18 gene expression measurement in RNAi transgenic plants
riA31	tggctgccttatctgagct	tggatgaatgcaaggtggct	For A31 gene expression measurement in RNAi transgenic plants
riA38	gagtaggagctaataccgat	gaacaacacgagacgtatct	For A38 gene expression measurement in RNAi transgenic plants
roA20	gctcaactcaactcaact	gaaatttcttccaacattg	For A20 gene expression measurement in over-expression transgenic plants
roA46	ccgctgtagttgtgttta	caaattcttagcagctgaca	For A46 gene expression measurement in over-expression transgenic plants
roA50	cactttatggcttcacgta	acaagccgctgtagaagt	For A50 gene expression measurement in over-expression transgenic plants
Act	gacattcagcgttccagcatgat	tggagcttccatgccgatgagagaa	For <i>Actin</i> gene control in RT-PCR
Hpt	gacagcgtctccgacctgat	catgcctcgtccagctcaat	For the detection of hygromycin-resistant gene fragment

Sequences that are bold and underlined represent the recognition sites of restriction endonucleases; ggatcc, aagctt, acgctg, and ctgcag are recognition sequences of endonucleases *Bam*HI, *Hind*III, *Mlu*I and *Pst*I, respectively

Results

Characterization of the γ -*rl* mutant phenotype

Compared with wild-type QHZ, the γ -*rl* mutants had the erect and inward rolling phenotype (Fig. 1a) as we described previously (Yi et al. 2007). Moreover, the γ -*rl* mutants also had increased germination rates and decreased growth. In the early stages of germination (24 h), the bud length was longer in the γ -*rl* mutants than in the control (Fig. 1b, c), whereas at 72 h after germination, both the shoot and root lengths were significantly shorter in the γ -*rl* mutants (Fig. 1d). A further examination suggested that the differences in shoot and root mean lengths diminished gradually 96 h after germination (Fig. 1e, f).

Gene mutations in the γ -*rl* locus

In our previous study, we had identified that the γ -*rl* locus was located in the short arm of chromosome 3, which also includes an FMO coding gene, *OsFMO*_(t). Despite the known function of *OsFMO*_(t) in IAA synthesis, it remains

unclear whether it is also involved in the regulation of leaf rolling phenotype (Yi et al. 2007, 2013). Moreover, the details of the genetic variation information of that locus are not available (Yi et al. 2014; Wang et al. 2016). To understand the genetic variations of this locus and to identify other genes related to the γ -*rl* mutant phenotype, we performed a CGH analysis on the third chromosome, which harbors the γ -*rl* locus.

We designed high density DNA probes using the reference genome of chromosome 3 from the indica type rice, and hybridized the genomic DNA from γ -*rl* and QHZ against the chip to compare signal differences (Fig. 2). A comparison between the wild type QHZ and γ -*rl* (S1 Table) revealed a total of 49 hybridization signal alterations including deletions and insertions in chromosome 3 (Fig. 2b, c) indicating the presence of a significant number of gene mutations in the γ -*rl* locus. We previously mapped the γ -*rl* locus in the region between 70,306 and 121,519 bp in the OJ1607A12, a bacterial artificial chromosome (BAC) clone with chromosome 3 from the rice variety, Nipponbare (Yi et al. 2007). Using the Blast program to search the genome sequence database, we matched a region

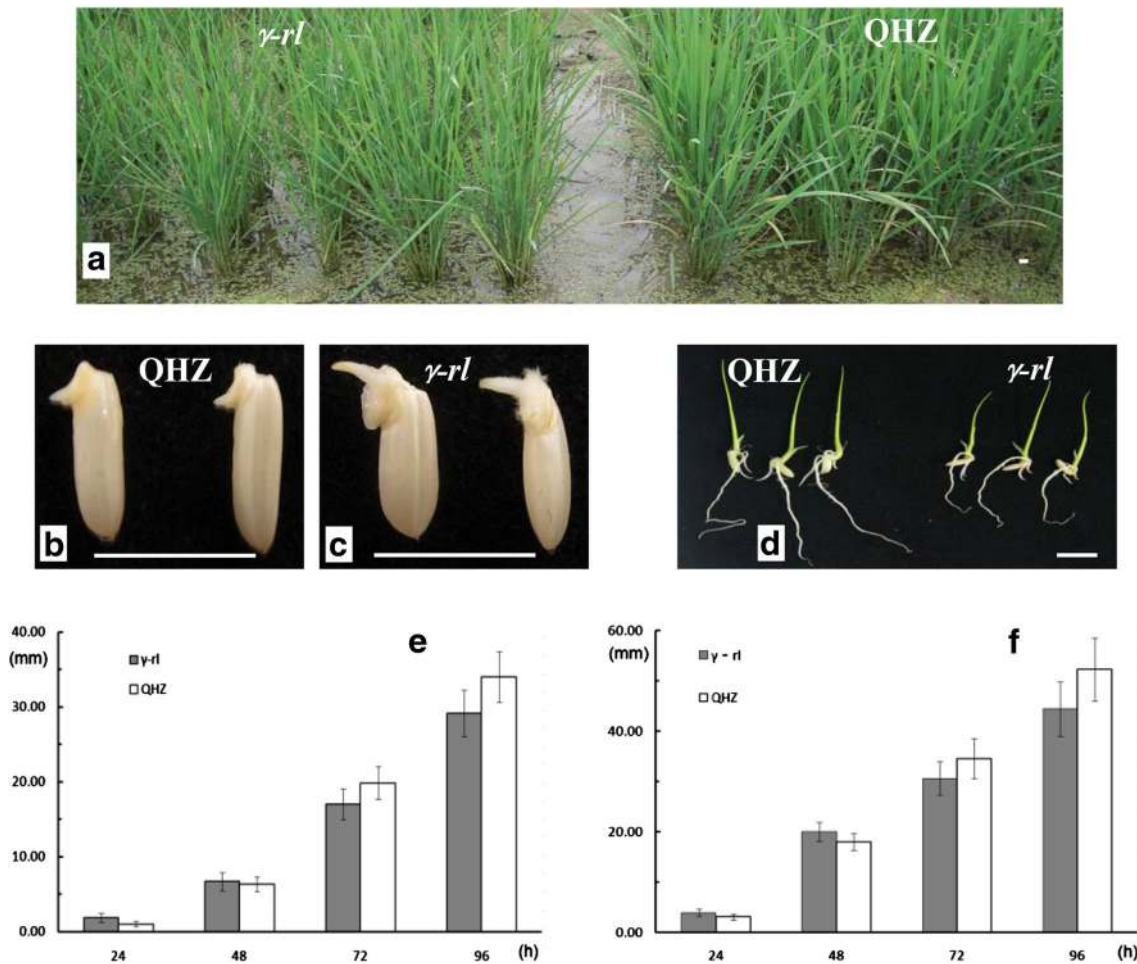


Fig. 1 Characterization of the γ -rl phenotype. **a** Gross morphology of tillering stage. *left: γ -rl, right: QHZ*. **b, c** 24 h after germination. **d** 72 h after germination. **e** Mean shoot length during different stages.

f Mean root length during different stages. Data represent mean (mm) \pm SD (n = 30). Bar = 1 cm

spanning 3,778,967–3,841,790 bp in the third chromosome of the rice variety, 93-11. The CGH data suggested that three major regions 5, 6 and 7 were significantly different in the γ -rl locus with a \log_2 ratio of 2.837, -3.111 and -1.592 , respectively (S1 Table). Region 5 covered the previously reported gene *OsFMO_(t)* (3,814,021–3,819,439 bp); thus, confirming the reliability of our data. Apart from the γ -rl locus, 46 additional regions had variations, suggesting the existence of massive genetic mutations in chromosome 3 after irradiation (S1 Table).

RNA-seq and qRT-PCR analysis

To further investigate the differences in gene expression between the γ -rl mutant and wild type QHZ, we used RNA-seq to analyze the mRNA transcriptome differences in earlier developmental phases (12 h) prior to significant phenotypic differences emerged (24 h) (Fig. 1b–f) during seeding growth period. RNA-seq analysis suggested that

12 h post germination, 1060 transcripts showed significant differences in expression between the wild type and mutant. Among these, 823 genes were up-regulated, and 237 genes were down-regulated in the γ -rl mutant. Analysis of 1060 genes via functional enrichment classification by GO in three categories including cellular location, molecular function, and biological process showed that these genes were involved in regulating the process of metabolism, cellular component biogenesis and organization, cell death, localization, reproduction and reproductive process, as well as anatomical structure formation (Fig. 3). Then, we searched the GenBank database via Blast and identified 72 annotated genes among the 49 hybridization signal alterations discovered in CGH. Forty-four of these 72 genes were also identified in the RNA-seq analysis (S1 Table). qRT-PCR validation of 44 genes showed that seven genes had significant changes in the mRNA levels in the γ -rl mutants. Among these, four were down-regulated [*OsI_10125* (LOC_Os03g06654), *OsI_14045*

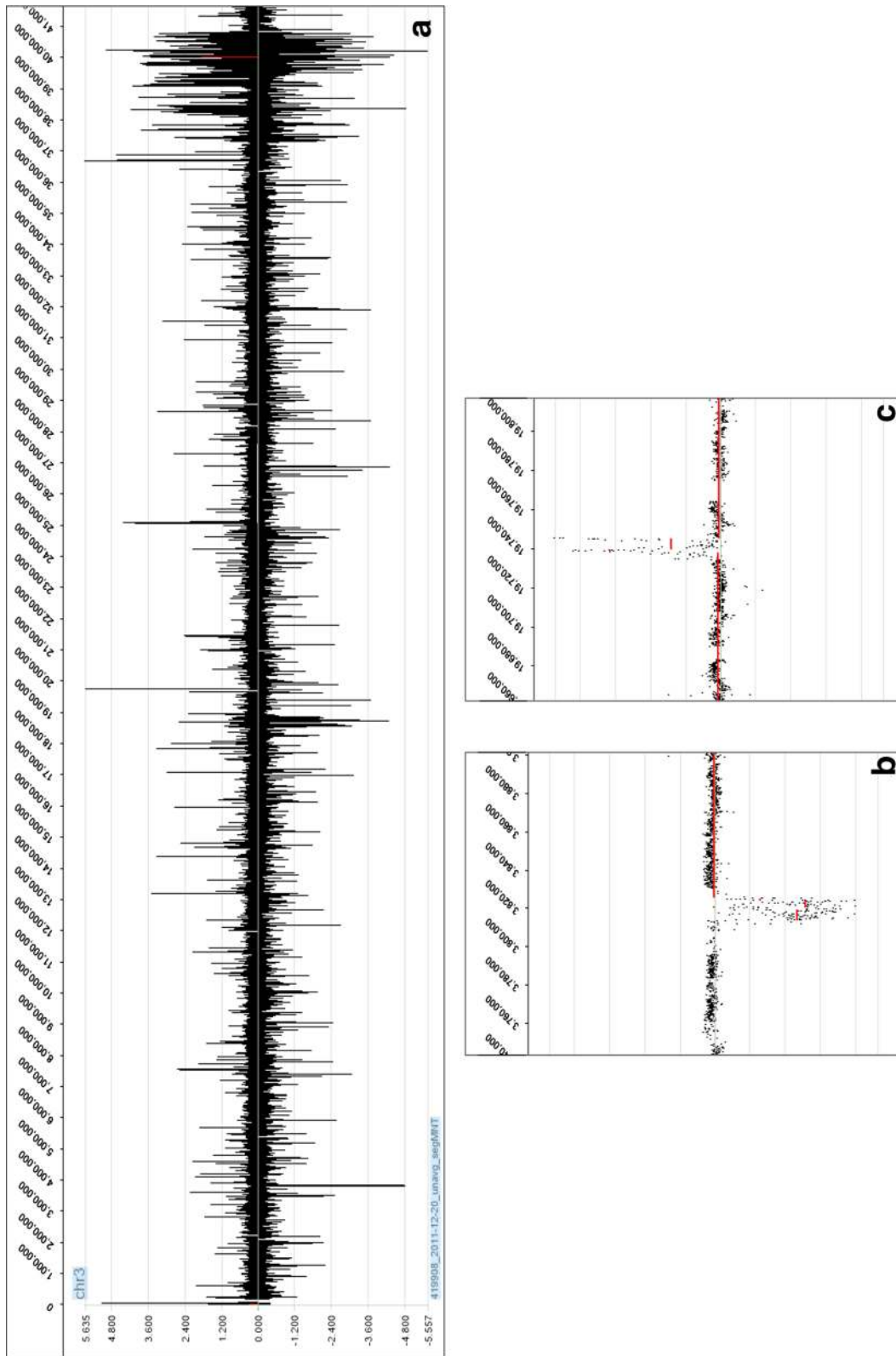
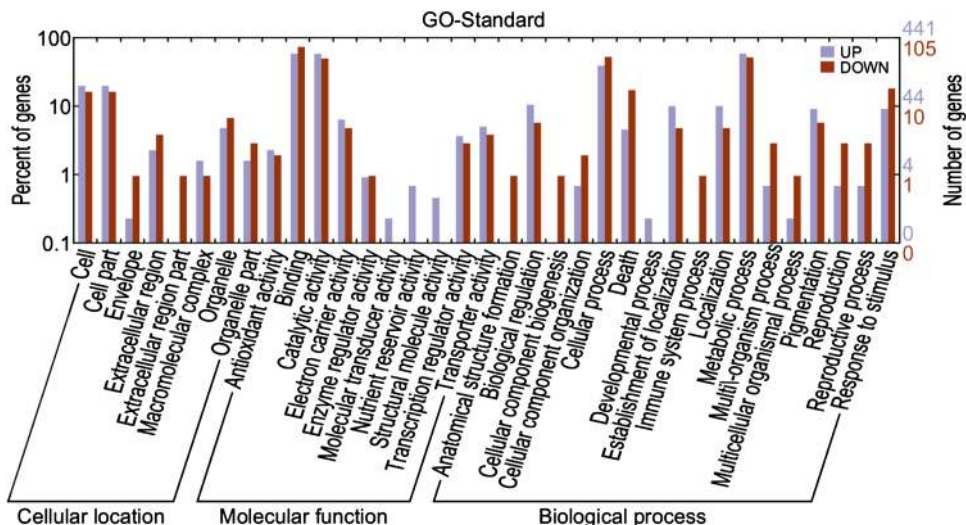


Fig. 2 Results of the CGH analysis using the SignalMap software. **a** Distribution of $\log_2(\gamma/r)/\text{QHZ}$ signals throughout chromosome 3. **b**, **c** Zoomed-in view of comprehensive regions with significant differences. **b** represents deletions with $\log_2(\gamma/r)/\text{QHZ}$ signal ratio ≤ -0.25 . **c** represents insertions with $\log_2(\gamma/r)/\text{QHZ}$ signal ratio ≥ 0.25 . Y axis represents the \log_2 of the signal ratio between γ - r and QHZ genomes; X axis represents genomic positions along chromosome 3

Fig. 3 Comparison of Gene Ontology (GO) classification analysis



(*LOC_Os03g62490*), *OsI_14279* (*LOC_Os03g62620*) and *OsI_14326* (*LOC_Os03g63250*), whereas the remaining three were up-regulated [(*OsI_10794* (*LOC_Os03g14950*), *OsI_11412* (*LOC_Os03g21250*) and *OsI_14152* (*LOC_Os03g61360*)] (S1 Table). Because *OsFMO₍₁₎* (*LOC_Os03g06654*) was identified previously (Yi et al. 2013), we will limit its discussion in this study. DNA sequencing of the above remaining six genes showed that in the γ -*rl* mutant, SNP and InDel variations occurred in the upstream 2.0 kb region and the first intron of *LOC_Os03g62490*. SNP variations also occurred in the upstream 2.0 kb region, the 5'-UTR and the first exon of *LOC_Os03g62620* (S2 Figure). Therefore, the above base mutations were probably the key factors leading to changes in gene expression level in the γ -*rl* mutant. These variations were not found in the sequences of the remaining four genes in the γ -*rl* mutant, suggesting the possibility of other unknown mutations in the mutant's genome, but the exact mechanism remains to be explored.

RNAi and over-expression analysis

In the RNAi group, knocking down *LOC_Os03g62620* (A31) significantly impaired gene expression and resulted in plants with rolled leaves, lower seed-setting rates, and decreased seed growth (Fig. 4a–h, o). As compared to the control, the average leaf rolling index (LRI) increased 138.00, 114.00, 70.00% while seed-setting rate decreased 40.52, 34.34, 33.74%, in three independent transgenic lines, A31-6, A31-29, A31-35, respectively (Fig. 4f–g). At 60 h after germination, both the shoot and root lengths of seeds of the three lines were significantly shorter than the control (Fig. 4h). Interestingly, plants transformed with the RNAi vector, *LOC_Os03g62490* (A18) and *LOC_Os03g63250* (A38) did not show any phenotypic

changes. Transgenic plants over-expressed with *LOC_Os03g14950* (A46) showed dwarf and shortened leaf phenotypes (Fig. 4i–n, p). The mean plant height was declined by 24.85, 31.25 and 21.18%, meanwhile, the mean flag-leaf length reduced 42.70, 48.93 and 36.72%, in three independent transgenic lines, A46-2, A46-15 and A46-17, respectively. However, ectopic expression of *LOC_Os03g21250* (A20) and *LOC_Os03g61360* (A50) did not result in phenotypic changes in the transgenic plants.

Discussion

Mutants are essential for functional genomics studies, and can be obtained via chemical or physical-induced mutagenesis. Chemical mutagenesis can be used to introduce point mutations while physical mutagenesis causes extensive mutations, which may give rise to multiple mutant phenotypes (Liu et al. 2009). For example, *mr21*, a mutant rice variety identified by γ irradiation, has both leaf color change and low temperature-sensitive phenotypes (Chen et al. 2010). Similarly, Ueguchi-Tanaka and his colleagues found numerous phenotypes including dwarfism and sterility in a ⁶⁰Co- γ -irradiated mutant rice variety, *gid1-4* (Ueguchi-Tanaka et al. 2005). In a previous study, we isolated the γ -*rl* mutant via ⁶⁰Co- γ irradiation of dry QHZ seeds (Yi et al. 2007). Besides well-known characters such as rolled reduced leaf width to increased tillers, we identified phenotypes with increased germination rate and decreased seed growth suggesting that γ irradiation can cause multiple phenotypes such as in the γ -*rl* mutant strain. We performed CGH analysis to investigate the genome-wide variations in chromosome 3, which contains the γ -*rl* locus. In addition to the previously characterized locus, we found 46 extra regions with significant mutations, which

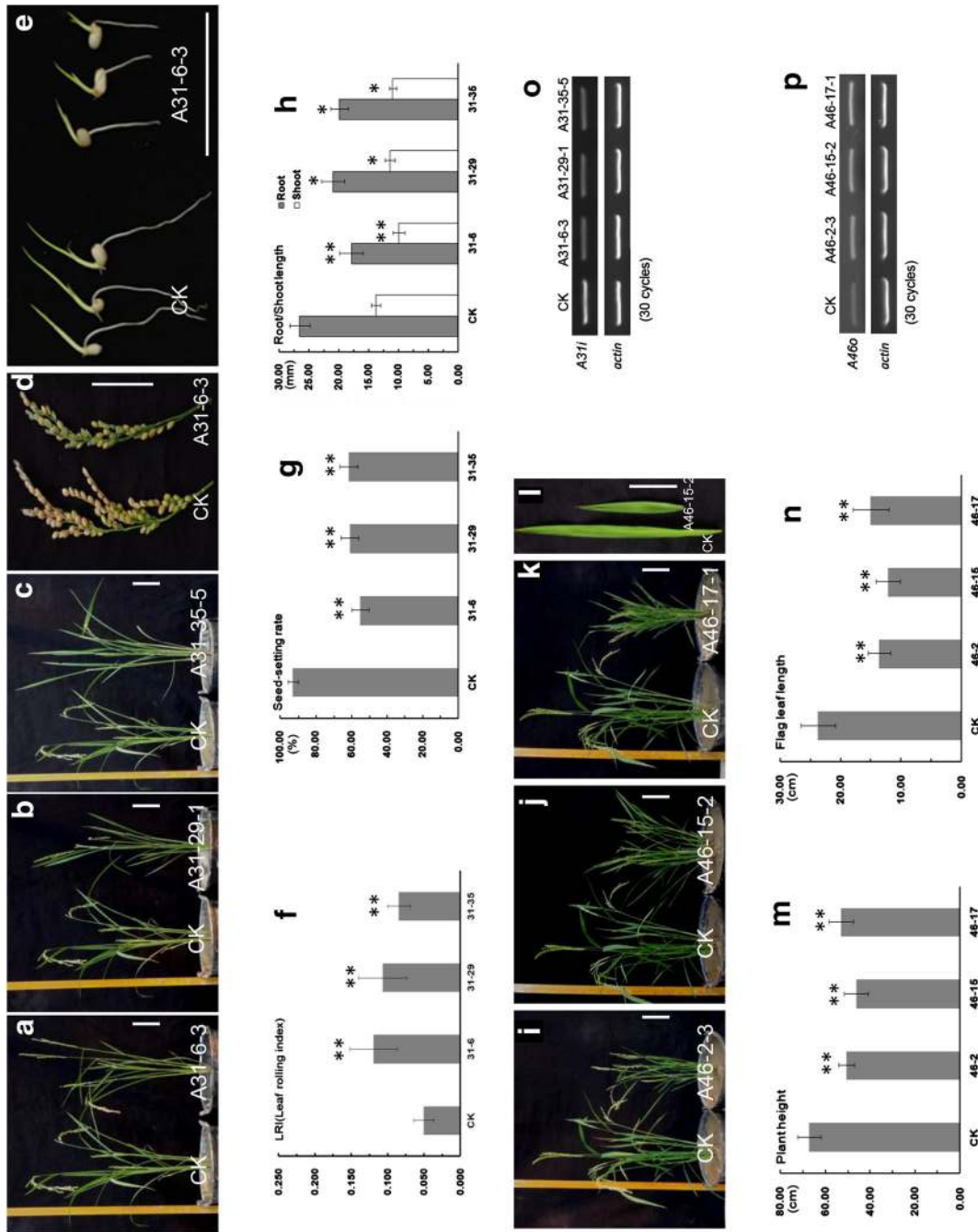


Fig. 4 Characterization of the transgenic plant phenotypes. **a, b, c, d, e** Representative gross morphology, spike and germinating seeds phenotype of *LOC_Os03g02620* (A31) RNAi transgenic plants (T_2) and controls. A31-6-3, A31-29-1 and A31-35-5 represented typical phenotypes of three independent RNAi transgenic lines, A31-6, A31-29 and A31-35 of *LOC_Os03g02620*. **f** Mean leaf rolling index (LRI) of the three transgenic lines, A31-6, A31-29, A31-25, and the control at grain filling stage ($n = 30$). **g, h** Mean seed-setting rate and root/shoot length 60 h after germination of the three transgenic lines at ripening stage. **i, j, k, l** Representative gross morphology and leaf phenotype of *LOC_Os03g14950* (A46) over-expression transgenic plants (T_2) and controls. A46-2-3, A46-15-2 and A46-17-1 represented typical phenotypes of three independent over-expression transgenic lines, A46-2, A46-15 and A46-17 of *LOC_Os03g02620*. **m, n** Mean plant height and flag leaf length of the three transgenic lines, A46-2, A46-15, A46-17, and the control at grain filling stage ($n = 30$). **o, p** RT-PCR analysis of independent transgenic plants with typical phenotypes transfected with A31i or A46o. The PCR products are 249 bp and 356 bp in length, respectively. The rice *Actin* gene was used as the internal positive control (300 bp). CK: Transgenic rice generated through empty vector control. Data represent mean (mm) \pm SD. * $P \leq 0.05$, ** $P \leq 0.01$, Student's *t* test. *Bar* = 5 cm

provides molecular insights into these phenotypes and suggests that γ -irradiation has a broad impact on the stability of rice genome. While the 46 extra regions explain the multiple phenotypes observed in the γ -*rl* mutants, further investigation is required to understand whether similar mutations occur in other chromosomes as well.

CGH analysis of genome-wide genomic alterations between mutant and wild type are extensively applied for the identification of novel genes and unraveling the molecular mechanisms regulating these phenotypes (Myllykangas et al. 2008; Yang et al. 2008). Here, we analyzed the genomic variations on the third chromosome between γ -*rl* and QHZ using CGH. We identified 49 regions with significant differences including genomic deletion and insertions and 72 annotated genes within the regions. RNA-seq technology can detect differences in mRNA expression level, and combining it with CGH can help identify mutation-regulating genes (Choi et al. 2014). Thus, we used RNA-seq to analyze the mRNA transcriptome differences between γ -*rl* and QHZ during early developmental phases (12 h) before phenotypic differences were manifested (24 h) during the seed growth period (Fig. 1b–f). This analysis led to the identification of 1060 differentially expressed genes. We further narrowed down the candidate gene number to seven by cross-referencing with CGH data, and validating with qRT-PCR and resequencing based on PCR amplification. The in vivo functional analysis in transgenic plants showed that two genes were related to the γ -*rl* mutant phenotype: *LOC_Os03g62620* and *LOC_Os03g14950*. Moreover, we found that the knock down of *LOC_Os03g62620* (A31) resulted in rolled leaf, lower seed-setting rates, and decreased seed growth. *LOC_Os03g62620* encodes a late embryogenesis abundant protein (LEA), which plays essential roles in seeding development and tolerance to abiotic stresses (Duan and Cai 2012). Finally, we showed that plants with ectopic expression of *LOC_Os03g14950* (A46)—a lipid transfer protein-coding gene—displayed dwarf and shortened leaf phenotypes. Recently, *LOC_Os03g14950* was shown to play a critical role in regulating lipid metabolism, signal transduction, reproduction and pollen development although the underlying mechanisms remain unclear (Lev et al. 2013). Thus, our data suggested that the two genes, *LOC_Os03g62620* encoding a late embryogenesis abundant protein, and *LOC_Os03g14950* encoding a lipid transfer protein, are essential for seed growth and leaf development. Given that all these genes are related to the γ -*rl* mutant phenotype, it would be interesting to investigate their individual molecular mechanisms further.

Acknowledgements This research was funded by Grants from the National Natural Sciences Foundation of China (Nos. 30671279 and

31071070). We would like to thank Mr. Jodelric Cruz for his kindly help in English modification and correction.

References

- Brady PD, DeKoninck P, Fryns JP, Devriendt K, Deprest JA, Vermeesch JR (2013) Identification of dosage-sensitive genes in fetuses referred with severe isolated congenital diaphragmatic hernia. *Prenat Diagn* 33:1283–1292
- Burland TG (2000) DNASTAR's lasergene sequence analysis software. *Methods Mol Biol* 132:71–91
- Byeon JH, Shin E, Kim GH, Lee K, Hong YS, Lee JW et al (2014) Application of array-based comparative genomic hybridization to pediatric neurologic diseases. *Yonsei Med J* 55:30–36
- Cahan P, Li Y, Izumi M, Graubert TA (2009) The impact of copy number variation on local gene expression in mouse hematopoietic stem and progenitor cells. *Nat Genet* 41:430–437
- Chen JY, Zhao J, Liu X, Li C, Lin DZ, Dong YJ et al (2010) Genetic analysis and molecular mapping of a new thermo sensitive leaf-color mutant in *Oryza sativa*. *Chin Bull Bot* 45(4):419–425
- Cho SH, Yoo SC, Zhang H, Pandeya D, Koh HJ, Hwang JY et al (2013) The rice narrow leaf2 and narrow leaf3 loci encode WUSCHEL-related homeobox 3A (OsWOX3A) and function in leaf, spikelet, tiller and lateral root development. *New Phytol* 198(4):1071–1084
- Choi S, Jung J, Jeon CO, Park W (2014) Comparative genomic and transcriptomic analyses of NaCl-tolerant *Staphylococcus* sp. OJ82 isolated from fermented seafood. *Appl Microbiol Biotechnol* 98:807–822
- Duan J, Cai W (2012) OsLEA3-2, an abiotic stress induced gene of rice plays a key role in salt and drought tolerance. *Plos ONE* 7(9):e45117
- Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altschuler DM et al (2006) Copy number variation: new insights in genome diversity. *Genome Res* 16(8):949–961
- Fujino K, Matsuda Y, Ozawa K, Nishimura T, Koshiba T, Fraaije MW et al (2008) NARROW LEAF 7 controls leaf shape mediated by auxin in rice. *Mol Genet Genomics* 279(5):499–507
- Glaesener S, Honecker F, Veltman IM, Gillis AJ, Rohlfing T, Streichert T et al (2008) Comparative proteome, transcriptome, and genome analysis of a gonadal and an extragonadal germ cell tumor cell line. *J Proteome Res* 7:3890–3899
- Hegde MR, Chin EL, Mülle JG, Okou DT, Warren ST, Zwick ME (2008) Microarray-based mutation detection in the dystrophin gene. *Hum Mutat* 29:1091–1099
- Hu J, Zhu L, Zeng D, Gao Z, Guo L, Fang Y et al (2010) Identification and characterization of NARROW AND ROLLED LEAF 1, a novel gene regulating leaf morphology and plant architecture in rice. *Plant Mol Biol* 73:283–292
- Jung SH, Shin SH, Yim SH, Choi HS, Lee SH, Chung YJ (2009) Integrated analysis of copy number alteration and RNA expression profiles of cancer using a high-resolution whole-genome oligonucleotide array. *Exp Mol Med* 41(7):462–470
- Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F et al (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258:818–821
- Kalra S, Puniya BL, Kulshreshtha D, Kumar S, Kaur J, Ramachandran S et al (2013) De novo transcriptome sequencing reveals important molecular networks and metabolic pathways of the plant *Chlorophytum borivilianum*. *Plos ONE* 8:e83336
- Khoury T, Hu Q, Liu S, Wang J (2014) Intracystic papillary carcinoma of breast: interrelationship with in situ and invasive carcinoma and a proposal of pathogenesis: array comparative

- genomic hybridization study of 14 cases. *Mod Pathol* 27(2):194–203
- Lambert SR, Mladkova N, Gulati A, Hamoudi R, Purdie K, Cerio R et al (2014) Key differences identified between actinic keratosis and cutaneous squamous cell carcinoma by transcriptome profiling. *Br J Cancer* 110:520–529
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25
- Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BM et al (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29(8):1035–1043
- Lev S, Kim S, Kedan A, Selitrennik M, Keinan O (2013) The role of Lipid transfer proteins in membrane trafficking and signal transduction. *FASEB J* 27:331–336
- Liu LX, Guo HJ, Zhao LS, Li JH, Gu JY, Zhao SR et al (2009) Current status and outlook perspectives of induced mutations for plant improvement. *J Nucl Agric Sci* 23(6):1001–1007
- Liu Z, Ji T, Yin L, Shen J, Shen F, Chen G (2013) Transcriptome sequencing analysis reveals the regulation of the hypopharyngeal glands in the honey bee *Apis mellifera carnica* Pollmann. *Plos ONE* 8:e81001
- Ma GH, Yuan LP (2015) Hybrid rice achievements, development and prospect in China. *J Integr Agric* 14(2):197–205
- Marques-Bonet T, Girirajan S, Eichler EE (2009) The origins and impact of primate segmental insertions. *Trends Genet* 25:443–454
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621–628
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8(19):4321–4325
- Myllykangas S, Junnila S, Kokkola A, Autio R, Scheinin I, Kiviluoto T et al (2008) Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes. *Int J Cancer* 123:817–825
- Naseer MI, Faheem M, Chaudhary AG, Kumosani TA, Al-Quaiti MM, Jan MM et al (2015) Genome wide analysis of novel copy number variations duplications/deletions of different epileptic patients in Saudi Arabia. *BMC Genomics* 16(Suppl):S10
- Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* 12(2):87–98
- Shao YJ, Pan CH, Chen ZX, Zuo SM, Zhang YF, Pan XB (2005) Fine mapping of an incomplete recessive gene for leaf rolling in rice (*Oryza sativa* L.). *Chin Sci Bull* 50(21):2466–2472
- Shi Z, Wang J, Wan X, Shen G, Wang X, Zhang J (2007) Overexpression of rice OsAGO7 gene induces upward curling of the leaf blade that enhanced erect-leaf habit. *Planta* 226(1):99–108
- Ueguchi-Tanaka M, Ashikari M, Nakajima M, Itoh H, Katoh E, Kobayashi M et al (2005) GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* 437:693–698
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63
- Wang FH, Mei MT, Yi JC (2016) Analysis of genomic deletion of a rice rolled-leaf mutant mutagenized by $^{60}\text{Co-}\gamma$ ray. *J Nucl Agric Sci* 30(3):413–421
- Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y (2009) Rice Micro RNA effectors complexes and targets. *Plant Cell* 21(11):3421–3435
- Xiang JJ, Zhang GH, Qian Q, Xue HW (2012) *SEMI-rolled leaf1* encodes a putative glycosylphosphatidylinositol- anchored protein and modulates rice leaf rolling by regulating the formation of bulliform cells. *Plant Physiol* 159(4):1488–1500
- Yan CJ, Yan S, Zhang ZQ, Liang GH, Lu JF, Gu MH (2006) Genetic analysis and gene fine mapping for a rice novel mutant (*rl9_(rl)*) with rolling leaf character. *Chin Sci Bull* 51(1):63–69
- Yang TL, Chen XD, Guo Y, Lei SF, Wang JT, Zhou Q et al (2008) Genome-wide copy-number-variation study identified a susceptibility gene, UGT217, for osteoporosis. *Am J Hum Genet* 83(6):663–674
- Yi JC, Zhuang CX, Wang XJ, Cao YP, Liu YG, Mei MT (2007) Genetic analysis and molecular mapping of a rolling leaf mutation gene in rice. *J Integr Plant Biol* 49:1746–1753
- Yi J, Liu L, Cao Y, Li J, Mei M (2013) Cloning, characterization and expression of *OsFMO_(rl)* in rice encoding a flavinmonooxygenase. *J genet* 92:471–480
- Yi JC, Cao YP, Mei MT (2014) Characterization of a $^{60}\text{Co-}\gamma$ mutated rolled-leaf mutant in rice. *J Nucl Agric Sci* 28(5):757–764
- Yu L, Jiang J, Zhang C, Jiang L, Ye N, Lu Y et al (2010) Glyoxylate rather than acerbate is an efficient precursor for oxalate biosynthesis in rice. *J Exp Bot* 61(6):1625–1634