


ORIGINAL ARTICLE

Open Access



# Production Assessment and Genome Comparison Revealed High Yield Potential and Novel Specific Alleles Associated with Fertility and Yield in Neo-Tetraploid Rice

Hang Yu<sup>1,2,3†</sup>, Muhammad Qasim Shahid<sup>1,2,3†</sup>, Qihang Li<sup>1,2,3</sup>, Yudi Li<sup>1,2,3</sup>, Cong Li<sup>1,2,3</sup>, Zijun Lu<sup>1,2,3</sup>, Jinwen Wu<sup>1,2,3</sup>, Zemin Zhang<sup>1,2,3</sup> and Xiangdong Liu<sup>1,2,3,4\*</sup> 

## Abstract

**Background:** Neo-tetraploid rice (NTR) is a new tetraploid rice germplasm that developed from the crossing and directional selection of different autotetraploid rice lines, which showed high fertility and promising yield potential. However, systematic yield assessment, genome composition and functional variations associated with fertility and yield remain elusive.

**Results:** Two seasons' field trials of 15 NTRs and 27 autotetraploid rice (ATR) lines revealed that the improvement of YPP (yield per plant, 4.45 g increase) were significantly associated with the increase of SS (seed setting, 29.44% increase), and yield and seed setting of NTRs improved significantly compared to parental lines. Whole genome resequencing of 13 NTR sister lines and their parents at about 48.63 depth were conducted and genome compositions were illustrated using inherited chromosomal blocks. Interestingly, 222 non-parental genes were detected between NTRs and their low fertility parental lines, which were conserved in 13 NTRs. These genes were overlapped with yield and fertility QTLs, and RNA-Seq analysis revealed that 81 of them were enriched in reproductive tissues. CRISPR/Cas9 gene knockout was conducted for 9 non-parental genes to validate their function. Knockout mutants showed on an average 25.63% and 4.88 g decrease in SS and YPP, respectively. Notably, some mutants showed interesting phenotypes, e.g., *kin7l* (kinesin motor gene) and *kin14m* (kinesin motor gene), *bzr3* (BES1/BZR1 homolog) and *nrfg4* (neo-tetraploid rice fertility related gene) exhibited 44.65%, 24.30%, 24.42% and 28.33% decrease in SS and 8.81 g, 4.71 g, 5.90 g, 6.22 g reduction in YPP, respectively.

**Conclusion:** Comparative genomics provides insights into genome composition of neo-tetraploid rice and the genes associated with fertility and yield will play important role to reveal molecular mechanisms for the improvement of tetraploid rice.

**Keywords:** *Oryza sativa*, Neo-tetraploid rice, Non-parental variation, Fertility, Yield

\* Correspondence: [xdliu@scau.edu.cn](mailto:xdliu@scau.edu.cn)

†Hang Yu and Muhammad Qasim Shahid contributed equally to this work.

<sup>1</sup>State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou 510642, China

<sup>2</sup>Guangdong Provincial Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou 510642, China

Full list of author information is available at the end of the article

## Background

Rice is an important main food source for world's population, and the yield of rice has long been the focus of breeders to feed growing population. The enrichment of genetic diversity can largely facilitate the improvement of rice cultivars (Huang et al. 2012; Wang et al. 2018). The phenomenon of whole-genome duplication is widespread throughout plant evolution (Soltis et al. 2007; Wood et al. 2009), and polyploid plants showed increased genetic variation that may provide a way to extend genetic diversity and breeding of elite cultivars (Wendel 2000; Soltis et al. 2009; Van de Peer et al. 2017).

Autotetraploid rice is a useful germplasm that generated by artificial genome duplication of diploid rice, which had biological advantages and wide adaptability (Shahid et al. 2011; Shahid et al. 2012; Wu et al. 2013; Tu et al. 2014; Yang et al. 2014). However, low fertility limited the yield performance of autotetraploid rice and its commercial utilization (Shahid et al. 2010; He et al. 2011; Shahid et al. 2013; Wu et al. 2014; Wu et al. 2015; Li et al. 2017; Chen et al. 2018b; Li et al. 2018; Ghouri et al. 2019). To solve this issue, rice scientists had made unremitting efforts for many years, and bred some new autotetraploid rice lines with high fertility (Tu et al. 2003; Luan et al. 2007; Cai et al. 2007; Guo et al. 2016; Guo et al. 2017). New male sterile lines, restorer lines and their F<sub>1</sub> autotetraploid rice hybrids produced more than 70% seed setting (Tu et al. 2003; Luan et al. 2007). Cai et al. (2007) developed two polyploid rice lines with more than 65% seed setting. Two neo-tetraploid rice lines, Huaduo1 and Huaduo2 with high fertility (> 80% seed setting), were devolved from the crossing of autotetraploid rice by our research group in 2012 (Guo and Liu 2014). Other neo-tetraploid rice lines or hybrids, including Huaduo3, Huaduo4, Huaduo5, Huaduo8 and three F<sub>1</sub> hybrids with high yield, including T449 × H1, T485 × H8 and H1 × H8, were also developed in recent years (Guo et al. 2017; Bei et al. 2019; Chen et al. 2019; Ghaleb et al. 2020). In contrast to autotetraploid and allotetraploid rice, neo-tetraploid rice (NTR) will play an important role in polyploid rice breeding because of high seed setting (> 70%), wide compatibility genes and stable fertility in next generations, like neo-tetraploid *Arabidopsis* (Yu et al. 2009; Guo and Liu 2014; Bei et al. 2019; Chen et al. 2019). Therefore, the systematic investigation of yield improvement, and the detection of specific DNA variations of neo-tetraploid rice will be helpful to decipher the high yield mechanism and optimize the directions of tetraploid rice breeding.

Next-generation sequencing is widely used for the detection of genomic variations and genome comparison analysis. Genome re-sequencing of three cultivars (sensitive cultivar IR64, drought resistance cultivar Nagina22,

and salinity resistance cultivar Pokkali) with contrasting drought and salinity stress responses displayed 232 variant genes in known QTLs that may regulate abiotic stress (Jain et al. 2013). Genome re-sequencing of bulked populations facilitate rapid mapping of QTLs, and QTL-seq can be used to identify QTLs for important agronomic traits using F<sub>2</sub> populations (Hiroki et al. 2013). Recently, the analysis of genome re-sequencing data of 3010 Asian cultivated rice accessions revealed about 30 million genomic variations, and the pan-genome was constructed, which provided a resource for rice genomics research and breeding (Wang et al. 2018). Deep genome re-sequencing provided high quality sequencing reads and reliable genomic variations, which allow the genome comparisons method to detect variation transmission patterns in pedigrees and identify the inherited blocks (Zhou et al. 2016). Whole genome sequencing of 30 cultivars revealed 28 chromosomal blocks inherited from ancestral lines that shared by all high-yielding cultivars (Huang et al. 2018). High resolution graphical genotype map was constructed for new upland African rice, which indicated similar genome constitution and common genetic events, and showed potential genes related to important agronomic traits by genome re-sequencing (Yamamoto et al. 2018).

In the present study, 18 morphological traits of 15 neo-tetraploid (NTRs) and 27 autotetraploid rice lines (ATRs) were observed during two seasons to illustrate the traits improvement between NTRs and their parental lines. Moreover, genome re-sequencing was employed to analyze the global DNA variations among 13 neo-tetraploid rice lines and their parental lines. Together with gene function and gene expression analysis, we planned to investigate the chromosomal composition map and to mine the functional variations that associated with fertility in neo-tetraploid rice. The fertility related genes in neo-tetraploid rice were knockout using CRISPR/Cas9 system, which validated the gene function and provided fundamental mutants for functional genomics research in tetraploid rice.

## Results

### Production Assessment of Neo-Tetraploid Rice (NTR)

Firstly, we evaluated yield and yield related traits in 15 NTRs (Figure S1) and 27 autotetraploid rice (ATRs) lines, and the results showed that seed setting (i.e. average 65.32% in NTRs and 35.88% in ATRs), yield per plant (i.e. average 10.25 g in NTRs and 5.80 g in ATRs) and grain numbers per panicle (i.e. average 115.57 in NTRs and 93.31 in ATRs) improved significantly in NTRs (Table S1a and b). In order to confirm the factor that responsible for yield improvement in NTRs, correlation analysis between yield-related traits and yield per plant of 15 NTRs and 27 ATRs were conducted. The

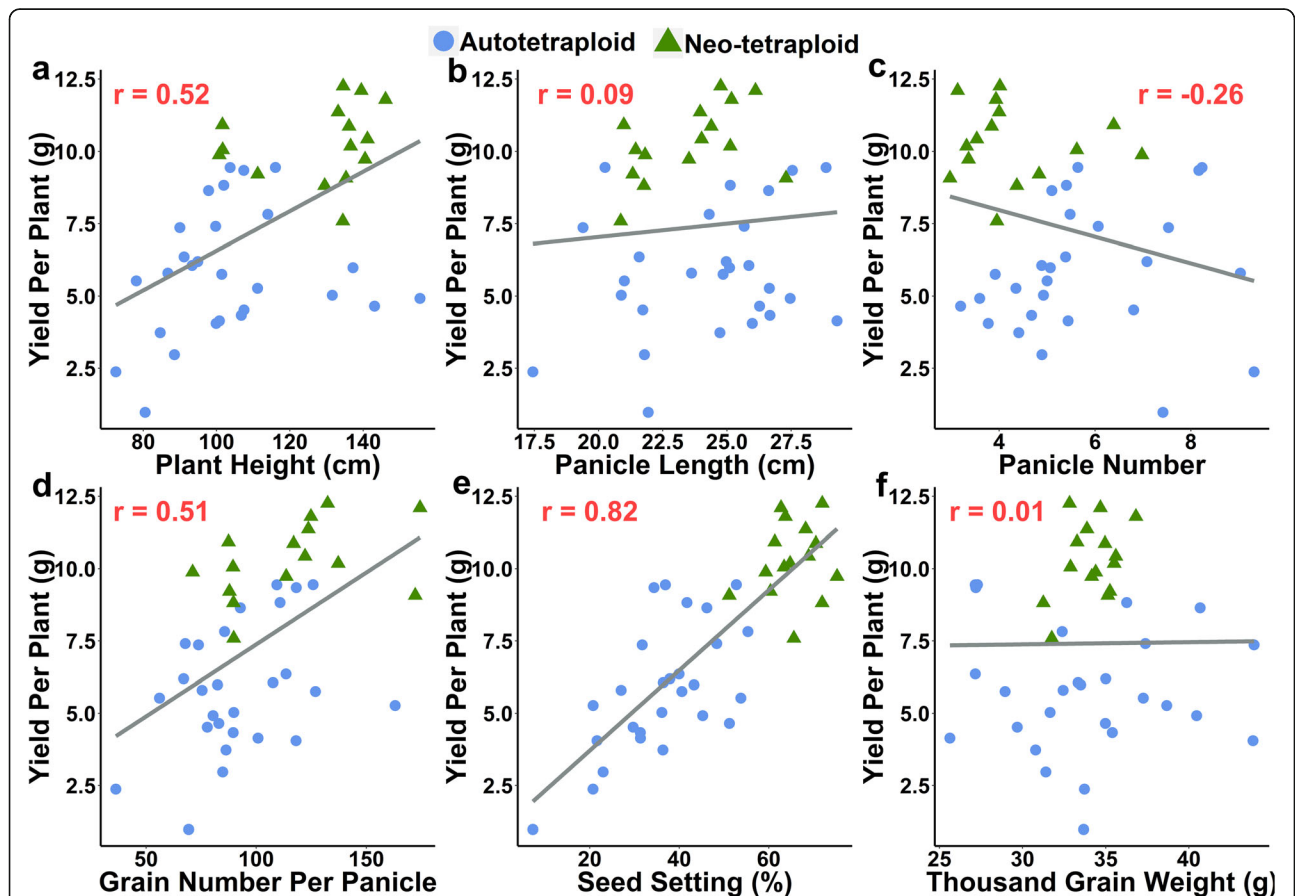
correlation coefficient was ranged from -0.26 to 0.82, and seed setting (SS) showed the highest correlation coefficient (0.82), which indicated that SS contributed greatly to improve the yield in NTRs (Fig. 1).

Secondly, we compared NTRs with a famous diploid rice, Huanghuazhan (E285) and its autotetraploid rice, Huanghuazhan-4x (T485), for 18 agronomic traits during two seasons in the field. The results demonstrated that seed setting (SS) was the most improved trait in NTRs, and the average SS was significantly higher in NTRs than that in T485, and the highest SS (76.96%) was produced by Huaduo25100 during early season in 2018 (Figure S1 and Figure S2, Table S1a). The grain yield was significantly higher in NTRs than T485, and the average yield per plant (YPP) and yield per block (YPB) of NTRs were 10.25 g and 1.05 kg, while T485 produced 6.06 g and 0.59 kg, respectively. NTRs also showed improvement in grain number per panicle (115.57 in NTRs and 109.14 in T485), dry weight (32.52 g in NTRs and 29.83 in T485) and harvest index (0.32 in NTRs and 0.21 in T485). The yield of NTRs was not higher than diploid rice E285, however, yield of two

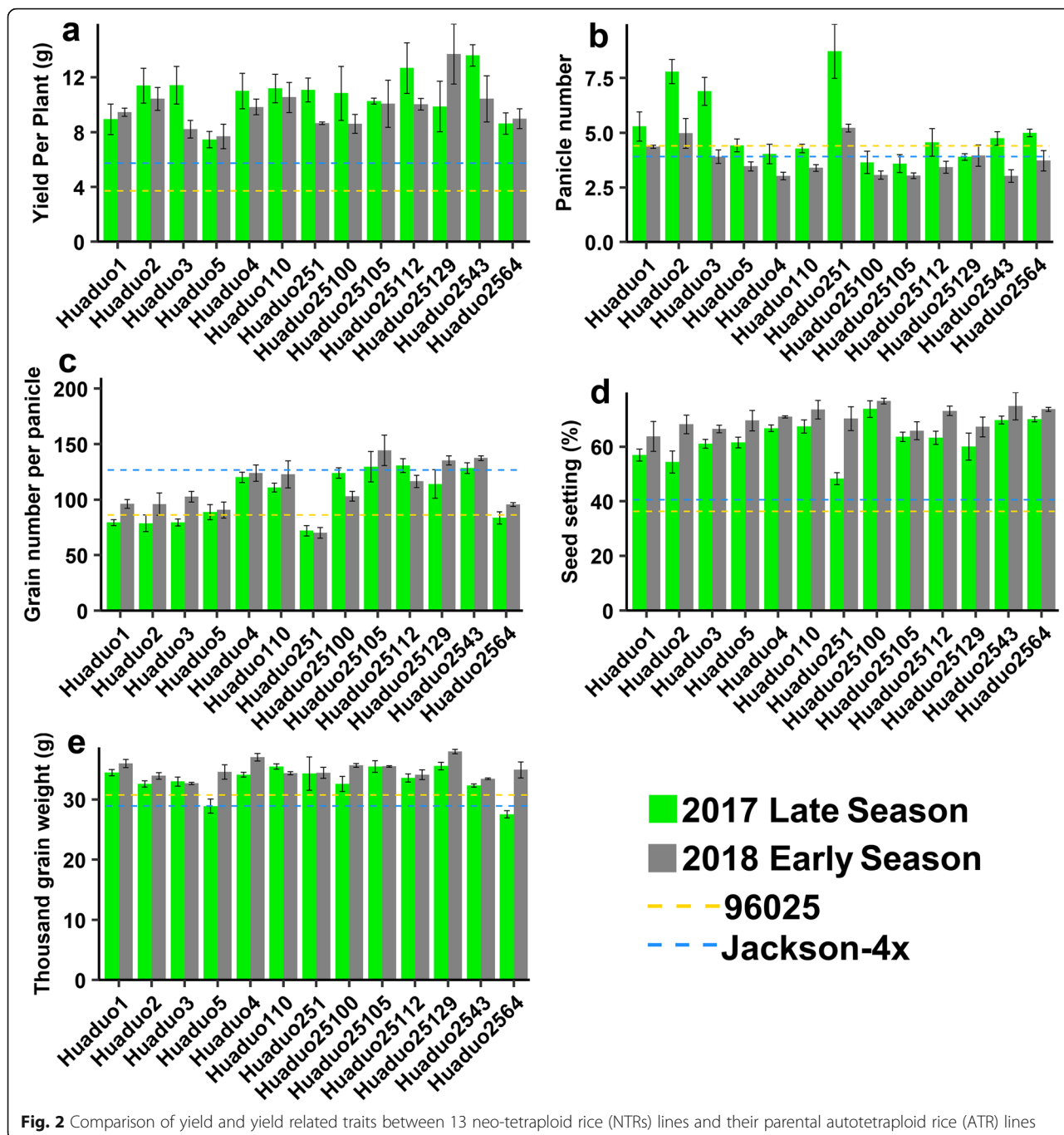
hybrids, 4001-4x/Huaduo8 and Yangeng48-4x/Huaduo8, reached 26.52 g and 20.60 g, which was higher than E285 (20.44 g) (Table S1a, c).

Moreover, we also compared the field performance of NTRs under two different planting methods, viz. one seedling per hole (OSPH) and three seedlings per hole (TSPH). Under the planting method of TSPH, NTRs showed an earlier heading (2.89 days), higher plant height (2.67 cm), more number of panicles (0.44), higher dry weight (1.46 g), average yield per plant (0.5 g) and higher average yield per block (39.6 g) than OSPH (Figure S2).

Then, we compared the yield and yield components between 13 NTR sister lines and their parental lines (Figure S3), and NTRs displayed significant improvement in yield per plant, seed setting and thousand grain weight. The average grain yield per plant increased by 5.652 g in NTRs than their parental lines, which was mainly contributed by the improvement of seed setting (34.225% higher SS than their parental lines). Additionally, NTRs also increase in thousand grain weight by 4.726 g compared to their parental lines (Fig. 2). The



**Fig. 1** Distribution and correlation analysis of yield and yield related traits of 15 neo-tetraploid and 27 autotetraploid rice lines. Pearson correlation coefficient (*r*) is denoted by red font. The values of traits in this figure are the mean values of late season (2017) and early season (2018)



significant improvement between NTRs and their parental lines prompted us to mine the genomic variations among them.

**Genome Re-Sequencing and Variation Detection in Neo-Tetraploid Rice**

A total of 13 neo-tetraploid sister lines and their parental autotetraploid rice lines, Jackson-4x and 96025, were deeply sequenced by about 48.63 depth using genome re-sequencing, which generated about one billion high

quality pair-end sequencing reads with an average ratio of Q30 score of 94.54% (Table S2), and the average genome coverage ratios for all samples were 95.48% of Nipponbare (MSU7) reference genome (Table S3). After removing of low-quality variations, with an average of about 1,849,640 variations were detected between these materials and reference genome (Table S4).

To test the reliability of these variations, two strategies were used to validate the variations. Firstly, the variations were validated by our previous genome



sequencing data. Eleven individuals of tetraploid lines used in this study have been sequenced previously, the variation data were compared with the newly sequenced individuals, and the variation accordance rate was ranged from 83.14% to 94.39% with an average of 91.02% (Table S5). Secondly, a total of 255 variations were validated using Sanger sequencing and 245 variations were in accordance with the re-sequencing data, and the accordance rate was 96.08% (Table S6). These results indicated that sequencing data and the detected variations were qualified and the variations are reliable.

#### Genome Composition and Detection of Non-parental Chromosomal Blocks in Neo-Tetraploid Rice

Genome composition analysis provided the information about inherited chromosome segments and the novel variations in neo-tetraploid rice lines. Graphical map showed the inherited genome composition of 13 NTR lines (Fig. 3, Figure S4). Most of the chromosomal blocks were inherited from either maternal (green) or paternal (gold) blocks in 12 chromosomes of NTRs. Interestingly, in chromosome 5, almost all chromosomal blocks were inherited from maternal line (96025) (Figure S4e), but in chromosome 10, almost all chromosomal blocks were inherited from paternal (Jackson-4x) (Figure S4i). We also observed non-parental chromosome blocks in 13 NTRs and the distribution of these blocks were similar among NTRs. A total of 140 non-parental chromosomal blocks that shared by all 13 NTRs were detected, which covered 1.4 Mb of rice genome (Table S7). Particularly, similar distribution was detected on chromosome 6 and 11, and almost same non-parental continuous blocks were identified in all 13 NTRs (Fig. 3a, b).

#### Function and Expression Analysis of Non-parental Alleles in Neo-Tetraploid Rice

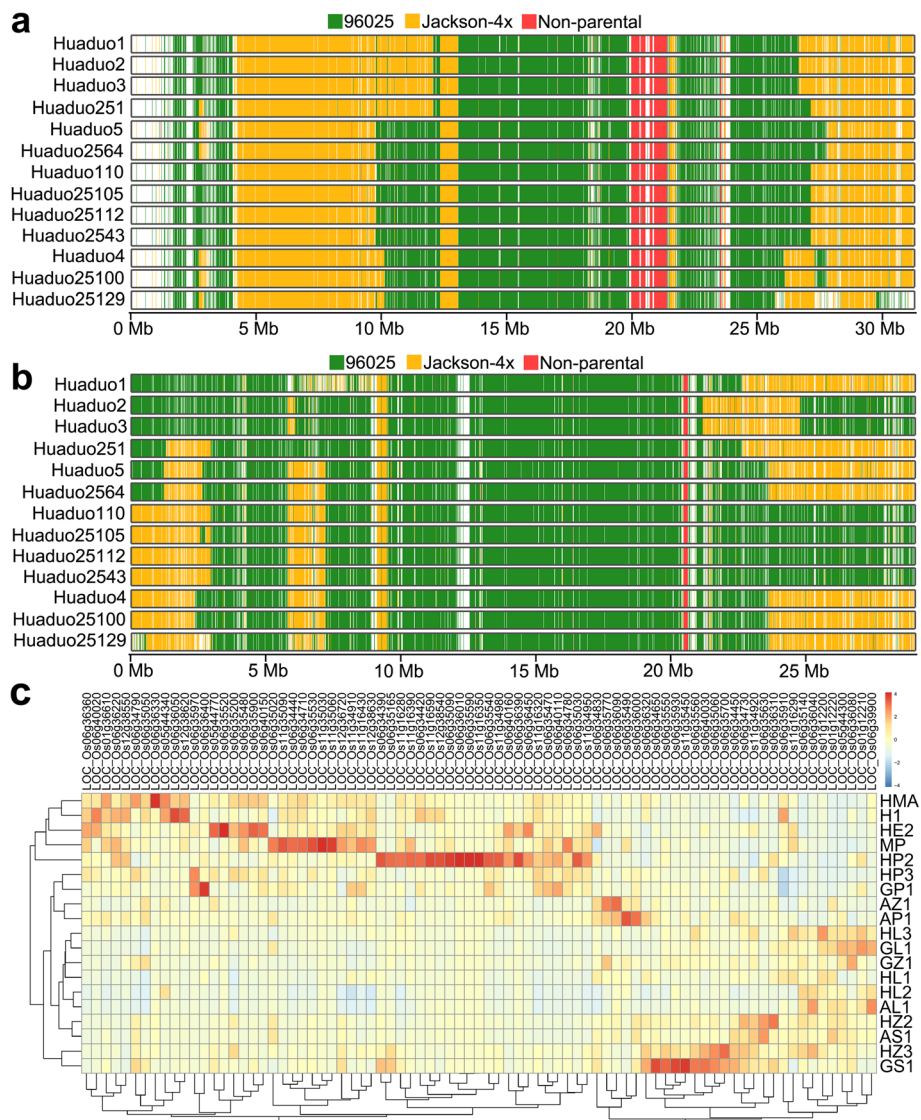
In these non-parental blocks, a total of 1080 nonsynonymous variations were detected in 222 genes. GO enrichment analysis revealed that these genes were enriched in FAD binding (Table S8). These variations were overlapped with the yield and fertility related QTLs. A total of 28 QTLs were overlapped with these variations, and 10 of them were yield related QTLs (such as biomass yield, spikelet number, panicle number and filled grain percentage) and 1 of them was sterility or fertility related QTL (Table S9). Results of QTL analysis indicated that non-parental variations maybe the key regions that controlled the yield and fertility in neo-tetraploid rice. Gene expression pattern analysis of these non-parental genes were conducted using two methods to further infer the function of these genes. Firstly, the expression levels of these genes were checked using RiceXPro database, and the expression information of 87

genes were recorded in this database, and 76 out of 87 genes were found to be expressed in the reproductive tissues including anther, pistil, ovary and embryo (Table S10). Secondly, using transcriptome data of neo-tetraploid rice line, Huaduo1, 81 genes were identified to be expressed in anther, ovary or flag leaf during meiosis or anthesis stage (Fig. 3c, Table S11). The genes that expressed in reproductive tissues may influence the reproductive process and seed setting rate of tetraploid rice. Interestingly, 64 genes were detected by both RiceXPro and our RNA-Seq data (Table S12).

#### Function Validation and Knockout Phenotypes of Non-parental Alleles in Neo-Tetraploid Rice

In order to validate the gene function of the non-parental alleles, nine non-parental genes that expressed in reproductive tissues with unknown functions were selected for gene knockout by CRISPR-Cas9 system using a neo-tetraploid rice line, Huaduo1, with high pollen fertility (94.45%) and high seed setting (80.21%). These nine genes were *BZR3* (*LOC\_Os06g35900*, [https://www.uniprot.org/uniprot/?query=LOC\\_Os06g35900&sort=score](https://www.uniprot.org/uniprot/?query=LOC_Os06g35900&sort=score)), *CLPC3* (*LOC\_Os11g16590*, [https://www.uniprot.org/uniprot/?query=LOC\\_Os11g16590&sort=score](https://www.uniprot.org/uniprot/?query=LOC_Os11g16590&sort=score)), *KIN7L* (*LOC\_Os11g35090*, [https://www.uniprot.org/uniprot/?query=LOC\\_Os11g35090&sort=score](https://www.uniprot.org/uniprot/?query=LOC_Os11g35090&sort=score)), *KIN14M* (*LOC\_Os06g36080*, <https://www.uniprot.org/uniprot/B9FTR1>), *LOC\_Os06g35970* (it was named as *NRFG2* in the present study, i.e. *neo-tetraploid rice fertility related gene 2*), *LOC\_Os06g36010* (it was named as *NRFG3* in the present study, i.e. *neo-tetraploid rice fertility related gene 3*), *LOC\_Os06g40020* (it was named as *NRFG4* in the present study, i.e. *neo-tetraploid rice fertility related gene 4*), *LOC\_Os06g34420* (it was named as *NRFG5* in the present study, i.e. *neo-tetraploid rice fertility related gene 5*) and *LOC\_Os06g40030* (it was named as *NRFG6* in the present study, i.e. *neo-tetraploid rice fertility related gene 6*) (Table S13). The pollen fertility of knockout mutant lines in T1 generation was decreased by 9.52% to 56.96%, with an average of 26.71%. Knockout of one kinesin motor domain containing protein gene, *KIN7L* produced the lowest pollen fertility (37.49%) (Fig. 4, Table 1).

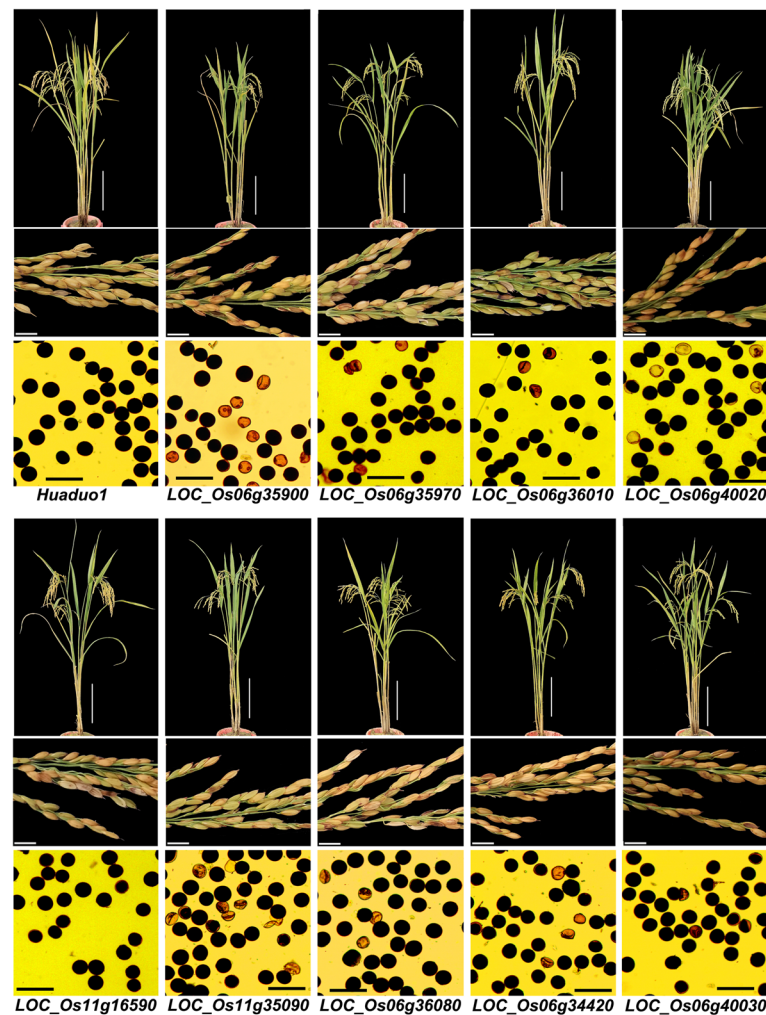
All knockouts showed an average of 25.63% reduction in seed setting, and the knockout of *KIN7L* showed the largest decrease in seed set (by 44.65%), while knockout of *NRFG5* showed the lowest decrease in seed set (by 13.08%). For yield per plant, nine knockouts showed an average of 4.87 g reduction in yield, and knockouts of four genes (*BZR3*, *NRFG3*, *NRFG4* and *KIN7L*) exhibited more than 5 g yield loss per plant. *kin7l* produced the highest lost in yield per plant by 8.81 g and *nrfg5* yielded the lowest lost in yield per plant by 2.20 g. In addition, other yield-



**Fig. 3** Inherited blocks inference of chromosome 6 (a) and chromosome 11 (b) of 13 NTR lines that developed from 96025 and Jackson-4x, and the expression pattern analysis of non-parental alleles in different tissues of neo-tetraploid rice, Huaduo1 (c)HMA, Anther and ovary in meiosis stage; H1, Anther in meiosis stage; HE2, ovary in flowering stage; MP, young panicle less than 5 mm; HP2, Anther in flowering stage; HP3, ovary at three days after flowering; GP1, sheath at five days after flowering; AZ1, branch and stem of panicle at flowering stage; AP1, sheath at flowering stage; HL3, flag leaf at three days after flowering; GL1, flag leaf at three days after flowering; GZ1, branch and stem of panicle at three days after flowering; HL1, flag leaf at meiosis stage; HL2 and AL1, flag leaf at flowering stage; H22, sheath at flowering stage; AS1, spikelet at flowering stage; HZ3, sheath at three days after flowering; GS1, spikelet at three days after flowering.

related traits were also measured, 9 knockout lines showed 2.93 cm shorter in panicle length, 7 (except for *NRF3* and *CLPC3*) knockout lines displayed 18.96 decrease in grain number per panicle, 8 (except for *CLPC3*) knockout lines showed 1.64 g reduction in thousand grain weight and 8 (except for *NRF5*) knockout lines demonstrated 5.01 cm decrease in plant height compared with wild type. Interestingly, 8 (except for *NRF3*) knockout lines showed an average of 0.72 increase in panicle number than wild type (Fig. 4, Fig. 5, Table 1, Table S15).

Of the 9 knockout genes, 3 genes exhibited important phenotype changes, such as knockouts of two kinesin motor domain containing gene, *kin7l* and *kin14m* decreased the seed setting by 44.65% and 24.30%, which cause yield loss of 8.81 g and 4.71 g per plant, respectively. Knockouts of *BES1/BZR1* homolog *BZR3*, showed shorter plants (8.19 cm shorter than WT Huaduo1), shorter panicles (2.85 cm shorter than WT Huaduo1), less grains per panicle (27.82 less than WT Huaduo1) and lower seed setting than WT (24.42% less than WT Huaduo1) (Fig. 4, Fig. 5, Table S15).



**Fig. 4** The non-parental allele knockout lines of neo-tetraploid rice showed different degrees of sterility. Knockout lines were developed by CRISPR-Cas9 system in neo-tetraploid rice line, Huaduo1. Scale bars: 20 cm for plant, 1 cm for panicle and 100  $\mu$ m for pollen

## Discussion

### Improvement of Seed Setting Is the Main Contributor for Yield Increase in Neo-Tetraploid Rice

Low fertility of autotetraploid rice has long been the obstacle for tetraploid rice breeding. The investigation of 40 autotetraploid rice genotypes and their diploid counterpart revealed 46.24% reduction in seed setting, which cause significant loss in yield (Wu et al. 2013). Another study about 10 autotetraploid rice lines and their diploid counterpart also indicated significant loss in seed setting by 49.25% and 44.45% during early season and late season, which lead to 3.70 t/ha and 4.95 t/ha yield reduction (Shahid et al. 2013). Similar tendency was observed in other previous reports (Wu et al. 2014; Chen et al. 2018b). The development of neo-tetraploid rice has improved the seed setting up to 80% and overcome this limitation (Guo and Liu 2014; Guo et al. 2017). In this study, we further emphasized the significance of fertility

improvement and thoroughly investigated the yield performance of neo-tetraploid rice. The correlation index between seed setting and yield per plant among 27 autotetraploid rice lines and 15 neo-tetraploid rice lines was 0.82, which indicating that seed setting is the main contributor for yield increase in neo-tetraploid rice. Other than fertility, correlation analysis also revealed that the grain number per panicle and plant height were also contributors for yield improvement, with correlation exponents of 0.51 and 0.52, respectively.

Furthermore, the systematic investigations of agronomic traits and field performance were conducted to evaluate the usage potential of neo-tetraploid rice and point out the direction of further tetraploid rice breeding. Guo et al. (2017) and Chen et al. (2019) evaluated 7 and 9 agronomic traits of neo-tetraploid rice lines and their hybrids. Another study investigated 7 agronomic traits of neo-tetraploid rice line Huaduo3, 66 and 134

**Table 1** Annotation of genes and phenotypes of CRISPR/Cas9 knockout mutants

Gene ID	Gene name	Gene Annotation	Pollen fertility (%)	Seed setting (%)	Yield per plant (g)
Huaduo 1	–	–	94.45 ± 1.67	80.21 ± 1.28	14.19 ± 0.78
LOC_Os06g35900	<i>BZR3</i>	BES1/BZR1 homolog protein, putative, expressed	64.33 ± 7.33*	55.79 ± 4.11**	8.29 ± 1.28**
LOC_Os06g35970	<i>NRFG2</i>	meiosis 5, putative, expressed	72.04 ± 12.75	60.87 ± 2.42**	10.77 ± 1.22*
LOC_Os06g36010	<i>NRFG3</i>	plastocyanin-like domain containing protein, putative, expressed	77.97 ± 5.95*	52.09 ± 4.43**	6.66 ± 0.92**
LOC_Os06g40020	<i>NRFG4</i>	DEAD-box ATP-dependent RNA helicase 52A, putative, expressed	84.93 ± 1.66**	51.88 ± 3.42**	7.97 ± 1.09**
LOC_Os11g16590	<i>CLPC3</i>	ATP-dependent Clp protease ATP-binding subunit clpA homolog CD4B, chloroplast precursor, putative, expressed	68.15 ± 11.31	57.43 ± 4.19**	11.74 ± 1.22
LOC_Os11g35090	<i>KIN7L</i>	kinesin motor domain containing protein, putative, expressed	37.49 ± 9.59**	35.56 ± 4.45**	5.38 ± 0.86**
LOC_Os06g36080	<i>KIN14M</i>	kinesin motor domain containing protein, expressed	77.81 ± 3.27*	55.91 ± 3.88**	9.48 ± 1.65*
LOC_Os06g34420	<i>NRFG5</i>	DEAD/DEAH box helicase domain protein, putative, expressed	54.05 ± 3.65**	67.13 ± 2.80**	11.99 ± 1.40
LOC_Os06g40030	<i>NRFG6</i>	S-locus-like receptor protein kinase, putative, expressed	72.92 ± 6.32*	54.58 ± 3.89**	11.59 ± 1.24

Huaduo1: a neo-tetraploid rice, wild type, control; \*: *p*-value of t-test between 0.01 and 0.05; \*\*: *p*-value of t-test less than 0.01

(Bei et al. 2019). In this study, 15 neo-tetraploid rice lines were used and 18 agronomic traits were systematically assessed in field experiments of two seasons. All the lines were cultivated in randomized complete block design (RCBD) with three replications. Neo-tetraploid rice showed advantage in many traits compared to autotetraploid rice, such as grain yield, seed setting, grain number per panicle, dry weight and harvest index. The hybrids showed high potential for tetraploid rice breeding based on the high heterosis effect between NTRs and autotetraploid rice. Based on these results, we concluded that neo-tetraploid rice has the characteristics of high fertility, high yield and high hybrid vigor.

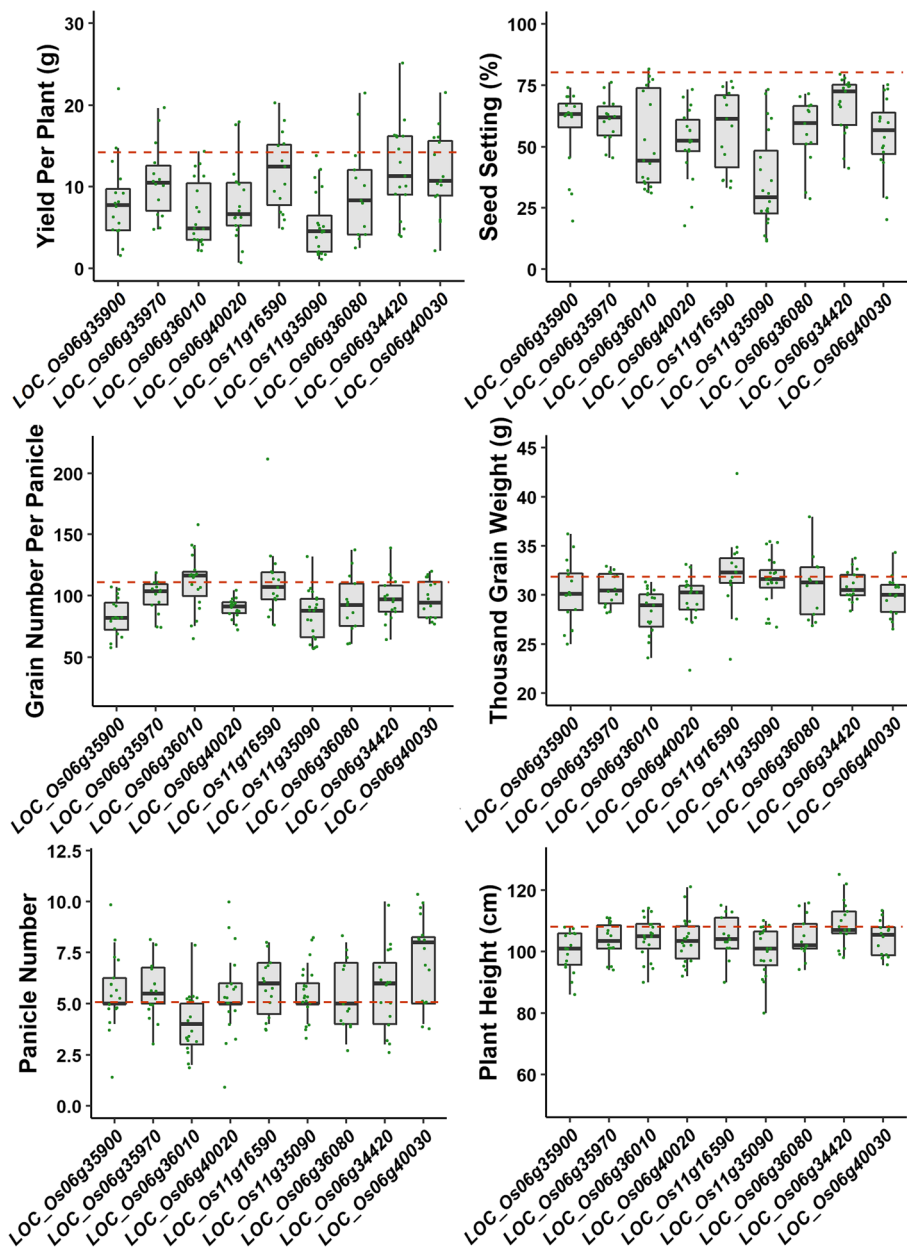
#### Non-parental Alleles Play Important Role in the Fertility Improvement of Neo-Tetraploid Rice

In recent years, genome comparison based on high depth next generation sequencing data has been used to identify inherited chromosome blocks and key trait loci in several studies (Zhou et al. 2016; Huang et al. 2018; Yamamoto et al. 2018). The newly developed NTRs showed remarkable improvement in yield performance especially in fertility than their parental ATR lines. We speculated that the improvement in yield and seed setting might be due to the generation and selection of novel superior alleles that associated with high fertility and yield during breeding process for many years or the interaction between parental alleles. Our previous study detected new mutations in three neo-tetraploid rice lines (Bei et al. 2019). Here, we used comparative genomics between 13 neo-tetraploid rice sister lines and their parental lines to precisely illustrate the genome composition map of

NTRs. The results of genome composition map provide useful information, such as most of the Chr5 regions were inherited from maternal line in NTRs, while most regions of Chr10 inherited from paternal line. This phenomenon of selective bias was previously reported in the HuangZaoSi-improved maize lines (Li et al. 2019) and Chr8 of “miracle rice” IR8 (Huang et al. 2018). Another promising result was the detection of 222 non-parental alleles and most of the non-parental genes were shared by 13 NTR sister lines, which indicated that these regions maybe the hotspot regions for genomic rearrangements as previously mentioned by Ramekar et al. (2015) in maize. Another potential reason for generation of novel mutations in NTRs is the enrichment or dosage effect of transposon elements (64 genes in non-parental regions were transposon element) in tetraploid rice genomes. These speculations may need further validation. The artificial selection of these novel variations in the following descendent generations showed that these variations were beneficial to the yield or fertility performance of NTR lines. These genes overlapped with yield and fertility related QTLs, and expressed in reproductive tissues, which indicated that they may regulate the fertility of polyploid rice.

Using CRISPR/Cas9 gene knockout system, the mutants of 9 non-parental genes were generated and they showed different extent of sterility, which validated the gene function. Kinesins are important intracellular transporter in microtubules, and *Pollen Semi-Sterility1* (*PSSI*) was previously reported to encode a kinesin-1-like protein that regulate male meiosis and fertility (Zhou et al. 2011). Two kinesin motor domain containing genes with non-parental and





**Fig. 5** Evaluation of yield, yield components and plant height of nine CRISPR/Cas9 gene knockout lines in T1 generation. The dash line represents the value of trait of wild type Huaduo1

nonsynonymous variations in NTRs were detected, and their functions were confirmed to be associated with fertility by using CRISPR/Cas9 gene knockout. The seed setting of pss1 mutant was 40%, in this study, and the seed setting of 2 kinesin motor domain containing genes knockout lines, *LOC\_Os11g35090* and *LOC\_Os06g36080*, were 35.56% and 55.91%. Brassinosteroid (BR) genes, *D11* and *OsBZR1*, were identified to promote anther and seeds, and over-expression of these genes resulted in higher grain yield (Zhu et al. 2015). A BES1/BZR1 homolog

protein gene, *LOC\_Os06g35900*, was identified to be non-parental allele in NTRs and its knockout showed defected mature pollen and low yield. Moreover, knockout of other non-parental genes, such as two DEAD-box ATP-dependent RNA helicase genes, and a S-locus-like receptor protein kinase gene, also showed low seed setting. These results suggested that genome comparison among various materials is a valid method to detect candidate genes, and these mutants will be useful materials for further functional genome research of tetraploid rice.

## Methods

### Field Experimental Design and Investigation of Agronomic Traits

Fifteen neo-tetraploid rice lines (NTRs) were used in the present study (Table S1d), and all the NTRs showed stable morphological traits for several generations. Field experiments were performed at the experimental farm of South China Agricultural University during late season (LS, August to November) in 2017 and early season (ES, March to July) in 2018 under conventional field management. All 15 NTRs were planted under randomized complete block design (RCBD) with three replications. A famous diploid rice variety (Huanghuazhan) and its autotetraploid counterpart (Huanghuazhan-4x), which was developed from Huanghuazhan by chromosome doubling using colchicine treatment by our research group in 2016, were used as controls (CK). A total of 18 traits were investigated, including heading date (HD), plant height (PH), flag leaf length (FLL), flag leaf width (FLW), flag leaf length width ratio (FLLWR), panicle number (PN), panicle length (PL), grain number per panicle (GNPP), total grain number (TGN), seed setting (SS), thousand grain weight (TGW), grain length (GL), grain width (GW), grain length width ratio (GLWR), yield per plant (YPP), dry weight (DW), harvest index (HI), and yield per block (YPB). A total of 27 autotetraploid rice lines were planted at the same location during both seasons to compare with NTRs (Table S1d). Yield related traits, such as plant height (PH), number of panicles (PN), panicle length (PL), grain number per panicle (GNPP), total grain number (TGN), seed setting (SS), thousand grain weight (TGW) and yield per plant (YPP), were investigated.

### Analysis of Field Experiment Data

Multiple trait correlation analysis was conducted and visualized using the corrplot package (Wei and Simko 2017). To validate the potential heterosis effect of NTRs, one NTR line, Huaduo8, was used as paternal line to cross with 2 *indica*, 3 *japonica* and 1 *javanica* autotetraploid rice lines. The yield and yield components of the hybrid lines were investigated, and the mid parent heterosis (MPH) and over parent heterosis (OPH) were estimated based on the method described previously (Guo et al. 2017).

### Sample Preparation, Library Construction and Genome Sequencing

Young leaves of 13 NTRs and their parental lines (Jackson-4x and 96025) were collected and their genomic DNA were extracted using the cetyltrimethylammonium bromide (CTAB) method (Cota-Sanchez et al. 2006). The sequencing libraries were constructed, and

sequenced based on the manufacturer's instructions of Illumina HiSeq (Yu et al. 2018). The low-quality data were trimmed using the following criterions: (a) sequencing adapters were removed; (b) sequencing reads with N percentage more than 50% were removed; (c) sequencing reads with low-quality base (base quality value lower than 10) percentage more than 50% were removed. After filtration, the data quality was evaluated using the FastQC (v0.11.6) software (Andrews 2010).

### Reads Mapping and Variation Calling

The sequencing reads that passed the quality control process were mapped onto the MSU7 (Nipponbare, *O. sativa japonica*) reference genome using BWA (0.7.17-r1188) (Li and Durbin 2009), and MarkDuplicates in Picard (2.12.1) (Broad Institute 2019) were used to eliminate data of PCR duplication. The SAM files were sorted, indexed and converted to BAM format using SAMtools (version 1.9) (Li 2011). The genomeCoverageBed in bedtools (v2.27.1) was used to estimate the reference genome coverage (Quinlan and Hall 2010). Genome Analysis Toolkit (GATK, version 3.8-0) was used to call the variations from the alignment file, and the analysis pipeline was constructed based on the GATK best practices (McKenna et al. 2010). The variations were annotated by SnpEff (4.3 s) based on the annotation GFF3 files of MSU7 reference genomes (Cingolani et al. 2012), and variations were compared by using VCFtools (0.1.16) software (Danecek et al. 2011). The software parameters for genome re-sequencing analysis are listed in Table S16.

### Genome Composition Analysis and Detection of Non-parental Alleles

A non-overlapped sliding windows method with step length of 10 kb was constructed using python script. Chromosome blocks were divided into three types: maternal, paternal and non-parental. Blocks with limited reliable markers (< 1 marker in one block) were excluded and there should be more than four markers in a block to determine non-parental block to increase the reliability. The blocks were retained if the block type was supported by at least 75% of markers in that block. Finally, the distributions of three types of blocks were plotted using ggplot2 (Wickham 2016). The non-parental alleles were detected using venn analysis provided by a toolkit of TBtools (Chen et al. 2018a). GO enrichment analysis was conducted using the agriGO v2.0 (Tian et al. 2017). During the update of Gramene database in 2010, the inferred location of rice QTLs were revealed by the sequences alignments of associated markers (Youens-Clark et al. 2011). QTL analysis was conducted by mapping the non-parental variations into the location of QTLs

using our Python script that available at GitHub repository under the name of “extractQTLByPos”.

### Expression Pattern Analysis of Candidates

First, the expression levels of non-parental alleles were checked in the rice expression database, RiceXPro (Yutaka et al. 2010). Second, the expression information was analyzed using our transcriptome data. The transcriptome data were mapped onto the MSU7 (Nipponbare, *O. sativa japonica*) reference genome using STAR (2.7.1a) software (Alexander et al. 2013), and the expression matrix was generated using RSEM (v1.3.1) software (Li and Colin 2011). The expression data were illustrated using heatmap (Raivo Kolde 2019).

### Validation of Variations Using Sanger Sequencing

Primers were designed using Primer Blast software in NCBI, and the segment of neo-tetraploid and autotetraploid containing variations were amplified and the PCR products were sequenced using Sanger sequencing. The sequences were aligned using ClustalW Multiple Alignment in BioEdit software.

### Construction of CRISPR/Cas9 Vector and Identification of Gene Knockout Mutants

Genome sequences of neo-tetraploid rice non-parental genes were used to design target-site primers using the targetDesign tool of CRISPR-GE (<http://skl.scau.edu.cn/home/>) (Table S13) (Xie et al. 2017). The target-site primers were located at the start of the exons to cause frameshift mutations. The CRISPR/Cas9 vectors were constructed as described previously (Butt et al. 2019; Ma et al. 2019), and CRISPR/Cas9 vectors were transferred into neo-tetraploid rice line, Huaduo1. The primers of gene-specific target site were used to select the transgenic plants, and the mutations were checked using Sanger sequencing. The sequencing data were decoded using DSDecodeM program of CRISPR-GE tool kit (Table S14) (Xie et al. 2017). The agronomic traits of about 15 plants for each gene mutant in T1 generation were measured, and the pollen fertility was observed by 1% I<sub>2</sub>-KI method using a microscope (Motic BA200).

### Conclusions

In this study, the agronomic traits of 15 neo-tetraploid rice lines were systematically investigated, and the improvement of yield and yield components, especially seed setting, were analyzed. Using comparative genomics between 13 neo-tetraploid sister lines and their parental lines, genome composition map of NTRs were illustrated and non-parental alleles were detected. The genes associated with reproductive tissues or development stages were detected and the gene functions were validated using CRISPR/Cas9

gene knockout. This study provided invaluable genomic resources, and the unique functional variations in NTRs will facilitate further marker assisted selection. The mutants of neo-tetraploid rice provide fundamental resources for functional genomics research in tetraploid rice.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12284-020-00387-3>.

**Additional file 1 Figure S1.** The performance of 15 neo-tetraploid rice (NTR) lines in field experiments. **Figure S2.** Yield performance and production assessment of neo-tetraploid rice lines compared with Huanghuazhan and Huanghuazhan-4x. **Figure S3.** Plant morphology of 13 NTRs and their parental lines. **Figure S4.** Inherited blocks inference in chromosome 1–5 (a–e), 7–10 (f–i), 12 (j) of 13 NTRs lines that developed from 96025 and Jackson-4x.

**Additional file 2 Table S1a.** Field performance of neo-tetraploid rice lines in 2017 late season (2017-LS) and 2018 early season (2018-ES). **Table S1b.** Average field performance of autotetraploid rice lines in 2017 late season and 2018 early season. **Table S1c.** Heterosis performance between neo-tetraploid and autotetraploid rice lines. **Table S1d.** List of the autotetraploid and neo-tetraploid rice materials used in this study. **Table S2.** Quality evaluation of genome sequencing data. **Table S3.** Mapping quality of sequencing reads to MSU7 reference genome. **Table S4.** Number of genomic variations against MSU7 reference genome. **Table S5.** Variation validation using previously sequenced individuals. **Table S6.** Validation of variations using Sanger sequencing. **Table S7.** Non-parental chromosome blocks that shared by 13 neo-tetraploid rice lines. **Table S8.** GO enrichment analysis of 222 non-parental genes in neo-tetraploid rice. **Table S9.** Reported QTLs that co-localized with non-parental variations in neo-tetraploid rice. **Table S10.** Non-parental genes that expressed in reproductive tissues using RiceXPro database. **Table S11.** Non-parental genes that expressed in reproductive tissues using RNA-Seq data of neo-tetraploid rice Huaduo1. **Table S12.** Non-parental genes that expressed in reproductive tissues in both RiceXPro and RNA-Seq data. **Table S13.** Gene annotation and primer sequences for CRISPR/Cas9 vector construction. **Table S14.** Genotype of CRISPR/Cas9 knockout lines of nine non-parental genes. **Table S15.** Pollen fertility and main agronomic traits of CRISPR/Cas9 knockout mutants. **Table S16.** Version, function and parameters of main software that used for genome re-sequencing and RNA-Seq data analysis.

### Abbreviations

NTR: Neo-tetraploid rice; ATR: Autotetraploid rice; SS: Seed setting; YPP: Yield per plant; RCBD: Randomized complete block design; QTL: Quantitative trait locus; OSPH: One seedling per hole; TSPH: Three seedlings per hole; GO: Gene ontology; *NTRFG*: Neo-tetraploid rice fertility related gene

### Acknowledgments

The authors are grateful to Ms. Shuhong Yu, Dr. Zhixiong Chen, Dr. Lan Wang and other lab members for assistance.

### Authors' Contributions

XDL conceived and designed the experiments. HY, MQS and XDL wrote the paper. HY, QHL, YDL, JWW, MQS, CL, and ZJL performed the experiments and analyzed the data. XDL bred neo-tetraploid rice. XDL and ZMZ contributed reagents/materials/analysis tools. The authors read and approved the final manuscript.

### Funding

This work was supported by the Guangzhou Science and Technology Key Program (201707020015), the National Natural Science Foundation of China (31571625), Guangdong Province Key Research and Development Program (2018B020202012) and South China Agricultural University Doctor Student Joint Training Project (2019LHPY013).

**Availability of Data and Materials**

The raw reads of whole-genome resequencing were deposited at the NCBI Sequence Read Archive with accession ID PRJNA526117. The sequences and annotations of rice *japonica* reference genome MSU7 are available from the website <http://rice.plantbiology.msu.edu/>. All data supporting the conclusions described here are provided in tables, figures, and additional files.

**Ethics Approval and Consent to Participate**

Not applicable.

**Consent for Publication**

Not applicable.

**Competing Interests**

The authors have declared that no competing interests exist.

**Author details**

<sup>1</sup>State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou 510642, China. <sup>2</sup>Guangdong Provincial Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou 510642, China. <sup>3</sup>College of Agriculture, South China Agricultural University, Guangzhou 510642, China. <sup>4</sup>Guangdong Laboratory for Lingnan Modern Agriculture, South China Agricultural University, Guangzhou 510642, China.

Received: 22 January 2020 Accepted: 28 April 2020

Published online: 03 June 2020

**References**

- Alexander D, Carrie AD, Felix S, Jorg D, Chris Z, Sonali J, Philippe B, Mark C, Thomas RG (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15–21
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Bei XJ, Shahid MQ, Wu JW, Chen ZX, Wang L, Liu XD (2019) Re-sequencing and transcriptome analysis reveal rich DNA variations and differential expressions of fertility-related genes in neo-tetraploid rice. *PLoS One* 14(4):e214953
- Broad Institute (2019) A set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats GitHub Repository: <http://broadinstitute.github.io/picard/>
- Butt H, Eid A, Momin AA, Bazin J, Crespi M, Arold ST, Mahfouz MM (2019) CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. *Genome Biol* 20:73
- Cai DT, Chen JG, Chen DL, Dai BC, Zhang W, Song ZJ, Yang ZF, Du CQ, Tang ZQ, He YC, Zhang DS, He GC, Zhu YG (2007) The breeding of two polyploid rice lines with the characteristic of polyploid meiosis stability. *Sci China Ser C-Life Sci* 37(2):217–226
- Chen CJ, Xia R, Chen H, He YH (2018a) TBtools, a toolkit for biologists integrating various HTS-data handling tools with a user-friendly interface. *bioRxiv preprint* <https://doi.org/10.1101/289660>
- Chen L, Shahid MQ, Wu JW, Chen ZX, Wang L, Liu XD (2018b) Cytological and transcriptome analyses reveal abrupt gene expression for meiosis and saccharide metabolisms that associated with pollen abortion in autotetraploid rice. *Mol Genet Genomics* 293(6):1407–1420
- Chen L, Yuan Y, Wu JW, Chen ZX, Wang L, Shahid MQ, Liu XD (2019) Carbohydrate metabolism and fertility related genes high expression levels promote heterosis in autotetraploid rice harboring double neutral genes. *Rice* 12:34
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain *w<sup>1118</sup>*; *iso-2*; *iso-3*. *Fly (Austin)* 6(2):80–92
- Cota-Sanchez JH, Remarchuk K, Ubayasena K (2006) Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Mol Biol Report* 24:161–167
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, De Pristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158
- Ghaleb MAA, Li C, Shahid MQ, Yu H, Liang JH, Chen RX, Wu JW, Liu XD (2020) Heterosis analysis and underlying molecular regulatory mechanism in a wide-compatible neo-tetraploid rice line with long panicles. *BMC Plant Biol* 20:83
- Ghouri F, Zhu J, Yu H, Wu J, Baloch FS, Liu X, Shahid MQ (2019) Deciphering global DNA variations and embryo sac fertility in autotetraploid rice line. *Turk J Agric For* 43:554–568
- Guo HB, Liu XD (2014) The research on autotetraploid rice. South China University of Technology Press, China, Guangzhou, pp 90–92
- Guo HB, Mendrikahy JN, Xie L, Deng JF, Lu ZJ, Wu JW, Li X, Shahid MQ, Liu XD (2017) Transcriptome analysis of neo-tetraploid rice reveals specific differential gene expressions associated with fertility and heterosis. *Sci Rep* 7: 1–11
- Guo HB, Shahid MQ, Zhao J, Li YJ, Wang L, Liu XD (2016) Agronomic traits and cytogenetic evaluation of newly developed autotetraploid rice line. *Pak J Agric Sci* 53(02):291–301
- He JH, Shahid MQ, Chen ZX, Chen XA, Liu XD (2011) Abnormal PMC microtubule distribution pattern and chromosome behavior resulted in low pollen fertility of an intersubspecific autotetraploid rice hybrid. *Plant Syst Evol* 291(3–4): 257–265
- Hiroki T, Akira A, Kentaro Y, Shunichi K, Satoshi N, Chikako M, Aiko U, Hiroe U, Muluneh T, Shohei T, Hideki I, Liliana MC, Sophien K, Ryohei T (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J* 74(1):174–183
- Huang J, Li J, Zhou J, Yang SH, Laurence DH, Wen-Hsiung L, Tian DC (2018) Identifying a large number of high-yield genes in rice by pedigree analysis, whole-genome sequencing, and CRISPR/Cas9 gene knockout. *Proc Natl Acad Sci U S A* 115(32):E7559–E7567
- Huang XH, Kurata N, Wei XH, Wang ZX, Wang AH, Zhao Q, Zhao Y, Liu KY, Lu HY, Li WJ, Guo YL, Lu YQ, Zhou CC, Fan DL, Weng QJ, Zhu CR, Huang T, Zhang L, Wang YC, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan XP, Xu Q, Dong GJ, Zhan QL, Li CY, Fujiyama A, Toyoda A, Lu TT, Feng Q, Qian Q, Li JY, Han B (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490(7421):497–501
- Jain M, Moharana KC, Shankar R, Kumari R, Garg R (2013) Genomewide discovery of DNA polymorphisms in rice cultivars with contrasting drought and their functional relevance. *Plant Biotechnol J* 12:253–264
- Li B, Colin ND (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323
- Li C, Song W, Luo Y, Gao S, Zhang R, Shi Z, Wang X, Wang R, Wang F, Wang J, Zhao Y, Su A, Wang S, Li X, Luo M, Wang S, Zhang Y, Ge J, Tan X, Yuan Y, Bi X, He H, Yan J, Wang Y, Hu S, Zhao J (2019) The HuangZaoSi maize genome provides insights into genomic variation and improvement history of maize. *Mol Plant* 12:1–8
- Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27(21):2987–2993
- Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25(14):1754–1760
- Li X, Shahid MQ, Xia J, Lu ZJ, Fang N, Wang L, Wu JW, Chen ZX, Liu XD (2017) Analysis of small RNAs revealed differential expressions during pollen and embryo sac development in autotetraploid rice. *BMC Genomics* 18:1–18
- Li X, Yu H, Jiao YM, Shahid MQ, Wu JW, Liu XD (2018) Genome-wide analysis of DNA polymorphisms, the methylome and transcriptome revealed that multiple factors are associated with low pollen fertility in autotetraploid rice. *PLoS One* 13(8):e201854
- Luan L, Tu SB, Long WB, Wang X, Liu YH, Kong FL, He T, Yan WG, Yu MQ (2007) Cytogenetic studies on two F<sub>1</sub> hybrids of autotetraploid rice varieties showing extremely high level of heterosis. *Plant Syst Evol* 267:205–213
- Ma K, Han JL, Yao Y, Yang ZF, Chen JY, Liu YG, Zhu QL, Chen LT (2019) An effective strategy to establish a male sterility mutant mini-library by CRISPR/Cas9-mediated knockout of anther-specific genes in rice. *J Genet Genomics* 46(5):273–275
- Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303
- Quinlan AR, Hall IM (2010) Bedtools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842
- Raivo Kolde (2019) Pheatmap: pretty Heatmaps. R package version 1.0.12 <https://CRAN.R-project.org/package=pheatmap>



- Ramekar RV, Sa KJ, Woo SY, Lee JK (2015) Non-parental banding patterns in recombinant inbred line population of maize with SSR markers. *Genet Mol Res* 14(3):8420–8430
- Shahid MQ, Li YJ, Saleem MF, Wei CM, Naeem M, Liu XD (2013) Yield and yield components in autotetraploid and diploid rice genotypes (*indica* and *japonica*) sown in early and late seasons. *Aust J Crop Sci* 5(7):632–641
- Shahid MQ, Liu GF, Li JQ, Muhammad N, Liu XD (2011) Heterosis and gene action study of agronomic traits in diploid and autotetraploid rice. *Acta Agric Scand Sect B-Soil Plant Sci* 61(1):23–32
- Shahid MQ, Sun JF, Wei CM, Peng Z, Liu XD (2010) Study on the abnormality of embryo sac and pollen fertility in autotetraploid rice during different growing seasons. *Pak J Bot* 42:7–19
- Shahid MQ, Xu HM, Lin SQ, Chen ZX, Muhammad N, Li YJ, Liu XD (2012) Genetic analysis and hybrid vigor study of grain yield and other quantitative traits in autotetraploid rice. *Pak J Bot* 44(1):237–246
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng CF, Sankoff D, Pamphilius CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. *Am J Bot* 96(1):336–348
- Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS (2007) Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56(1):13–30
- Tian T, Liu Y, Yan HY, You Q, Yi X, Du Z, Xu WY, Zhen S (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res* 45(W1):W122–W129
- Tu SB, Kong FL, Xu QF, He T (2003) Study on new system of heterosis utilization in autotetraploid rice. *Bull Chin Acad Sci* 6:426–428
- Tu Y, Jiang AM, Gan L, Mokter H, Zhang JM, Peng B, Xiong YG, Song ZJ, Cai DT, Xu WF, Zhang JH, He YC (2014) Genome duplication improves rice root resistance to salt stress. *Rice* 7:15
- Van de Peer Y, Mizrahi E, Marchal K (2017) The evolutionary significance of polyploidy. *Nat Rev Genet* 18(7):411–424
- Wang WS, Ramil M, Hu ZQ, Dmytro C, Tai SS, Wu ZC, Li M, Zheng TQ, Roven RF, Zhang F, Locedie M, Dario C, Millicent S, Kevin CP, Xu JL, Sun C, Fu BY, Zhang HL, Gao YM, Zhao XQ, Shen F, Cui X, Yu H, Li ZC, Chen ML, Jeffrey D, Zhou YL, Zhang XY, Zhao Y, Dave K, Wang CC, Li R, Jia B, Lu JY, He XC, Dong ZT, Xu JB, Li YH, Wang M, Shi JX, Li J, Zhang DB, Seunghee L, Hu WS, Alexander P, Inna D, Victor JU, Frances NB, John RM, Jauhar A, Li J, Gao Q, Niu YC, Yue Z, Ma Elizabeth BN, Jayson T, Wang XQ, Li JJ, Fang XD, Yin Y, Jean-Christophe G, Zhang JW, Li JY, Ruaraidh SH, Rod AW, Ruan J, Zhang GY, Wei CC, Nickolai A, Kenneth LM, Li Z, Hei L (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557(7703):43–49
- Wei TY, Simko V (2017) R package "corrplot": visualization of a correlation matrix (version 0.84) Available from <https://github.com/taiyun/corrplot>
- Wendel JF (2000) Genome evolution in polyploids. *Plant Mol Biol* 42(1):225–249
- Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. *Proc Natl Acad Sci U S A* 106(33):13875–13879
- Wu JW, Hu CY, Shahid MQ, Guo HB, Zeng YX, Liu XD, Lu YG (2013) Analysis on genetic diversification and heterosis in autotetraploid rice. *SpringerPlus* 2:1–12
- Wu JW, Shahid MQ, Chen L, Chen ZX, Wang L, Liu XD, Lu YG (2015) Polyploidy enhances F<sub>1</sub> pollen sterility loci interactions that increase meiosis abnormalities and pollen sterility in autotetraploid rice. *Plant Physiol* 169(4):2700–2717
- Wu JW, Shahid MQ, Guo HB, Yin W, Chen ZX, Wang L, Liu XD, Lu YG (2014) Comparative cytological and transcriptomic analysis of pollen development in autotetraploid and diploid rice. *Plant Reprod* 27:181–196
- Xie XR, Ma XL, Zhu QL, Zeng DC, Li GS, Liu YG (2017) CRISPR-GE: a convenient software toolkit for CRISPR-based genome editing. *Mol Plant* 10:1246–1249
- Yamamoto N, Garcia R, Suzuki T, Solis CA, Tada Y, Venuprasad R, Kohli A (2018) Comparative whole genome re-sequencing analysis in upland new rice for Africa: insights into the breeding history and respective genome compositions. *Rice* 11(1):33
- Yang PM, Huang QC, Qin GY, Zhao SP, Zhou JG (2014) Different drought-stress responses in photosynthesis and reactive oxygen metabolism between autotetraploid and diploid rice. *Photosynthetica* 52(2):193–202
- Youens-Clark K, Buckler E, Casstevens T, Chen C, Decker G, Derwent P, Dharmawardhana P, Jaiswal P, Kersey P, Karthikeyan AS, Lu J, McCouch SR, Ren L, Spooner W, Stein JC, Thomason J, Wei S, Ware D (2011) Gramene database in 2010: updates and extensions. *Nucleic Acids Res* 39:D1085–D1094
- Yu H, Shahid MQ, Li RB, Li W, Liu W, Fozia G, Liu XD (2018) Genome-wide analysis of genetic variations and the detection of rich variants of NBS-LRR encoding genes in common wild rice lines. *Plant Mol Biol Report* 36(4):618–630
- Yu Z, Haage K, Streit VE, Gierl A, Ruiz RAT (2009) A large number of tetraploid *Arabidopsis thaliana* lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations. *Theor Appl Genet* 118:1107–1119
- Yutaka S, Baltazar AA, Nobukazu N, Hinako T, Hiroshi M, Kaori K, Kazuhiko S, Yuji S, Hirohiko H, Yoshiaki N (2010) RiceXPro: a platform for monitoring gene expression in *japonica* rice grown under natural field conditions. *Nucleic Acids Res* 39(Database issue):D1141–D1148
- Zhou D, Chen W, Lin Z, Chen H, Wang C, Li H, Yu R, Zhang F, Zhen G, Yi J, Li K, Liu Y, Terzaghi W, Tang X, He H, Zhou S, Deng XW (2016) Pedigree-based analysis of derivation of genome segments of an elite rice reveals key regions during its breeding. *Plant Biotechnol J* 14(2):638–648
- Zhou SR, Wang Y, Li WC, Zhao ZG, Ren YL, Wang Y, Gu SH, Lin QB, Wang D, Jiang L, Su N, Zhang X, Liu LL, Cheng ZJ, Lei CL, Wang JL, Guo XP, Wu FQ, Hiroshi I, Wang HY, Wan JM (2011) *Pollen Semi-Sterility1* encodes a kinesin-1-like protein important for male meiosis, anther dehiscence, and fertility in rice. *Plant Cell* 23(1):111–129
- Zhu XL, Liang WQ, Cui X, Chen MJ, Yin CS, Luo ZJ, Zhu JY, Lucas WJ, Wang ZY, Zhang DB (2015) Brassinosteroids promote development of rice pollen grains and seeds by triggering expression of carbon starved anther, a MYB domain protein. *Plant J* 82(4):570–581

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)