Genome-wide dissection of Quan9311A breeding and application

Qianlong Li
  Shanghai ZKW Molecular Breeding Technology Co., Ltd. Shanghai, 200234, China

Qi Feng
  Center of Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200233, China

Heqin Wang
  Win-all Hi-tech Seed Co., Ltd. /Key Laboratory for New Variety Creative of Hybrid Rice, Ministry of Agriculture and Rural Affairs, Hefei, 230088, China

Yunhai Kang
  Shanghai ZKW Molecular Breeding Technology Co., Ltd. Shanghai, 200234, China

Conghe Zhang
  Win-all Hi-tech Seed Co., Ltd. /Key Laboratory for New Variety Creative of Hybrid Rice, Ministry of Agriculture and Rural Affairs, Hefei, 230088, China

Ming Du
  Shanghai ZKW Molecular Breeding Technology Co., Ltd. Shanghai, 200234, China

Yunhu Zhang
  Win-all Hi-tech Seed Co., Ltd. /Key Laboratory for New Variety Creative of Hybrid Rice, Ministry of Agriculture and Rural Affairs, Hefei, 230088, China

Hui Wang
  Win-all Hi-tech Seed Co., Ltd. /Key Laboratory for New Variety Creative of Hybrid Rice, Ministry of Agriculture and Rural Affairs, Hefei, 230088, China

Jinjie Chen
  Win-all Hi-tech Seed Co., Ltd. /Key Laboratory for New Variety Creative of Hybrid Rice, Ministry of Agriculture and Rural Affairs, Hefei, 230088, China

Bin Han
  Center of Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200233, China

Yu Fang
  Shanghai ZKW Molecular Breeding Technology Co., Ltd. Shanghai, 200234, China

Ahong Wang (✉ ahwang@ncgr.ac.cn)
  Center of Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200233, China
Abstract

Germplasm resource innovation is the key to variety development. Quan9311A is a three-line cytoplasmic male sterile (CMS) line of rice used for hybrid rice breeding based on the three-line system. It is the first CMS line that has been successfully cultivated with rice restoration materials and widely applied as a female parent line for hybrid breeding in China. The variety improvement of Quan9311A is based on an ingenious design. It was screened from the cross progeny of the three-line sterile line Zhong9A and the two-line restorer line 93–11, which are the core germplasms used to breed Quan9311A. According to phenotype, China breeders removed the restorability of 93–11, and cultivated Quan9311A: it combines the desirable agronomic traits of both 93–11 and Zhong9A, such as a moderate plant height, suitable heading date, and more tillers. In this study, genome-wide analysis was performed to investigate how Quan9311A was bred. RiceNavi analysis technology was used to dissect its functional genes associated with desirable agronomic traits based on whole-genome sequencing. The results show that rice breeders removed the fertility restorer gene $Rf3$ in 93–11 but retained the sterility gene $WA352c$ of Zhong9A during the breeding process of Quan9311A. Not only are there many 93–11 chromosomal segments in the Quan9311A genome background but also multiple superior alleles of the parental 93–11 and Zhong9A lines are pyramided, conferring to Quan9311A excellent agronomic traits that were unavailable in earlier three-line sterile lines. We also found that Quanyou Simiao, a hybrid rice combination bred with Quan9311A as the female parent, harbors more superior alleles which could explain its high grain yield and wide adaptability. This study’s findings will promote the utilization of hybrid rice pairing groups and guide future breeding efforts. Meanwhile, it also provides fresh ideas for further promoting the closer integration of genomics with traditional breeding methods.

Background

Advances in plant breeding technology over the past few decades have driven a revolution in crop breeding, especially for rice, one of the world’s major food crop staples. Since the 1970s, China breeders have taken the lead in carrying out cutting-edge research into the utilization of rice heterosis, cultivating numerous hybrid rice varieties via the three-line and two-line hybrid rice breeding system, which have contributed to ensuring the food security of China and the world at large.

Given the complex geopolitical situation and ongoing climate change, along with changing planting patterns and consumer demands for rice, the in-depth utilization of rice heterosis, especially by applying genomics and other technologies to enhance rice design breeding, is the key to attain future food security. China is among the first countries in the world to pursue plant genomics research. In 1998, China participated in the “International Rice Genome Sequencing Project” as one of the main initiators (Sasaki et al. 1998; Sasaki et al. 2000). In 2002, Chinese scientists took the lead in completing the draft sequence of genome of *Oryza sativa* L. ssp. *indica* 93–11 (Yu et al. 2002) and the accurate sequencing of chromosome 4 of *O. sativa* ssp. *japonica* cv. Nipponbare (Feng et al. 2002). Over the next 20 years, China made great progress in rice functional genomics research, having cloned many genes related to crucial agronomic traits (Li et al. 2021; Guo et al. 2019). With the continuous development of quantitative
genetics, computational biology and genomics, a new era of genetics-guided breeding of rice has emerged. For example, Huang et al. (2016) identified the main genetic factors controlling the heterosis loci—including quantitative trait locus or genes in rice, and revealed the genetic mechanisms of three-line, two-line, and inter-subspecies heterosis. Later, Wei et al. (2021) completed their construction of a map for quantitative trait genes/causative gene critical variation in rice, developed a breeding navigation program, and established a new method for the molecular design breeding of rice variety. In order that such leading genomics knowledge or technologies can best serve breeding goals, not only must rice breeders come to know more molecular biology but the achievements of genomics should be made readily understandable to them. More and more efficient molecular biology tools should be developed to support breeding goals and efforts.

Here we briefly recount the history of the three-line CMS Quan9311A, whose application is now the largest in China, and present the breeding ideas underpinning it. Around 2003, there was a rapid development of the two-line varieties of rice. The breeders from Win-All High-tech Seed Co., Ltd. found that 93−11 was the male parent in most of the two-line hybrid rice varieties that were widely planted. As a conventional rice variety, 93−11 had also passed the national and multi-provincial approvals, indicating it was promising material with strong matching and cooperation capabilities (to date, 93−11 remains the most widely used two-line restorer line in China). Given that 93−11 cannot completely restore the fertility of most three-line male sterile lines (except the Honglian type, HL-CMS), as well as the recognized risks of existing two-line hybrid seed production and inherent limitations in the relationship between three-line restoration and sterility, breeders decided to transform 93−11 into a new three-line male sterile line. After years of efforts, breeders finally bred Quan9311A in 2011 (Wang et al. 2013). It is the first three-line sterile line that was successfully transformed by restoring genetic material and since its introduction it has been widely used, broadening the germplasm resources of three-line sterile lines in China. This achievement opened up a new method of breeding three-line sterile lines of hybrid rice. The successful breeding of Quan9311A arose from an ingenious design in the process of variety breeding, which fully aggregated the dominant traits together to a high level.

In this study, we take Quan9311A as an exemplar to analyze its genetic background and functional genes, and to explore the reasons for its successful breeding at the genetics and genomics level. Such an investigation also promotes the better utilization and improvement of Quan9311A; allows the scientific researchers engaged in genomics research to better understand the rice breeding process; encourages the closer integration of genomics and traditional breeding, and; provides ideas for improved guidance and strengthened support of rice breeding in the future.

**Results**

**Phenotypic characterization of Quan9311A**

Quan9311A is a three-line sterile line that is bred from crossing 93−11, Zhong9A and 222B, in which 93−11 and Zhong9A are used as the core germplasm for improvement (Fig. 1a). Its breeding process is
described in detail in Materials and methods, but is briefly summarized here: select 93 – 11 as the male parent, conduct sexual crosses with the three-line maintainers Zhong9B and 222B, and then screen for the desirable individual plants for their continuous backcrossing with sterile individuals. After multiple generations, a new three-line sterile line, Quan9311A, was chosen (Fig. 1b). That selected Quan9311A combined the dominant agronomic traits of both 93 – 11 and Zhong9A, such as moderate plant height, suitable whole growth period (days from sowing seed to plant producing mature grain), more tillers, longer panicle length, moderate thousand-grain weight (TGW), slender grain shape, and more transparent and polished rice grains (Fig. 1c, Supplementary Fig. 1). Its plant height is about 85 cm and similar to that of Zhong9A. The whole growth period lies between that of its two parents (ca. 110–115 days, 25 days shorter than that of 93 – 11 and 20 days longer than that of Zhong9A). Its number of tillers is similar to that of Zhong9A, at about 13. Its panicle length is about 25 cm, longer than that of either parent, and the thousand-grain weight is moderate, at about 25 g. The grain quality met the first-class standard of the Ministry of Agriculture of China. (Wang et al. 2013, NYT593-2013 of China). In practical applications, Quan9311A also has other desirable phenotypic characteristics, namely, a high stigma exertion rate, tolerance of high temperature, high outcrossing and seed-setting rate, high nitrogen-use efficiency, in addition to drought tolerance.

**Sequencing and genome-wide SNP analysis**

We obtained sequencing data of the five varieties with 37- ~ 140- coverage of the rice genome. For each, the Q30 value was greater than 94%. The next-generation sequencing data of each variety was then compared with the reference genome of *O. sativa* japonica cv. Nipponbare, and more than two million SNPs were identified between each variety and Nipponbare (Supplementary Table 1). The respective SNPs of 93 – 11 and Zhong9A were compared with those of Quan9311A, and the SNP loci consistent with Quan9311A were counted (Supplementary Table 2). The proportion of similar SNPs shared between them was 74.9% and 67.8%, respectively. Although Quan9311A retained the genetic background of 93 – 11, according to the identification threshold (genetic similarity < 96%) of the plant variety MNP (multiple nucleotide polymorphism) marker method in the *Joint Study of Chinese Agricultural and Rural Improved Rice Varieties* (GB/T 38551 – 2020, China), it can be defined as the new creation of germplasm resources. Next, the SNPs obtained from the comparing the two varieties were displayed for the whole genome, and the discrepancy in SNPs between the two parents on different chromosome segments also indicated (Fig. 2, Supplementary Table 3). These results revealed that Quan9311A and 93 – 11 had large similarity fragments on Chr.2, Chr.3, chr.6, Chr.7, Chr.9, and Chr.12 (Fig. 2a). But Quan9311A and Zhong9A differed greatly on multiple chromosomes or across the whole genome (Fig. 2b).

**Functional loci analysis of Quan9311A, 93 – 11 and Zhong9A**

The QTNpick module in the RiceNavi methodology (Supplementary Fig. 2) was used to conduct a genome ‘physical examination’ analysis. The variation of the 348 QTN loci in the whole genome was
judged, and those gene loci harboring functional trait variation were found (Supplementary Table 4). These results demonstrated that Quan9311A not only inherited the superior alleles from 93 – 11, such as OsMADS51, OsACS6, Waxy, Hd1, Sdr4, BG2, GW8, Sub1A, and STV11 (Kim et al. 2007; Matsushima et al. 2015; Su et al. 2011; Yano et al. 2000; Sugimoto et al. 2010; Nagasawa et al. 2013; Wang et al. 2012; Jung et al. 2010; Wang et al. 2014), but also several from Zhong9A, namely Xa1, Ghd7, and Xa3 (Yoshimur et al. 1998; Xue et al. 2009; Cao et al. 2007). However, a small number of superior genes such as PTB1 and Bph3 in 93 – 11 got lost during the breeding process (Fig. 3) (Li et al. 2013; Liu et al. 2015). It is noteworthy that in 93 – 11 some genes which unfavorable with respect to desirable rice crop traits, such as the Rf3 gene with a restorative function, the D2 gene responsible for plant height dwarfism and a greater tiller angle, and the BOC1 gene which easily browns the callus, were not inherited by Quan9311A (Qi et al. 2008; Hong et al. 2003; Lu et al. 2014). Furthermore, the Rf2 gene in Zhong9A was likewise not inherited by Quan9311A (Itabashi et al. 2011). Hence, regarding the excellent genes held by both parents, Quan9311A undoubtedly inherited most of them (Supplementary Table 5). Therefore, Quan9311A, as a three-line male sterile line, contains more superior allele genes than does Zhong9A.

Causal analysis of incomplete restoration of 93 – 11 to three-line male sterile lines (except the Honglian type) based on functional loci

It is well known that 93 – 11, as the male parent, is a strong restorer line in the two-line hybrid breeding system for rice, but it is not completely restored in the vast majority of wild-abortive three-line sterile lines (except for HL-CMS). The fertility restoration of wild abortive sterile lines is mainly governed by two pairs of dominant genes, Rf3 and Rf4, these located on chromosomes 1 and 10, respectively (Qi et al. 2008); of the two, the Rf4 restorer gene has a stronger genetic effect while Rf3 has a synergistic effect (Sattari et al. 2008). Fertility restoration of the Boro II sterile line (BT-CMS) is controlled by Rf1, which contains two closely linked restorer genes, Rf1a and Rf1b (Wang et al. 2006). Recovered fertility of the HL-CMS line is under the control of two restorer genes, Rf5 and Rf6, which are capable of restoring the BT-CMS line to normality (Huang et al. 2003; Zhang et al. 2017).

According to the genome sequencing information, we found that the wild-abortive sterile line restorer gene Rf3 in 93 – 11 was a mutant type with a restorative function, whereas Rf4 was a wild-type genotype lacking that function. The restorer gene Rf1b is a wild type and has no restorative function (the Rf1a genotype is unknown). The genotype of the restorer gene Rf5 of the HL-CMS line is unknown, but Rf6 is the genotype with a restorative function. In practical applications, 93 – 11 is able to restore the HL-CMS line. This genotype information may explain why the combination of wild-abortive sterile lines with 93 – 11 as the male parent typically shows a heritable pattern split between recovery and conservation.

Removal of the restorer genes and retention of sterile genes in 93 – 11’s genetic background

Conventional breeding of three-line male sterile lines usually adopts the method of ‘maintainer line begets maintainer line’; that is, crossing maintainer lines with each other, and selecting an excellent and stable individual plant as a new maintainer line from the progeny of multiple generations. In past breeding
practices of maintainer lines, it was challenging to breed a maintainer line with stable fertility from rice varieties with restorer genes. However, as a two-line restorer line, 93−11 showed incomplete recovery characteristics to most three-line male sterile lines (except the Honglian type), which indicated it was a good object for transformation.

After analyzing the functional genes of the three cultivars, we found that Quan9311A/B removed the mutant genotypes of \( Rf3 \) and \( RF6 \) in 93−11, but kept the genotypes of \( Rf3 \) and \( RF6 \) lacking the fertility-restoring function in Zhong9A/B, while the genotypes of other fertility restoring genes went unchanged (all having no restorative function). This led to Quan9311A lacking any recovery function. We also found that the infertility gene of Quan9311A originated from \( WA352c \) of Zhong9A, which was derived from the mitochondrial genome in cytoplasm (Luo et al. 2013). All rice varieties containing the wild-abortive male sterility-restoring gene can restore their fertility. Therefore, Quan9311A was deemed suitable as a sterile male line because it had the basic characteristics sought for in sterile lines: no restorer gene and a wild-abortive sterile gene.

Analysis of functional genes related to plant height, stigma exertion rate, tillering ability, flowering characteristics, and high outcrossing and seed-setting rate in Quan9311A

It was proposed that Quan9311A would serve for hybridization rice breeding in southern China, especially in the middle and lower areas along the Yangtze River. The geographical and ecological environment of these areas has positive impacts on the plant height of Quan9311A, which is not too tall. At the same time, it is also necessary to ensure rice plants have a sufficient biomass to sustain high and stable yields, and conveniently produce hybrid seeds. Therefore, when formulating the breeding plan of Quan9311A, the dominant agronomic traits of moderate plant height, suitable heading date, more tillers, and high stigma exertion rate were explicit breeding goals.

Through the RiceNavi analysis, we found that, with respect to plant height, Quan9311A inherited the \( Ghd7 \) mutant genotypes from Zhong9B (Xue et al. 2009). Because of the pleiotropic characteristic of \( Ghd7 \), Quan9311A not only has a lower plant height than 93−11, but also flowers earlier flowering and forms more tillers. The other genes that control plant height, such as \( OsGA2ox4 \) and \( OsSPY \), are wild types that do not increase plant height (Lo et al. 2008; Asako et al. 2006). According to the RiceNavi predictions, plant height under long-day conditions in the middle and lower Yangtze River areas was about 90 cm, this close to that of planted fields (85 cm).

In terms of tillering ability, the tiller genes \( SLB1 \) and \( SLB2 \) of Quan9311A were present in both parents (Catarina et al. 2014). Concerning the heading date, Quan9311A retained the mutant genotypes of \( Hd1 \) and \( OsMADS51 \) in 93−11, whose genetic effect delayed the heading date (Yano e al. 2000; Kim et al. 2007). In addition, the mutant genotype \( Ghd7 \) of Zhong9B was retained in Quan9311A, and the genetic effect suggested an advanced heading date. Finally, under the joint action of these genes, the heading date was earlier for Quan9311A than 93−11. The RiceNavi analysis also predicted that the heading date of Quan9311A under long-day conditions in the middle and lower Yangtze River was about 78 days, a duration consistent with the actual heading date (74–80 days) recording in the breeding of Quan9311A.
The higher stigma exertion rate renders the male sterile line adept at outcrossing, which underpins the reproduction of male sterile line offspring and high yields of hybrid seed production. However, in actual production applications, breeders based in China have found that the high stigma exertion rate is often negatively correlated with the closed-glume performance of hybrid seeds, resulting in a low germination rate and weak storage ability of hybrid seeds with poor resistance to glume closure. Therefore, due to this apparent trade-off, it is difficult to simultaneously improve the grain yield and grain quality of hybrid rice seeds. In the process of hybrid seed production, hybrid seeds propagated by Quan9311A were found to possess characteristics of high seed-setting rate, good glume closure, and high germination rate. Importantly, we found that Quan9311A contained the $GS3$ gene, originally held by both parents, which can increase the stigma exertion rate, whose mutation can also increase the grain size (Fan et al. 2006; Zhou et al. 2017). In terms of seed germination rate, Quan9311A contained the mutant gene type $Sdr4$ inherited from 93 – 11, which can improve seed germination success (Sugimoto et al. 2010). Further, it also harbored the gene $OsGSK2$ (Sun et al. 2018), which can lengthen the mesocotyl, as well as the gene $OsTPP7$, which can enhance the tolerance of rice plants to anaerobic germination (Kretzschmar et al. 2015).

**Analysis of functional genes related to grain quality, pest/disease resistance, and other advantageous characteristics in Quan9311A**

The genome of Quan9311A retained the $OsACS6$, $BG2$, and $GW8$ mutant genotypes from 93 – 11, which together can increase both grain size and grain width (Matsushima et al. 2015; Nagasawa et al. 2013; Wang et al. 2012). Accordingly, the grain length-to-width ratio and 1000-grain weight of Quan9311A was larger. Moreover, the favorable genotype of $Waxy$ from 93 – 11 was retained in Quan9311A, thus ensuring the amylose content of Quan9311A stayed within a suitable range (Su et al. 2011).

In terms of pathogen resistance, Quan9311A retained the bacterial blight resistance genes $XA7$ and $XA3/XA26$ of Zhong9B, which are resistance allelic genotypes (Yoshimur et al. 1998; Cao et al. 2007). With respect to rice blast resistance genes, Quan9311A inherited from both parents several resistance alleles including $pi25/pid$, $pia$, and $pib$ (Jun et al. 2009; Zeng et al. 2011; Wang et al. 1999), but it did not contain broad-spectrum rice blast resistance genes such as $pi2/pi9$ and $Pigm$ (Su et al. 2015; Deng et al. 2017). According to the RiceNavi analysis, despite no risk of susceptibility to bacterial blight one still exists to rice blast. We recommend improving resistance to rice blast or select a male parent resistant to rice blast for matching.

In addition, the genome of Quan9311A also contained the higher nitrogen-use efficiency genes $NRT1.1B$ and $OsNR2$ (Hu et al. 2019; Gao et al. 2019), as well as several other notable genes. These include $OsCERK1$ which improves phosphorus absorption efficiency (Huang et al. 2020), $Sub1A$ which confers submergence tolerance (from 93 – 11) (Jung et al. 2010), $STV11$ which strengthens resistance to the rice stripe virus (from 93 – 11) (Wang et al. 2014), $TOND1$ which increases nutrition tolerance (Zhang et al. 2015). Other genes consisted of $NAL1$ which reduces plant height and narrows leaves and increases yield, as well $LAX1$ which increases the number of grains (Komatsu et al. 2003); $OsHKT4$ which increases
the sodium content of roots to improve salt tolerance (Rong et al. 2015); and SCM2/APO1 which enhances lodging resistance ability (Ookawa et al. 2010)—in addition several more functional genes. The pleiotropic functions of some of these genes, the influence of their interactions, and their commercial value still need to be explored and characterized.

**Functional gene analysis of the hybrid rice QYSM with Quan9311A as the female parent**

Quan9311A can be matched with various types of restorer lines in hybrid rice breeding, showing a robust combining ability. The hybrid rice combination obtained with it as the female parent obviously has heterosis. To date, 133 hybrid rice varieties with Quan9311A as the female parent have been approved in China at the national or multi-provincial level (https://www.ricedata.cn/variety). These varieties are widely distributed, covering the entire rice cultivation areas in southern China. Quan9311A has since become the largest three-line sterile line grown in China. It is also one of the most widely used backbone materials in hybrid rice pairing and for the directional improvement of sterile lines. We selected a hybrid combination, QYSM, whose female parent is Quan9311A and its male parent is WSSM for their whole-genome sequencing analysis. This variety has passed the national approval for use as indica rice in the middle and lower areas of the Yangtze River, likewise, in the middle and upper areas of Yangtze River, and for double-cropped rice cultivation in southern China, resulting in yield increases of 3.3–9.43%.

Combined with RiceNavi technology, we analyzed the functional loci of Quan9311A, WSSM, and QYSM. Evidently, Quan9311A and WSSM each contained more causative allele genes, which got aggregated in the genome of QYSM. The female parent Quan9311A provided several causative genes such as LAX1 and LTG (Lu et al. 2014), GW5 (Liu et al. 2017), Sdr4, BG2, and Ghd7.1 (Yan et al. 2013), TIG1 (Zhang et al. 2019), Pi9, Pi-ta, Pi5, Pi56, and GS6 (Sun et al. 2013), as well as Ghd8 (Feng et al. 2014) and PTB1 and NOG1 (Huo et al. 2017). It was found that the fertility-restorer genes Rf3 and Rf4 of the wild-abortive male sterile line were mutants in the WSSM genome, whose genetic function was to restore fertility, indicating WSSM is a good restorer variety. In addition, Quan9311A and WSSM also shared a series of superior alleles (Fig. 4, Supplementary Table 6).

Based on the results of the analyzed functional loci, we tried to explain the reasons for the high yield and disease resistance of QYSM. We uncovered many excellent mutant genes in the genome of QYSM, some of which were heterozygous (i.e., genotype locus inherited from one of the parents), and others that were homozygous (i.e., genotypes present in both parents). For example, the rice blast resistance loci contained in QYSM included Pib, Pi2/Pi9, Pi25/Pid3,Pid2, Pi21, Pia, and Pi-ta, etc. Among them, the most critical broad-spectrum rice blast resistance locus, Pi2/ Pi9, was inherited from the male parent WSSM, while the other two rice blast resistance loci, Pi-ta and Pi21, were also derived from WSSM. This enables QYSM to possess broad-spectrum resistance to rice blast and the potential to grow in a variety of environments. This potential can compensate for the risk of Quan9311A being not resistant to rice blast. In terms of resistance to bacterial blight, the contribution of Quan9311A as the female parent was greater.
than that of the male parent WSSM, because Quan9311A contained multiple bacterial blight resistance mutant loci, such as \textit{Xa1, Xa4, Xa26}, whereas WSSM contained only the \textit{Xa4} locus. In terms of resistance to an insect pest, the rice planthopper, two resistance genes, \textit{Bph3} and \textit{Bph29} were provided by WSSM, which offset the deficiency of rice planthopper resistance genes in Quan9311A.

In terms of yield, QYSM inherited a series of excellent alleles contained in both parents. These included the genes \textit{Gn1a, LAX1, GNP1} that can increase the number of grains per panicle, along with the semi-dwarf gene \textit{sd1}, in addition to the \textit{NAL1} gene that narrow rice leaves while increasing the yield by reducing plant height (Heng et al. 2014). Other genes consisted of \textit{GS3} which increases grain shape and stigma exertion rate; \textit{SLB1} and \textit{SLB2} which increase tiller number; the \textit{SCM2/APO1} which increase the number of grains per panicle while rendering the rice plant more resistant to lodging resistant (among others). More importantly, in this hybrid rice combination, the female parent Quan9311A provided the superior genotypes of \textit{GW5} and \textit{GS6} that can increase the 1000-grain weight and grain width, and the gene \textit{Ghd7.1} that can alter the heading date and yield; the male parent WSSM provided the gene \textit{NOG1} that increases the number of grains per panicle and the gene \textit{PTB1} which increases the seed-setting rate. In this way, the hybrid combination can have high and stable yields.

In terms of abiotic stress, both parents also possessed some excellent allelic variation sites. These include the \textit{BET1} gene for boron toxicity tolerance, the \textit{OsHKT1.1} gene for improving salt tolerance, the \textit{Sub1A} gene conferring submergence tolerance, and the \textit{OsTPP7} gene for augmenting tolerance to anaerobic germination. In terms of nutrient absorption, both parents contained the genotypes \textit{NRT1.1B} and \textit{OsNR2} that enhance nitrogen-use efficiency, and the variant genotype \textit{TOND1} that was more resistant to stress incurred for low nitrogen availability. They also contain a variant genotype of \textit{OsGSK2} capable of increasing the mesocotyl length, which should make seeds more likely to germinate.

In addition, Quan9311A contained an excellent variant of the \textit{Sdr-4} genotype for seeds’ germination tolerance on panicle, as well as the \textit{TIG1} genotype for a reduced tillering angle, but neither was present in WSSM. However, the unfavorable genotype \textit{hbd2} (weakened cold tolerance) contained in Quan9311A was absent from WSSM. Conversely, the other two unfavorable genotypes \textit{D2} (diminished plant height and wider tiller angle after the mutation) and \textit{OsSRO1c} (less drought-tolerant) contained in WSSM were absent from Quan9311A (Hong et al. 2003; You et al. 2012). Hence, their hybrid-rice combination QYSM gained complementarity in these prominent traits, fostering favorable phenotypes while weakening the unfavorable phenotypes.

From the results of this extensive genotype analysis, besides a series of excellent genotypes shared by Quan9311A and WSSM, there was strong evidence of achieved complementarity in multiple functional genes, such as those for resistance to rice blast, bacterial blight, rice planthopper in addition to bolstered yield, among others. The hybrid rice combination QYSM also harbored multiple heterozygous mutations in its whole genome. This complementation and aggregation of dominant genes were also intrinsic causes of heterosis. Endowed with the desirable functions of these genes, QYSM should have extensive adaptability and high-productivity when cultivated across various ecological regions.
Discussion

Since the development of hybrid rice, it has become necessary to fully utilize genetic varieties of different backgrounds for parental improvement, to avoid adverse effects of in-breeding, to widen the genetic distance between rice traits, and to achieve breakthroughs in rice breeding. But the improvement of sterile lines in three-line hybrid rice is more difficult and takes a longer time. Therefore, finding innovative ways to breed excellent male sterile lines is an imperative task which can hasten the development of adaptable, productive hybrid rice varieties. Quan9311A is an improved male sterile line, this accomplished by the two-line restorer line 93−11 as the core germplasm. In the breeding process, the ability to restore the fertility of the sterile line was purposefully removed, so that the three-line sterile line Quan9311A was endowed with the following sought-after traits: better grain quality, complete pollen abortion, stable sterility, strong tillering ability, good recoverability, and strong combining ability. This sterile line retains the excellent characters of the restorer line 93−11 and the maintainer line Zhong9A. Quan9311A is also the first super high-quality three-line sterile line to have been bred in China. In 2011, it was tested by the Rice and Product Quality Supervision, Inspection and Testing Center of the Ministry of Agriculture, who concluded the rice quality met their first-class standard (NYT593-2013, China). Through the RiceNavi analysis, we can clearly infer that during the breeding process of Quan9311A, the $Rf3$ mutant genotype with restoring function in 93−11 had been eliminated, whereas the $Rf3$ genotype without restoring function in Zhong9B was retained. At the same time, the $Ghd7$ mutant genotype of Zhong9B was retained, rendering it a male sterile line with shorter plant height and earlier heading date.

The genotypes and genetic sources related to excellent traits, namely grain yield, plant height, heading date, disease resistance, stigma exertion rate, and nitrogen-use efficiency in Quan9311A were all confirmed by the RiceNavi analysis. It is also clear that Quan9311A contains a number of superior allelic loci, such as $Gn1a$, $Sd1$, $GS3$, $Ghd7$, $Ghd7.1$, $Xa4$, $Xa1$, $NAL1$, $LAX1$, $pi25/pid3$, $pia$, $NRT1.1B$, $OsNR2$, and $Sdr4$. These results not only provide timely guidance for the optimal matching of Quan9311A with other lines, but they also reveal how Quan 9311A was pioneered using a restorer line to create male sterile lines in its breeding process, thus providing new ideas and genomic support methods for improving the performance of male sterile lines of rice.

Beyond that, by comparing the functional genes contained in QYSM, WSSM, and Quan9311A, we found that QYSM not only harbors a series of homozygous excellent alleles but also contains many heterozygous loci. The aggregation of these homozygous superior alleles conferred to QYSM obvious phenotypic advantages. Those heterozygous superior alleles also formed complementary genotypes in QYSM, and confer the dominant phenotypes, so that the excellent phenotypes of the two parents were partially expressed in QYSM. Moreover, the unfavorable alleles contained in both parents were neutralized, and the unfavorable phenotypes they exhibited were also weakened. Under the combined action of multiple genes, QYSM featured a moderate tiller angle, better leaf morphology, adequate defense against diseases and pests, lodging resistance, high yield, and good adaptability, traits that altogether constitute the best manifestation of heterosis and dominant gene aggregation.
Yet we did find that although Quan9311A had excellent field traits, such as better stigma vigor, good glume-closure, and rapid yellowing of leaf color at maturity, it lacked corresponding genotypic support. Even for one of its most prominent advantages, the strong matching ability, it was difficult to find one or more genes to support it. Further, it seems the commercial value of some genes in Quan9311A remains to be explored, such as the submergence tolerance gene \textit{Sub1A}, the mycorrhizal symbiosis gene \textit{OsCERK1}, both salt tolerance genes \textit{OsHKT4} and \textit{OsHKT1.1}, and the nitrogen-deficiency tolerance gene \textit{TOND1}.

In fact, the successful breeding of Quan9311A depended on more than the ingenious ideas of its breeders. It required that breeders invest much time and effort, perseverance during the selection process of each generation of huge populations. For example, during the breeding process, a population of about 5000 rice plants was built in the $F_5$ generation. In the $F_6$ generation, 215 individual plants were selected for the cross with Zhong9A according to their field phenotype. Then the ensuing hybrids’ seeds and preferred individual plant seeds were sown, and then their phenotypes observed in the field, one by one, with the pollen of each individual plant collect for indoor microscopic examination. These tedious efforts also showed that breeding efforts warrant support, advanced technologies, including genomics, to reduce the workload and speed up breeding.

Obviously, there are some problems with the current development of genomic breeding technology. First, molecular theoretical research of functional genes tends to evaluate the traits of a single gene or a single locus; hence, it cannot truly reflect the breeding value of related loci which distributes on separate chromosomes. Some genes are capable of epistatic interactions, and target traits can only be expressed phenotypically by forming a regulatory network with other interdependent genes. A sole locus is unable to convey the actual effect of the breeding combinations. Secondly, the current genomic breeding tools are imperfect and cannot completely solve the practical problems encountered in the breeding process, and a certain distance has yet to be bridged between the use cost and the expected outcomes. Thirdly, genomics methods are generally difficult to understand, leaving many breeders reluctant to accept their use or findings. Moreover, many breeding institutions in China still lack the proper equipment and technical support to carry out genome breeding, which limits the development of genome based breeding for rice varieties.

**Conclusions**

Chinese breeders have shown new ideas in the selection and breeding of rice varieties, which is paramount for ensuring China’s food security. In the future, whole-genome sequencing and RiceNavi analysis methods will enable breeders to better understand their own breeding materials than traditional breeding techniques; to know of the advantages, as well as disadvantages, of rice breeding materials; to comprehensively understand the inherited genotype proportion of different parental lines in their offspring, in the process of variety breeding; and to meet the development needs of system generating diverse varieties. More precisely, we should be able to determine whether differing superior genes are actually introduced into breeding material. Through the RiceNavi analysis, the optimal variety improvement plan can be given to appropriately reduce the breeding threshold and improve the breeding
efficiency. With the development of modern biotechnology and under the guidance of genomics
techniques, it ought to be possible to achieve a level of high efficiency that traditional breeding struggles
to attainment, search more genetic backgrounds, rapidly aggregate more superior genes, and cultivate
breakthrough large varieties, to realize the vision of targeted breeding of rice.

Design breeding still faces some challenges. Fortunately, more molecular biology experts are increasingly
paying attention to these and assisting in breeding efforts, while more breeders are beginning to embrace
or use genomic breeding techniques. We are optimistic that for most research-based rice field settings,
molecular design breeding methods with Chinese characteristics could be developed soon.

Materials And Methods

Breeding process of Quan9311A/B

Quan9311A was bred by cross-breeding among different rice varieties and screening the hybrid offspring
based on dominant agronomic characters. A more important point for its successful breeding is the
selection of suitable female and male parents during the crossbreeding process. The female and male
parents used in each generation of crosses are different from those used in the previous generation
(Fig. 1b). Spring 2003 in Hainan, 222B and 93 – 11 were selected as the female parent and the male
parent respectively for crossbreeding to obtain F\textsubscript{1} seeds. Summer 2003 in Hefei, Zhong9B was used as
the female parent, the hybrid F\textsubscript{1} of 222B and 93 – 11 was used as the male parent to cross again, and the
three-parent hybrid F\textsubscript{1} was obtained. After that, the three-parent hybrid F\textsubscript{1} seeds and their selfing-progeny
were successively sown until the F\textsubscript{5} generation. In each generation, individual plants with desirable
agronomic traits were selected for inbreeding. By the F\textsubscript{5} generation, 215 individuals with desirable
agronomic traits were screened. The desirable agronomic traits here means that in addition to the specific
phenotypes of maintainer lines and sterile lines (short plant height, more tillers, moderate whole growth
period, and high stigma exertion rate), it should also have dominant agronomic characters of 93 – 11, such as stout stems, lodging resistance, longer panicle length, heavier 1000-grain weight. Spring 2006 in
Hainan, 215 groups of new F\textsubscript{1} seeds were obtained with Zhong9A as the female parent and 215
individuals as the male parent. In the summer of 2006, these 215 groups of new F\textsubscript{1} seeds and their male
parents(215 individuals screened on the last generation ) were planted in Hefei. The phenotypic
requirements for screening the sterile line are: the hybrid progeny of sterile and maintainer lines differ
only in fertility phenotype, and other agronomic traits are consistent with those of the maintainer line.
Accordingly, 73 sets of sterile lines and maintainer lines were screened. However, the genomes of these
lines are heterozygous, and their phenotype cannot be stabilized. Therefore, these maintainer lines are
planted every year for selfing, and individuals with desirable agronomic traits are screened from them (as
the male parent), and sterile individuals are screened from the offspring obtained by crossing the
corresponding last generation maintainer lines with sterile individuals(as the female parent), carry out
hybridization, then sow the hybrid seeds to continue to screen sterile individuals with desirable agronomic
traits as the female parent, and cross with the corresponding maintainer line male parent, do cross like
this until Spring 2010. The maintainer line was inbred until the F\textsubscript{13} generation, and the sterile line was backcrossed until the BC\textsubscript{7}F\textsubscript{1} generation, and their dominant agronomic phenotypes were stable. So far, the sterile line was named Quan9311A, and the corresponding maintainer line was named Quan9311B. Therefore, from Quan9311B to Quan9311A, it is also a purposeful improvement process. The sterile lines and maintenance lines based on the three-line hybrid system differed only in the fertility phenotype and were consistent in other agronomic traits. Their genomic origin and genomic background are basically the same. Therefore, we choose Quan9311A to represent the set of Quan9311A and Quan9311B, and Zhong9A to represent the set of Zhong9A and Zhong9B.

**Plant materials**

The 93 – 11 is an indica cultivar that is often used as a two-line restorer line in hybrid rice breeding. Zhong9A is an indica-type three-line sterile line that was used in the early years of rice breeding. Quan9311A is an indica-type three-line sterile line, it bred by crossing 93 – 11 and Zhong9A. As the female parent, 222B was chosen to cross with 93 – 11 because of its many forming tillers (so it may improve the tillering ability of 93 – 11 or Zhong9A). However, since its genotype was heterozygous at the early stage of Quan9311A breeding, there were no samples available for sequencing and doing the RiceNavi analysis. Wushan Simiao (WSSM) is an indica cultivar, often used as a three-line restorer line for hybrid rice. Quanyou Simiao (QYSM) is an indica-type three-line hybrid rice combination, with Quan 9311A as its female parent and Wushan Simiao as its male parent. All rice varieties used in this study came from the Win-All High-Tech Seed Industry Co., Ltd., based in Anhui, China.

**DNA extraction and genome sequencing**

Genomic DNA was extracted from fresh young leaves of each individual by using the Hi-DNAsecure Plant Kit according to the manufacturer's instructions (Tiangen Biotech, Beijing, China). All samples were fragmented into 500-bp lengths to construct the Illumina sequencing libraries, using the KAPA HyperPrep Kit (Kapa Biosystems, Roche, USA) and following the manufacturer's recommendations. The libraries were sequenced on Illumina HiSeq 4000 platform under its paired-end 150-bp mode. A total of 183 Gb raw data were obtained for five varieties.

**High-throughput genotyping**

In this study, Quan9311A, 93 – 11, Zhong9A, WSSM, and QYSM were each sequenced by whole-genome high-throughput sequencing technology. Trimmomatic-0.38, Bwa-0.7.1, Samtools-1.9, GATK v4.1.4.1 (The Genome Analysis Toolkit) software tools were used to detect variants in the original sequencing data and the reference genome Nipponbare (MSU7.0); the variation site information of each variety relative to the Nipponbare reference genome was thus obtained. Next, we used the awk command to link the respective SNP loci of 93 – 11 and Zhong9A with the SNP locus of Quan9311A, to obtain mutually consistent SNP loci, and then we calculated the similarity between them. Meanwhile, the whole genome of rice was divided into 3740 intervals of 100 kb each, and all SNPs found in each interval were counted. An interval was considered to be indifferent to breeding heritability if it had fewer than 200 SNPs. Based on this
information, graphs were drawn to show the differential distribution of SNPs in the genome between the two parents.

**Functional loci analysis by RiceNavi**

Based on the SNP information, the QTNpick module in RiceNavi software was used to genotype the variants present at each QTN (quantitative trait nucleotide) site, to obtain the QTN genotype of each rice variety (https://github.com/xhhuanglab/ RiceNavi or http://www.xhhuanglab.cn/tool/RiceNavi.html). The rice materials’ phenotype was interpreted according to the genetic effect of QTN. These QTN loci—348 in total, where for each gene variation can lead to changes in gene function and related phenotypic changes—entailed 225 reported QTL genes in rice, all of which are naturally mutated genes; hence, there were artificially induced mutations or reversed mutations in the data. Therefore, these genes present in common rice varieties could be transferred or acquired by crosses between different cultivars during the breeding of rice (Wei et al. 2021).

**Abbreviations**

CMS: cytoplasmic male sterile; HL-CMS: Honglian type cytoplasmic male sterility; BT-CMS:Boro II cytoplasmic male sterile line; QTL: quantitative trait loci; SNPs: single nucleotide polymorphisms; QTN: Quantitative Trait Nucleotide; RiceNavi: rice genome navigation; MNP: multiple nucleotide polymorphism; WSSM: Wushan Simiao; QYSM: Quanyou Simiao. TGW: Thousand-grain weight.

**Declarations**

**Acknowledgements**

The authors thank Ms. Qin Zhang, President of the National Rice Commercial Molecular Breeding Technology Innovation Alliance for support our cooperation platform. We also thank Dr. Xin Wei and Dr. Jie Xu, who developed the RiceNavi program.

**Author contributions**

WAH and FY designed the experiment and suggested the structure and content of the manuscript. CJJ, WHQ, ZCH and ZYH are the breeder of Quan9311A. WAH and LQL analyzed the data. FQ performed genome sequencing and revise the manuscript. LQL, KYH, DM and WH were responsible for field experiment. HB supervised the project and supplied useful suggestions. LQL, FY and WAH wrote the paper. All authors read and approved the final manuscript.

**Funding**

This work was supported by the National Natural Science Foundation of China (Grant No. 32001516), Project on Science and Technology for Agricultural Development of Shanghai (NYNCW-SH-2019-1-3), and
Rice Industry of China Agriculture Research System (CARS-01-03).

Availability of data and materials

The data supporting the conclusions of this article are included within the article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


Figures
Figure 1

Phenotypic characteristics and breeding process of Quan9311A.

a Plant morphology of 93-11, Zhong9A, Quan9311A. 93-11 (authorized name: Yangdao6) has higher plant height(110-115cm), longer panicle length(24cm), longer whole growth period(135-144days), fewer tiller number(8), heavier thousand-grain weight(30g), suitable panicle extension and panicle shape, a large number of secondary branches and grain number, high seed-setting rate, fewer glume-lemma hairs, yellow glume tip and yellow glume color, longer grain length, intermediate grain width, slender grain shape, white seed coat color, and strong restore ability. (https://www.ricedata.cn/variety/varis/600611.htm). Zhong9A is a three-line male sterile line of Indonesian Shui tiangu (ID) type, with lower plant height(80-83cm), shorter panicle length(19.4cm), shorter whole growth period(84-95days), suitable tiller number(13) and thousand-grain weight(24.4g), a large number of grain number, plant and leaf shape favorable for photosynthesis, and small tiller angle, intermediate-length flag leaves, concentrated flowering time, white stigma, high stigma exsertion rate and outcrossing rate; slender grain shape, thinner paddy husk, white seed coat color. (https://www.ricedata.cn/variety/varis/601141.htm).

b Quan9311A/B breeding pedigree.
Phenotypes of 93-11, Quan9311A and Zhong9A. Plant height, Whole growth period, Tillers number, Panicle length, Thousand-grain weight.

Figure 2

SNPs discrepancy between the two parents on the whole genome.

a SNPs distribution between Quan9311A and 93-11.

b SNPs distribution between Quan9311A and Zhong9A.

The x-axis refers to the corresponding physical position of each chromosome in a and b. According to the physical division of an interval of 100Kb, 3740 intervals are distributed on 12 chromosomes in the whole genome. The y-axis is the number of SNPs in each interval in a and b.
Figure 3

Superior genes inherited by Quan9311A from both parents 93-11 and Zhong9A. The unique genes in 93-11 and Zhong9A are represented by yellow and brown bar respectively. The genes with arrows in circles are inherited by Quan9311A. Genes without circles and arrows were not inherited by Quan9311A. Dark blue bar indicates the superior genes simultaneously contained in the three varieties, and each bar represents a superior gene.
Figure 4

Superior genes for QYSM aggregated from Quan9311A and WSSM. The unique genes in Quan9311A and WSSM are represented by blue and yellow bar respectively. Purple bar indicates the superior genes simultaneously contained in the three varieties, and each bar represents a superior gene.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFig.1.tif
- SupplementaryFig.2.tif
- 20220526supplementaryDataset.xlsx