

# Genetic Diversity and Population Structure of the Parental lines of Hybrid Japonica Rice in Northern China Revealed by 8K SNP-Chips

**Hong Gao**

Liaoning Academy of Agricultural Sciences

**Zuobin Ma**

Liaoning Academy of Agricultural Sciences

**Yuanzheng Wang**

Liaoning Academy of Agricultural Sciences

**Changhua Wang**

Liaoning Academy of Agricultural Sciences

**Zhiqiang Tang**

Liaoning Academy of Agricultural Sciences

**Liyang Zhang**

Liaoning Academy of Agricultural Sciences

**Liang Fu**

Liaoning Academy of Agricultural Sciences

**Na He**

Liaoning Academy of Agricultural Sciences

**Hui Wang**

Liaoning Academy of Agricultural Sciences

**Yuanjun Bai**

Liaoning Academy of Agricultural Sciences

**Guomin Sui**

Liaoning Academy of Agricultural Sciences

**Wenjing Zheng** (✉ [zwj27@126.com](mailto:zwj27@126.com))

Liaoning Academy of Agricultural Sciences <https://orcid.org/0000-0001-8692-7090>

---

**Original article**

**Keywords:** Hybrid japonica rice, the restorer lines, the sterility lines, gene diversity, population structure, indica blood

**Posted Date:** January 7th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-139546/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

The hybrid rice varieties have made a significant contribution to food security. Although there has been rapid development of hybrid *indica* rice variety, the development of hybrid *japonica* rice has been relatively slow. This study aimed to understand the genetic background of representative parental lines of hybrid *japonica* rice in northern China that were beneficial for increasing efficiency to find a superior breeding combination using a restorer line and a sterile line.

## Results

We selected 137 parental lines of hybrid *japonica* rice, including 90 restorer lines, 47 sterile lines, which broadly represented the recent rice breeding trends in China. These lines were genotyped using 8K SNP-Chips (China Golden Marker Biotechnology Co. Ltd.) to understand the genetic diversity, population structure, phylogenetic evolution, and *indica* blood content. The genetic diversity of total parental lines averaged 0.264, with values for the restorer line and sterility line as 0.287 and 0.148, respectively. Based on the model-based population structure analysis and distance-based clustering, these 137 lines were divided into 14 groups, including seven independent restorer lines groups and seven mixture groups. There were 70% restorer lines in the above-mentioned seven independent restorer lines groups, and the *indica* blood content was 0.348, while 30% restorer lines were genetically similar to the sterile lines and constituted the other seven mixture groups, where the *indica* blood content in the restorer lines and the sterility lines was 0.142 and 0.121, respectively. The results of distance-based clustering revealed that the restorer lines, Group 1 and Group 2 (containing only restorer lines), had longer genetic distances with groups containing mainly sterile lines (the genetic distance ranged from 0.672 to 0.788), which served as a potential heterotic for hybrid rice breeding. This observation was consistent with the breeding strategy of high yield hybrid *japonica* rice.

## Conclusions

The typical *japonica* sterile lines were crossed with the restorer lines containing high *indica* components indicating a strong heterosis pattern was a feasible scheme for heterosis utilization of *indica-japonica* subspecies. Thus, the effective ways to further improve the rice quality of hybrid *japonica* rice in northern China included maintaining moderate genetic distance and *indica* components between the parental lines along with the excellent quality of both the parental lines.

## Background

The hybrid rice varieties have made a significant contribution to food security as they produce approximately 15–20% higher yield than the inbred rice varieties. The development of hybrid *indica* rice was rapid, and the planting area accounted for > 80% of *indica* rice and more than 50% of rice. On the contrary, the development of hybrid *japonica* rice was relatively slow, and its planting area accounted for

< 3% of *japonica* rice (Chen et al. 2015; Sui 2018). This limited development was mainly attributed to the fact that the heterosis in hybrid *japonica* rice was weaker compared with hybrid *indica* rice, which was due to the narrow genetic distance between the parental lines. Therefore, it was necessary to analyze the genetic background of the parental lines of hybrid *japonica* rice to select elite combinations, and it constituted an important technical strategy to improve its breeding efficiency (Pu et al. 2015; Xie and Peng 2016; Yang et al. 2016; Lin et al. 2020; Zheng et al. 2020). The genetic diversity and population structure of parental lines is an effective reference for breeding selection. In the past, genealogical information along with the origin or the phenotypic variation has been used to evaluate the genetic diversity and structure populations; however, the results had low reliability (AlmanzaPinzón et al. 2003; Fufa et al. 2005). The molecular marker technology provides an effective tool to study the genetic differences between germplasm resources. The results showed that in China, the genetic resources for three-line hybrid rice production with low genetic differences were scarce, and the resources with large genetic differences were difficult to be used for production, which further restricted the development of hybrid rice (Duan et al. 2002; He et al. 2006; Zhao et al. 2010). Moreover, *japonica* rice has low genetic diversity than *indica* rice. The *japonica* rice lacks the restorer gene, which is introduced from the *indica* rice. Thus, a single source of the restorer gene and a single type of male sterile line are the bottlenecks of *japonica* hybrid rice breeding in northern China (Hong et al. 2009; Wang et al. 2010; Jin et al. 2014). Although there are several reports on molecular marker-based genetic diversity analysis of parental lines of *indica* hybrid rice, there are only a few reports on the genetic diversity analysis and population structure of the parental lines of *japonica* hybrid rice in northern China (Qiu et al. 2005; Chen et al. 2009; Wang et al. 2009; Cheng et al. 2012; Zhang et al. 2014; Chen et al. 2017). Additionally, previous studies were mostly focused on restriction fragment length polymorphisms (RFLP) and simple sequence repeats (SSR) markers. There are no published reports on the use of single nucleotide polymorphism (SNP), a more efficient molecular marker, to study the genetic diversity of the parental lines of *japonica* hybrid rice (Sun et al. 2012; Xu et al. 2016; Gao et al. 2018; Wang et al. 2018). Here, we analyzed the genetic diversity, population structure, phylogenetic evolution, and *indica* blood of 137 representative parental lines of hybrid *japonica* rice in northern China with important breeding value using 8K SNP-Chips, developed by China Golden Marker Biotechnology Co. Ltd. The results would help us understand the genetic differences among these varieties and provide guidance and reference for the breeding of strong heterosis combinations to effectively utilize restorer lines and sterile lines to breed hybrid *japonica* rice.

## Results

### Characterization of SNPs and genetic diversity of the parental lines of hybrid *japonica* rice in northern China

Of the 137 parental lines of hybrid *japonica* rice, 4172 high-quality SNPs were obtained for genetic background analyses, from which 8344 alleles were detected, each with two expected alleles (Fig. 1a). The average PIC value of the 137 parental lines was 0.215 within a range of 0.014–0.375, and the PIC values for the restorer lines and the sterile lines were 0.231 and 0.124, respectively, within a range of 0.00–0.375 (Fig. 1c, Table 1). The average gene diversity of the 137 parental lines was 0.264 within a range of

0.014-0.500; the gene diversity of the restorer lines and the sterile lines were 0.287 and 0.148, respectively, within a range of 0.0-0.500 (Fig. 1b, Table 1). Thus, the average values of PIC and genetic diversity of the restorer lines were higher than those of the sterile lines. Of the 8K SNP-Chips, we obtained 3190 SNPs, with differences between Nipponbare and 93 - 11 genomes, for *indica* blood analyses. The average *indica* blood of the 137 parental lines was 0.237 within a range of 0.062–0.623; for the restorer lines was 0.297 within a range of 0.065–0.623; for the sterile lines was 0.121 within a range of 0.062–0.295 (Fig. 1d, Table 1). Thus, the *indica* blood of the restorer lines was higher than the sterile lines.

Table 1  
Summary statistic of genetic diversity and indica blood

<b>Groups</b>	<b>No. of lines</b>	<b>Gene Diversity</b>	<b>Polymorphism information content</b>	<b>Range of <i>indica</i> blood</b>	<b>Average of <i>indica</i> blood</b>
All lines	137	0.264	0.215	0.062–0.623	0.237
Restorer lines	90	0.287	0.231	0.065–0.623	0.297
Sterile lines	47	0.148	0.124	0.062–0.295	0.121
Gp_1	11	0.196	0.164	0.313–0.623	0.513
Gp_2	6	0.211	0.170	0.451–0.591	0.526
Gp_3	10	0.192	0.160	0.267–0.442	0.388
Gp_4	7	0.115	0.091	0.332–0.410	0.371
Gp_5	4	0.192	0.153	0.295–0.484	0.354
Gp_6	12	0.156	0.130	0.202–0.352	0.285
Gp_7	11	0.095	0.079	0.171–0.245	0.207
Gp_8	7	0.118	0.097	0.132–0.247	0.169
Gp_9	6	0.057	0.048	0.137–0.244	0.160
Gp_10	6	0.127	0.106	0.124–0.372	0.192
Gp_11	9	0.147	0.124	0.087–0.279	0.156
Gp_12	17	0.118	0.099	0.067–0.200	0.131
Gp_13	9	0.063	0.053	0.072–0.137	0.093
Gp_14	22	0.129	0.112	0.062–0.218	0.103

## Population structure of the parental lines of hybrid japonica rice in northern China

In the model-based grouping analysis, when  $K = 2$ , the parental lines of hybrid *japonica* rice showed apparent differentiation; 42.2% of the restorer lines (38 lines) were in a single cluster and had high *indica* blood averaging 0.442 (Fig. 1f, Additional files 1: Table S2). Also, 57.8% of restorer lines (52 lines) and sterile lines (47 lines) were divided into another cluster, where *indica* blood of the restorer lines and the sterile lines were 0.191 and 0.121, respectively (Fig. 1g, Additional files 1: Table S2). With an increase in the  $K$  value, further population differentiation occurred in the restorer lines, and the entropy criterion was used to choose the number of ancestral populations that could explain the genotypic data. Thus, we had a clear minimum at  $K = 14$ , suggesting the presence of 14 genetic clusters in the data (Fig. 1e), including seven independent restorer line groups (Group 1, Group 2, Group 3, Group 4, Group 6, Group 7, and Group 9) with a total of 63 restorer lines and seven mixture groups (Group 5, Group 8, Group 10, Group 11, Group 12, Group 13, and Group 14) containing varied proportions of restorer and sterile lines (Fig. 1h). Group 5 was dominated by the restorer lines (three restorer lines and one sterile line (53A)) derived from glutinous rice and had high *indica* blood. Group 8 and Group 10 both included two restorer lines and five and four sterile lines, respectively. Group 11 and Group 13 were dominated by sterile lines, and each group contained a single restorer line with low *indica* blood. Group 12 and Group 14 had 17 and 22 lines, respectively, and the ratios of sterile lines to restorer lines were 9:8 and 6:5, respectively. Thus, 70% of the restorer lines had independent genetic structure and high *indica* blood averaging 0.348, and 30% of restorer lines had a close genetic relationship with sterile lines and the *indica* blood was 0.142 (Additional files 1: Table S3).

## Phylogenetic analysis of the parental lines of hybrid japonica rice in northern China

Next, we constructed a bootstrap NJ tree based on 4172 SNPs variations to further understand the impact of *indica* blood on the genetic structure of the parental lines of hybrid *japonica* rice in northern China. Most sterile lines were clustered together in one region of the tree except for one (53A) and were intermixed with 30% of the restorer lines, which were more genetically similar to the Nipponbare genome. Moreover, 70% of restorer lines were clustered in another region of the NJ tree, which was more genetically similar to the 93-11 genome. The NJ tree, combined with the genetic structure analysis, revealed that *japonica*-type sterile lines in northern China had gradually separated from most sterile lines (Fig. 2).

## Genetic difference and breeding potential of the parental lines of hybrid japonica rice in northern China

The genetic diversity derived from model-based grouping analysis was as follows: Restorer lines > All lines > Group 2 > Group 1 > Group 3 > Group 5 > Group 6 > Sterile lines > Group 11 > Group 14 > Group 10 > Group 8 > Group 12 > Group 4 > Group 7 > Group 13 > Group 9. Group 1 and Group 2 comprised restorer lines and exhibited a high average genetic diversity as well as *indica* blood (Table 1). The genetic diversity of Group 4 (0.115), Group 7 (0.095), and Group 9 (0.057), which were also comprised of restorer lines, was low and should be considered to improve genetic diversity in breeding practice. Group 11, which mainly contained sterile lines, had the highest genetic diversity (0.147), and thus, could be used as an important sterile line reference for parental lines in the future. The genetic diversity of group 14,

composed of 12 sterile lines and 10 restorer lines, was 0.129 with low *indica* blood averaging 0.103, indicating that a neo-diversity of *japonica*-type restorer lines could have gradually arisen through breeding selection by the introduction of limited *indica* lineages in northern China (Table 1, Additional files 1: Table S2).

The overall pairwise comparisons between the 14 groups classified based on model-based grouping indicated significant differentiation between the two model-based groups (Group 1 and Group 2) and Group 8, Group 10, Group 11, Group 12, Group 13, and Group 14, mainly with sterile lines and genetic distance values ranging from 0.672 to 0.788 (Fig. 3). Group 3, Group 4, Group 6, and Group 7 (containing only restorer lines) and Group 5 (contains three restorer lines and one sterile line) showed a modest degree of differentiation with the above-mentioned six model-based groups (containing mainly sterile lines) with the pairwise genetic distance values ranging from 0.225 to 0.534. Lower levels of differentiation were observed in the pairwise comparisons of Group 9 (containing only restorer lines) with the above-mentioned groups (containing mainly sterile lines) with pairwise genetic distance values ranging from 0.153 to 0.271. The abundant genetic differences among the parental lines produced strong heterosis; thus, high levels of differentiation were observed in the pairwise comparisons of Group 1 and Group 2 with the other groups, specifically the groups containing mainly sterile lines. This indicated that there was a large space for heterosis utilization.

## Discussion

### Genetic background of the parental lines of hybrid japonica rice in northern China

The hybrid rice yield largely depends on the diversity of germplasm resources. Thus, it was vital to fully understand the genetic diversity of the parental lines of hybrid rice in China. The discovery of sterile lines has promoted the large-scale production of hybrid rice; however, most of the sterile lines have a similar genetic background and, thus, low genetic diversity, which has created new bottlenecks in the development of hybrid rice (Peng et al. 2008; Shen et al. 2015). In this study, the sterile lines had low genetic diversity, even lower than the restorer lines. The main reason was the low genetic diversity of *japonica* rice compared with *indica* rice. Also, the sterile cytoplasm of *japonica* hybrid rice in northern China only possessed BT-type of pollen abortion, leading to a more singular-type of genetic background. Therefore, it was necessary to breed a new type of cytoplasmic sterile lines to broaden the genetic diversity of *japonica*-type sterile lines in northern China. Additionally, since the *japonica* rice lacked restorer genes, they were derived from *indica* rice. The introduction of *indica* lineage resulted in the presence of *indica*-like components, which increased the genetic diversity of restorer lines compared with sterile lines, which was consistent with the results of previous studies (Cha et al. 2007; Wang et al. 2010; Chen et al. 2017). Therefore, breeding restorer lines with *indica* genetic background was an effective way to expand genetic diversity and the heterosis of progenies. However, we needed to coordinate the appropriate *indica* components with the ecological adaptability of northern hybrid *japonica* rice in low-temperature areas. The progenies from *indica-japonica* crossing generally show early flowering, long growth periods, and poor tolerance of low temperatures. Thus, under the influence of the genetic

background of *indica* rice, hybrid *japonica* rice often exhibits premature senescence and differentiation between strong and weak (filled and partially filled) grains. Additionally, during the later stages of growth, low temperature results in the incomplete filling of the weak grains, which prevents maturation, reducing the yield potential (Jin et al. 2014; Zheng et al. 2020). Previous research showed that the presence of too many *indica* components hindered its adaptation to the ecological conditions of northern China, and too few *indica* components restricted the expansion of the genetic distance between the parental lines, resulting in weak yield heterosis (Wang et al. 2010; Lin et al. 2012; Xie et al. 2015). Northern restorer line breeding is based on the cross between *japonica* rice varieties and restorer lines to reduce *indica* components to regulate premature senescence and differentiation of strong and weak (filled and partially filled) grains. Additionally, it resulted in a small genetic difference between the restorer lines and the sterile lines (Hong et al. 2009; Liu et al. 2018). Here, 70% of the restorer lines had independent genetic structure and differentiated with the male sterile lines, while 30% of the restorer lines were still closely related to the male sterile lines and were present in the same group. These results were not consistent with the previous results of genetic structure analysis or group classification of the parental lines of hybrid *indica* rice, where the *indica* restorer lines and male sterile lines were clustered into different groups (Wang et al. 2006; Yang and Tan 2009; Cheng et al. 2012; Wang et al. 2014; Zhang et al. 2014). There was a minimal genetic difference between partial japonicaclinous restorer lines and sterile lines to avoid the negative effects of the introduction of the genetic background of *indica* rice. Also, with the improvement of germplasm, the population genetic diversity was expected to decline to a certain extent, but in this study, although the japonicaclinous restorer lines had fewer *indica* components, the genetic diversity did not decrease significantly. Thus, the breeding of japonicaclinous restorer lines in northern China was successful in coordinating moderate *indica* composition and genetic diversity.

### **Genetic difference and breeding potential of parents of hybrid japonica rice in northern China**

It is necessary to understand the genetic distance between the parental lines of hybrid rice to breed excellent hybrid combinations. The parental lines that are clustered in different groups produce greater heterosis (Qiu et al. 2005; Zhao et al. 2009; Xie et al. 2012; Xu et al. 2017). However, the results from previous studies have been inconsistent: either the correlation between genetic distances based on molecular markers and heterosis has not been significant or small, or the degree of correlation has shown variation with research materials and markers used (Xu et al. 2009; Singh et al. 2011). Thus, these results indicate that the generation of heterosis has a very complex biological basis. There are few studies that have studied the relationship between genetic distance and heterosis of *japonica* hybrid rice in the northern climatic conditions; however, there has been a lack of in-depth and detailed analysis. Here, we analyzed 137 representative parental lines of hybrid *japonica* rice in northern China, with important breeding value, with 4172 high-quality SNPs. The results showed that the parental lines in Group 1 and Group 2 (consisting only restorer lines) crossed with sterile lines (present in Group 5, Group 8, Group 10, Group 11, Group 12, Group 13, and Group 14) could produce strong heterosis based on the genetic distance, which has been supported by practical application. Recently, Liaoyou 9906 released by the Liaoning rice research institute, a hybrid *japonica* rice popularized in northern China in recent years, resulted from a cross between the restorer line C2106 (in Group 1) and the sterile line 99A (in Group 13)

and had a high value of genetic distance between the two groups (0.745). Also, the *indica* blood of the restorer parental line and the sterile parental line was 56.41% and 7.28%, respectively. The breeding strategy involved the use of a typical *japonica* male sterile line as the female parent, and crossed with a restorer line containing an appropriate number of *indica* components, which could retain the characteristics and ecological adaptability of the female parent, meanwhile, increasing the genetic distance between the parental lines as much as possible to ensure the seed setting rate. Based on the current research progress, this strategy is still a feasible scheme for heterosis. Also, it is an effective way to further utilize the heterosis of *indica-japonica* subspecies by constructing intermediate materials, expanding the genetic distance between the parental lines, and utilizing restorer lines with wide compatibility and a wide restoring spectrum. However, it should be noted that although the northern hybrid *japonica* rice, under the influence of the *indica* genetic background, showed strong biological advantages, it often exhibited premature senescence and reduced seed setting rate and quality at low temperatures during the later stages of growth. Jingyou 653, a new hybrid good-quality *japonica* rice also released by the Liaoning rice research institute, is a successful case of coordinate consanguinity of *indica*. The parental lines of Jingyou 653 were restorer line C315 (in Group 7) and sterile line 65A (in Group 11) and had a moderate value of genetic distance between two groups (0.273). The *indica* blood of the restorer parental line and the sterile parental line were 24.54% and 11.05%, respectively. Although there was no significant difference in the *indica* blood and the genetic distance between the groups, the quality, resistance, and maturity were greatly improved while maintaining the yield. The breeding strategy basically followed the technical route of "double fitness and double excellence," i.e., the genetic distance and *indica* components between the parental lines were moderate, and the quality of both parental lines was excellent. Therefore, the quality of hybrid *japonica* rice could be improved by trans breeding the sterile line from the abundant good-quality *japonica* resources and crossing them with restorer lines with good appearance quality.

## Conclusions

The introduction of *indica* lineage increased the genetic diversity of restorer lines compared with sterile lines, breeding restorer lines with *indica* genetic background was an effective way to expand genetic diversity and the heterosis of progenies. However, under the influence of the genetic background of *indica* rice, northern hybrid *japonica* rice often exhibits premature senescence and differentiation between strong and weak (filled and partially filled) grains in low-temperature areas. Thus, the effective ways to further improve the rice quality of hybrid *japonica* rice in northern China included maintaining moderate genetic distance and *indica* components between the parental lines along with the excellent quality of both the parental lines.

## Materials And Methods

### Plant Materials

We selected 137 parental lines of hybrid *japonica* rice for genetic analysis(Additional files 1: Table S1), including 90 restorer lines, 47 sterile lines, which were mainly collected from Liaoning Province in northern China. Nipponbare (*japonica* reference genome) and 93-11 (*indica* reference genome) were used to calculate the *indica* and *japonica* blood of the parental lines of hybrid *japonica* rice.

## Genomic DNA Extraction and SNP Genotyping

Twenty days post-transplantation, a leaf sample was collected from one plant of each accession, followed by DNA extraction using the modified cetyltrimethylammonium bromide (CTAB) method. The SNP genotyping was done using the 8K SNP-Chips, developed by China Golden Marker Biotechnology Co. Ltd. We excluded the monomorphic SNPs or the ones with >50% missing data. A total of 4172 high-quality SNP loci were obtained for further analyses.

## Data Analysis

PowerMarker v3.25 (Liu and Muse 2005) was used to compute the polymorphic information content (PIC) as well as other diversity parameters of the samples. An R package for landscape and ecological association studies (LEA) was used for inference of individual admixture coefficients using sNMF to infer population structure and assign individuals to them (Frichot and Francois. 2015). We considered models that had approximately 1-20 groups (K) with admixture and correlated allele frequencies and applied 100 iterations for each K. The number of groups was determined by the cross-entropy method (Alexander and Lange. 2011; Frichot et al. 2014). The genetic distance matrix was analyzed using the neighbor-joining (NJ) algorithm with MEGA-X (Kumar S. 2018).

Next, the *indica* and *japonica* blood of the parental lines of hybrid *japonica* rice were calculated based on 3190 SNP loci with the difference between Nipponbare(*japonica* reference genome) and 93-11 (*indica* reference genome): the loci consistent with "Nipponbare" were recorded as jj; the loci consistent with "93-11" were recorded as ii; the heterozygous loci were marked as ij; the number of loci was N. The *indica* blood (Fi) and *japonica* blood (Fj) of each sample was calculated as follows:

$$F_i = \left[ 2 \sum_1^x X_{ii} + \sum_1^x X_{ij} \right] / 2N$$

$$F_j = \left[ 2 \sum_1^x X_{jj} + \sum_1^x X_{ij} \right] / 2N$$

## Abbreviations

**PIC:** Polymorphic information content

**SNP:** Single nucleotide polymorphism

## Declarations

### Acknowledgements

We sincerely thank funding supports and all the participants.

### Authors' Contributions

WJZ and GMS designed and supervised the researches. HG, ZBM, YZW, CHW, ZQT, LYZ, LF, NH, HW, YJB performed the experiments. HG, ZBM and YZW analyzed the data. HG and ZBM wrote the manuscript. All authors read and approved the final manuscript.

### Funding

This research was funded by Liaoning Key Agricultural Program (2019JH1/10200001–2), Liao Ning Revitalization Talents Program (XLYC1802119), Hybrid Rice Breeding Subject (2019DD133719), and Natural Science Foundation of Liaoning Province (2019-MS-203).

### Availability of Data and Materials

All data supporting the conclusions of this article are provided within the article (and its Additional files).

### Ethics Approval and Consent to Participate

This study complied with the ethical standards of China, where this research work was carried out.

### Consent for Publication

All authors are consent for publication.

### Competing Interests

The authors declare that they have no competing interests.

## References

Alexander DH, Lange K (2011) Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12: 246

AlmanzaPinzón MI, Khairallah M, Fox PN, Warburton ML (2003) Comparison of molecular markers and coefficients of parentage for the analysis of genetic diversity among spring bread wheat accessions. *Euphytica* 130 (1): 77-86

- Chen J, Peng YL, Zhang SH, Zhuo CF, Luo F, Tang Y, Yang GT, Wang XC, Chen YJ, Lv QM, Hu YG (2017) Genetic diversity of parents of hybrid rice from upper Yangtze River. *Chinese Journal of Applied & Environmental Biology* 23(4): 595-601
- Chen J, Qiu L, Tu SB, Luan L, Wang X, Long WB (2009) Genetic variation of three-line parents of indica hybrid rice detected with SSR markers. *Chin J Appl Environ Biol* 15 (5): 585-590
- Chen LY, Lei DY, Tang WB, Deng HB, Xiao YH, Zhang GL (2015) Challenges and strategies of hybrid rice development in China. *Hybrid Rice* 30(5): 1-4
- Cheng BS, Luo BX, Yang JY (2012) Analysis of genetic similarity among 44 parents of japonica hybrid rice by simple sequence repeats (SSR). *Chinese Journal of Tropical Crops* 33(9): 1625-1629
- Duan SH, Mao JN, Zhu YG (2002) Genetic variation of main restorer lines of hybrid rice in China was revealed by microsatellite markers. *Journal of Genetics and Genomics* 29(3): 250-254
- Frichot E, Francois O (2015) LEA: An R package for Landscape and Ecological Association studies. *Methods in Ecology and Evolution* 6: 925-929
- Frichot E, Mathieu F, Trouillon T, Bouchard G and Francois O (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196:973–983.
- Fufa H, Baenziger PS, Beecher BS, Dweikat I, Graybosch RA, Eskridge KM (2005) Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. *Euphytica* 145 (1): 133-146
- Gao H, Xu Q, Zhong M, Xu ZJ, Sui GM (2019) A Neo-diversity group of *Oryza Sativa* L. SSP. japonica in northeastern china evaluated by large-effect SNPs. *Pak. J. Bot.* 51(2): 583-600
- He HH, Luo XJ, Zhu CL, He XP, Fu JR, Sun JL, Zhang HL, Li ZC (2006) Classification for some sterile lines and their restorers of Hhybrid rice with SSR markers. *Acta Agronomica Sinica* 32(2): 169-175
- Hong RK, Jin SL, Tan YL, Wen JC, Xu HY, Guo A, Huang DJ, Zhang ZL, Tao GX, Chen LJ, Wang X, Li ZY, Tan XL (2009) Innovations advances of techiques and germplasm in CMS-D1 based on beeding of hybrid japonica rice. *Southwest Chin J Agric Sci* 22(4): 1170-1174
- Jin CH, Gao MJ, Zhou GC, Quan DX, Meng WR, Piao HM (2014) Research Aadvance on restorer lines of japonica hybrid rice in northern China. *Modern Agricultural Science and Technology* (23): 63-64
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547-1549
- Lin JR, Song XW, Wu MG (2012) Biological characteristics and heterosis utilization of four indica-japonica intermediate type restorer lines with wide compatibility. *Chin J Rice Sci* 26(6):656-662

- Lin Z, Qin P, Zhang X, Fu C, Deng XW (2020) Divergent selection and genetic introgression shape the genome landscape of heterosis in hybrid rice. *Proc Natl Acad Sci U S A* 117(9): 4623-4631
- Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129
- Liu SF, Chen Q, Hong GC, Jun HU, Qin XJ (2018) Research progress of cytoplasmic male sterility and fertility restoration in rice (*Oryza sativa*). *Plant Physiology Journal* 54(1):1-9
- Peng ST, Wang HG, Wei XH, LV JZ, Zhang XL, Yuan XP, Yang WD (2008) SSR diversity and genetic structure of main male sterile lines of three-Line hybrid rice in China. *Chin J Rice Sci* 22(4): 365-369
- Pu HC, Zhou ZL, Xu DY (2015) History and problems of three-line japonica hybrid rice development. *Jiangsu Agric Sci* 43(1): 74-77
- Qiu FL, Zhuang JY, Hua ZT, Wang YR, Cheng SH (2005) Inspect of genetic differentiation of main parents of japonica hybrid rice in the northern China by simple sequence repeats(SSR). *Chin J Rice Sci* 19(02):101-104
- Shen GW, Xu SK, Li HL, Shen GH, Wang L (2015) Progress and technical strategy of BT type japonica hybrid rice breeding. *China Seed Industry* (1): 9-11
- Singh VK, Upadhyay P, Sinha P, Mall AK, Ellur RK, Singh A, Jaiswal SK, Biradar S, Ramakrishna S, Sundaram RM, Ahmed I, Viraktamath BC, Kole C, Singh S (2011) Prediction of hybrid performance based on the genetic distance of parental lines in two-line rice (*Oryza sativa* L.) hybrids. *Journal of Crop Science and Biotechnology* 14: 1-10
- Sui GM (2018) Research progress and development strategy of japonica hybrid rice. *Liaoning Agric Sci* (1): 51-55
- Sun J, Liu D, Wang JY, Ma DR, Tang T, Gao H, Xu ZJ, Chen WF (2012) The contribution of inter subspecific hybridization to the breeding of super-high-yielding japonica rice in northeast China. *Theor Appl Genet* 125: 1149-1157
- Tang X, Tan XL (2009) SSR fingerprint and genetic diversity of major parents of Dian-type japonica hybrid rice. *Hybrid Rice* 24(6): 54-58
- Wang CH, Zhang YZ, Hua ZT, Xu ZJ, Zheng WJ, Zhao JM, Liu X, Wang H, Zhou YH (2009) Genetic basis of parents of japonica hybrid rice in the Northern China. *Chin J Rice Sci* 23 (5): 489-494
- Wang K, Qiu FL, Paz MAD, Zhuang JY, Xie FM (2014) Genetic diversity and structure of improved indica rice germplasm. *Plant Genetic Resources* 12(2): 248-254

Wang SJ, Wan JM, Lu ZM (2006) Parental cluster analysis in indica hybrid rice (*Oryza sativa* L.) by SSR analysis. *Acta Agron Sin* 10: 1437-1443

Wang WS, Mauleon R, Hu ZQ, Chebotarov D, Tai SS, Wu ZC, Li M, Zheng TQ, Fuentes RR, Zhang F, Mansueto L, Copetti D, Sanciangco M, Palis KC, Xu JL, Sun C, Fu BY, Zhang HL, GaoYM, Zhao XQ, Shen F, Cui X, Yu H, Li ZC, Chen ML, Detras J, Zhou YL, Zhang XY, Zhao Y, Kudrna D, Wang CC, Li R, Jia B, Lu JY, He XC, Dong ZT, Xu JB, Li YH, Wang M, Shi JX, Li J, Zhang DB, Lee SH, Hu WS, Poliakov A, Dubchak I, Ulat VJ, Borja FN, Mendoza JR, Ali J, Li J, Gao Q, Niu YC, Yue Z, Naredo ME, Talag J, WangXQ, Li JJ, Fang XD, YinY, Glaszmann JC, Zhang JW, Li JY, Hamilton RS, Wing R, Ruan J, Zhang GY, Wei CC, Alexandrov N, McNally K, Li ZK, Leung H (2018) Genomic variation in 3,010 diverse accessions of asian cultivated rice. *Nature* 557(7703): 43-49

Wang YR, Qiu FL, Hua ZT, Dai GJ (2010) Relationship between indica-japonica index of parents and hybrid yield and rice quality traits in japonica hybrid rice in Northern China. *Chin J Rice Sci* 24(2): 145-150

Xie FM, Guo LB, Ren GJ, Hu PS, Wang F, Xu JL, Li XQ, Qiu FL, MA Dela Paz (2012) Genetic diversity and structure of indica rice varieties from two heterotic pools of southern China and IRRI. *Plant Genetic Resources* 10(3): 186-193

Xie FM, Peng SB (2016) History and prospects of hybrid rice development outside of China. *Chinese Sci Bulletin* 61: 3858-3868

Xie LJ, Pan XJ, Duan M, Huang SJ (2015) Research progress and breeding strategies of sterile lines in japonica rice. *Zhejiang Agric Sci* 56(5): 625-628, 630

Xu ML, Jin ZX, Li XG, Zhang ZC, Liu HY, Zhang FZ, Zhao SY, Zhang HB (2009) Comparison of genetic distance among seven cultivars of japonica rice based on SSR and SRAP and its relationship with heterosis of yield traits. *Molecular Plant Breeding* 7(6): 1084-1092

Xu Q, Yuan XP, Wang S, Feng Y, Yu HY, Wang YP, Yang YL, Wei XH, Li XM (2016) The genetic diversity and structure of indica rice in China as detected by single nucleotide polymorphism analysis. *BMC genomics* 17: 53. DOI 10.1186/s12863-016-0361-x

Xu SL, Yan TZ, Zhong RC, Qin P, Qang K, Yang YZ (2017) Analysis on genetic differences between parents of two-line medium indica hybrid rice by SSR markers. *Agricultural Science & Technology* 18(10): 1800-1804

Yang ZY, Li ZB, Dong L, Zhu W, Hua ZT (2016) Development and prospect of hybrid japonica rice in China. *China Sci Bull* 61: 3770-3777

Zha RM, Yang ZL, Zhao FM, Zhong BQ, Peng T, Xie R, He GH (2007) Genetic difference changes of main restorer and sterile lines in different stages of hybrid rice (*Oryza sativa* spp. indica L.). *Acta Agronomica Sinica* 33 (4): 572-577

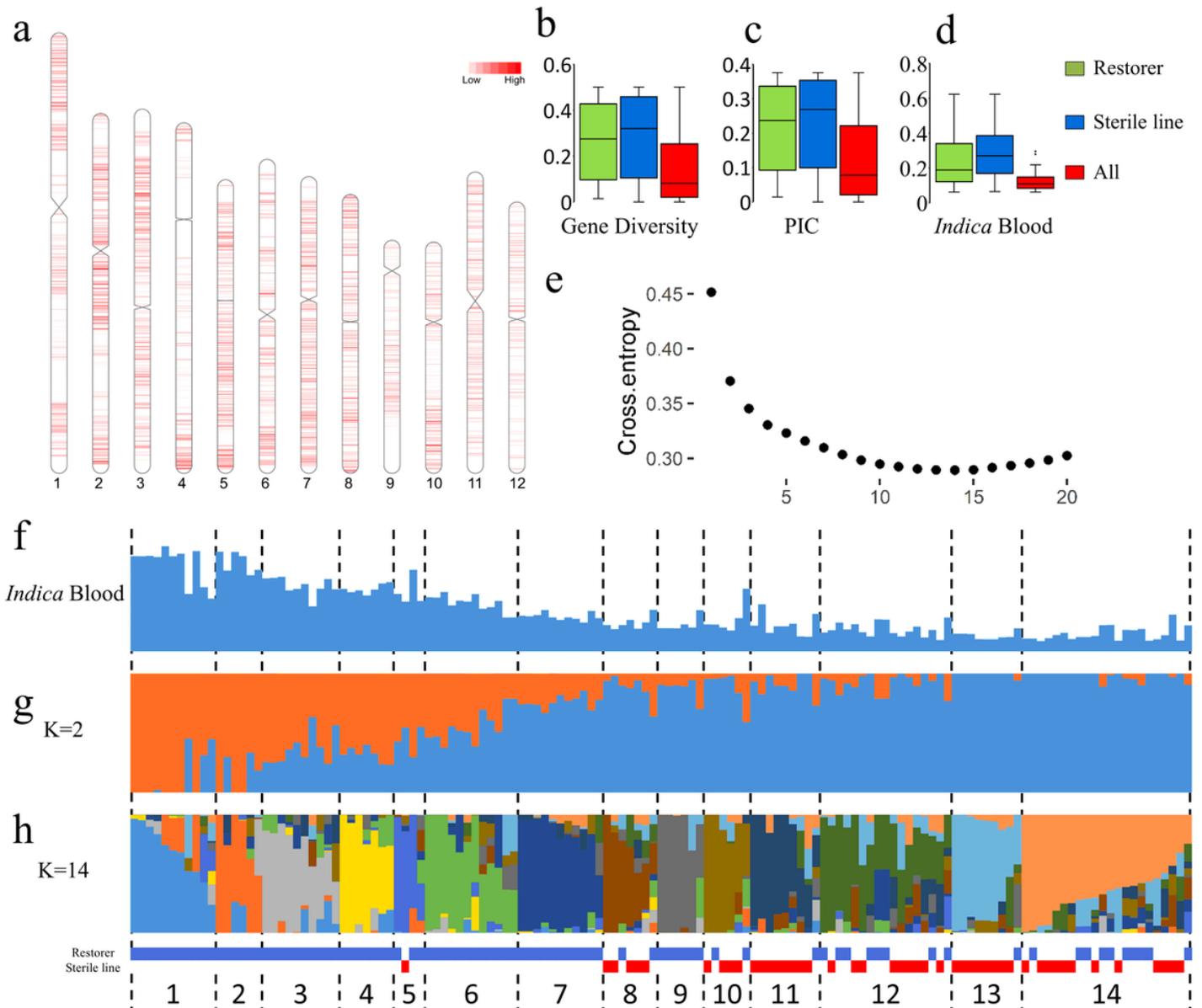
Zhang T, Yang J, Jiang KF, Cao YingJ, Yang L, Yang QH, Wan XQ, You SM, Luo J, Gao L, Li ZX, Zheng JK (2014) Genetic diversity of the main chinese three-line hybrid rice parents based on functional genetic markers related to yield. *Sci Agric Sin* 47(1):11-23

Zhao QY, Zhang YD, Zhu Z, Zhao L, Chen T, Zhou LH, Wang CL (2010) Genetic diversity of parental lines in japonica hybrid rice based on cluster analysis of SSR markers and phenotypic characters. *Hybrid Rice* 25(04): 68-74

Zhao QY, Zhu Z, Zhang YD, Zhao L, Chen T, Zhang QF, Wang CL (2009) Analysis on correlation between heterosis and genetic distance based on simple sequence repeat markers in japonica rice. *Chin J Rice Sci* 23(2): 141-147

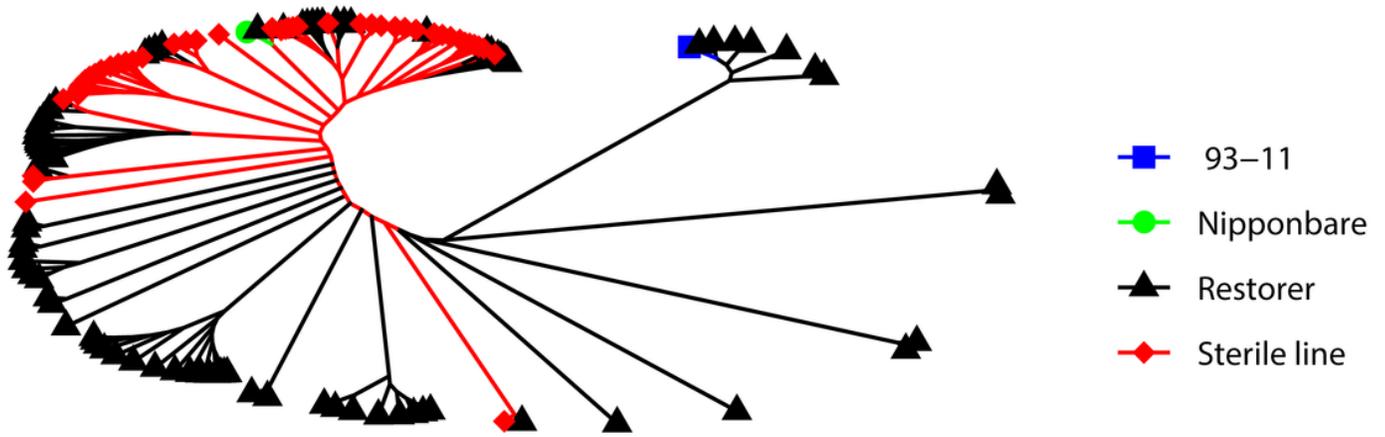
Zheng WJ, Ma ZB, Zhao MZ, Xiao MG, Zhao JM, Wang CH, Gao H, Bai YJ, Wang H, Sui GM (2020) Research and development strategies for hybrid japonica rice. *Rice* 13: 36.  
<https://doi.org/10.1186/s12284-020-00398-0>

## Figures



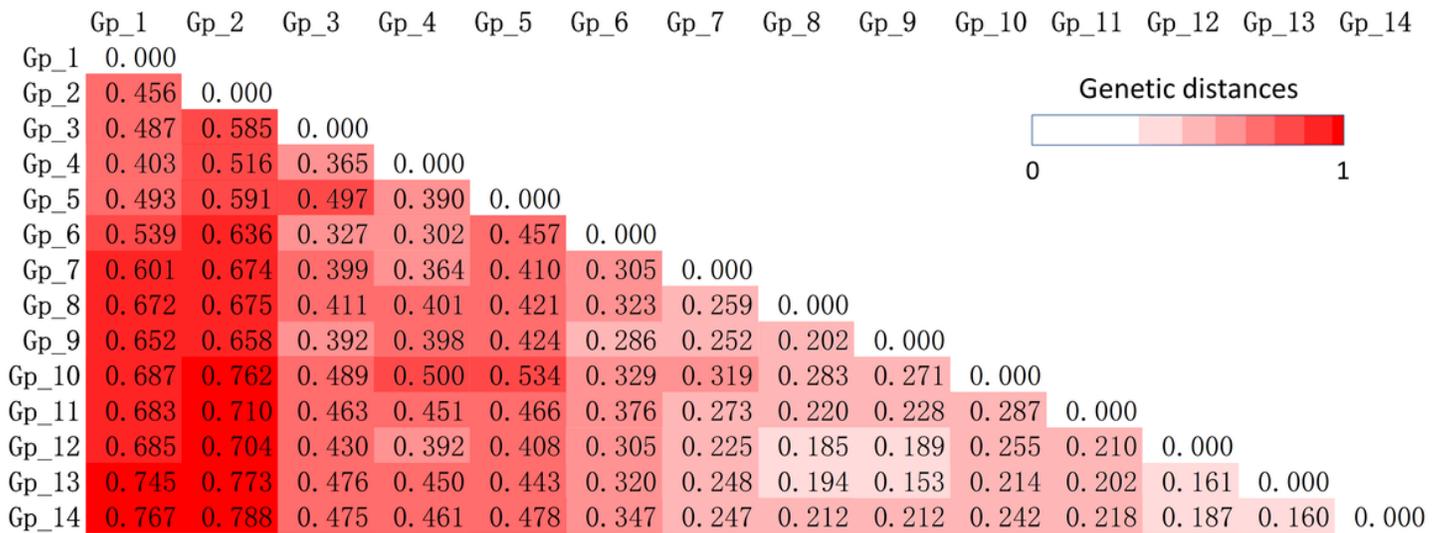
**Figure 1**

Genetic background analyses of the parental lines of hybrid japonica rice in northern China. a Distribution of SNP markers on 12 chromosomes. The red line represents the position of the SNP on the chromosome, and the staining indicates the polymorphism information content (PIC) of the SNP in the parental lines of hybrid japonica rice. The deeper the color, the higher the PIC. b Genetic diversity of the parental lines of hybrid japonica rice. c The PIC of the parental lines of hybrid japonica rice. d The indica blood of the parental lines of the hybrid japonica rice. Green indicates all parental lines, blue indicates restorer lines, red indicates sterile lines in b, c, and d. e The cross-entropy statistics for each given K. f The indica blood of each individual parental line of hybrid japonica rice. g, h Clustering analysis of the 137 parental lines via the model-based method when K = 2 and K = 14. The 14 groups are distinguished by dotted lines; the bars indicate the restorer lines (blue) and the sterile lines (red).



**Figure 2**

Neighbor-joining phylogenetic tree based on the Chord genetic distance



**Figure 3**

Genetic distances between groups based on model clustering

## Supplementary Files

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

- [SupplementaryTable13.xlsx](#)