Dissecting the Genetic Basis of Superior Traits in Thermosensitive Genic Male Sterile Line 1892S Through Genome-Wide Analysis

Wei Zhang, Liang Zhou, Dewen Zhang*

Anhui Province Key Laboratory of Rice Genetics and Breeding (Rice Research Institute, Anhui Academy of Agricultural Sciences), Hefei 230031, China.

Corresponding author: Dewen Zhang
E-mail: zhangdewen0551@163.com

ABSTRACT. Hybrid rice has revolutionized food security by leveraging heterosis, a phenomenon where offspring outperform their parents. Sterile lines, crucial for controlled cross-pollination in hybrid breeding, have played a central role in this success. This study delves into the superior alleles of 1892S, a two-line sterile rice line widely used as a female parent in central China. By integrating extensive hybridization data, high-throughput genome sequencing, and bioinformatic analysis (RiceNavi), we elucidate the genetic underpinnings of 1892S's exceptional adaptability. The results reveal its remarkable compatibility as a female parent in over 114 hybrid varieties, potentially due to the influence of *Japonica* characteristics contributing to strong hybrid vigor. Furthermore, some favorable alleles were identified that were associated with lodging resistance, high yield potential, and improved nitrogen use efficiency etc. The comprehensive characterization of 1892S provides valuable insights for future hybrid rice breeding programs, ultimately facilitating the development of superior rice varieties.

Key words: 1892S; Rice; Superior alleles; Thermosensitive genic male sterile line.

INTRODUCTION

Heterosis, or hybrid vigor, plays a critical role in boosting crop yield and quality(Paril et al., 2024; Wu et al., 2021). The discovery and successful application of thermosensitive genic male sterile lines (TGMS Lines) in rice revolutionized hybrid rice breeding for self-pollinating crops (Wang et al., 2005; Yuan, 2014; Ashraf et al., 2020). TGMS lines like 1892S offer several advantages over traditional sterile lines used in three-line breeding systems (Yang Lian-song, 2012; Lian-song et al., 2016). These advantages include wider cross-compatibility, simpler sterile material reproduction, higher hybrid seed production yield, and easier sterility gene transfer (Lian-song et al., 2016; Yang Lian-song, 2012).

The risk and slow rate of popularization of two-line hybrid rice are mainly due to the unstable sterility and poor combining ability of some *Indica* sterile lines, and the limited production and application of the F_1 hybrid (Xu et al., 2023; Cao, 2014). However, the two-line sterile line 1892S stands out for its exceptional characteristics. It exhibits high combining ability, good propagation traits, strong resistance to diseases, and excellent grain quality. Additionally, 1892S-based hybrids demonstrate strong heterosis, tolerance to fertilizers, resistance to lodging, and broad adaptability (Khan et al., 2024; Lian-song et al., 2016).

Despite its success, a comprehensive analysis of 1892S's genetic makeup is lacking. This knowledge gap hinders the wider application of 1892S and limits its potential to optimize breeding programs in the era of molecular design breeding, which relies heavily on in-depth varietal analysis for the efficient development of new lines (Collard et al., 2008).

This study aims to address this critical gap by employing a multi-pronged approach. We will leverage RiceNavi, a powerful rice breeding tool that provides detailed quantitative trait gene information, to analyze the excellent alleles present within the entire genome of 1892S (Wei et al., 2021; Qianlong et al., 2023). Additionally, we will assemble the genome of 1892S to support the existence of these beneficial genes. Furthermore, we will analyze the *Indica-Japonica* properties of 1892S using indel markers to gain insights into its heterosis characteristics(Shen et al., 2004).

The comprehensive analysis empowers breeders with a detailed genetic analysis of 1892S. it reveals favorable alleles associated with crucial agronomic traits like lodging resistance, high yield potential, and improved nitrogen use efficiency. Armed with this knowledge, breeders could make strategic decisions about crossing partners, breeding process optimization, and prioritizing desired traits in the development of hybrid rice varieties.

MATERIAL AND METHODS

Rice materials

The Rice Research Institute of the Anhui Academy of Agricultural Sciences bred the TGMS line 1892S. The passing of the technical appraisal in 2004 and the application for new plant variety rights in the same year, followed by authorization in 2007, the variety right number is CNA20040612.4 (https://www.ricedata.cn).

DNA Extraction and Genome Sequencing

The genomic DNA extraction from the young leaves of 1892S plants using the CTAB method followed by sequencing library construction. Utilizing the HiSeq2500 next-generation sequencing platform with a sequencing depth of 100 ensures thorough coverage of the genome. The generation of 157,494,513 RawReads reflects the extensive data obtained from the sequencing process, enabling in-depth genome-wide analysis and identification of key genetic features and variations associated with 1892S.

Functional locus analysis using RiceNavi

Use Trimmomatic to trim and filter the raw reads based on quality scores and adapter sequences. Then run FastQC on the trimmed read files to assess their quality and identify any remaining issues or biases. These software bowtie2, samtools, sambamba, GATK3, GATK4, Manta, bam2fastq are called by RiceNavi (Wei et al., 2021). The rice genome (MSU v7) was downloaded from http://rice.uga.edu/. The whole genome clean data was Aligned to the reference genome. Utilize RiceNavi-QTNpick mode to calculate genotyping causal variant sites based on the alignment results. The mode provides information on QTN sites and their associated characteristics based on rice accessions in QTNlib with different alleles of each causative site. The detailed information on the QTN site comes from the data from Wei's article (Wei et al., 2021).

Genome assembly and local alignment of gene sequences

The genomic clean data were used to assemble the genome by SOAPdenovo2 (Luo et al., 2012). Utilize GapCloser to fill the gaps in the assembled genome using the paired relationships of short reads. Run BUSCO on the assembled genome to assess its integrity and completeness by comparing it to conserved single-copy orthologs (Manni et al., 2021). Perform collinearity analysis between the assembled genome and the reference genome using Mummer software which was used to identify large-scale structural variations and similarities between the assembled genome and the reference genome (Marcais et al., 2018). Align the assembled genome to the reference genome using Minimap2 (Li, 2021). The gene sequence of the reference genome is downloaded from http://rice.uga.edu/. Muscle was used to align the gene sequences locally and assembled contig sequences to identify similarities, differences, and structural variations at the gene level (Edgar, 2022).

Indica-Japonica attribute identification

Some Bacterial Artificial Chromosome (BAC) sequences that come from Shen's article (Shen et al., 2004) were downloaded from the NCBI database. These sequences are likely to contain the InDel markers of interest. The downloaded BAC sequences were aligned to the genome of rice variety 93-11 using the Minimap2. Based on the mapping results, the corresponding sequences were extracted. These sequences should contain the InDel markers. Forward and Reverse Primers specific to the InDel markers were used to extract the sequences from both *Indica* and *Japonica* rice varieties separately. The extracted InDel sequences from *Indica* and *Japonica* were mapped back to the 93-11 genome to provide detailed information about their positions and contexts within the genome. Contigs of 1892S were also mapped to the 93-11 genome using Minimap2, which was used to find corresponding regions in 1892S that align with the InDel markers identified in 93-11. By combining the mapping information from the previous steps, the sequences of the InDel markers

from 1892S were obtained. These sequences were used for comparison with those from *Indica and Japonica*. The sequences of InDel markers from *Indica, Japonica*, and 1892S were aligned together using MUSCLE. it is determined whether each sequence of InDel marker from 1892S is more like the *Indica or Japonica* variety.

Data analysis of hybrid rice combinations with 1892S

To analyze the hybrid rice combinations involving 1892S as the female parent, we visited the National Rice Data Center website (https://www.ricedata.cn/) to access the data on hybrid rice combinations for searching for hybrid rice varieties which 1892S was listed as the female parent. The relevant information was extracted, such as the names, approval numbers, and characteristics of the hybrid rice varieties. We also Gathered information on TGMS lines that 1892S was the female parent.

RESULTS

Phenotypic Characterization of 1892S

1892S exhibited several notable phenotypic characteristics upon cultivation. Such as 1892S displayed visibly stronger stems compared to other varieties. Its hybrid variety has potentially higher yields with increased grain number per spike. Its high stigma exsertion rate which could improves pollination efficiency by ensuring better exposure of the stigma to pollen grains (Y.N.J. et al., 2022). Additionally, as an *Indica-type* variety, 1892S demonstrates compatibility with other indica varieties, facilitating hybridization to leverage complementary traits and achieve superior agronomic advantages. This highlights 1892S's potential as a valuable genetic resource for breeding programs focused on developing high-yielding and resilient rice varieties suited to specific environmental conditions.

Functional Locus Analysis of 1892S

High-throughput sequencing of 1892S was performed, achieving comprehensive coverage of the genome (110x) and an impressive Q30 value exceeding 92.69%. Clean data was used for further analysis with the rice reference genome Nipponbare (MSU 7.0) as a reference. RiceNavi software was employed to identify quantitative trait nucleotides (QTNs) associated with various phenotypic traits (Wei et al., 2021). This analysis identified a total of 319 QTN loci across the 1892S genome. The predicted effects of these QTNs on the phenotype of 1892S were then interpreted (Table S1). This comprehensive analysis revealed 32 potentially superior genes within 1892S (Table 1). This information provides valuable insights into the genetic basis of 1892S's traits and paves the way for further investigation and breeding efforts to improve rice varieties.

Genome Assembly and Assessment

The assembled genome of 1892S has an estimated size of 324.7 Mb with a GC content of 43.00%. The longest contig fragment is 147,639 bp, while the Contig N50 and N90 values are 16,196 bp and 3,277 bp, respectively (Table 2). The completeness of the assembly was assessed using BUSCO, revealing that 93.8% of the core genes were successfully identified, with only 2%

Table 1. Summary of superior genes associated with crucial agronomic traits in 1892S.

| ID	RAP_Locus	MSU_Locus	Gene	Chr	Start position	End positon	Trait	Trait group	1892S_tait	Local alignment of gene sequences	Gene
1	Os01g0197100	LOC_Os01g10040	D2/CYP90D2/SMG11	Chr1	5236623	5244023	tiller angle	plant architecture	The tiller angle becomes smaller	√	D2/CYP90D2/SMG11
2	Os01g0197700	LOC_Os01g10110	Gn1a/OsCKX2	Chr1	5270449	5275585	grain production	yield components	The number of grains per panicle increased	√	Gn1a/OsCKX2
3	Os01g0201700	LOC_Os01g10504	Rf3/OsMADS3	Chr1	5559532	5568924	fertility restoration	yield components	Restore wild abortion	√	Rf3/OsMADS3
4	Os01g0718300	LOC_Os01g52050	D61/OsBRI1	Chr1	29927587	29931452	flag leaf angle	plant architecture	The blade angle becomes smaller	√	D61/OsBRI1
5	Os01g0831000	LOC_Os01g61480	LAX1	Chr1	35558148	35559225	grain number and drought tolerance	yield components	Increased drought resistance	√	LAX1
6	Os01g0869800	LOC_Os01g64960	OsPsbS1	Chr1	37697024	37699582	nonphotochemical quenching	others	Increased non-photochemical quenching ability	√	OsPsbS1
7	Os02g0131800	LOC_Os02g03900	NRAT1	Chr2	1658576	1662644	aluminum tolerance	abiotic stress	Resistant to aluminium stress	√	NRAT1
8	Os02g0196600	LOC_Os02g10290	OsHMA4	Chr2	5404703	5410606	copper accumulation	abiotic stress	High copper content	√	OsHMA4
9	Os02g0770800	LOC_Os02g53130	OsNR2	Chr2	32513739	32517155	nutrition	others	Nitrogen use efficiency was improved	√	OsNR2
10	Os03g0171700	LOC_Os03g07540	OsbHLH153/ILI3	Chr3	3845564	3846658	flag leaf angle	plant architecture	The blade horn is small	√	OsbHLH153/ILI3
11	Os03g0726700	LOC_Os03g51660	TAC3	Chr3	29583117	29584833	tiller angle	plant architecture	The tillering angle is small	√	TAC3
12	Os03g0856700	LOC_Os03g63970	GNP1/OsGA20ox1	Chr3	36150664	36152355	grain number and yield	yield components	The number of grains per panicle increased, and the plant height increased	√	GNP1/OsGA20ox1
13	Os04g0477300	LOC_Os04g40140	BET1	Chr4	23885498	23889561	boron-toxicity tolerance	abiotic stress	Increased boron tolerance	√	BET1
14	Os04g0518800	LOC_Os04g43840	An-2/OsLOGL6/LABA1	Chr4	25959399	25963504	awn length and grain production	seed morphology	Mangs become shorter, The number of grains per panicle increased	√	An-2/OsLOGL6/LABA1
15	Os04g0615000	LOC_Os04g52479	NAL1/SPIKE//LSCHL4/GPS	Chr4	31205267	31214632	grain productivity	yield components	Increase leaf width and increase yield	√	NAL1/SPIKE//LSCHL4/GPS
16	Os04g0653000	LOC_Os04g55920	OsJAZ1	Chr4	33306468	33310169	drought tolerance	abiotic stress	Increase root length and root weight	√	OsJAZ1
17	Os05g0207500	LOC_Os05g11730	OsGSK2	Chr5	6657481	6661493	mesocotyl length	plant architecture	The mesocotyls become elongated	√	OsGSK2
18	Os06g0107800	LOC_Os06g01860	BPH29	Chr6	484346	485308	brown planthopper resistance	biotic stress	Rice planthopper resistance is enhanced	√	BPH29
19	Os06g0213100	LOC_Os06g11010	S5	Chr6	5759685	5761518	compatibility	yield components	Wide affinity	√	S5
20	Os06g0665400	LOC_Os06g45460	APO1/SCM2	Chr6	27480082	27481453	lodging resistance, number of grains	plant architecture	Anti-lodging	√	APO1/SCM2
21	Os06g0701700	LOC_Os06g48810	OsHKT2;1	Chr6	29538938	29541203	potassium use efficiency	abiotic stress	Potassium use efficiency is improved	√	OsHKT2;1
22	Os07g0211500	LOC_Os07g11020	Rc/qSD7-1/qPC7	Chr7	6062889	6069317	red pericarp, seed germination	seed morphology	The seed coat is white and weakly dormant	√	Rc/qSD7-1/qPC7
23	Os07g0232900	LOC_Os07g12900	OsHMA3	Chr7	7405745	7409553	cadmium accumulation	abiotic stress	Does not enrich cadmium	√	OsHMA3
24	Os07g0569700	LOC_Os07g38240	OsSAP16	Chr7	22930745	22933283	low-temperature germination	abiotic stress	Improved low-temperature germination ability	√	OsSAP16
25	Os08g0101500	LOC_Os08g01120	OsMOT1;1/qGMo8	Chr8	86388	87854	Mo accumulation	abiotic stress	Molybdenum accumulation increases	√	OsMOT1;1/qGMo8
27	Os08g0432300	LOC_Os08g33530	TIG1	Chr8	20931199	20932112	plant architecture	plant architecture	The tiller angle becomes smaller	√	TIG1
28	Os08g0538300	LOC_Os08g42580	OsCERK1	Chr8	26909127	26913494	blast disease resistance	biotic stress	The phosphorus absorption efficiency was improved and the resistance to rice blast was enhanced	√	OsCERK1
29	Os10g0400200	LOC_Os10g26060	OsGluA2	Chr10	13497363	13499401	grain protein content	taste quality	Reduces protein content	√	OsGluA2
30	Os10g0403800	LOC_Os10g26410	OsbHLH174	Chr10	13721970	13722965	flag leaf angle	plant architecture	The blade horn is small	√	OsbHLH174
31	Os10g0554200	LOC_Os10g40600	NRT1.1B	Chr10	21757771	21762202	nitrate-use divergence	others	Nitrogen use efficiency was improved	√	NRT1.1B
32	Os12g0630100	LOC_Os12g43440	TOND1	Chr12	26956203	26956894	tolerance to nutrition	others	Resistant to low nitrogen stress	√	TOND1

Table 2. Preliminary genome assembly of 1892S.

Type	Contigs
Total length (bp)	324669546
GC (%)	43.00
N50 (bp)	16196
N90 (bp)	3277
Longest (bp)	147639
contigs	45720
contigs (>= 10000 bp)	10273
contigs (>= 25000 bp)	2671
contigs (>= 50000 bp)	375

Table 3. Quantitatively assessed of 1892S genome assembly integrity.

Type	BUSCOs num	Percentage(%)
Complete BUSCOs (C)	1514	93.8
Complete and single-copy BUSCOs (S)	1489	92.3
Complete and duplicated BUSCOs (D)	25	1.5
Fragmented BUSCOs (F)	68	4.2
Missing BUSCOs (M)	32	2
Total BUSCO groups searched	1,614	100
Database	embryophyta_odb10	

missing (Table 3). Synteny analysis confirmed a high degree of similarity between the 1892S genome and the reference genome MSU 7.0. Overall, the 1892S genome assembly demonstrates high reliability, completeness, and synteny with the reference, providing a solid foundation for further studies (Figure 1).
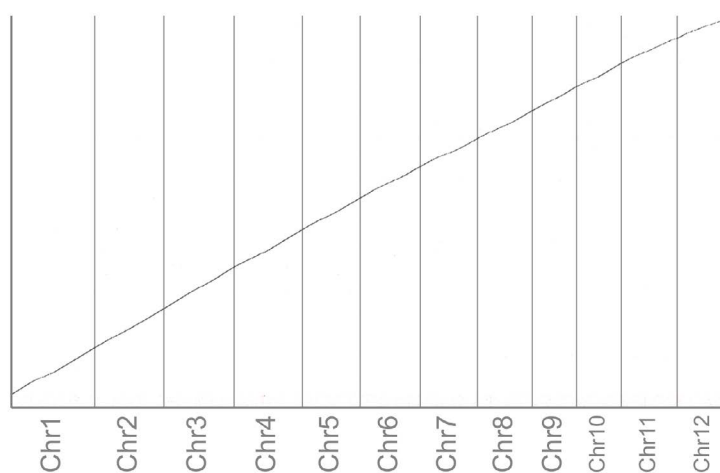
Figure 1. Sequence collinearity analysis among MSU 7.0 and 1892S.

Local Alignment of Gene Sequences

To determine the location of contigs within the reference genome and gain context about the assembled sequences, the draft genome of 1892S was mapped to the reference genome using Minimap2. Extracted gene sequences and their corresponding contig sequences from the draft assembly were then aligned and compared using MUSCLE software. Based on the alignment results, genes within the 1892S draft genome exhibiting high similarity with known superior alleles were identified as excellent genes (Table 1). This successful comparison confirms the presence of these excellent genes within the draft genome of 1892S.

Indica-Japonica Classification Using InDel Markers

Indel markers were employed to distinguish between the *Indica and Japonica* subspecies of rice and understand the genetic relationship between 1892S and these varieties. Briefly, BAC clone sequences were aligned with the genome of rice variety 93-11 to identify matching regions. Primer sequences flanking indel markers were designed to amplify and sequence these regions. A total of 43 indel markers specific to the BAC clone were identified and mapped to the 93-11 genome. The 1892S contig sequence was then aligned to the 93-11 genome, and corresponding sequences were extracted based on the marker positions. Analysis of these alignments revealed 30 indel markers present in all three varieties (Table 4). Six of these markers indicated shared sequences between 1892S and Nipponbare *(Japonica)*, while the remaining 24 matched 93-11 *(Indica)*. These findings suggest that 1892S possesses a mixed genetic background with characteristics of both *Indica* and *Japonica* subspecies.

Favorable Traits and Breeding Potential of 1892S

This section combines findings from several subsections to highlight the key advantages and breeding potential of the 1892S rice variety.

Wide Affinity: The wide affinity *S5* locus in 1892S, as indicated in Table 1, aligns with previous studies highlighting its advantageous traits. Specifically, the *S5-n* gene plays a crucial role in controlling the wide affinity of rice intersubspecific hybrids. Loss of sequence in the *S5-n* gene results in loss of its function, yet hybrid-wide affinity remains unaffected with both *Indica and Japonica* rice varieties (Chen et al., 2008). To confirm the presence of the *S5-n* gene in 1892S, the assembled contig sequence of its genome was locally compared with the corresponding site in the reference genome. The comparison revealed a deletion in the N-terminal region, providing evidence that 1892S indeed possesses the *S5-n* gene.

The extensive use of 1892S, with its wide affinity gene *S5-n*, is evident in the approval of combination varieties where it serves as the female parent. A total of 114 hybrid rice varieties with 1892S as the female parent have received approval. Notably, since its technical appraisal in Anhui Province in 2004, the first variety, "Wandao 153," incorporating 1892S as the female parent, obtained approval the following year, marking a significant milestone in its contribution to national rice variety selection (Liansong, 2006). Over the years, from 2005 to 2023, there has been a consistent increase in the approval of hybrid rice varieties with 1892S as the female parent, culminating in a record-setting 24 approvals in 2021. These varieties are widespread and distributed across 9 provinces, with 42 provincial rice varieties receiving national approval. Among these, the number of varieties approved in Anhui ranks second only to those with national approval. This analysis underscores the broad affinity and significant contribution of 1892S to rice variety development.

Table 4. The comprehensive information on the InDel markers designed for diverse combinations of japonica and indica.

Chromosome	Marker Name	BAC Accession	Forward Primer (5# 3#)	Reverse Primer (5# 3#)	Nipponbare/Indica
1	R1M7	AP002482	ATTCCTGGTTCTACAT-TACTTA	CGCCTCACTAGAATATCGGA	Indica
1	R1M47	AP003442	AATAGAATTACTGAT-GAAACCTTA	GCCCGTTACCGCTTATGT	Indica
2	R2M24	AP005414	GGGCAACAACGGCTCTG	AGGGAATAAGGCGATACGG	Indica
2	R2M26	AP005696	GCAGCAAAGTGCGGAGTA	CAGGTGAATTGCCAATTT	Indica
2	R2M50	AP004888	CCTGAAGGAAATGATAG-CAATAG	GTTTTGTATGCTCTTCACTT-GTC	Indica
3	R3M23	AC099323	TGCTTACAAGGGTCCAAT	GGAGGTGCCTACCAAGAG	Indica
3	R3M30	AC091234	AGGCTAAGTGAAGAAATA-ATAAG	CTCCGTATTCATTACTGGTTG	Indica
3	R3M53	AC091123	ACACTGGCTACGGCAAAG	TTTGTTCGGGAATAATGATGC	Indica
4	R4M13	AL606597	TACACGGTAGACATCCAACA	ATGATTTAACCGTAGATTGG	Nipponbare
4	R4M17	AL731585	AGTGCTCGGTTTTGTTTTC	GTCAGATATAATTGATG-GATGTA	Indica
4	R4M30	AL662979	GCTTCTCCTGGTTGTATGC	AAAATAGGGAGGCA-GATAGAC	Nipponbare
4	R4M43	AL662938	CTTGAACCTGAGTGAGTGG	CGATGAAAATGATGTCTA	Indica
4	R4M50	AL606639	TTTTGTGAAACTTGACCCTC	GCGTCCATGTCTTTATTGTG	Indica
5	R5M13	AC132493	GAGAAAGAGTGGAAGGAG	AGTATCGTCAGGAGGGTC	Indica
5	R5M20	AC137622	CTCGCTGTTTACTGACTGG	TTTGATGTACTGCCTGCTCT	Nipponbare
5	R5M30	AC134930	CTCAATTTCACCCATCCC	CGCTCCGTCTCCAACCTC	Indica
5	R5M43	AC121365	AGCGTGACTTGAGTTCCA	ATGACTTTCCCACCGTAT	Nipponbare
6	R6M14	AP004725	AAATGTCCATGTGTTTGCTTC	CATGTGTGGAATGTGGTTG	Indica
6	R6M30	AP005929	CACAAGCCGTAGCAGAGC	TCACGAAAAAGACCCCAAG	Indica
6	R6M44	AP005386	TTAGGAATAAAGGCTGGATA	TTACCGTTAATAGGTGGAA	Nipponbare
8	R8M33	AP003881	CGAAAGAGGAGAGGGGTAGT	CGAAAACGAGAAACAAATA	Indica
9	R9M20	AP005879	ACTGCTTTGATGGCTTGTG	CTCCCCAAACTGAATCC	Indica
9	R9M30	AP005397	CTCACCTACCTA-AAACCCAAC	CCACCCAAATCTGATACTG	Indica
9	R9M42	AC108757	CTATAAGACCAAAAC-GAAAACT	GAAAACCATTGTGTCACT-GTA	Indica
10	R10M17	AC090486	TGAACAATAAACCA-CAGAAGCA	CCCTTTATTCCCTCCTTTG	Nipponbare
10	R10M40	AC091122	GTCCCTAGGCCATCTCTTG	GCGAATAGGGGTGGACAG	Indica
11	R11M40	AC125780	AAGAAAAATATCTATTGAG-GAGTG	GGAGGACCATAAATGACGG	Indica
12	R12M10	AL954158	ATCATTTCAGCCTGTGCC	AGCTTAATAGGGGGGACG	Indica
12	R12M27	AL713927	ATTTCATTGCCATCAGTT	GTAATCTTCTATCCGTTCA	Indica
12	R12M33	AL731888	TTGATGATAGTATTTGCTGATG	AGATAGTGTCGGCGGTGG	Indica

Lodging Resistance: 1892S possesses the APO1 site linked to lodging resistance (Table 1). This finding aligns with observations from approved hybrid combinations using 1892S, which exhibit anti-inversion characteristics. Additionally, the presence of semi-dwarf trait genes (SBI/Sd1/Ghd7 and SD1/OsSPL14) in 1892S contributes to its compact growth habit (Liu et al., 2018; Asano et al., 2007; Jiao et al., 2010), a factor not only enhancing lodging resistance but also allowing for higher planting density and potentially greater yield. Real-world examples like Wandao 153 further solidify 1892S's contribution to lodging resistance. This variety demonstrates exceptional resilience, characterized by short stature, robust stems, and superior root architecture (Miaomiao et al., 2013).

High Yield Potential: The genetic makeup of 1892S includes genes associated with a high number of grains per panicle, including *Gn1a/OsCKX2* (Ashikari et al., 2005), *GNP1/OsGA20ox1* (Wu et al., 2016), and *An-2/OsLOGL6/LABA1* (Table 1) (Gu et al., 2015; Hua et al., 2015). These genes play a crucial role in regulating grain number and panicle development, ultimately impacting yield per unit area. Varieties derived from 1892S exhibit a significantly higher number of grains per spike and demonstrably higher yields compared to controls.

Improved Nitrogen Use Efficiency: The presence of *OsNR2* and *NRT1.1B* sites in 1892S suggests enhanced potential for nitrogen use efficiency (NUE) (Table 1) (Gao et al., 2019; Hu et al., 2015). The *Indica OsNR2* variant offers superior traits compared to its *Japonica* counterpart, leading to increased chlorate sensitivity, improved nitrate uptake, and ultimately, greater grain yield (Gao et al., 2019). *NRT1.1B* also plays a significant role in nitrate signaling and NUE. By harboring these genes, 1892S holds promise for optimizing nitrogen utilization in rice crops.

Sterile Line Development: 1892S serves as a valuable resource for breeding two-line sterile rice varieties. As a two-line sterile line itself, 1892S has been used to develop many additional sterile lines. Notably, 1892S exhibits a high stigma exposure rate, a characteristic crucial for efficient hybridization (Y.N.J. et al., 2022). Gene sequence analysis confirmed the presence of stigma exsertion genes (gs3/gw8/gs9) within 1892S, further supporting its role in sterile line development (Zhu et al., 2023).

DISCUSSION

The success of molecular design breeding relies heavily on the accurate understanding of the genetic makeup of the crop, the traits of interest, and the complex interactions between genes and the environment (Ahmar et al., 2020; Begna, 2022). It can significantly enhance the efficiency and effectiveness of crop improvement programs (Pasala et al., 2024). This approach has significant advantages in crop breeding, where molecular design breeding allows breeders to work at the genetic level to select and combine genes with specific traits very precisely to create new varieties with the desired traits (Jeon et al., 2023). Before proceeding with molecular breeding, it is essential to understand the characteristics of the variety (Mumm, 2008). This includes an in-depth understanding of the genetic background, agronomic traits, growth habits, stress tolerance, quality characteristics of the target crop, etc. This information can help determine breeding goals, select suitable parents, and predict trait performance in crossbred offspring. The success of molecular breeding depends to a large extent on the accurate understanding and utilization of the characteristics of the variety (Ahmar et al., 2020).

As a rice variety widely used in actual production, the association between the characteristics of 1892S and specific functional gene loci may not be fully understood. Although modern molecular breeding techniques have come a long way, there are still some challenges (Lamichhane, 2022). To better understand the association between the characteristic properties of 1892S and specific functional gene loci, this study analyzed the *Indica-Japonica* properties of 1892S using an improved indel maker analysis strategy. Due to the distant genetic relationship between *Indica rice and Japonica rice, Indica-Japonica* rice hybridization often produces many genetic recombination and segregation types, which provides breeders with more selection opportunities (Xu et al., 2020). The results showed that 1892S had a certain proportion of *Japonica* rice attributes, and the characteristics of 1892S, as the female parent had strong stress resistance, tillering ability, and yield advantages, confirmed the results of this analysis (Table 3).

This study implemented an enhanced RiceNavi approach. First, RiceNavi was used to predict the characteristic features of 1892S. Second, the assembled 1892S genome sequence (second-generation) was used to confirm the presence of gene sequences predicted by RiceNavi. This approach enhances the accuracy of variety characterization and analysis. For example, analyzing the wide affinity of 1892S solely through sequence alignment might only suggest the presence of the S5-n gene due to a deletion in the N-terminal region. However, our analysis not only predicted a wide affinity gene in 1892S but also confirmed this through real-world production data. This is supported by the fact that 114 varieties utilizing 1892S as the female parent have been nationally approved across nine provinces. These varieties have diverse paternal parents originating from different ecological regions and genetic backgrounds, further supporting the analysis of wide affinity genes in 1892S.

Furthermore, this study predicted lodging resistance in 1892S, and analysis confirmed the presence of loci related to this trait, such as APO1 site SCM2 (Ookawa et al., 2010), SBI/, and SD1 (Guha et al., 2024). Hybrid varieties with 1892S as the female parent also demonstrate strong lodging resistance, exemplified by Wandao 153 – a widely grown variety in the middle and lower Yangtze River region known for its exceptional lodging resistance.

While next-generation sequencing technologies offer high-resolution data for genome-wide analysis, limitations exist, including challenges in covering specific genomic regions and detecting large structural variants or repeats (Satam et al., 2023). Advancements in sequencing technology, particularly PacBio SMRT and Oxford Nanopore platforms, offer longer read lengths and higher accuracy, facilitating a more comprehensive analysis of genomic structures, including complex regions and variations in repeats (Udaondo et al., 2021). Utilizing data from multiple generations alongside novel analytical methods will allow for more precise identification and verification of functional gene loci in 1892S, providing deeper support for molecular design breeding. This data can empower breeders to gain a clearer understanding of 1892S's genetic background, refine cross-breeding strategies, and enhance the efficiency and success rate of breeding new varieties. Additionally, integrating multi-omics data offers a more comprehensive view of gene regulation across different levels, leading to more precise guidance for molecular design breeding (Zhang et al., 2022). As technology advances and costs decline, next-generation sequencing and other multi-omics technologies will play an increasingly significant role in propelling molecular design breeding forward (Mahmood et al., 2022; Yang et al., 2021).

Overall, this study conducted a comprehensive analysis of 1892S's characteristic features, revealing its dominance in heterosis, wide affinity, lodging resistance, and other traits. Additionally, we identified some potentially disadvantaged genes in 1892S. This comprehensive characterization not only provides theoretical support for the wider adoption and application of 1892S but also establishes a solid foundation for using 1892S as breeding or research material in future endeavors.

ACKNOWLEDGMENTS

REFERENCE

Ahmar S, Gill RA, Jung KH, et al. (2020). Conventional and Molecular Techniques from Simple Breeding to Speed Breeding in Crop Plants: Recent Advances and Future Outlook. *Int J Mol Sci*. 21: 2590.

Asano K, Takashi T, Miura K, et al. (2007). Genetic and Molecular Analysis of Utility of sd1 Alleles in Rice Breeding. *Breeding Science.* 57: 53-58.

Ashikari M, Sakakibara H, Lin S, et al. (2005). Cytokinin oxidase regulates rice grain production. *Science.* 309: 741-745.

Ashraf MF, Peng G, Liu Z, et al. (2020). Molecular Control and Application of Male Fertility for Two-Line Hybrid Rice Breeding. *Int J Mol Sci.* 21: 21.

Begna T. (2022). Challenges and Progresses of Molecular Plant Breeding in Crop Improvement. *International Journal of Research in Agricultural Sciences.* 7: 2348-3997.

Cao L, Zh X. (2014). Chinese Experiences in Breeding Three-Line, Two-Line and Super Hybrid Rice in Rice – Germplasm. *Genetics and Improvement.*

Chen J, Ding J, Ouyang Y, et al. (2008). A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proc Natl Acad Sci U S A* 105: 11436-11441.

Collard BC, Vera Cruz CM, McNally KL, et al. (2008). Rice molecular breeding laboratories in the genomics era: Current status and future considerations. *Int J Plant Genomics.* 524847.

Edgar RC. (2022). Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun.* 13: 6968.

Gao Z, Wang Y, Chen G, et al. (2019). The indica nitrate reductase gene OsNR2 allele enhances rice yield potential and nitrogen use efficiency. *Nat Commun.* 10: 5207.

Gu B, Zhou T, Luo J, et al. (2015). An-2 Encodes a Cytokinin Synthesis Enzyme that Regulates Awn Length and Grain Production in Rice. *Mol Plant.* 8: 1635-1650.

Guha PK, Magar ND, Kommana M, et al. (2024). Strong culm: a crucial trait for developing next-generation climate-resilient rice lines. *Physiol Mol Biol Plants.* 30: 665-686.

Hu B, Wang W, Ou S, et al. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nat Genet.* 47: 834-838.

Hua L, Wang DR, Tan L, et al. (2015). LABA1, a Domestication Gene Associated with Long, Barbed Awns in Wild Rice. *Plant Cell.* 27: 1875-1888.

Jeon D, Kang Y, Lee S, et al. (2023). Digitalizing breeding in plants: A new trend of next-generation breeding based on genomic prediction. *Front Plant Sci.* 14: 1092584.

Jiao Y, Wang Y, Xue D, et al. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat Genet.* 42: 541-544.

Khan AA, Iqbal B, Jalal A, et al. (2024). Advanced Molecular Approaches for Improving Crop Yield and Quality: A Review. *Journal of Plant Growth Regulation.* 43: 2091–2103.

Lamichhane S, Thapa A. (2022). Advances from Conventional to Modern Plant Breeding Methodologies. *Plant Breeding and Biotechnology.* 10: 1-14.

Li H. (2021). New strategies to improve minimap2 alignment accuracy. *Bioinformatics.* 37:4572-4574.

Lian-song Y, Yi-song B, Zheng Q. (2016). Promotion and Application of Male Sterile Line 1892S Breeding Series □esistant Combination. *Journal of Anhui Agri. Sci.* 44: 63-66.

Liansong Y, Yisong B. (2006). Wandao153,a New Medium Indica Two-line Hybrid Rice Combination. *HYBRID RICE.* 21: 87-88.

Liu C, Zheng S, Gui J, et al. (2018). Shortened Basal Internodes Encodes a Gibberellin 2-Oxidase and Contributes to Lodging Resistance in Rice. *Mol Plant.* 11: 288-299.

Luo R, Liu B, Xie Y, et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience.* 1: 18.

Mahmood U, Li X, Fan Y, et al. (2022). Niu et al. Multi-omics revolution to promote plant breeding efficiency. *Front Plant Sci.* 13: 1062952.

Manni M, Berkeley MR, Seppey M, et al. (2021). BUSCO: Assessing Genomic Data Quality and Beyond. *Curr Protoc.* 1: e323.

Marcais G, Delcher AL, Phillippy AM, et al. (2018). MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol.* 14: e1005944.

Miaomiao P, Yachun Y, Li L, et al. (2013). Analysis of the Lodging □esistance of WD153□A Two-line Hybrid Indica □ice. *Journal of Anhui Agri. Sci.* 41: 11620-11621.

Moose SP, Mumm RH. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* 147: 969-977.

Ookawa T, Hobo T, Yano M, et al. (2010). New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nat Commun.* 1: 132.

Paril J, Reif J, Fournier-Level A, et al. (2024). Heterosis in crop improvement. *Plant J.* 117: 23-32.

Pasala R, Chennamsetti M, Patil B, et al. (2024). Revolutionizing Crop Production: The Imperative of Speed Breeding Technology in Modern Crop Improvement. Crop Breeding, *Genetics and Genomics.* 6: 2.

Qianlong L, Qi F, Heqin W, et al. (2023). Genome-Wide Dissection of Quan 9311A Breeding Process and Application Advantages. *Rice Science.* 30: 552-566.

Satam H, Joshi K, Mangrolia U, et al. (2023). Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology (Basel)* 12: 7.

Shen YJ, Jiang H, Jin JP, et al. (2004). Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiol.* 135: 1198-1205.

Udaondo Z, Sittikankaew K, Uengwetwanit T, et al. (2021). Comparative Analysis of PacBio and Oxford Nanopore Sequencing Technologies for Transcriptomic Landscape Identification of Penaeus monodon. *Life (Basel)* 11: 8.

Wang Y, Xue Y, Li J. (2005). Towards molecular breeding and improvement of rice in China. *Trends Plant Sci.* 10: 610-614.

Wei X, Qiu J, Yong K, et al. (2021). A quantitative genomics map of rice provides genetic insights and guides breeding. *Nat Genet.* 53: 243-253.

Wu X, Liu Y, Zhang Y, et al. (2021). Advances in Research on the Mechanism of Heterosis in Plants. *Front Plant Sci.* 12: 745726.

Wu Y, Wang Y, Mi XF, et al. (2016). The QTL GNP1 Encodes GA20ox1, Which Increases Grain Number and Yield by Increasing Cytokinin Activity in Rice Panicle Meristems. *PLoS Genet.* 12: e1006386.

Xu D, Zhu Y, Chen Zf, et al. (2020). Yield characteristics of japonica/indica hybrids rice in the middle and lower reaches of the Yangtze River in China. *Journal of Integrative Agriculture.* 19: 2394-2406.

Xu Y, Yu D, Chen J, et al. (2023). A review of rice male sterility types and their sterility mechanisms. *Heliyon.* 9: e18204.

Y.N.J, Z. X.Z C, T. D.W, et al. (2022). Mapping QTLs for rice stigma exsertion rate using recombinant inbred line population. *Molecular Plant Breeding.* 20: 1198-1205.

Yang Lian-song, Yi-song B. (2012). Study on the Breeding of Indica Rice PGMS Line 1892S and Its Application. *Journal of Anhui Agri. Sci.* 40: 12808-12810.

Yang Y, Saand MA, Huang L, et al. (2021). Applications of Multi-Omics Technologies for Crop Improvement. *Front Plant Sci.* 12: 563953.

Yuan LP. (2014). Development of Hybrid Rice to Ensure Food Security. *Rice Science.* 21: 1-2.

Zhang R, Zhang C, Yu C, et al. (2022). Integration of multi-omics technologies for crop improvement: Status and prospects. *Front Bioinform.* 2: 1027457.

Zhu X, Gou Y, Heng Y, et al. (2023). Targeted manipulation of grain shape genes effectively improves outcrossing rate and hybrid seed production in rice. *Plant Biotechnol J.* 21: 381-390.