agronomy

MDPI

Article

Genetic Analysis of *S5* Regulating the Hybrid Sterility between *Indica* and *Japonica* Subspecies in Rice

Jie Guo [1,2,3], Xinqiao Zhou [1,2,3], Ke Chen [1,2,3], Chanjuan Ye [1,2,3], Juan Liu [1,2,3], Kangli Sun [4], Guang Tang [4], Shaokui Wang [4], Guiquan Zhang [4], Youding Chen [1,2,3], Dagang Chen [1,2,3,*] and Chuanguang Liu [1,2,3,*]

1 Key Laboratory of Genetics and Breeding of High Quality Rice in Southern China (Co-Construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, Guangzhou 510640, China; guojie2503@126.com (J.G.)
2 Guangdong Key Laboratory of New Technology in Rice Breeding, Guangzhou 510640, China
3 Guangdong Rice Engineering Laboratory/Rice Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China
4 Guangdong Provincial Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou 510640, China; tangguang2012@163.com (G.T.)
* Correspondence: chendg@gdaas.cn (D.C.); guyliu@126.com (C.L.); Tel.: +86-20-87596275 (C.L.)

Abstract: Hybrid sterility is the major obstacle to the utilization of inter-subspecific heterosis in hybrid rice breeding. The *S5* locus, composed of three adjacent genes *ORF3*, *ORF4*, and *ORF5*, plays a crucial role in regulating *indica*/*japonica* hybrids' female sterility. Through a series of crosses involving 38 parents, three alleles of *S5*, *ORF3+ORF4−ORF5n*, *ORF3+ORF4+ORF5n*, and *ORF3−/ORF4−/ORF5n*, all could be regarded as wide-compatibility alleles, and when crossed with *indica* or *japonica* rice, they all showed significantly high fertility. Then, in order to explore the genes' function, we further knocked out genes by using CRISPR/Cas9-based genome editing. Our results demonstrate that the *ORF3+* was not just the protector in the killer-protector system, and knocking out *ORF3* of the *indica* allele seriously affected the rice's normal development. We observed the concrete enhancing hybrid spikelet fertility from the crosses between the *ORF4+* knockout *japonica* materials with *indica* varieties. By conducting the comparative RNA-Seq analysis of young spikelets, we found that the *ORF4+/ORF4−* could modulate the hybrid fertility by affecting the expressions of genes related to the function of the Golgi apparatus. This study indicated that knocking out the *ORF4+* of the *japonica* allele or using the alleles carrying *ORF5n* would provide effective approaches to overcome *indica*/*japonica* hybrid female sterility in rice breeding.

Keywords: hybrid sterility; genetic effect; transcriptome analysis; CRISPR/Cas9; heterosis

Citation: Guo, J.; Zhou, X.; Chen, K.; Ye, C.; Liu, J.; Sun, K.; Tang, G.; Wang, S.; Zhang, G.; Chen, Y.; et al. Genetic Analysis of *S5* Regulating the Hybrid Sterility between *Indica* and *Japonica* Subspecies in Rice. *Agronomy* **2023**, *13*, 1094. https://doi.org/10.3390/agronomy13041094

Academic Editor: Matthew Hegarty

Received: 24 February 2023
Revised: 24 March 2023
Accepted: 3 April 2023
Published: 11 April 2023

1. Introduction

Rice is one of the most important crops in the world. With the rapid growth of the world population and the dramatically decreasing cultivated lands globally, high yield is the unchangeable theme and challenge for rice breeding in the world [1]. The Asian cultivated rice (*Oryza sativa* L.) is divided into two subspecies, known as *Xian*/*indica* and *Geng*/*japonica* [2,3]. Due to their greater genetic diversity, *indica*/*japonica* crosses contained stronger heterosis than intra-subspecific crosses, mainly in plant height, effective spike, grain weight, and other essential agronomic traits [4,5]. So, the *indica-japonica* interspecific hybrid rice, known as fifth generation (5G) rice, could be represented as a potential approach to increasing rice productivity [6–8]. However, this heterosis of inter-subspecific crosses is hardly used in breeding programs because of the serious hybrid sterility [9–11].

Hybrid sterility is the common form of postzygotic reproductive isolation [9,12]. To unveil the molecular mechanism of hybrid sterility among *indica/japonica* subspecies remained a hugely complicated work. Subsequently, about 50 loci controlling inter-specific or inter-subspecific hybrid sterility have been identified or mapped, among which 12 of

them have been cloned [7,11,13,14]. However, in *indica/japonica* hybrids, only three loci *Sa*, *Sc*, *pf12*, and *DPL1/DPL2* were associated with hybrid pollen sterility [15–18], and three loci *S5*, *HSA1*, and *S7* relating to hybrid embryo sac sterility [19–21] were cloned. Several genetic models have been proposed to explain the genetic mechanisms of hybrid sterility, among them, two major genetic models focusing on F_1 hybrid sterility were more influential, the one-locus sporo-gametophytic interaction model, and the duplicate gametic lethal model [22–24]. Previous reports of five cloned loci conformed to the one-locus sporo-gametophytic interaction model but they have unique molecular mechanisms [7].

Among all these hybrid sterility loci, *S5* is a major reproductive barrier regulator in cultivated rice. Chen et al. [19] cloned this gene and found that it encodes an aspartic protease, hereby proposing a tri-allelic system. Their research showed that the two different nucleotides (C819A and C1412T) in *S5*-i and *S5*-j have caused hybrid sterility as a heterozygous pair, while the *S5*-n allele with a 136 bp DNA-sequence deletion triggered the malfunction of genes conferring fertility with either allele [25–28]. Based on these results, Yang et al. [29] proposed a killer-protector system in *indica/japonica* hybrid rice. It indicated that the gene *S5* contained three tightly linked open reading frames (ORFs). From protein function prediction, the *ORF3* encodes a heat shock protein Hsp70, while the *ORF4* encodes a membrane protein and the *ORF5* encodes an aspartate protease. The genetic evidence showed that the *ORF5*+ in combination with the *ORF4*+ acts as a "killer" and selectively eliminates the female gametes without the "protector" *ORF3*+. In contrast, the *ORF3*+ prevented the gametes from being killed and subsequently rescued the hybrid fertility [29]. In addition, another study [30] showed the processes underlying a reproductive barrier induced by the *S5* gene at the transcriptomic level, revealing that the interaction of different *ORF* combinations induced endoplasmic reticulum (ER) stress, and then led to programmed cell death (PCD), eventually resulted in abortive female gametes. So, diverse *ORF* combinations of *S5* locus in hybrids in different genetic backgrounds would show a different level of hybrid sterility [30,31]. In summary, *S5* gene is the key reason responsible for sterility of *indica/japonica* subspecies hybrids.

The discovery of wide-compatible genes (WCGs) and wide-compatible varieties (WCVs) provides the possibility to overcome interspecific hybrid sterility. The *S5* is the first identified WCG, and the *S5*-n allele showed wide compatibility and could produce normal or higher fertility whether crossed with *indica* or *japonica* rice [29,31]. It is an essential resource to dispatch the sterility of *indica/japonica* subspecies hybrids. In order to subdue the *indica/japonica* hybrid sterility, the *S5*-n allele could be crossed with neutral alleles of some hybrid pollen sterility loci to develop wide compatibility lines (WCLs), which had compatibility with both *indica* and *japonica* varieties [32–34]. Furthermore, the *S5*-n allele could be pyramided with pollen sterility allele *S-i* in the *japonica* genetic background to develop *indica*-compatible *japonica* lines (ICJLs), which when crossed with *indica* rice could produce hybrid rice with normal or near normal fertility [35,36].

The clustered regularly interspaced short palindromic repeats (CRISPR) based gene editing technology, such as CRISPR/Cas9 system, have become powerful tools for genetic improvement and molecular breeding of crops [37–39]. It is capable of rapidly generating novel neutral or advanced alleles at the hybrid sterility loci to alter the expression level of the gene, which subsequently leads to conquering the inter-subspecies reproductive barrier [16,40]. The *S5* gene is a vital locus associated with hybrid sterility [19,29]. In this study, we analyzed the genetic effect of the *S5* genes on hybrid sterility by using *indica* or *japonica* testers and Huajingxian74 (HJX74) single-segment substitution lines (SSSLs). Moreover, we created new alleles of *ORF3* and *ORF4* to overcome the *S5*-induced hybrid sterility by knockout the allele of *ORF4* in the *japonica* genotype. Finally, we performed RNA-sequencing (RNA-seq) to quantify the gene expression levels in hybrids F_1 (*ORF4*+/*ORF4*+) in comparison to the hybrids F_1 (*ORF4*+/*ORF4*−). Our study provided significant implications for understanding the molecular mechanisms of hybrid sterility and promoted the application of heterosis between *indica* and *japonica* cultivars in the future.

2. Materials and Methods

2.1. Plant Materials

The plant material included 38 rice varieties, four breeding lines HJX74-SSSLs, and one pyramiding line TISL-Dbc-Gde. (A) HXJ74, a recipient of the SSSLs and an elite variety planted widely in South China, was used as the donor of genetic background. (B) Three SSSLs carrying *S5* locus in the substituted segments were selected from the SSSL library in HJX74 genetic background, and the genotypes and chromosome substitution segments were analyzed in previous research [34]. (C)TISL-Dbc-Gde was a pyramiding line that carried the *S*-i allele at the *Sb* and *Sc* loci from Dijiaowujian and the *S*-i allele at the *Sd* and *Se* loci from Guangluai4, and it was also a *japonica* line with a genotype and phenotype similar to Taichung 65 (T65) [36]. (D) Other testers were selected from a previously reported study [36]. Among them, 35 varieties were used for sequencing the coding sequence (CDS) of the *S5* locus, and 38 varieties were used as testers to analyze the genetic effect of the *S5* locus. The cultivars belong to typical *indica* or typical *japonica* based on available information. All rice plants were cultivated under natural growth conditions in Guangzhou, China.

2.2. Plasmid Construction and Plant Transformation

Three vectors were constructed in this paper. (A) The full length of *ORF4* cDNA (*Os06g0213000*) was cloned from rice elite cultivar HJX74 into a plant binary vector pCAMBIA1300 under the control of the ubiquitin (UBI) promoter to generate the *ORF4* overexpression construct *Ubi:ORF4*. Then the vector was transferred into *Agrobacterium tumefaciens* strain EHA105 by electroporation and consequently delivered into 10-06 as previously described [41]. (B) The *ORF4* gene (*Os06g0213000*) in SSSL 10-06 was used as a reference to identify proper editing target sites. Two gRNAs, 5′-ACTACCTACATCACCTGCCG-3′, and 5′-CCGCGGCAGCTACGCGGCTT-3′ located within the 1st exon in *ORF4* were selected using the CRISPR2 website (http://crispr.hzau.edu.cn/CRISPR2/, accessed on 1 May 2018). These gRNAs were cloned separately in the pBWA-Hu-cas9pl-U3 vector using this protocol [42]. The *ORF4*-gRNA1 and Cas9 expression cassettes were cloned into pCAMBIA1300 to construct the intermediate vector. Then, the *ORF4*-gRNA2 was cloned into this vector. Finally, the single gene *ORF4* knock-out expression vector pBWA-Hu-cas9pl-U3-*ORF4* plasmid was obtained. It was transferred into EHA105 and subsequently delivered into HJX74 SSSL 10-06 and TISL-Dbc-Gde, respectively. (C) The *ORF3* (*Os06g0212900*) and *ORF4* (*Os06g0213000*) gene in SSSL 14-06 was used as a reference, and the CRISPR/Cas9 gene plasmids for *ORF3* and *ORF4* were constructed as [43]. Six target site sequences of *ORF3* GAACCAGTGATCGAAATCGATGG, GATTTCGATCACTGGTTCCGTGG, and CAGAGGGAACAGCAGCAGCTCGG) and *ORF4* (CACCTGCCGCGGCAGCTACGCGG, CCGCGGCAGCTACGCGGCTTCGG, and GCTGCCGCGGCAGGTGATGTAGG) were cloned into the single guide RNA (sgRNA) expression cassette pBWA-Hu-cas9pl-U3 accordingly. Then, all integrated sgRNA expression cassettes were amplified and cloned into the CRISPR/Cas9 vector to construct pBWA-Hu-cas9pl-U3-*ORF3*-*ORF4* plasmids. Then, this plasmid was transferred into EHA105 by electroporation and later delivered into HJX74 SSSL 14-06. All the primer sequences used in this study are listed in Table S9. All transgenic plants were cultivated in natural conditions in Guangzhou, China.

2.3. PCR Amplification and DNA Analysis

The genomic DNA was extracted from fresh-frozen leaves of each plant as described by Zheng et al. [44] with minor alterations. For DNA sequencing, the PCR was conducted using the method with minor modifications [45]. Two primers (*ORF3*-G and *ORF4*-G) listed in Table S9 were used to amplify sequences with the target sites accordingly to identify the genotype of the transgenic plants. All the sequence data were assembled and aligned by using the Lasergene Seqman software (DNASTAR Lasergene 7.0, Madison, WI, USA).

2.4. Fertility Examination

Three panicles per plant were harvested to examine spikelet fertility and scored as the ratio of the number of filled grains to the total spikelet. These mature flowers were collected from the upper one-third portion of the panicle and fixed in FAA solution (ethanol, formaldehyde, and acetic acid at a ratio of 89:6:5) to examine pollen fertility. More than 300 pollen grains were scanned by microscope using the I_2-KI (containing 0.1% (w/v) iodine and 1% (w/v) potassium iodide) staining method randomly per plant.

2.5. Agronomic Traits and Statistical Analysis

All data on agronomic traits were collected from 10 plants per line. Multiple repeated mean values of each trait were statistically analyzed. The data are presented as the means ± standard deviation (SD) and transferred by $\arcsin^{-1}$ prior if they were provided as a percentage. Student's *t*-test and Dunnett's *t*-test were used to calculate the significant differences. The Chi-square (χ^2) test was used to detect the distorted segregation of three genotypes in F_2 populations from the Mendelian ratio of 1:2:1. The results were considered statistically significant when $p < 0.05$ (*) and extremely significant when $p < 0.01$ (**). The least significant difference (LSD) method or Duncan's method was used for multiple-range tests among multiple groups. SPSS version 17 (Chicago, IL, USA), Microsoft Office Excel 2016 (Redmond, WA, USA), and Origin Pro 9.0 (Northampton, MA, USA) were used for data statistical analysis and figure making.

2.6. Sample Collection and Total RNA Extraction

Panicles were separately collected from three hybrids, 04-06/14-06 (as controls), 04-06/10-06 and 14-06/10-06 at the different panicle developed stages, such as 3.0–4.5-mm-long young panicles (meiosis stages, Sp1), 10.2–11.5-mm-long young panicles (Sp2), and mature panicles before heading (Sp3), correspondingly named Sp1-ck, Sp1-1, Sp1-2, Sp2-ck, Sp2-1, Sp2-2, Sp3-ck, Sp3-1, Sp3-2 with two replicates. Collected samples were frozen in liquid nitrogen and stored at −80 °C. Then, RNA was extracted using a Promega RNA extraction kit (Promega, Beijing, China) according to the manufacturer's instructions. Total RNA was qualified and quantified using a NanoDrop and Agilent 2100 bioanalyzer (Thermo Fisher Scientific, Waltham, MA, USA).

2.7. Library Construction and RNA-Seq Analysis

In this paper, 1 μg RNA was used as raw material for the RNA sample preparations. Sequencing libraries were generated using NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) following the manufacturer's recommendation with index codes adding to attribute sequences to each sample. The sequencing data was filtered through in-house perl scripts, the reads containing sequencing adapter; low-quality read, and reads with ploy-N were removed, afterward clean reads were obtained and stored in FASTQ format [46] The Q20, Q30, and GC content of the clean reads were shown in Table S5. The rice reference genome and annotation files were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_001433935.1/ accessed on 5 June 2020). The clean reads were mapped to the reference genome using HISAT2 (v2.0.4) [47] and aligned to the reference coding gene using Bowtie2 (v2.2.5) [48].

2.8. DEG Identification and Analysis

Differential expression analysis of two biological replicates per group was performed using the DESeq2 (v1.4.5, R Foundation for Statistical Computing, Vienna, Austria). The genes with an adjusted $Q \leq 0.05$ were assigned as differentially expressed. The heatmap was drawn by pheatmap (v1.0.8, Raivo Kolde, Boston, MA, USA) according to the gene expression in different samples. To take insight into the functions of differentially expressed genes (DEGs), Gene Ontology (GO, http://www.geneontology.org/ accessed on 2 August 2021) and Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/ accessed on 1 August 2021) enrichment analysis of annotated DEGs was performed by Phyper (https://en.wikipedia.org/

wiki/Hypergeometric_distribution/ accessed on 3 July 2021) based on Hypergeometric test. The significant levels of terms and pathways were corrected by Q value with a rigorous threshold (Q value $\leq$ 0.05) by Bonferroni [49]. The model of the protein was developed by using SWISS-MODEL (https://swissmodel.expasy.org/ accessed on 4 May 2021) and AlphaFold Protein Structure Database (https://www.alphafold.ebi.ac.uk/ accessed on 4 May 2021).

2.9. Real-Time Quantitative PCR Analysis

In order to confirm the genetic effect of *S5*, Real-time quantitative PCR (qRT-PCR) reactions were performed with three biological replicates using RNA samples from young panicles of SSSL carrying different genotypes. To validate the DEGs identified in the RNA-seq analysis, qRT-PCR analysis also was conducted on the selected genes. Real-time PCR reactions were performed using the SYBR® Premix Ex TaqTM Kit (Takara Biomedical Technology, Beijing, China) on an ABI PRISM 7900HT platform (Applied Biosystems, Waltham, MA, USA), according to the manufacturer's instruction. Relative gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method [50]. The primers used for qRT-PCR are listed in Table S9.

3. Results

3.1. Nucleotide Variations and Genetic Effect of S5 Locus

Samples of 35 Oryza accessions (13 *indica* cultivars, 22 *japonica* cultivars) were taken to investigate the natural variations of *ORF3*, *ORF4,* and *ORF5* at the *S5* locus. The Coding sequences in *ORF3-5* amplified by primers in Table S9 were obtained for each accession. Our results showed that two variants of *ORF3* (one SNP and one InDel), four variants of *ORF4* (three SNPs and one InDels), and seven variants of *ORF5* (five SNPs and two InDels) in the coding regions (Figures 1A and S1–S3). The analysis of all genetic sequences at the *S5* locus, *ORF3*− and *ORF3*+, *ORF4*− and *ORF4*+, *ORF5*−, *ORF5*+, and *ORF5n*, was distinguished by predicted proteins according to previous reports [29]. Six allele combinations were found in these 35 selected varieties, *ORF3+/ORF4+/ORF5+*, *ORF3+/ORF4–/ORF5+*, *ORF3+/ORF4−/ORF5n*, *ORF3+/ORF4+/ORF5n*, *ORF3−/ORF4−/ORF5n,* and *ORF3–/ORF4+/ORF5–*, and the last three allele combinations were only discovered in 16 *japonica* varieties (Table S1).

To assay the genetic effect of the *S5* locus, these *Oryza* accessions were regarded as parental testers in pairs to develop 62 crosses. In addition, the genotypes of the *S5* locus for 15 other varieties were obtained from a previously reported study [36]. The spikelet fertility in all F_1 hybrids varied from 3.28% to 92.48%, mainly due to the different alleles at the *S5* locus in the crossing parental. The hybrids in heterozygous of *S5* with common *indica*/*japonica* genotypes (*ORF3+ORF4−ORF5+/ORF3−ORF4+ORF5−*) showed full sterility, varied from 3.28% to 8.21%, with an average of 5.50%. The distorted segregation from the Mendelian segregation ratio was observed in the selected F_2 populations. In contrast, the average spikelet fertility of the hybrids in the genotype of *ORF3−ORF4+ORF5−/ORF3+ORF4+ORF5+* was significantly higher, reached to 39.51% (Figure 1B, Table S2). Combined with the genotypic analysis of the parents, this result may be related to the allelic interactions (+/−) of the *ORF4* gene. In contrast, three different wide-compatibility alleles (*ORF3+/ORF4−/ORF5n*, *ORF3+/ORF4+/ORF5n*, and *ORF3−/ORF4−/ORF5n*) could rescue the fertility to 68.89%, 79.67%, and 80.29%, when crosses with either *japonica* or *indica* genotypes, and no segregation distortion was observed in the F_2 offspring from these crosses (Figure 1B, Table S2). Therefore, the presence of *ORF5n* may have no impact on hybrid embryo sac sterility whatever alleles of *ORF3* or *ORF4* it combined with. Three wide-compatibility alleles could rescue the reproductive isolation between *indica* and *japonica*. Besides, the fertility of these hybrids was lower than expected, possibly due to the lower pollen fertility controlled by pollen sterility genes.

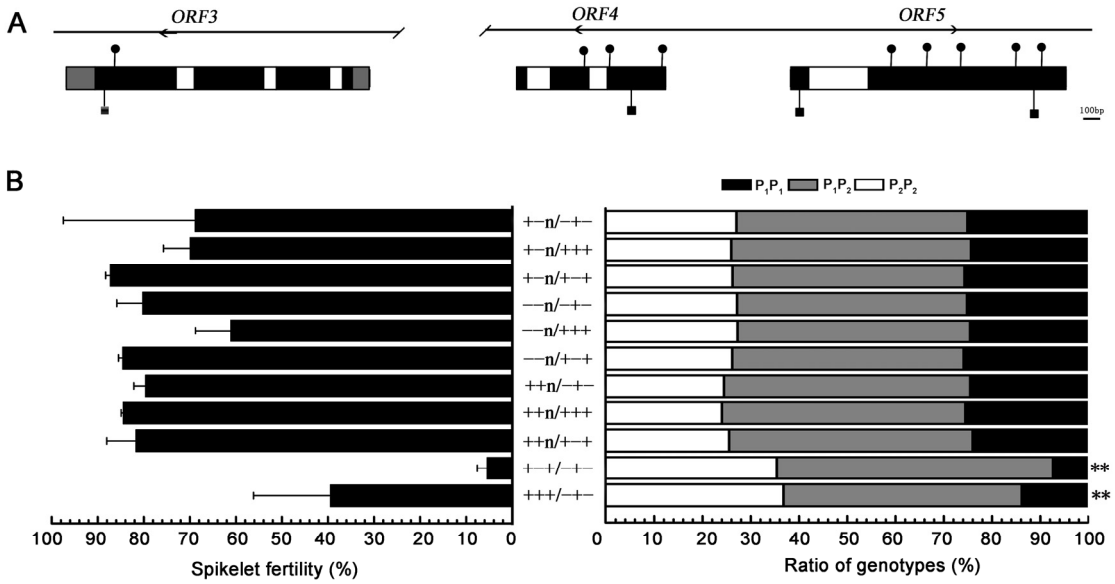

Figure 1. Genetic effect of *ORF3*, *ORF4*, and *ORF5* at the *S5* locus in different *indica/japonica* crosses. (**A**) Schematic drawing of polymorphisms in CDS of three genes at *S5* locus. Black boxes and white boxes represent the exons and introns respectively. Single nucleotide polymorphisms are indicated by lines with solid circles in the gene model. Insertions/deletions (InDels) are indicated by lines with a solid block in the gene model. (**B**) Spikelet fertility of F_1 hybrids in different genotypes at the *S5* locus (left) and the ratio of genotypes at the *S5* locus in the F_2 populations from different *indica/japonica* crosses (right). +−n/−+−, +−n/+++, +−n/+−+, −−n/−+−, −−n/+++, −−n/+−+, ++n/−+−, ++n/+++, ++n/+−+, +−+/−+−, and +++/−+− represented the hybrid F_1 in the genotype of *ORF3+/ORF4−/ORF5n/ORF3−/ORF4+/ORF5−*, *ORF3+/ORF4−/ORF5n/ORF3+/ORF4+/ORF5+*, *ORF3+/ORF4−/ORF5n/ORF3+/ORF4−/ORF5+*, *ORF3−/ORF4−/ORF5n/ORF3−/ORF4+/ORF5−*, *ORF3−/ORF4−/ORF5n/ORF3+/ORF4+/ORF5+*, *ORF3−/ORF4−/ORF5n/ORF3+/ORF4−/ORF5+*, *ORF3+/ORF4+/ORF5n/ORF3−/ORF4+/ORF5−*, *ORF3+/ORF4+/ORF5n/ORF3+/ORF4+/ORF5+*, *ORF3+/ORF4+/ORF5n/ORF3+/ORF4−/ORF5+*, *ORF3+/ORF4−/ORF5+/ORF3−/ORF4+/ORF5−*, and *ORF3+/ORF4+/ORF5+/ORF3−/ORF4+/ORF5−* respectively. P_1P_1, Genotype of *S5* in female parents. P_1P_2, Heterozygous genotype. P_2P_2, Genotype of *S5* in male parents. Error bars represent SD, $n = 10$. ** Significantly different at 0.01 probability level.

3.2. Allelic Interactions of ORF4 in the HJX74-SSSLs

Three SSSLs carrying *S5* locus in the substituted segments were selected from the SSSL library in HJX74 genetic background, and the genotypes and chromosome substitution segments were analyzed in previous research [34]. In the *S5* locus, the genotypes of 04-06, 10-06, and 14-06 were *ORF3+ORF4−ORF5+/ORF3+ORF4−ORF5+*, *ORF3−ORF4+ORF5−/ORF3−ORF4+ORF5−* and *ORF3+ORF4+ORF5+/ORF3+ORF4+ORF5+* respectively [34]. The *ORF3+ORF4−ORF5+* is the typical genotype in *indica* rice. Three *ORF*s all showed low chromatin accessibility during the spikelet development periods via Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) (Figure S4). Then, we evaluated the expression of *ORF3*, *ORF4*, and *ORF5* in young panicles. The results showed that the transcript levels of *ORF3* in 04-06 were the highest, showing that *ORF3+* and *ORF3−* have divergent functions. However, the transcript levels of *ORF4* were obviously lower in 04-06 and 10-06 (Figure 2A). This result might due to the genetic structure of *ORF4−* (Figure S5). Three different *ORF* combinations at the *S5* locus in F_1 hybrids were obtained from the crosses between the SSSLs. The results indicated that the F_1 hybrids in all genotypes showed normal pollen fertility, the fertility rates range from 92.18% to 95.02%. The genotypes of *ORF3+ORF4+ORF5+/ORF3+ORF4−ORF5+* as control showed normal spikelet fertility. However, we found that the spikelet fertility in the *ORF3+ORF4+ORF5+/ORF3−ORF4+ORF5−* genotype was 76.55%, higher than 68.72% in the *ORF3+ORF4−ORF5+/ORF3−ORF4+ORF5−* genotype, which was only different in the genotype

of *ORF4*, *ORF4+/ORF4+* or *ORF4+/ORF4−* (Figure 2B,C). Interestingly, we discovered that the remarkable distorted segregation was detected both in two F_2 populations, but the population of the *japonica* genotype was significantly lower in *ORF3+ORF4−ORF5+/ORF3−ORF4+ORF5−* population (Figure 2D). Taken together, the allele interaction of *ORF4+/ORF4−* could lead to lower fertility in F_1 hybrids of *indica/japonica* crosses.

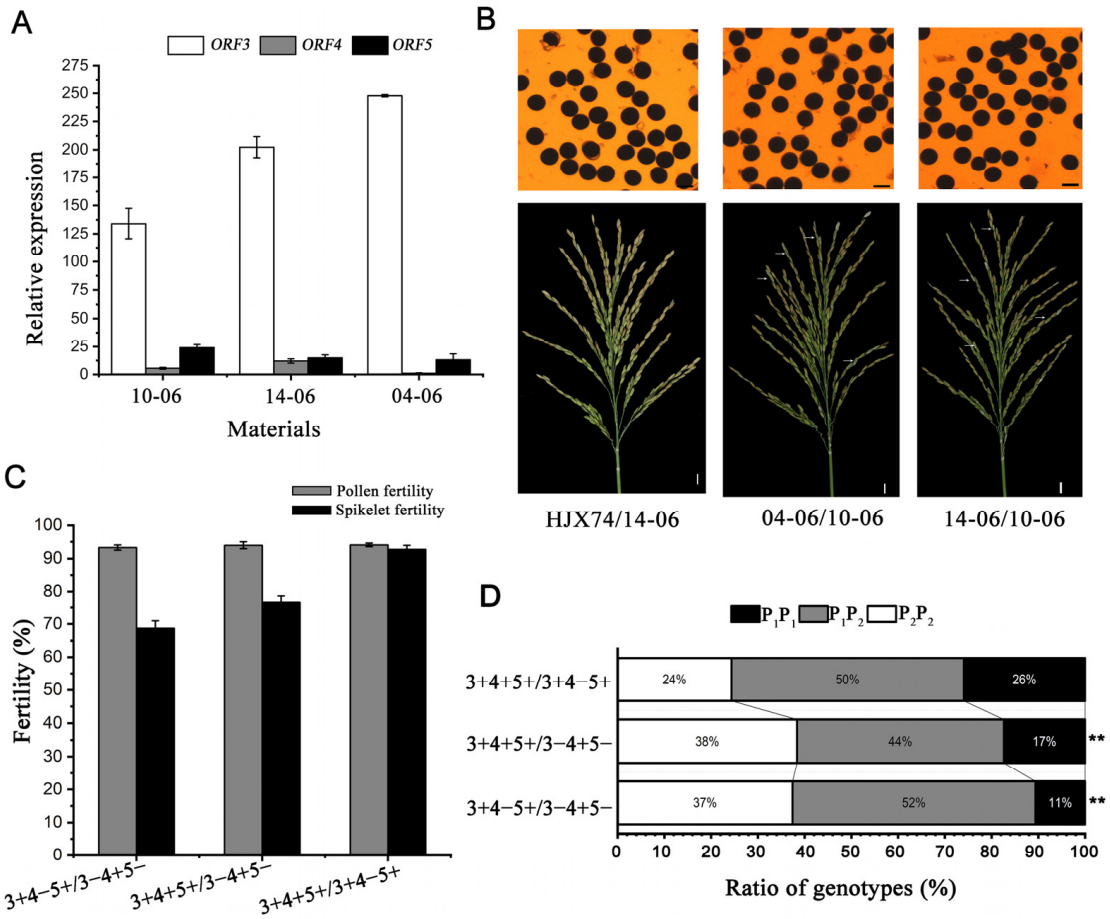

Figure 2. Genotypes and genetic effects of three genes at the *S5* locus in HJX74 SSSLs (**A**) Expression level of three adjacent genes at *S5* locus in young panicles of three SSSLs. The reference gene was Profilin. (**B**) Pollen grains stained by I_2-KI solution (top) and spikelet fertility in the panicles (bottom) in F_1 hybrids. Scale bar in pollen grains, 30 μm. Scale bar in panicles, 1 cm. The arrow points to the empty grain. (**C**) Pollen fertility and spikelet fertility in hybrids F_1 from the crosses of the SSSLs carrying the *S5* gene in different genotypes. *ORF3*, *ORF4*, and *ORF5* were abbreviated as "3", "4", and "5". Error bars represent SD, n = 10. (**D**) The ratio of genotypes at the *S5* locus in the F_2 populations from the crosses of SSSLs with *S5* gene and testers. P_1P_1, Genotype of the left SSSL. P_1P_2, Heterozygous genotypes. P_2P_2, Genotypes of the right SSSL. **, significantly different at 0.01 probability level.

3.3. CRISPR/Cas9-Mediated Editing of ORF4 Gene in HJX74-SSSL

Previously, we collected 171 accessions of *O. sativa* and used them to identify the genotypes of the *S5* gene in previous studies, but the combination *ORF3−ORF4−ORF5−/ORF3−ORF4−ORF5−* was not detected in *O. sativa* [36]. The *ORF4* acted as an essential gene to assist the *ORF5+* in eliminating gametes purposefully to control hybrid fertility. The *ORF4* encodes a transmembrane structural domain protein, and *ORF4−* lost its function because of an 11 bp deletion, resulting in the premature termination of the encoded protein and leading to the loss of the transmembrane structural domain [29].

To obtain new alleles of the *S5* locus, two adjacent targets sites of *ORF4* located in exon 1 were selected and constructed into the CRISPR/Cas9 gene editing vectors, and transformed into the inbred HJX74-SSSL 10-06 (Figure 3A). Six T_0 generation plants were amplified by PCR-specific primers, and positive plants were screened. Thus, two *ORF4* knockout mutant lines E403-9 and E403-12 were obtained. The E403-9 harbored a 1-base pair(bp) insertion in the T_1 editing site and a 41-bp deletion in the T_2 editing site, while E403-12 harbored a 1 bp insertion of the sequence A in the T_1 editing site and a 2 bp insertion of the sequence A in T_2 editing site. Subsequently, the protein sequence analysis showed that the 1 bp base "A" insertion led to the early termination of the *ORF4*+ protein. In comparison with the reference *ORF4*+ protein, 234 amino acids (aa), the E403-9 and E403-12 were both predicted to produce the same truncated proteins of 103 amino acids, named *ORF4*(−) (Figures 3B and S6). In addition, several agronomic traits of edited lines were investigated, and no significant difference between SSSL 10-16 and two gene editing mutants was found in the majority of investigated traits (Figure S7).

Hence, we examined the expression of three *ORF*s at the *S5* locus in the young spikes of E403-9 mutants. The results revealed that the expression level of the *ORF4* gene was significantly decreased in the two mutants in comparison to the control SSSL 10-06. There was no significant difference in the expression level of the *ORF3* gene, while the expression of the *ORF5* was significantly lower in mutants than that of the WT plant. Our results suggested that the editing of *ORF4* caused a significant reduction in the expression of both *ORF4* and *ORF5* (Figure 3C).

3.4. ORF4(−) Improved Indica/Japonica Hybrid Fertility in HJX74 Background

To analyze the effect of *ORF4*(−) on hybrid sterility, two mutants, E403-9 and E403-12, were crossed with HJX74, SSSL 04-06, and 14-06. The F_1 hybrid from HJX74/E403-9 and 04-06/E403-9 (*ORF3*+*ORF4*−*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*−) showed normal spikelet fertility, compared with 65.39% of F_1 hybrid from HJX74/10-06. While the F_1 hybrid of 14-06/E403-9 (*ORF3*+*ORF4*+*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*−) showed lower spikelet fertility, about 69.46%, compared with 78.51% of the 14-06/10-06 plant. (Figure 3D,E, Table S3). In addition, no distorted segregation of the S5 locus was detected in the F_2 population of HJX74/E403-9. Previous studies showed that the majority of current indica rice varieties owned the genotype *ORF3*+*ORF4*−*ORF5*+/*ORF3*+*ORF4*−*ORF5*+ at the *S5* locus, while the genotype of japonica rice was *ORF3*−*ORF4*+*ORF5*−/*ORF3*−*ORF4*+*ORF5*− [29,36]. In conclusion, our results revealed that *ORF4* knockout in the japonica genotype might significantly enhance the hybrid spikelet fertility.

3.5. Overexpression of ORF4 Gene in HJX74-SSSL

Previous studies showed that the expression of *ORF3* and *ORF5* genes had no significant effect on the fertility of hybrid F_1 lines [29,30,51]. In this study, we constructed an *ORF4* overexpression vector and transferred it into HJX74-SSSL 10-06 background to obtain positive plant line E343. The results exhibited that the E343 had significantly higher expression of *ORF4*, *ORF3*, and *ORF5* genes in young spikelets in comparison to 10-06 (Figure 3C). Then, we used E343 crossing with the 04-06 and 14-06 to obtain hybrid F_1. Analysis of the fertility revealed that the combinations 04-06/E343 and 14-06/E343 showed normal pollen fertility, and lower spikelet fertility, 77.52% and 79.98%, respectively (Figure 3D,E). In sum, our results confirmed that overexpression of the *ORF4* gene in *japonica* rice parents significantly improved the fertility of hybrid F_1 of *ORF4*+/*ORF4*− combinations but had no significant effect on *ORF4*+*ORF4*/+ combinations.

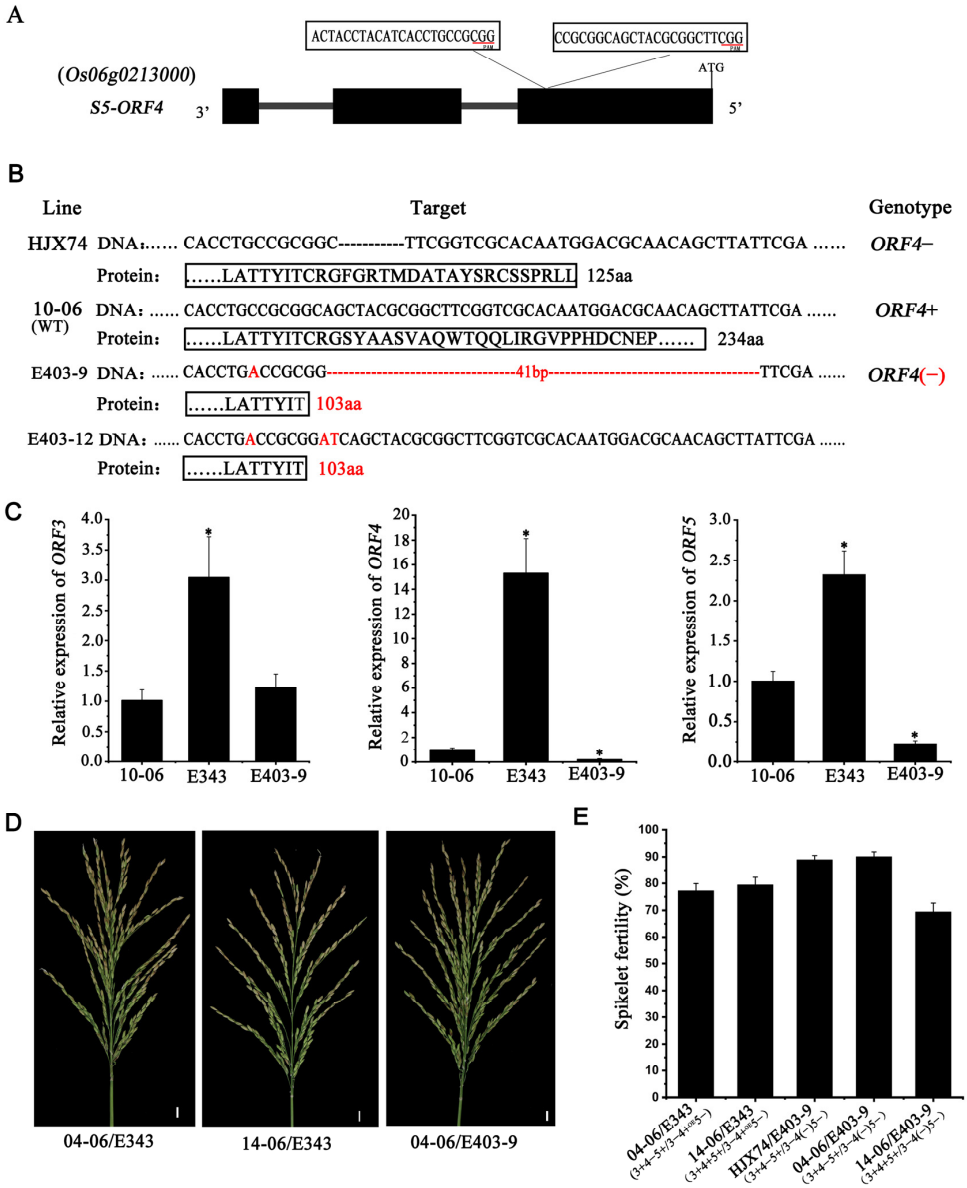

Figure 3. Knockout of *ORF4* in SSSL *S5*-10 improved the hybrid fertility (**A**) The sites of target sequences in *ORF4*. The black boxes represent the exons. Two target sites are located on the first exon of the *ORF4* gene. The target site is located in the box. The PAM sequence is underlined and shown in red. (**B**) Identification of mutations in *ORF4* by sequencing of the target site in T_2 generation. The positions highlighted in red show the targeted mutations and deletions indicated by the red dash "−". *ORF4* (−) indicated the *ORF4* gene loss of function. (**C**) qRT-PCR of *ORF3*, *ORF4*, and *ORF5* expression levels of RNA isolated from young panicles in different lines. 10-06, wild-type. E343, *ORF4*+ overexpression plant in the 10-06 background. E403-9, *ORF4*+ edited mutant line in the 10-06 background. *Profilin* was the reference gene. Results were presented relative to the expression level of three *ORF*s in 10-06. * Significantly different at 0.05 probability level. (**D**) The panicles in F_1 hybrids from three crosses. Scale bar in panicles, 1 cm. (**E**) Spikelet fertility of hybrids F_1 from the crosses of *indica*-SSSLs with E343 or E403-9. The genotypes of the parents in the crosses were in parentheses, *ORF3*, *ORF4*, and *ORF5* were abbreviated as "3", "4", and "5". Error bars represent SD, $n = 10$.

3.6. CRISPR/Cas9-Mediated Editing of ORF3 and ORF4 Gene in HJX74-SSSL

To obtain new *indica* genotypes and confirm the function of the *ORF4* gene, three target sites were selected in exon 4 of the *ORF3* gene and exon 1 of the *ORF4* gene respectively to develop two CRISPR/Cas9-based editing constructs and introduced in the

HJX74-SSSL 14-06 (Figure 4A). In the T_2 generation, we selected three homozygous mutant lines for further study. Sequence alignments of these two homozygous mutants Y184-2 and Y184-3 against that of the control 14-06 further indicated that 1 bp deletion appeared at the first target site, inducing gene frameshift of *ORF3* resulting in a deletion of amino acid residues and no difference in *ORF4*. However, compared to the control 14-06, mutant Y184-4 existed 4 bp deletion between the PAM sites in the *ORF3* gene which resulted in a premature translation but remained intact in the *ORF4* gene. This allele of *ORF3* was named *ORF3*($-^{\beta}$) (Figure 4B). In the T_3 generation, we focused on three homozygous plants (Y184-18, Y184-22, and Y184-6) with different editing patterns in both *ORF3* and *ORF4* proteins. However, only the last ten amino acids were altered in the C termino of the *ORF3* protein in Y184-6, we suspected that it might not affect the function of *ORF3* (Figures 4B, S8 and S9).

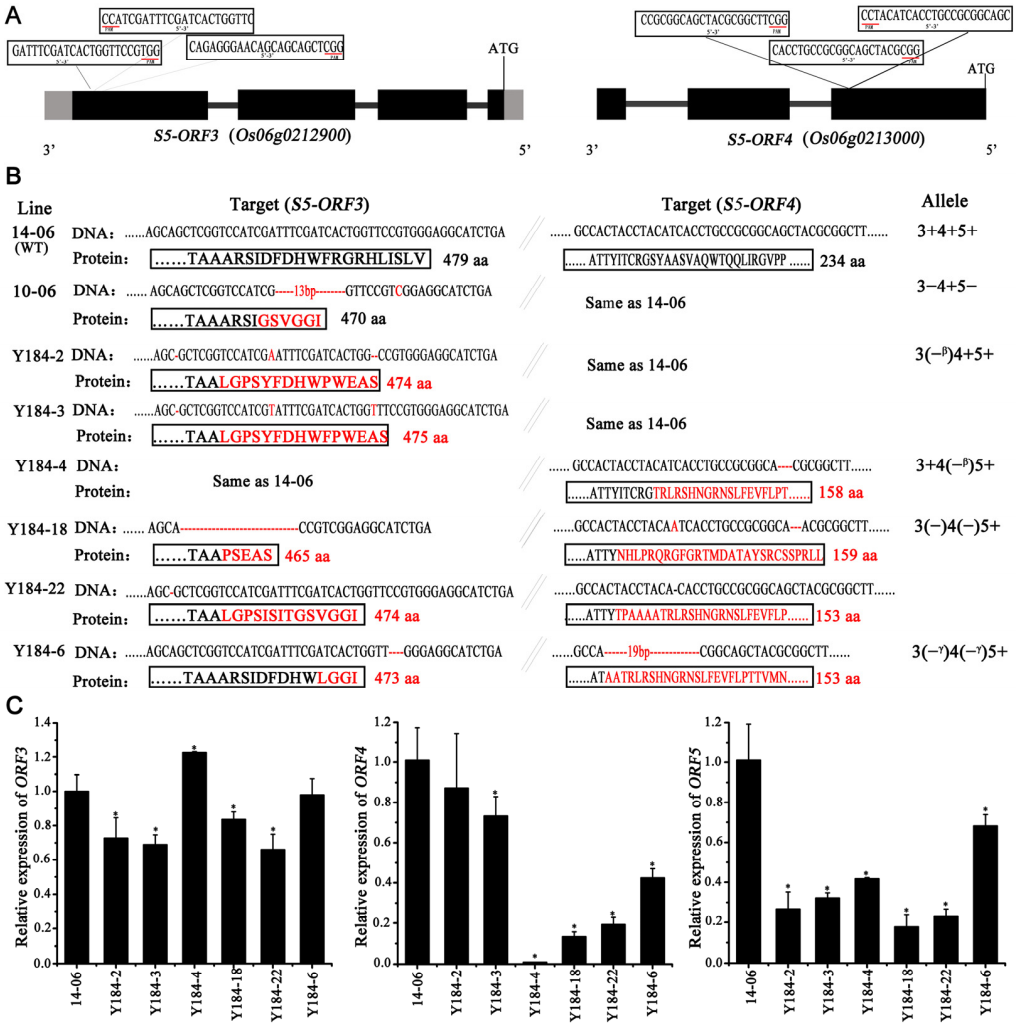

Figure 4. CRISPR/Cas9-mediated editing of *ORF3* and *ORF4* gene in HJX74-SSSL (**A**) Target sites in *ORF3* or *ORF4* for the CRISPR/Cas9 editing. The black boxes represent the exons. Two target sites are located on the four exons of the *ORF3* or the first exon of *ORF4*. The target sites were located in frames. (**B**) Sequencing of the CRISPR/Cas9-targeted sites in *ORF3* (left) and *ORF4* (right). The PCR fragments containing the targeted sites were directly sequenced. The wild-type sequences in 14-06 were as references. The positions highlighted in red show the targeted mutations and deletions indicated by the red dash "−". (**C**) qRT-PCR analysis of *ORF3*, *ORF4*, and *ORF5* in young panicles of different gene-edited lines. *Profilin* was the reference gene. Results were presented relative to the expression level of three *ORF*s in 14-06. * Significantly different at 0.05 probability level.

The expression level of the three genes at the *S5* locus was analyzed in the young spikelet of the six mutants by qRT-PCR. Compared with it in WT, the expression of the gene clusters in the double knockout gene mutants Y184-18 and Y184-22 significantly declined. However, the expression of *ORF3* in double knockout gene mutant Y184-6 was no obvious changes. In contrast, the expression of *ORF4* and *ORF5* in the mutants Y184-4 were significantly down-regulated, and its relative expression of *ORF3* was up-regulated (Figure 4C). Previous reports discussed that the promoter region of *ORF4* and *ORF5* overlapped and frequently co-expressed [29]. In addition, lower expressions of *ORF3* and *ORF5* were shown in Y184-2 and Y184-3 (Figure 4C). In conclusion, based on the mutation sites and the expression of *ORF*s, the allele of *S5* in Y184-18 and Y184-22 was named *ORF3*(−)*ORF4*(−)*ORF5*+, while it in Y184-6 was named *ORF3*($-^{\gamma}$)*ORF4*($-^{\gamma}$)*ORF5*+. The allele of *S5* in Y184-4 was named *ORF3*+*ORF4*($-^{\beta}$)*ORF5*+, while it in Y184-2 and Y184-3 was named *ORF3*($-^{\beta}$)*ORF4*+*ORF5*+. The result also indicated that *ORF3*($-^{\gamma}$), *ORF3*(−), *ORF3*($-^{\beta}$) and *ORF3*+ alleles may have diverse functions.

Besides fertility, we also investigated other major relevant agronomic traits of 14-06 and six mutants. As expected, we discovered that no significant difference was present between 14-06 and Y184-6, Y184-4 in the majority of traits. In comparison to the 14-06, the panicle number, grain length, 1000 grain weight, and plant height of Y184-2, Y184-3, and Y184-22 consistently decreased to a relatively lower status. In total, the results indicated the editing of *ORF3* would affect plant growth and development (Figure S10).

3.7. Strong Influence on Hybrid Sterility by Editing of ORF3 and ORF4 in HJX74 Background

To analyze the function of new alleles, we chose six mutants crossed with HJX74-SSSL 10-06 and mutants E403-9. We analyzed the fertility of mutants in heterozygous or homozygous genotypes and the hybrid F_1 lines. The results suggested that the mutants in the homozygous genotype had normal fertility, except for mutant Y184-2 and Y184-3 (genotype of *ORF3*($-^{\beta}$)*ORF4*+*ORF5*+/*ORF3*($-^{\beta}$)*ORF4*+*ORF5*+), which had lower spikelet fertility, 33.24%, and 14.55%, but normal pollen fertility (Figure 5, Table S4). These data indicated that unlike common *ORF3*+, *ORF3*($-^{\beta}$) showed no function on hybrid sterility, and could not prevent the gametes from being eliminated.

Furthermore, the spikelet fertility of most mutants in the heterozygous genotype showed about 25% abortive, while the mutant Y184-4 (genotype of *ORF3*+*ORF4*($-^{\beta}$)*ORF5*+/*ORF3*+*ORF4*+*ORF5*+) and Y184-6 (genotype of *ORF3*($-^{\gamma}$)*ORF4*($-^{\gamma}$)*ORF5*+/*ORF3*+*ORF4*+*ORF5*+) showed normal spikelet fertility, revealing the *ORF3*+ and *ORF3*($-^{\gamma}$) acted as the same role and prevented the gametes from being killed.

In this case, we identified the spikelet fertility of 14-06/10-06 hybrid (genotype of *ORF3*+*ORF4*+*ORF5*+/*ORF3*−*ORF4*+*ORF5*−), served as control, was lower than normal level, about 78.51%. The hybrids F_1 from the combinations Y184-2/10-06 (genotype of *ORF3*($-^{\beta}$)*ORF4*+*ORF5*+/*ORF3*−*ORF4*+*ORF5*−), Y184-2/E403-9 (genotype of *ORF3*($-^{\beta}$)*ORF4*+*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*−) and Y184-18/10-06 (genotype of *ORF3*(−)*ORF4*+*ORF5*+/*ORF3*−*ORF4*+*ORF5*−) showed the lowest spikelet fertility 19.89%, 21.56% and 25.14%, respectively (Figure 5, Table S4). It revealed that the sequence difference in *ORF3(−)* and *ORF3*($-^{\beta}$) probably destroyed the *ORF3*+ protein structure resulting in functional changes in the hybrid sterility in the case of combining with *japonica* gametes.

Besides, compared with 14-06/10-06 hybrid (*ORF3*+*ORF4*+*ORF5*+/*ORF3*−*ORF4*+*ORF5*−, 78.51%) and 13-06/10-06 (*ORF3*+*ORF4*−*ORF5*+/*ORF3*−*ORF4*+*ORF5*−, 66.50%), the hybrid F_1 of Y184-4/10-06 showed semi-sterility 49.62% in genotype *ORF3*+*ORF4*($-^{\beta}$)*ORF5*+/*ORF3*−*ORF4*+*ORF5*−, remarkably less than the controls. Previous research showed 14-06/E403-12 produce hybrid F1 with 65.25% fertility in genotype *ORF3*+*ORF4*+*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*− (Figure 5, Table S4). In conclusion, all these results revealed there might exist a correlation between *ORF4*+, *ORF4*−and *ORF5*+ when *indica*-allele crossed with *japonica*-allele. This existence would regulate the hybrid fertility and *ORF4* knockout in the *indica* allele would destroy this correlation leading to reduce the hybrid fertility.

In addition, the spikelet fertility of Y184-4/E404-9 in the genotype of *ORF3*+*ORF4*($-^{\beta}$)*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*− and Y184-18/E403-9 or Y184-22/E403-9 in the genotype of *ORF3*(−)*ORF4*(−)*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*− was normal, about 88.45%, 87.19% and 90.25% respectively (Figure 5, Table S4). These results revealed that the *ORF4* knockout in two parents might hugely restore the hybrid F_1 spikelet fertility.

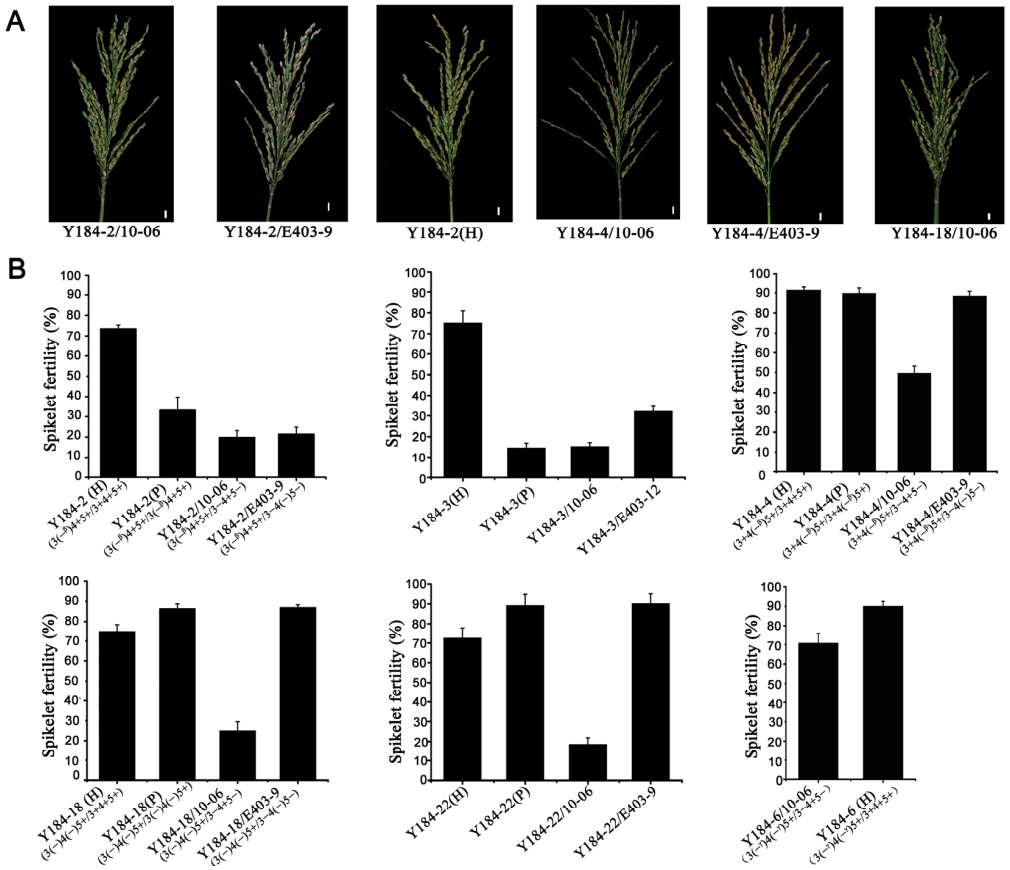

Figure 5. Knockout of both *ORF3* and *ORF4* by CRISPR/Cas9 resulted in lower hybrid fertility (**A**) Spikelet fertility in the panicles in F_1 hybrids from six crosses. Scale bar in panicles, 1 cm. (**B**) Spikelet fertility in hybrids F_1 from the crosses of gene-edited mutations in the SSSL 14-06 background with 10-06. The genotypes of the parents in the crosses were in parentheses, *ORF3*, *ORF4*, and *ORF5* were abbreviated as "3", "4", and "5". Error bars represent SD, n = 10.

3.8. Konckout the ORF4 Gene to Overcome Hybrid Sterility Caused by S5-Dependent in Indica/Japonica Background

The TISL-Dbc-Gde was a pyramiding line that carried the *S-i* allele at the *Sb* and *Sc* loci from Dijiaowujian and the *S-i* allele at the *Sd* and *Se* loci from Guangluai4, and it was also a *japonica* line with a genotype and phenotype similar to T65 [36]. In this study, we introduced *ORF4*-sgRNA/Cas9 editing constructs into the pyramiding line TISL-Dbc-Gde. We obtained three homozygous mutants, T304-3, T304-6, and T304-12. We identified the *ORF4* gene premature terminated in all mutants by sanger sequencing, named *ORF4* ($-^{t}$) (Figures 6A and S11). All these mutants have excluded transgenic elements of the Cas9 vectors. To further explore whether mutations in the *ORF4* gene in the *japonica* background affect the main agronomic traits, we measured the main agronomic traits of three mutant lines, including plant height, 1000-seed weight, etc. The results showed almost no significantly different between TISL-Dbc-Gde and mutant lines (Figure S12).

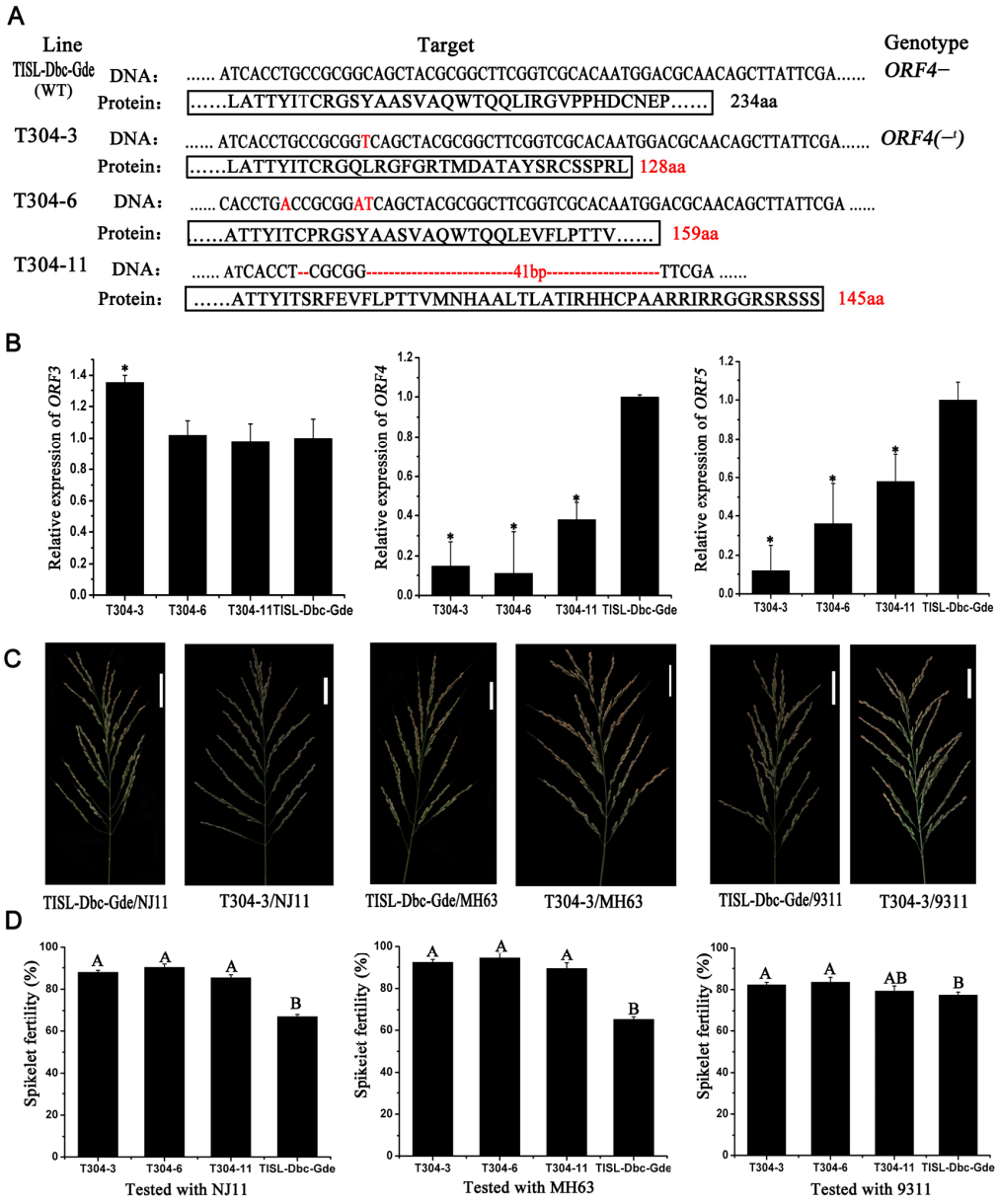

Figure 6. *ORF4* knockout overcame the *S5*-dependent hybrid sterility in *indica*/*japonica* background (**A**) Identification of mutations in *ORF4* by sequencing of the target site in T_2 generation. The positions highlighted in red show the targeted mutations and deletions indicated by the red dash "–". *ORF4* ($-^t$) indicated the *ORF4* gene loss of function. (**B**) qRT-PCR of *ORF3*, *ORF4*, and *ORF5* expression levels in RNA isolated from young panicles of different lines. *Profilin* was the reference gene. Results were presented relative to the expression level in TISL-Dbc-Gde. * Significantly different at 0.05 probability level. (**C**) Spikelet fertility in the panicles in F_1 hybrids from three crosses. Scale bar in panicles, 1 cm. (**D**) Spikelet fertility in hybrids F_1 from the crosses of TISL-Dbc-Gde with *indica* testers Minghui63 (MH63), Nanjing11 (NJ11), and 9311. Capital letters indicate statistical differences at 0.01 probability level, $n = 10$.

To evaluate the effect of *ORF4* ($-^t$), the mutants were test-crossed with a set of testers including typical *indica* varieties Nanjing11 (NJ11), Minghui63 (MH63), and 9311. The genotype of the *S5* locus in NJ11 and MH63 was *ORF3+ORF4−ORF5+*/*ORF3+ORF4−ORF5+*, while its genotype in 9311 was *ORF3+ORF4+ORF5+*/*ORF3+ORF4+ORF5+*. Compared with TISL-Dbc-Gde, three edited lines showed higher spikelet fertility of F_1 hybrids when tested with NJ11 or MH63, with an average of 87.83% or 91.99% (Figure 6C,D). In the

F_2 populations from the T304-3/NJ11 and T304-3/MH63 crosses, no distorted segregation at the *S5* gene was found, and the three genotypes' segregated ratio were 56:113:52 and 63:142:59. Thus, these results indicated that the *ORF4* knockout *japonica* lines could produce F_1 hybrids with high spikelet fertility when it crossed with *indica* rice.

3.9. Comparative Transcriptome Analysis of Young Spikelet in Different Hybrids by RNA-Seq

To better understand the effects of the *S5* locus in the transcription levels, the hybrid F_1 lines from three hybrid combinations, 04-06/10-06 (genotype of *ORF3*+OR4−*ORF5*+/*ORF3*−*ORF4*+*ORF5*−), 14-06/10-06 (genotype of *ORF3*+OR4+*ORF5*+/*ORF3*−*ORF4*+*ORF5*−), and 04-06/14-06 (genotype of *ORF3*+OR4−*ORF5*+/*ORF3*+*ORF4*+*ORF5*+, as control), were selected at different periods of spikelet development, young spikelet length 3.0–4.5 cm (Sp1-1, Sp1-2, Sp1-ck), young spikelet length 10.2–11.5 cm (Sp2-1, Sp2-2, Sp2-ck), and the day before heading (Sp3-1, Sp3-2, Sp3-ck) for transcriptome analysis. Totally 1237.04 M clean reads were obtained. The Q30 base percentage was 95.00% (Table S5). The DEGs in different spikelet development stages were screened out based on the expression of fold change ≥ 2 and p-value ≤ 0.05. Then we verified the expression of six randomly selected DEGs by qRT-PCR, and the results showed similar expression trends between RNA-Seq data and transcriptome analysis (Figure S14).

In the Sp1 stage, the oocyte meiosis stage, genes related to ER stress as well as PCD such as *OsBip5* [52] and *OsAP25* [53] were not significantly differentially expressed in Sp1-1 and Sp1-2 compared with Sp1-ck, revealing the ER stress and PCD did not contribute to *S5*-dependent fertility. These results were probably due to the presence of *ORF3*+ genes in both hybrid combinations (Figure S13, Table S6). This conclusion was consistent with the results previous [30]. In the Sp2 stage, the young spikelet development, 1938 DEGs (495 up- and 1443 down-regulated) were identified in the Sp2-1 vs. Sp2-ck comparison (Figure 3B); while 1232 DEGs (175 up- and 1057 down-regulated), were identified in the Sp2-2 vs. Sp2-ck comparison. Furthermore, in the Sp3 stage, a total of 6294 and 4330 genes were differentially expressed in 04-06/10-06, 14-06/10-06 compared with 04-06/14-06, respectively. Among them, 3382 and 2296 genes were up-regulated, while 2912 and 2034 genes were down-regulated respectively at this stage (Figure 7A,B).

In general, the number of DEGs in Sp2-1/Sp2-ck and Sp2-2/Sp2-ck was 804 while it in S3-1/Sp3-ck and Sp3-2/Sp3-ck was 3212. Subsequently, to explore the metabolic regulatory network of rice hybrid sterility, KEGG enrichment analysis of these 804 DEGs and 3212 DEGs showed that these DEGs were enriched in these pathways relative to the ck, such as metabolic pathways, global and overview maps pathway, secondary metabolite biosynthesis pathway, phenylpropanoid biosynthesis synthesis, photosynthesis, TGF-beta signaling pathway, starch and sucrose metabolism and flavonoids biosynthesis (Figure S15). Interestingly, the secondary metabolism processes are important in plant stress response including phenylpropanoids and flavonoids metabolism [54,55], suggesting that the *ORF* interaction at the *S5* locus might affect the metabolism and the activities in two hybrids involved in reducing fertility.

In particular, some genes related to fertility were selected from these DEGs. Among them, the expression of genes related to ER stress was significantly up-regulated in 04-06/10-06 and 14-06/10-06 compared with the control 04-06/14-06, such as *OsHSP70*, and *OsBip5*. However, there was no significant difference in PCD-related genes except *OsAP37* [56]. The genes related to oocyte meiosis and phenylpropanoid biosynthesis were also significantly different from the control, with some genes significantly up-regulated in expression at the Sp2 stage and significantly down-regulated at the Sp3 stage (Figure 7C, Table S7). These results indicated that *ORF* interaction at the *S5* locus induced the ER stress but not led to premature PCD, and it might produce some phenylpropanoids and flavonoids to response the stress and up-regulated oocyte meiosis genes expression to induce embryo-sac-producing abortive female gametes.

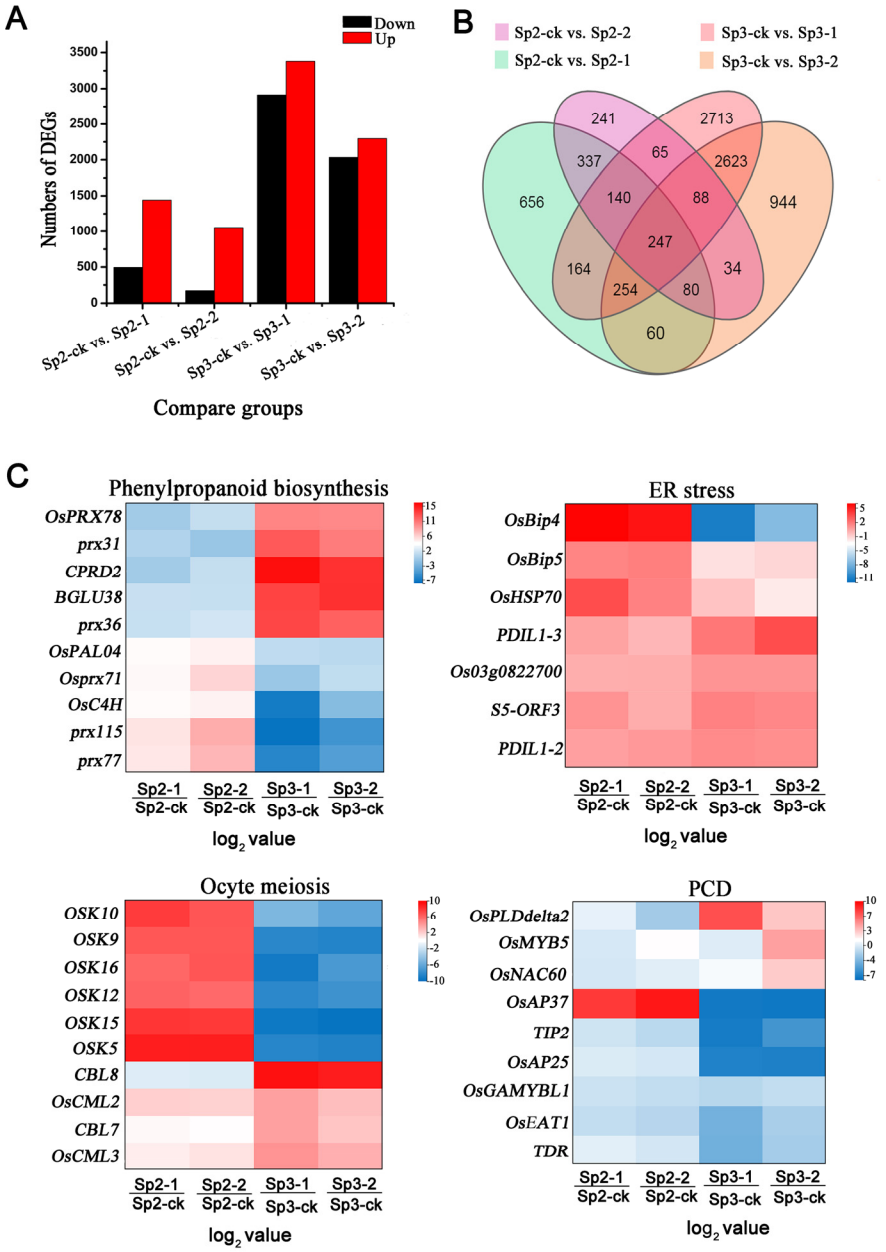

Figure 7. Transcriptome analysis of the panicles of WT (fertility) and F_1 hybrids (sterility) (**A**) Number of the up- and down-regulated DEGs at Sp2 and Sp3 stages. (**B**) The Venn diagram shows the numbers of shared and unique DEGs. (**C**) The ratio of the expression of genes selected in the DEGs. The expression level based on log2 (FPKM + 0.01) is represented by a heat map. Sp2-ck, Sp2-1 and Sp2-2 indicated the florets in length ranges of 10.2–11.5 mm of the hybrid F_1 from 04-06/14-06, 04-06/10-06, and 14-06/10-06 respectively. Sp3-ck, Sp3-1, and Sp3-2 indicated the florets one day before heading off the hybrid F_1 from 04-06/14-06, 04-06/10-06, and 14-06/10-06 respectively.

3.10. Analysis of Genes Differentially Expressed in Hybrids F_1 (ORF4+/ORF4+) Compare to Hybrids F_1 (ORF4+/ORF4−)

The *ORF4* gene at the *S5* locus encodes a transmembrane structural domain protein [29]. The hybrid F_1 spikelet fertility from the cross of 14-06/10-06 was higher than it of 04-06/10-06. Generally speaking, these two combinations had the same background. To further understand the molecular mechanism of hybrid sterility, we performed a transcriptomic analysis of these two hybrids. Only 6, 107, and 984 DEGs (2, 48, and 543 up-regulated and 4, 59, and 441 down-regulated) were identified in the 04-06/10-06 vs. 14-06/10-06 in

three different stages (Figure 8A). The GO standardized classification system was used to analyze these DEGs and classify 1097 DEGs into three categories: biological process, molecular function, and cellular components. Under molecular function, cellular process, binding and catalytic activity were the largest subcategories. The membrane and membrane part were two of the predominant subcategories in the cellular component category. Major subcategories under the biological process were the metabolic process, response to stimulus, and biological regulation (Figure 8B).

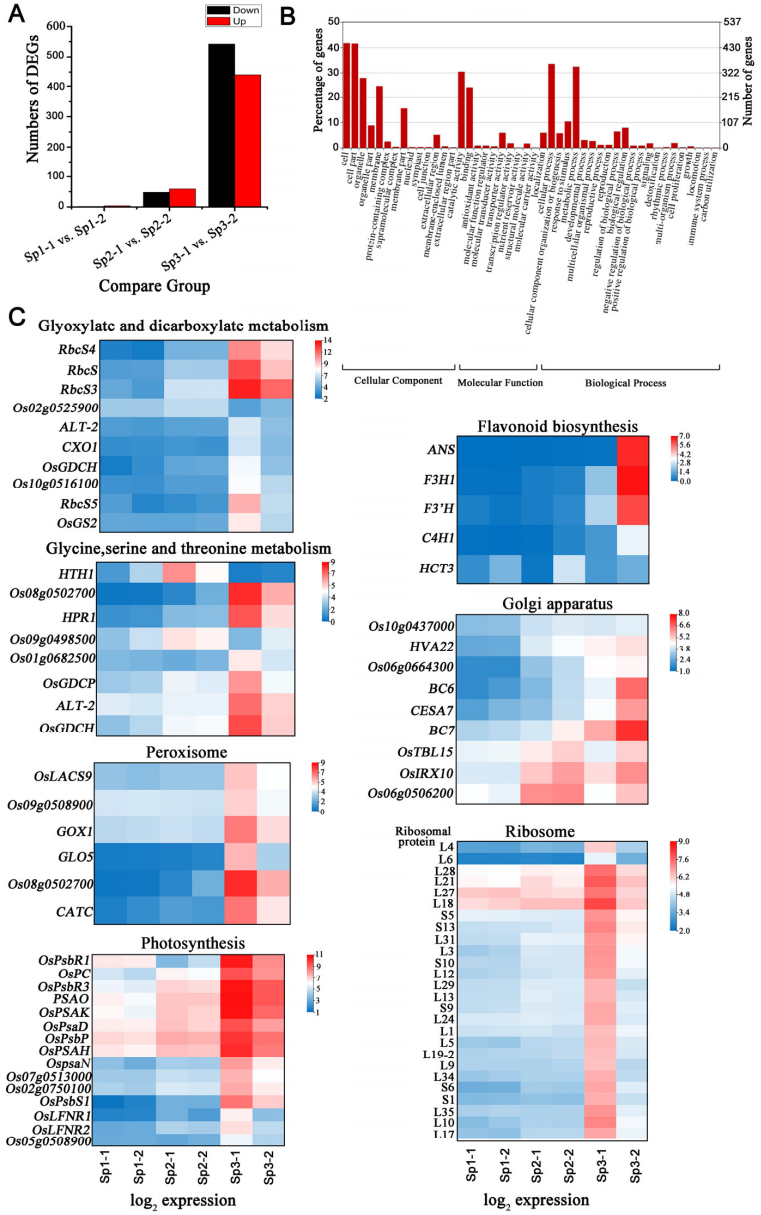

Figure 8. Analysis of genes differentially expressed in hybrids F_1 (*ORF4+*/*ORF4+*) compared with hybrids F_1 (*ORF4+*/*ORF4−*). (**A**) Number of the up- and down-regulated DEGs of hybrid F_1 in different genotypes of *ORF4*. (**B**) The significant GO annotations of DEGs during three spikelet development stages. (**C**) Heat map of expression change of genes selected in the DEGs. The color in each frame represents the level of expression change based on the $\log_2$ fold. Sp1-1 and Sp1-2 indicated the florets in length ranges of 3.0–4.5 mm of the hybrid F_1 from 04-06/10-06 and 14-06/10-06 respectively; Sp2-1 and Sp2-2 indicated the florets in length ranges of 10.2–11.5 mm of the hybrid F_1 from 04-06/10-06 and 14-06/10-06 respectively; Sp3-1 and Sp3-2 indicated the florets one day before heading of the hybrid F_1 from 04-06/10-06 and 14-06/10-06 respectively.

Later on, we analyzed these differentially expressed DEGs in the three periods (Figure S16). The genes related to glyoxylate and dicarboxylic acid metabolism, such as *RBCS* multigene family were significantly higher expressed compared with the hybrid F_1 from 14-06/10-06 combination in Sp3 stage, leading to an accumulation of Rubisco holoenzyme [57,58]. Some genes related to glycines, serine and threonine metabolisms, and photosynthesis, such as *Os-GDCH*, *OsLFNR1*, and the genes encoded by Ribosome protein were remarkable up-regulation (Figure 8C). In addition, comparing with the hybrid F_1 from 14-06/10-06 combination, there were several noteworthy down-regulated genes such as *F3H*, flavanone 3-hydroxylase that is involved in flavonoid biosynthesis to response to stress [59], and *ANS*, important structural gene to response to abiotic stress [60]. *ORF4+* localized to the plasma membrane and Golgi. The expression of *CESA7* and *BC7*, encoding cellulose synthase catalytic subunit [61] were significantly decreased in the hybrid F_1 from 04-06/10-06 combination. These genes may play crucial roles in hybrid sterility caused by *ORF4+*/*ORF4−* and *ORF5+*.

4. Discussion

Hybrid sterility is the major obstacle to the utilization of the strong vigor in *indica/japonica* hybrids. *S5* was identified as the major locus contributing to inter-subspecific hybrid spikelet sterility [62–64]. It was a complex gene cluster, composed of three adjacent genes, *ORF3*, *ORF4*, and *ORF5* [29]. Here, we explored the functions of three *ORF*s and analyzed the genetic effect of different *ORF* combinations on hybrid sterility via knocking out of these *ORF*s with CRISPR/Cas9-based genome editing or by transcriptomic analysis.

The *S5* locus conformed to the one-locus sporo-gametophytic interaction model [65]. In the hybrids, the *S5*-n was regarded as a wide-compatibility allele and does not produce female sterility in combination with *S5*-i or *S5*-j alleles in the triallelic system [19] and further studies [19,25,27,28,65]. A killer-protector system was put forward, and it showed *ORF5+* together with *ORF4+* could selectively eliminate the female gametes without protector *ORF3+* [29]. Previous studies showed that four different alleles containing *ORF5n* were all identified in *Oryza sativa* L. [31,36,66]. Three common alleles of them were used to cross with *S5*-i and *S5*-j alleles in this paper, and the higher fertility of F_1 hybrids and no segregation distortion in the F_2 population indicated that the presence of *ORF5n* has no effect on embryo sac sterility regardless of the alleles of *ORF3* and *ORF4*. For instance, the hybrid with *ORF3+ORF4+ORF5+*/*ORF3−ORF4−ORF5n* showed higher fertility than it with *ORF3+ORF4+ORF5+*/*ORF3−ORF4+ORF5−*. The *ORF5+*, *ORF4+*, and *ORF3+* all existed in these two hybrids, but they showed different fertility, which was not completely consistent with previous studies [29]. So, according to the one-locus sporo-gametophytic interaction model, *ORF3+ORF4−ORF5n*, *ORF3+ORF4+ORF5n*, and *ORF3−*/*ORF4−*/*ORF5n* might be the *S5*-n as previous [17], and they all could be regarded as wide-compatibility alleles. The superior wide compatible varieties were useful in obtaining an F_1 hybrid with high fertility crossed with *indica* or *japonica* varieties [67,68]. Therefore, the germplasm carrying *ORF5n* would be an important resource to break the reproductive barrier between *indica* and *japonica* subspecies in rice breeding.

The HSP70s in plants were encoded by multiple gene families. The *ORF3+* encoded a heat shock protein and localized to the ER [29], and most closely resembles the BiP (the luminal binding protein) protein. Five *BiP* genes were previously identified and the expression of all these genes could be up-regulated by ER stress [52,69,70]. The *ORF3+* was regarded as a protector because it may help to fold the signal proteins to suppress the *ORF5+* and *ORF4+* induced ER stress and PCD [29,30]. Here, we provided some evidence supporting this point. We created new alleles of *ORF3* with CRISPR/Cas9-based genome editing technology. *ORF3+* and *ORF3−*, *ORF3*(*−*), and *ORF3*($-^{\beta}$) had several amino acids difference at the C terminal, but showed two different functions on hybrid sterility. without *ORF3+*, the hybrid in genotype *ORF3*($-^{\beta}$)*ORF4+ORF5+*/*ORF3*($-^{\beta}$)*ORF4+ORF5+*, *ORF3*($-^{\beta}$)*ORF4+ORF5+*/*ORF3−ORF4+ORF5−*, *ORF3*($-^{\beta}$)*ORF4+ORF5+*/*ORF3−ORF4*(*−*)*ORF5−*, and *ORF3*(*−*)*ORF4*(*−*)*ORF5+*/*ORF3−ORF4+ORF5−* showed normal pollen fertility but lowest spikelet fertility. These results confirmed the function of *ORF3* in the killer-protector system [29].

Furthermore, we found that all the *ORF3*-knockout mutants in the *indica* background showed dwarf, sterility, and low yield. However, the *japonica* varieties grow normally in the genotype *ORF3*−*ORF4*+*ORF5*−/*ORF3*−*ORF4*+*ORF5*−. These results indicated there may exist a gene in the *indica* background that interacts with the *ORF3*+ gene, and knocking out *ORF3*+ might cut off the relationship and affect plant growth and development. *ORF3* was an important gene that not only could protect the gametes to produce hybrid offspring with normal fertility, but also protect plant growth normally in *indica* background.

The *ORF4*+ was localized to the plasma membrane and Golgi, and it could help *ORF5*+ selectively eliminate the female gametes [29]. Typical *japonica* varieties carry the *ORF3*−/*ORF4*+/*ORF5*− haplotype, while typical *indica* varieties carry the *ORF3*+/*ORF4*−/*ORF5*+ haplotype. Here, we generated a new hybrid-compatible allele by knocking out the *ORF4* in the *japonica* genotype. The hybrid in *ORF3*+*ORF4*−*ORF5*+/*ORF3*−*ORF4*(−)*ORF5*− genotype showed normal spikelet fertility from CRISPR/Cas9 knockout mutants crossed with typical *indica* varieties, showing *ORF4*(−) did break down the reproductive barrier induced by *S5* in the hybrids. Besides, the hybrid F_1 in the genotypes of *ORF3*+*ORF4*+*ORF5*+/*ORF3*−*ORF4*+*ORF5*− showed higher fertility than it in *ORF3*+*ORF4*−*ORF5*+/*ORF3*−*ORF4*+*ORF5*− genotype, which was only different in the interaction of *ORF4*+/*ORF4*−. The comparative RNA-Seq analysis of young spikelets was then conducted. Compared with the hybrid F_1 from 14-06/10-06 combination, *F3H*, and *ANS* were dramatically down-regulated which were important in Flavonoid biosynthesis to respond to abiotic stress [59,60]. Meanwhile, the expression of *CESA7* and *BC7* encoded cellulose synthase catalytic subunit was also significantly decreasing in the hybrid F_1 from 04-06/10-06 combination, which was related to the formation or function of the Golgi apparatus [61].

The three *ORFs* of the *S5* locus together regulated inter-specific hybrid sterility in rice. In the *indica*/*japonica* hybrid, *ORF3*+ may help the folding of the signal proteins to suppress the ER stress and PCD induced by *ORF5*+ and *ORF4*+, and *ORF4*+ may be a receptor to receive a signal from *ORF5*+. Without *ORF4*+ in both parents would not induce female gamete abortion. If *ORF4*+ and *ORF4*− both existed, the genes related to the function of the Golgi apparatus were dramatically down-regulated and may affect the Golgi apparatus's reprocessing of signal proteins from ER, inducing the lower fertility of hybrid in *ORF4*+/*ORF4*− combinations.

The CRISPR/Cas9 system was a powerful and highly efficient genome editing tool [37–39]. It could rapidly generate new neutral alleles at the hybrid sterility loci. Shen et al. [16] used CRISPR/Cas9 system to knock out one or two of the three *Sc*-i copies and rescued the expression level to overcome *Sc*-induced hybrid sterility. Xie et al. [40] created neutral *Sa* alleles by knocking out the *SaF* and *SaM* alleles by CRISPR/Cas9 system, which could overcome the male sterility caused by the *Sa* gene in hybrids. In this paper, TISL-Dbc-Gde was the pyramiding line in the *japonica* background. Knocking out *ORF4* in TISL-Dbc-Gde showed wide compatibility and could produce high pollen fertility and spikelet fertility in their F_1 hybrids with *indica* rice.

In conclusion, our study analyzed the function of three *ORF*s of *S5*. The allele carrying *ORF5n* showed wide compatibility and could produce high fertility in their F1 hybrids with both *indica* and *japonica* alleles. The *ORF3*+ was an important gene, and knockout in *indica* background could affect rice normal development. The allele interaction of *ORF4* could regulate hybrid fertility by affecting the expressions of genes related to the function of the Golgi apparatus. Finally, we provide effective approaches to overcome *indica*/*japonica* hybrid sterility by knocking out the *ORF4*+ in the *japonica* genotype or using the alleles carrying *ORF5n*. These results of the present study would have laid a good foundation for the utilization of inter-subspecific heterosis in hybrid rice breeding.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy13041094/s1, Figure S1: Alignment of the protein sequences encoded by *ORF3* in 35 varieties; Figure S2: Alignment of the protein sequences encoded by *ORF4* in 35 varieties; Figure S3: Alignment of the protein sequences encoded by *ORF5* in 35 varieties; Figure S4: The chromatin accessibility of the three genes during different spikelet development periods via ATAC-seq; Figure S5: Structure of *ORF3*+, *ORF3*−, *ORF4*+ and *ORF4*−; Figure S6: Sequence alignment of the sgRNA target region amplified by marker *ORF4*−G in two different *ORF4* gene edited lines compared

with the HJX74 SSSL 10-06; Figure S7: Phenotype of HJX74-SSSL 10-06 as control and two ORF4 gene edited mutants E403-9 and E403-12 in T_2 generation, and *ORF4* over-expression mutants E 343 and E342; Figure S8: Sequence alignment of the sgRNA target region amplified by marker *ORF3*-G in six different *ORF3-ORF4* genes edited lines; Figure S9: Sequence alignment of the sgRNA target region amplified by marker *ORF4*-G in six different *ORF3-ORF4* genes edited lines; Figure S10: Phenotype of HJX74-SSSL 14-06 as control and six *ORF3-ORF4* genes edited mutants in T_2 generation; Figure S11: Sequence alignment of the sgRNA target region amplified by marker *ORF4*-G in three different *ORF4* gene edited lines compared with the TISL-Dbc-Gde; Figure S12: Phenotype of TISL-Dbc-Gde as control and three *ORF4* genes edited mutants in T_2 generation; Figure S13: Transcriptome analysis of the panicles of the hybrids from 04-06/10-06 and 14-06/10-06 (sterility) vs. 04-06/14-06 (fertility) at meiosis stage; Figure S14: The comparisons of key genes expressions between qRT-PCR and RNA-Seq data; Figure S15: Differentially expressed genes enriched GO categories and KEGG pathway (the spikelet of hybrids F_1 from 04-06/10-06 and 14-06/10-06 vs. 04-06/14-06 in Sp2 and Sp3 stage); Figure S16: Differentially expressed genes enriched KEGG pathway (the spikelet of hybrids F_1 from 04-06/10-06 vs. 14-06/10-06 in three development stage); Table S1: Genotype of *ORF3*, *ORF4*, and *ORF5* at *S5* locus for 38 parents used in diallel crosses; Table S2: Pollen fertility and spikelet fertility in hybrids F_1 and the F_2 distorted segregation of the *S5* locus in a diallel set of crosses involving 38 parents (%±SD, n = 10); Table S3: The spikelet fertility in hybrids F_1 from the crosses of *indica*-SSSLs with *ORF4*+ overexpression mutations or *ORF4*+ edited mutations in the *japonica*-SSSL 10-06 background; Table S4: The spikelet fertility in hybrids F_1 from the crosses of gene-edited mutations in the SSSL 14-06 background with HJX74-SSSL 10-06 or *ORF4*-edited mutations in 10-06 background; Table S5: Sequencing data quality assessment and reads mapping with reference genome; Table S6: The differentially expressed genes in combination from 04-06/14-06 (fertility) and 04-06/10-06 or 14-06/10-06 (sterility) at meiosis stage; Table S7: The differentially expressed genes in combination from 04-06/14-06 (fertility) and 04-06/10-06 or 14-06/10-06 (sterility) at different period of immature inflorescences; Table S8: The differentially expressed genes in combination from 04-06/10-06 (low) and 14-06/10-06 (high) at different period of immature inflorescences; Table S9: Primer sequences used in this paper.

Author Contributions: Conceptualization and methodology, C.L., D.C., J.G. and G.Z.; validation, G.T., X.Z. and K.C.; formal analysis and investigation, C.Y., J.L. and K.S.; resources, S.W. and G.Z.; data curation, C.L., X.Z. and D.C.; writing—original draft preparation, J.G.; writing—review and editing, C.L., D.C. and Y.C.; supervision, Y.C. funding acquisition, J.G. and C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (grant number 31801449), the Natural Science Foundation of Guangdong Province, China (grant number 2018A030313867), the Special fund for scientific innovation strategy-construction of high-level Academy of Agriculture Science (grant number R2019PY-QF007), and the Guangdong Key Laboratory of New Technology in Rice Breeding (grant number 2020B1212060047).

Data Availability Statement: The raw data of RNA-seq reads were deposited in the NCBI database under the accession number PRJNA912986.

Acknowledgments: We thank BioRun biotechnology Inc. Ltd. (Wuhan, China) for performing the transgenic experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khush, G.S.; Gupta, P. Strategies for increasing the yield potential of cereals: Case of rice as an example. *Plant Breed.* **2013**, *132*, 433–436. [CrossRef]
2. Garris, A.J.; Tai, T.H.; Coburn, J.; Kresovich, S.; McCouch, S. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **2005**, *169*, 1631–1638. [CrossRef]
3. Wang, X.; Guan, Z.; Gong, Z.; Yan, J.; Yang, G.; Liu, Y.G.; Yin, P. Crystal structure of *WA352* provides insight into cytoplasmic male sterility in rice. *Biochem. Biophys. Res. Commun* **2018**, *501*, 898–904. [CrossRef] [PubMed]
4. Yang, G.P.; Maroof, M.A.; Xu, C.G.; Zhang, Q.; Biyashev, R.M. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. *Mol. Gen. Genet.* **1994**, *245*, 187–194. [CrossRef] [PubMed]
5. Ying, Y.G.; Wang, B.X.; Chen, T.M.; Yang, B.; Liu, Y.; Xu, B.; Sun, Z.G.; Chi, M.; Li, J.; Li, J.M.; et al. Heterosis analysis of yield and quality characters of two-line *indica-japonica* hybrid combination. *Jiangsu Agric. Sci.* **2021**, *49*, 59–63.
6. Cheng, S.H.; Zhuang, J.Y.; Fan, Y.Y.; Du, J.H.; Cao, L.Y. Progress in research and development on hybrid rice: A super-domesticate in China. *Ann. Bot.* **2007**, *100*, 959–966. [CrossRef] [PubMed]

7. Zhang, G.Q. Prospects of utilization of inter-subspecific heterosis between *indica* and *japonica* rice. *J. Integr. Agric.* **2020**, *19*, 1–10. [CrossRef]
8. Zhang, G.Q. The Next Generation of Rice: Inter-subspecific *indica-japonica* hybrid rice. *Front. Plant Sci.* **2022**, *13*, 857896. [CrossRef]
9. Kato, A. On the affinity of rice varieties as shown by the fertility of rice plants. *Centr. Agric. Inst. Kyushu Imp. Univ.* **1928**, *2*, 241–276.
10. Ikehashi, H.; Araki, H. Varietal screening of compatibility types revealed in F_1 fertility of distant crosses in rice. *Jpn. J. Breed.* **1984**, *34*, 304–313. [CrossRef]
11. Ouyang, Y.D.; Zhang, Q.F. The molecular and evolutionary basis of reproductive isolation in plants. *J. Genet. Genom.* **2018**, *45*, 613–620. [CrossRef]
12. Ouyang, Y.D.; Liu, Y.G.; Zhang, Q.F. Hybrid sterility in plant: Stories from rice. *Curr. Opin. Plant Biol.* **2010**, *13*, 186–192. [CrossRef] [PubMed]
13. Ouyang, Y.; Chen, J.J.; Ding, J.H.; Zhang, Q.F. Advances in the understanding of inter-subspecific hybrid sterility and wide-compatibility in rice. *Chin. Sci. Bull.* **2009**, *54*, 2332–2341. [CrossRef]
14. Li, J.; Zhou, J.; Zhang, Y.; Yang, Y.; Tao, D. New insights into the nature of interspecific hybrid sterility in rice. *Front. Plant Sci.* **2020**, *11*, 555572. [CrossRef] [PubMed]
15. Zhou, P.; Wang, Z.; Zhu, X.; Tang, Y.; Ye, L.; Yu, H.; Li, Y.; Zhang, N.; Liu, T.; Wang, T.; et al. A minimal genome design to maximally guarantee fertile inter-subspecific hybrid rice. *Mol. Plant* **2023**, *16*, 726–738. [CrossRef]
16. Shen, R.; Wang, L.; Liu, X.; Wu, J.; Jin, W.; Zhao, X.; Xie, X.; Zhu, Q.; Tang, H.; Li, Q.; et al. Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat. Commun.* **2017**, *8*, 1310. [CrossRef]
17. Mizuta, Y.; Harushima, Y.; Kurata, N. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20417–20422. [CrossRef]
18. Long, Y.; Zhao, L.; Niu, B.; Su, J.; Wu, H.; Chen, Y.; Zhang, Q.; Guo, J.; Zhuang, C.; Mei, M.; et al. Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18871–18876. [CrossRef]
19. Chen, J.; Ding, J.; Ouyang, Y.; Du, H.; Yang, J.; Cheng, K.; Zhao, J.; Qiu, S.; Zhang, X.; Yao, J.; et al. 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. USA* **2008**, *105*, 11436–11441. [CrossRef]
20. Yu, Y.; Zhao, Z.; Shi, Y.; Tian, H.; Liu, L.; Bian, X.; Xu, Y.; Zheng, X.; Gan, L.; Shen, Y.; et al. Hybrid sterility in rice (*Oryza sativa* L.) involves the tetratricopeptide repeat domain containing protein. *Genetics* **2016**, *203*, 1439–1451. [CrossRef]
21. Kubo, T.; Takashi, T.; Ashikari, M.; Yoshimura, A.; Kurata, N. Two tightly linked genes at the *hsa1* locus cause both F_1 and F_2 hybrid sterility in rice. *Mol. Plant* **2016**, *9*, 221–232. [CrossRef] [PubMed]
22. Kitamura, E. Genetic studies on sterility observed n hybrids between distantly related varieties of rice, *Oryza sativa* L. *Rep. Chugoku Exp. Stat.* **1962**, *8*, 141–205.
23. Oka, H.I. Genic analysis for the sterility of hybrids between distantly related varieties of cultivated rice. *J. Genet.* **1957**, *55*, 397–409. [CrossRef]
24. Oka, H.I. Analysis of genes controlling F_1 sterility in rice by the use of isogenic lines. *Genetics* **1974**, *77*, 521–534. [CrossRef]
25. Wei, C.M.; Wang, L.; Yang, Y.X.; Chen, Z.X.; Shahid, M.Q.; Li, J.Q.; Liu, X.D.; Lu, Y.G. Identification of an *S5*n allele in *Oryza rufipogon* Griff. and its effect on embryo sac fertility. *Chin. Sci. Bull.* **2010**, *55*, 1255–1262. [CrossRef]
26. Du, H.; Ouyang, Y.; Zhang, C.; Zhang, Q. Complex evolution of *S5*, a major reproductive barrier regulator, in the cultivated rice *Oryza sativa* and its wild relatives. *New Phytol.* **2011**, *191*, 275–287. [CrossRef]
27. Tong, J.F.; Li, Y.H.; Yang, Y.X.; Shahid, M.Q.; Lu, Y.G. Molecular evolution of rice *S5*n and functional comparison among different sequences. *Chin. Sci. Bull.* **2011**, *56*, 2016–2024. [CrossRef]
28. Peng, H.; Shahid, M.Q.; Li, Y.H.; Tong, J.F.; Zhou, X.W.; Liu, X.D.; Lu, Y.G. Molecular evolution of *S5* locus and large differences in its coding region revealed insignificant effect on *indica/japonica* embryo sac fertility. *Plant Syst. Evol.* **2015**, *301*, 639–655. [CrossRef]
29. Yang, J.; Zhao, X.; Cheng, K.; Du, H.; Ouyang, Y.; Chen, J.; Qiu, S.; Huang, J.; Jiang, Y.; Jiang, L.; et al. A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* **2012**, *337*, 1336–1340. [CrossRef] [PubMed]
30. Zhu, Y.F.; Yu, Y.M.; Cheng, K.; Ouyang, Y.D.; Wang, J.; Gong, L.; Zhang, Q.H.; Li, X.H.; Xiao, J.H.; Zhang, Q. Processes underlying a reproductive barrier in *indica-japonica* rice hybrids revealed by transcriptome analysis. *Plant Physiol.* **2017**, *174*, 1683–1696. [CrossRef] [PubMed]
31. Mi, J.; Li, G.; Xu, C.; Yang, J.; Yu, H.; Wang, G.; Li, X.; Xiao, J.; Song, H.; Zhang, Q.; et al. Artificial selection in domestication and breeding prevents speciation in rice. *Mol. Plant* **2020**, *13*, 650–657. [CrossRef]
32. Mi, J.; Li, G.; Huang, J.; Yu, H.; Zhou, F.; Zhang, Q.; Ouyang, Y.; Mou, T. Stacking *S5*-n and *f5*-n to overcome sterility in *indica-japonica* hybrid rice. *Theor. Appl. Genet.* **2016**, *129*, 563–575. [CrossRef] [PubMed]
33. Ma, X.; Wei, X.; Gao, G.; Jiang, H. Development and evaluation of improved lines based on an elite rice variety 9311 for overcoming hybrid sterility in rice. *Mol. Breed.* **2020**, *40*, 102. [CrossRef]
34. Guo, J.; Li, Y.; Xiong, L.; Yan, T.; Zou, J.; Dai, Z.; Tang, G.; Sun, K.; Luan, X.; Yang, W.; et al. Development of wide-compatible *indica* lines by pyramiding multiple neutral alleles of *indica-japonica* hybrid sterility loci. *Front. Plant Sci.* **2022**, *13*, 890568. [CrossRef] [PubMed]

35. Chen, L.; Zhao, Z.; Liu, X.; Liu, L.; Jiang, L.; Liu, S.; Zhang, W.; Wang, Y.; Liu, Y.; Wan, J. Marker-assisted breeding of a photoperiod-sensitive male sterile *japonica* rice with high cross-compatibility with *indica* rice. *Mol. Breed.* **2011**, *27*, 247–258. [CrossRef]
36. Guo, J.; Xu, X.; Li, W.; Zhu, W.; Zhu, H.; Liu, Z.; Luan, X.; Dai, Z.; Liu, G.; Zhang, Z.; et al. Overcoming inter-subspecific hybrid sterility in rice by developing *indica*-compatible *japonica* lines. *Sci. Rep.* **2016**, *6*, 26878. [CrossRef]
37. Afzal, S.; Sirohi, P.; Singh, N.K. A review of CRISPR associated genome engineering: Application, advances and future prospects of genome targeting tool for crop improvement. *Biotechnol. Lett.* **2020**, *42*, 1611–1632. [CrossRef]
38. Gao, C. Genome engineering for crop improvement and future agriculture. *Cell* **2021**, *184*, 1621–1635. [CrossRef] [PubMed]
39. Zhang, R.X.; Li, B.B.; Yang, Z.G.; Huang, J.Q.; Sun, W.H.; Bhanbhro, N.; Liu, W.T.; Chen, K.M. Dissecting plant gene functions using CRISPR toolsets for crop improvement. *J. Agric. Food Chem.* **2022**, *70*, 7343–7359. [CrossRef] [PubMed]
40. Xie, Y.; Niu, B.; Long, Y.; Li, G.; Tang, J.; Zhang, Y.; Ren, D.; Liu, Y.G.; Chen, L. Suppression or knockout of *SaF/SaM* overcomes the *Sa*-mediated hybrid male sterility in rice. *J. Integr. Plant Biol.* **2017**, *59*, 669–679. [CrossRef]
41. Hiei, Y.; Komari, T.; Kubo, T. Transformation of rice mediated by Agrobacterium tumefaciens. *Plant Mol Biol* **1997**, *35*, 205–218. [CrossRef] [PubMed]
42. Mao, Y.F.; Zhang, H.; Xu, N.F.; Zhang, B.T.; Gou, F.; Zhu, J.K. Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant* **2013**, *6*, 2008–2011. [CrossRef] [PubMed]
43. Ma, X.; Zhang, Q.; Zhu, Q.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.; Li, H.; Lin, Y.; et al. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* **2015**, *8*, 1274–1284. [CrossRef] [PubMed]
44. Zheng, K.; Ning, H.; Bennett, J.; Khush, G.S. PCR-based marker-assisted selection in rice breeding. *IRRI Discuss. Pap.* **1995**, *12*.
45. Panaud, O.; Chen, X.; McCouch, S.R. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* **1996**, *252*, 597–607. [CrossRef] [PubMed]
46. Li, R.Q.; Li, Y.R.; Kristiansen, K.; Wang, J. SOAP: Short oligonucleotide alignment program. *Bioinformatics* **2008**, *24*, 713–714. [CrossRef]
47. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef]
48. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]
49. Abdi, H. The Bonferonni and Šidák Corrections for Multiple Comparisons. *Encycl. Meas. Stat.* **2007**, *3*, 2007.
50. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}$Ct Method. *Methods* **2001**, *25*, 402–408. [CrossRef]
51. Ji, Q.; Zhang, M.; Lu, J.; Wang, H.; Lin, B.; Liu, Q.; Chao, Q.; Zhang, Y.; Liu, C.; Gu, M.; et al. Molecular basis underlying the S5-dependent reproductive isolation and compatibility of *indica/japonica* rice hybrids. *Plant Physiol.* **2012**, *158*, 1319–1328. [CrossRef] [PubMed]
52. Hayashi, S.; Wakasa, Y.; Takahashi, H.; Kawakatsu, T.; Takaiwa, F. Signal transduction by IRE1-mediated splicing of *bZIP50* and other stress sensors in the endoplasmic reticulum stress response of rice. *Plant J.* **2012**, *69*, 946–956. [CrossRef] [PubMed]
53. Niu, N.; Liang, W.; Yang, X.; Jin, W.; Wilson, Z.A.; Hu, J.; Zhang, D. *EAT1* promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. *Nat. Commun.* **2013**, *4*, 1445. [CrossRef] [PubMed]
54. Dixon, R.A.; Paiva, N.L. Stress-induced phenylpropanoid metabolism. *Plant Cell* **1995**, *7*, 1085–1097. [CrossRef]
55. Winkel-Shirley, B. Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant Biol.* **2002**, *5*, 218–223. [CrossRef]
56. Mao, C.; Ding, J.; Zhang, B.; Xi, D.; Ming, F. *OsNAC2* positively affects salt-induced cell death and binds to the *OsAP37* and *OsCOX11* promoters. *Plant J.* **2018**, *94*, 454–468. [CrossRef]
57. Ogawa, S.; Suzuki, Y.; Yoshizawa, R.; Kanno, K.; Makino, A. Effect of individual suppression of *RBCS* multigene family on Rubisco contents in rice leaves. *Plant Cell Environ.* **2012**, *35*, 546–553. [CrossRef] [PubMed]
58. Suzuki, Y.; Makino, A. Availability of rubisco small subunit up-regulates the transcript levels of large subunit for stoichiometric assembly of its holoenzyme in rice. *Plant Physiol.* **2012**, *160*, 533–540. [CrossRef]
59. Chen, S.; Sun, B.; Shi, Z.; Miao, X.; Li, H. Identification of the rice genes and metabolites involved in dual resistance against brown planthopper and rice blast fungus. *Plant Cell Environ.* **2022**, *45*, 1914–1929. [CrossRef]
60. Reddy, I. Rice flavonoid pathway genes, *OsDfr* and *OsAns*, are induced by dehydration, high salt and ABA, and contain stress responsive promoter elements that interact with the transcription activator, *OsC1-MYB*. *Plant Sci.* **2004**, *166*, 1505–1513.
61. Zhang, B.; Deng, L.; Qian, Q.; Xiong, G.; Zeng, D.; Li, R.; Guo, L.; Li, J.; Zhou, Y. A missense mutation in the transmembrane domain of *CESA4* affects protein abundance in the plasma membrane and results in abnormal cell wall biosynthesis in rice. *Plant Mol. Biol.* **2009**, *71*, 509–524. [CrossRef] [PubMed]
62. Song, X.; Qiu, S.Q.; Xu, C.G.; Li, X.H.; Zhang, Q. Genetic dissection of embryo sac fertility, pollen fertility, and their contributions to spikelet fertility of intersubspecific hybrids in rice. *Theor. Appl. Genet.* **2005**, *110*, 205–211. [CrossRef] [PubMed]
63. Wang, G.W.; He, Y.Q.; Xu, C.G.; Zhang, Q. Identification and confirmation of three neutral alleles conferring wide compatibility in inter-subspecific hybrids of rice (*Oryza sativa* L.) using near-isogenic lines. *Theor. Appl. Genet.* **2005**, *111*, 702–710. [CrossRef] [PubMed]
64. Li, G.; Li, X.; Wang, Y.; Mi, J.; Xing, F.; Zhang, D.; Dong, Q.; Li, X.; Xiao, J.; Zhang, Q.; et al. Three representative inter and intra-subspecific crosses reveal the genetic architecture of reproductive isolation in rice. *Plant J.* **2017**, *92*, 349–362. [CrossRef]

65. Ikehashi, H.; Araki, H. Genetics of F_1 sterility in remote crosses of rice. In *Rice Genetics I*; IRRI: Manila, Philippines, 1986; pp. 119–130.
66. Ouyang, Y.; Li, G.; Mi, J.; Xu, C.; Du, H.; Zhang, C.; Xie, W.; Li, X.; Xiao, J.; Song, H.; et al. Origination and establishment of a trigenic reproductive isolation system in rice. *Mol. Plant* **2016**, *9*, 1542–1545. [CrossRef]
67. Kallugudi, J.; Singh, V.J.; Vinod, K.K.; Krishnan, S.G.; Nandakumar, S.; Dixit, B.K.; Ellur, R.K.; Bollinedi, H.; Nagarajan, M.; Kumar, A.; et al. Population dynamics of wide compatibility system and evaluation of intersubspecific hybrids by *indica-japonica* hybridization in rice. *Plants* **2022**, *11*, 1930. [CrossRef]
68. Priyadarshi, R.; Arremsetty, H.P.S.; Singh, A.K.; Khandekar, D.; Ulaganathan, K.; Shenoy, V.; Sinha, P.; Singh, V.K. Marker-assisted improvement of the elite maintainer line of rice, IR 58025B for wide compatibility (*S5*(n)) gene. *Front. Plant Sci.* **2018**, *9*, 1051. [CrossRef]
69. Oono, Y.; Wakasa, Y.; Hirose, S.; Yang, L.; Sakuta, C.; Takaiwa, F. Analysis of ER stress in developing rice endosperm accumulating beta-amyloid peptide. *Plant Biotechnol. J.* **2010**, *8*, 691–718. [CrossRef]
70. Ohta, M.; Wakasa, Y.; Takahashi, H.; Hayashi, S.; Kudo, K.; Takaiwa, F. Analysis of rice ER-resident J-proteins reveals diversity and functional differentiation of the ER-resident Hsp70 system in plants. *J. Exp. Bot.* **2013**, *64*, 5429–5441. [CrossRef]