



Article

# Genome Resequencing for Autotetraploid Rice and Its Closest Relatives Reveals Abundant Variation and High Potential in Rice Breeding

Yachun Zhang<sup>1,2,†</sup>, Anping Du<sup>1,†</sup>, Liqi Tong<sup>2</sup>, Gui Yan<sup>2</sup>, Longxiang Lu<sup>1,3</sup>, Yanni Yin<sup>2</sup>, Xingyue Fu<sup>1,3</sup>, Huixin Yang<sup>1,3</sup>, Hui Li<sup>1</sup>, Weizao Huang<sup>1</sup>, Detian Cai<sup>2</sup>, Zhaojian Song<sup>2</sup>, Xianhua Zhang<sup>2</sup>, Yuchi He<sup>2,\*</sup> and Shengbin Tu<sup>1,3,\*</sup>

- <sup>1</sup> Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610213, China; zhangyachun1225@163.com (Y.Z.); duap@cib.ac.cn (A.D.); llxdewyix@163.com (L.L.); fuxingyue1@163.com (X.F.); huixin98\_yang@163.com (H.Y.); lihui@cib.ac.cn (H.L.); huangwz@cib.ac.cn (W.H.)
- <sup>2</sup> State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan 430062, China; tong451027@163.com (L.T.); 202111107010086@stu.hubu.edu.cn (G.Y.); yinyannider@163.com (Y.Y.); dtcai8866@163.com (D.C.); zjsong99@126.com (Z.S.); 20150072@hubu.edu.cn (X.Z.)
- <sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China
- \* Correspondence: hyc@hubu.edu.cn (Y.H.); tusb@cib.ac.cn (S.T.)
- † These authors contributed equally to this work.

**Abstract:** Polyploid rice and its reverted diploid show rich phenotypic variation and strong heterosis, showing great breeding value. However, the genomic differences among tetraploids, counterpart common diploids, tetraploid-revertant diploids, and hybrid descendants are unclear. In this work, we bred a new excellent two-line hybrid rice variety, *Y Liang You Duo Hui 14* (HTRM12), using Haitian tetraploid self-reverted diploid (HTRM2). Furthermore, we comparatively analyzed the important agronomic traits and genome-wide variations of those closest relatives, Haitian diploid (HT2), Haitian tetraploid (HT4), HTRM2, and HTRM12 in detail, based on multiple phenotypic investigations, genome resequencing, and bioinformatics analysis. The results of agronomic traits analysis and genome-wide variation analysis of single nucleotide polymorphism (SNP), insertion-deletion (InDel), and copy number variation (CNV) show that HT4 and HTRM2 had abundant phenotypic and genomic variations compared to HT2. HTRM2 can inherit important traits and variations from HT4. This implies that tetraploid self-reverted diploid has high potential in creating excellent breeding materials and in breeding breakthrough hybrid rice varieties. Our study verifies the feasibility that polyploid rice could be used as a mutation carrier for creating variations and provides genomic information, new breeding materials, and a new way of application for tetraploid rice breeding.

**Keywords:** rice; tetraploid; tetraploid revertant diploid; genome resequencing; genome variation



**Citation:** Zhang, Y.; Du, A.; Tong, L.; Yan, G.; Lu, L.; Yin, Y.; Fu, X.; Yang, H.; Li, H.; Huang, W.; et al. Genome Resequencing for Autotetraploid Rice and Its Closest Relatives Reveals Abundant Variation and High Potential in Rice Breeding. *Int. J. Mol. Sci.* **2024**, *25*, 9012. <https://doi.org/10.3390/ijms25169012>

Academic Editor: Abir U.

Igamberdiev

Received: 5 July 2024

Revised: 8 August 2024

Accepted: 12 August 2024

Published: 19 August 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Rice, the main food source for half of the world's population, has a slow growth rate compared with corn. Increasing its yield has become a worldwide goal [1]. Twice "Green Revolutions" in rice breeding, dwarf breeding, and heterosis utilization have significantly increased rice yields [2]. Polyploidy plays an essential role in species domestication, and numerous studies have shown that 70% of angiosperms have had one or more polyploidy events [3,4]. The main cereals, cotton, and oil crops cultivated around the world, such as wheat, cotton, and oilseed rape, have evolved from diploids to allopolyploids [5,6]. From the perspective of systematic evolution, allopolyploids have generated heterozygosity between genomes, increased genome capacity, widened the range of genetic variation, and enhanced tolerance to adverse factors, resulting in a high and stable yield [7,8]. With the

combination of the heterologous genome and the occurrence of polyploidy, crop yield has increased significantly. Creatively, scientists have put forward the idea of using the dual advantages of distant hybridization and polyploidy to cultivate new varieties and improve yield in rice [9].

Compared with other crops, rice has the smallest genome, being 40 times smaller than that of wheat and 10 times smaller than that of barley. In addition, the number of chromosomes in rice is also very small [10,11]. From the theory of species ploidy level evolution and chromosome size suitability, it is possible to combine the advantages of polyploidy and genomic interactions to greatly improve (i.e., increase the yield by 50%) or even double the rice yield and obtain new rice varieties with high quality, high yield, and resistance to multiple stressors [9,12,13].

Autotetraploid rice, as a new germplasm, is generated by artificial chromosome duplication of diploid rice, which has attracted extensive attention in rice polyploid breeding in recent years [14]. Compared with diploid rice, autotetraploid rice has some characteristics of other polyploid plants, such as strong plant morphology, large grain size, high 1000-grain weight, increased nutrient content and stress resistance, and higher genetic diversity, which show great potential in rice breeding [9,14–16]. However, tetraploid rice generally has low fertility, which affects its yield and further limits the popularization and application [17–22]. Currently, some high-fertility tetraploid materials have been developed, such as high-fertility tetraploid rice restorer lines [23], polyploid meiosis stability lines (PMeS) [24], and neo-tetraploid [25–29], which led to an important breakthrough in the utilization of tetraploid rice. Furthermore, the important progress in the construction of three-line hybrid and two-line hybrid tetraploid rice systems for tetraploid heterosis utilization had been made by using traditional breeding and modern biotechnology breeding methods. Tu et al. explored the feasibility of the three-line hybrid tetraploid rice system and produced autotetraploid rice materials fitting in the cytoplasmic male sterility (CMS) system, including male sterile, maintainer, and restoring lines [23]. Afterwards, some important new tetraploid rice materials have been developed for two-line hybrid tetraploid rice systems, such as photoperiod- and thermo-sensitive genic male sterile lines (PTGMS) and thermo-sensitive male sterile lines (TGMS) [30,31].

Although the tetraploid rice heterosis utilization systems were explored and a number of new autotetraploid rice materials had been created, there is no report of commercial rice tetraploid variety developed directly from autotetraploid rice materials. In the genetic process of polyploidization, polyploid rice can revert to diploid rice by self-crossing, which may be caused by uneven chromosome distribution during the formation of male and female gametes in meiosis, and by the anther culture technique, creating rich available materials for breeding [32]. Research has shown that some excellent characteristics of such polyploids may be inherited by diploids. Using rice polyploid recovery to create excellent diploid breeding materials is of great practical significance in improving the genetic characteristics of existing crops and in creating new crop varieties with greater yield potential and better quality [33]. Excellent diploids have been selected from the ploidy-reverted progeny of autopolyploids by self-crossing from extensive test mating groups, and some hybrids have shown strong heterosis, indicating that this method can also provide excellent parent materials for the breeding of hybrid rice with high yield, high quality, and strong resistance [33]. Thus, this suggests a new way for the breeding and utilization of tetraploid rice.

At present, there are some studies on the genomic variation of diploids and tetraploids [21,22,27,28,34,35], but there are almost no studies on the genomic variation among diploids, tetraploids, tetraploid self-reverted diploids, and hybrid descendants of tetraploid self-reverted diploids. A series of analyses and the creation of excellent material for polyploid rice breeding, for example, creating superior revertant mutants using abundant genetic recombination during polyploid rice meiosis, can provide guidance for the selection of new hybrid varieties.

In this study, a new excellent two-line hybrid rice variety, *Y Liang You Duo Hui 14* (HTRM12), was bred using a systematically selected Haitian tetraploid self-reverted diploid (HTRM2) as the male parent. The hybrid variety recently obtained a national certificate. Furthermore, we comparatively analyzed the important agronomic traits and genome-wide variations of Haitian diploid (HT2), Haitian tetraploid (HT4), HTRM2, and HTRM12 in detail, based on multiple phenotypic investigations, genome resequencing technology, and bioinformatics analysis. Our results showed that HT4 and HTRM2 had abundant phenotypic and genomic variations compared to HT2. HTRM2 can inherit important traits and variations from HT4, thus the tetraploid self-reverted diploid has high potential in creating excellent breeding materials and in breeding breakthrough hybrid rice varieties. This study provides insights into parental utilization of tetraploid rice and also provides genomic information and new breeding materials for tetraploid rice breeding.

## 2. Results

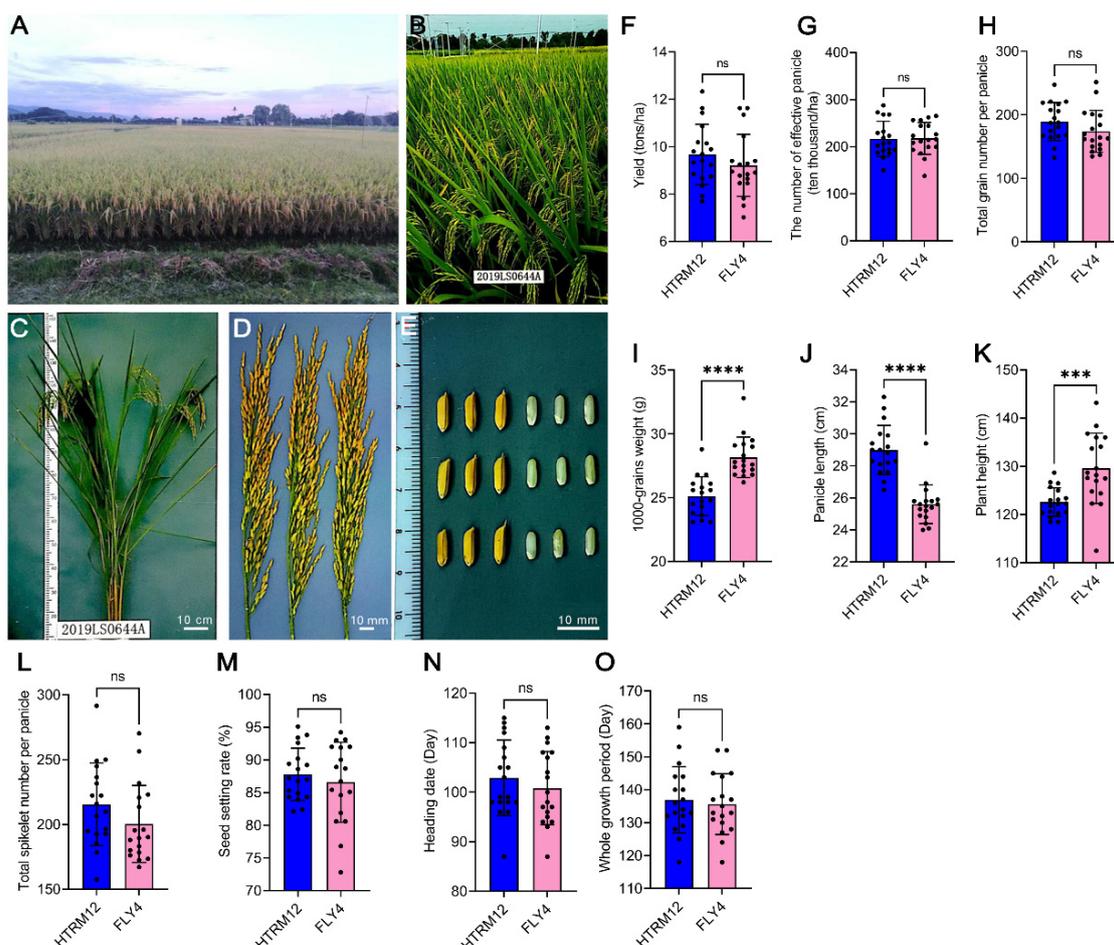
### 2.1. Breeding and Field Performance of Superior Hybrid Rice HTRM12

The variety HT2 is a superior restorer diploid line for three-line hybrid rice breeding. HT4 is an autotetraploid rice developed from HT2 by chromosome doubling with colchicine treatment, and HTRM2 is a self-reverted diploid developed from the self-fertile offspring of tetraploid HT4. We crossed excellent rice material *Y58S* and HTRM2 to breed the two-line hybrid rice variety HTRM12. Generally, HTRM12 performed better than the control variety *Feng Liang You 4 Hao* (FLY4) in field trials in the area of the lower and middle reaches of the Yangtze River. HTRM12 had outstanding performance for planting population, plant type, panicle type, and rice grain type (Figure 1A–E) and passed the national variety validation. Although HTRM12 showed a significantly smaller 1000-grain weight than the control variety FLY4 in the yield constitution factors, the yield of HTRM12 was 9.67 tons/ha, which was slightly greater than FLY4 (Figure 1F–I). HTRM12 showed a longer panicle and a significantly smaller plant height than the control variety FLY4 (Figure 1J,K), and its total spikelet number per panicle, seed setting rate, heading date, and whole growth period did not differ significantly from those of FLY4 at different trial sites (Figure 1L–O; Table S1). This variety had a balanced resistance to biotic and abiotic stress. Its comprehensive rice blast resistance index was 5.2, and it had moderate resistance to rice blast and bacterial blight but was highly susceptible to brown planthoppers based on the national standard test (Table 1). HTRM12 had a stronger tolerance to high temperatures than FLY4 at the pollination stage (Table S2). This variety had a normal rice quality according to national standards. Its head rice ratio reached 67.8%, and its lowest chalkiness degree was 5.3%. The amylose content of HTRM12 reached 19.5%, and its gel consistency reached 45 mm. The alkali spreading value of HTRM12 was 7-grade, and its smallest grain length-to-width ratio was 3.0 (Table S3).

**Table 1.** The disease and pest resistance test of hybrid variety HTRM12.

Variety	Susceptibility Index of Rice Blast	Susceptibility Index of Bacterial Leaf Blight	Susceptibility Index of Brown Planthopper
HTRM12	5.2	5	9
FLY4 (CK)	6.5	5	9

Note: The numbers of the index range from 0 to 9 grade. A smaller number means more resistance, and a larger number means more susceptible (for the test method, see Section 4).

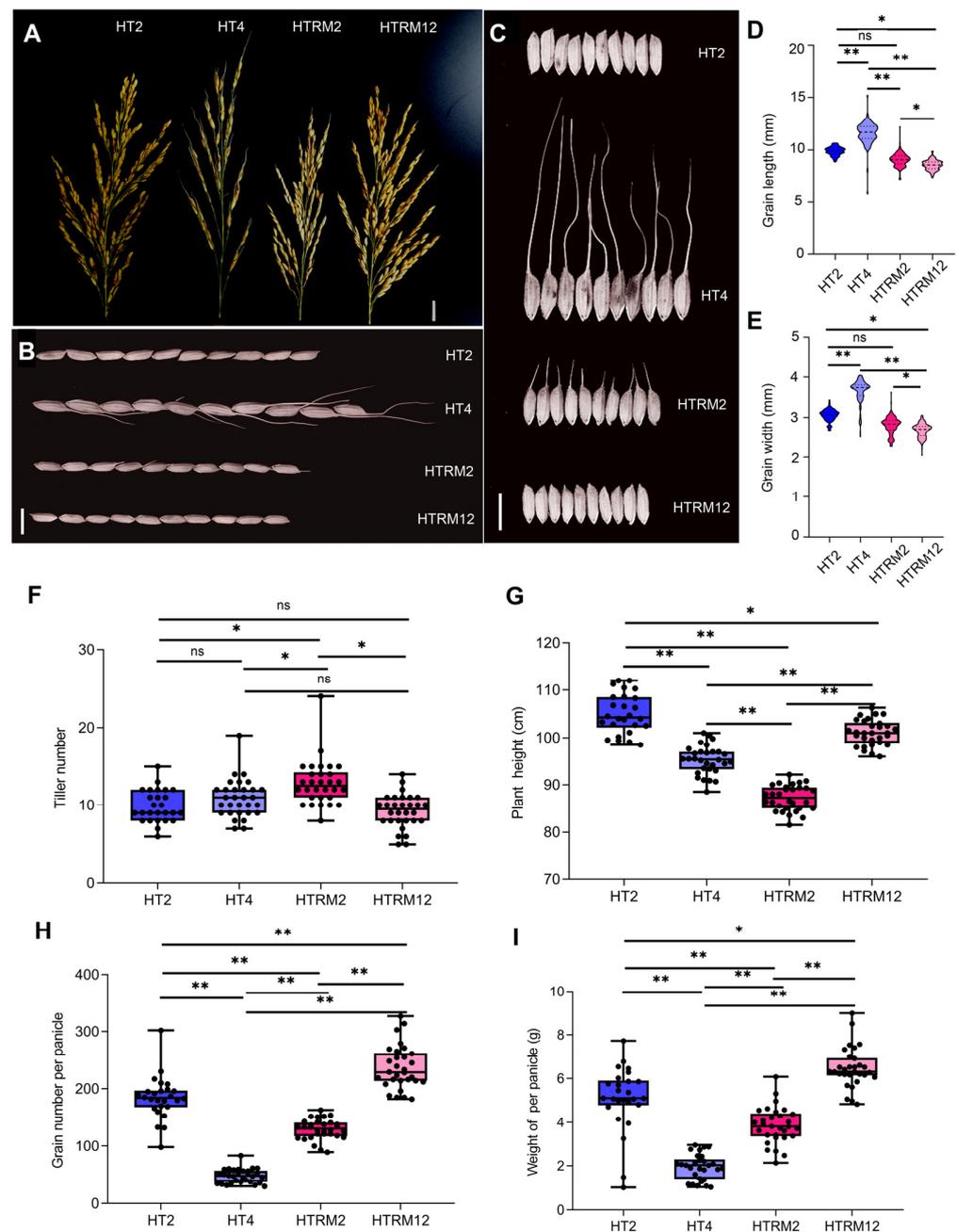


**Figure 1.** The field performance of hybrid rice HTRM12. (A,B) the field performance of HTRM12 population in production trial at maturity stage (A) and grain filling stage (B); (C) the plant phenotype of HTRM12 at grain filling stage; (D) the panicles phenotype of HTRM12 at grain filling stage; (E) the caryopses and brown grains phenotype of HTRM12 at maturity stage; (F–I) comparison of the yield and yield constitution factors between HTRM12 and control variety FLY4; (J) comparison of the panicle length between HTRM12 and control variety FLY4; (K) comparison of the plant height between HTRM12 and control variety FLY4; (L–O) comparison of total spikelet number per panicle, seed setting rate, heading date, and whole growth period between HTRM12 and control variety FLY4. Data collected from different regional trial sites were used for plotting, and values were shown as means  $\pm$  SD ( $n = 18$  trial sites). Student's  $t$ -test was used for difference significance analysis (two-tailed). Specially for (K), the Welch's  $t$ -test (two-tailed) was used for difference significance analysis. For (H–J), the Mann–Whiney U test (two-tailed) was used for difference significance analysis. ns presents no significant difference; \*\*\* presents  $p < 0.001$ ; \*\*\*\* presents  $p < 0.0001$ .

## 2.2. Comparison of Agronomic Traits of HT2, HT4, HTRM2, and HTRM12

To determine the differences in agronomic traits and genetic characteristics among HT2, HT4, HTRM2, and HTRM12, the four rice materials were planted together, and several important traits were investigated. Field observations showed that tetraploid HT4 had the fewest branches per panicle, and the number of branches was similar between HT2 and HTRM12. The number of branches in HTRM2 was less than that in HT2 and HTRM12 (Figure 2A). The grains of HT4 had longer awns, and HTRM2 inherited the awn trait from HT4 but had shorter awns (Figure 2B,C). The awn protects seeds from being eaten by animals and can improve the spread of seeds, making it an excellent protective trait [36]. The grain length and grain width of HT4 were significantly higher than those of the diploids, namely HT2, HTRM2, and HTRM12, and the length and width of HT2 and HTRM2 were

similar and significantly higher than those of HTRM12 (Figure 2B–E). The tiller number of HT2, HT4, and HTRM12 did not differ much, but the tiller number of HTRM2 was 13, on average, which was significantly higher than that of the other three materials (Figure 2F). However, the plant height of HTRM2 was significantly lower than that of HT4. The plant height of HTRM2 was 11.49% lower than that of HT4, and the plant height of HTRM12 was between that of HT2 and that of HT4, with a significant difference (Figure 2G). Plant height is closely related to photosynthetic efficiency and lodging resistance and directly affects the biomass of rice plants. Therefore, an appropriate reduction in plant height is beneficial for improving the lodging resistance of rice [37]. Although HTRM12 had smaller grains, the number of grains per panicle and the yield per panicle were significantly higher than those of the other three materials; the number of grains per panicle was 4.1 times higher than that of HT4, and the yield per panicle was 3.4 times that of HT4 (Figure 2H,I).



**Figure 2.** Agronomic traits of HT2, HT4, HTRM2, and HTRM12. (A) Comparison of spike shape of each sample. Bar, 1 cm; (B) comparison of seed length of each sample. Bar, 1 cm; (C) comparison of

seed width of each sample. Bar, 1 cm; (D) grain length of each sample; (E) grain width of each sample; (F) tiller number of each sample; (G) plant height of each sample; (H) grain number per panicle of each sample; (I) weight of per panicle of each sample. The values in (D,E) show the mean  $\pm$  SD ( $n > 300$  seeds). The values in (F–I) show the mean  $\pm$  SD ( $n = 30$  plants). Significant differences: ns, no significance; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$  (Student's  $t$ -test).

These results indicate that the reverted diploid HTRM2 inherited some excellent traits, such as plant height and awn trait, from tetraploid HT4 and that some traits, such as tiller number, were even better than those of HT2. The number of grains per panicle and yield per panicle of the  $F_1$  generation of HTRM2 were better than those of its female parent, and it had strong heterosis. The excellent self-reverted diploid selected from progeny of the autopolyploid by self-crossing provided excellent parental material for breeding high-quality, high-resistance hybrid rice, which verified the feasibility that polyploid rice could be used as a mutation carrier for creating variations and breeding excellent varieties of diploid rice.

### 2.3. Genome-Wide Polymorphism Detection in HT2, HT4, HTRM2, and HTRM12

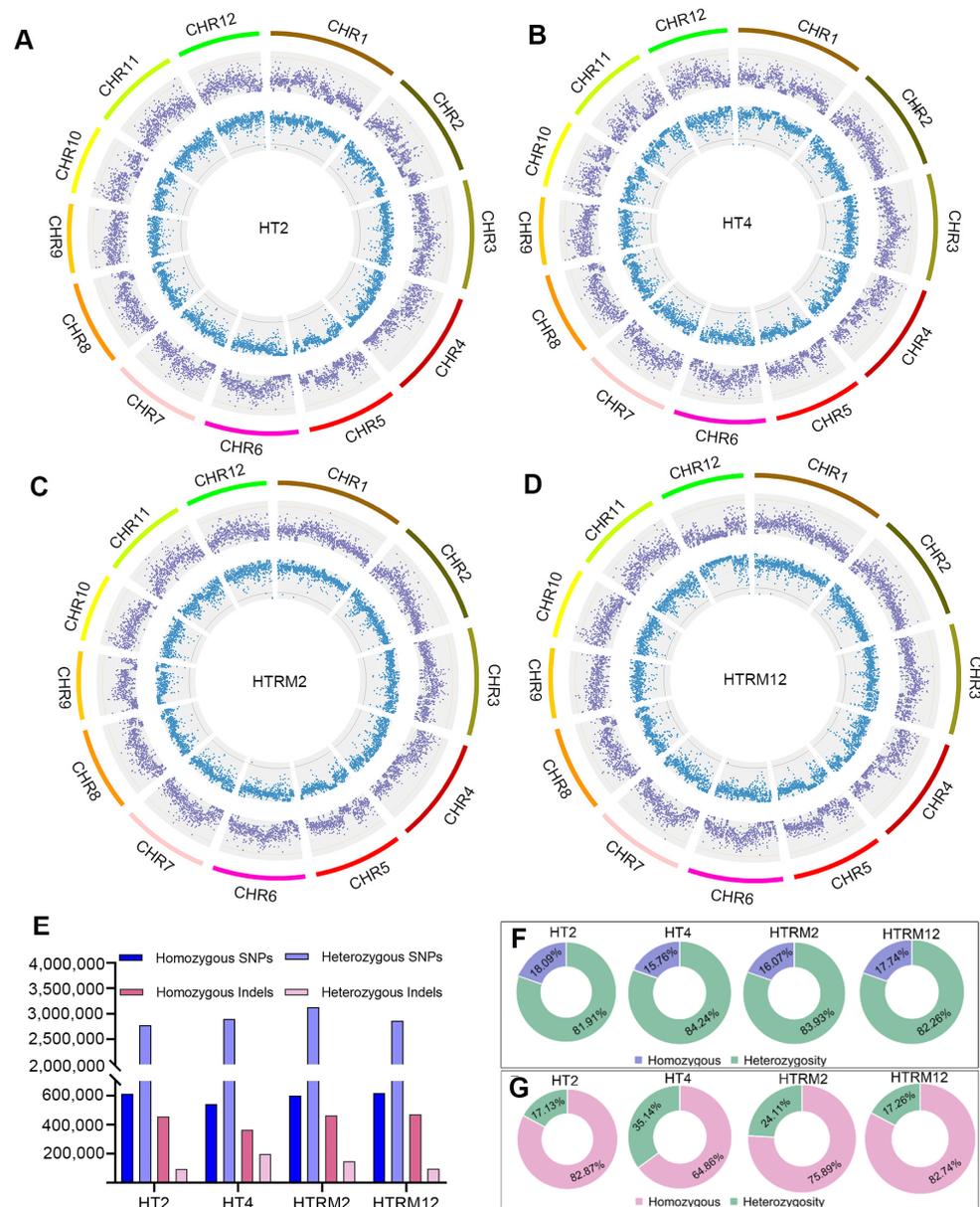
To study the genomic variation differences among HT2, HT4, HTRM2, and HTRM12, whole-genome sequencing was performed using Illumina HiSeq2500, and 23, 24, 20, and 22 Gb of resequencing data, respectively, were obtained. The raw sequence data were deposited in the NCBI Short Read Archive under accession number PRJNA1017988, and the sequencing quality was high ( $Q_{20} \geq 98\%$ ,  $Q_{30} \geq 93\%$ ) (Table S4). The GC distribution was normal, and the samples were not contaminated, indicating that the library was constructed and sequenced successfully and could be used for subsequent research (Figures S1 and S2).

After quality control and filtering of the sequencing results, there were 133,508,370, 146,785,840, 159,363,942, and 151,883,348 valid reads (total reads) in HT2, HT4, HTRM2, and HTRM12, respectively. The valid reads were then mapped to the *Nipponbare* reference genome. A total of 111,540,478, 121,670,405, 133,024,246, and 125,869,239 reads of HT2, HT4, HTRM2, and HTRM12, respectively, were mapped to the reference genome, and the average coverage depths of HT2, HT4, HTRM2, and HTRM12 were  $44\times$ ,  $48\times$ ,  $52\times$ , and  $50\times$ , respectively (Table S5). The detailed genome mapping quality and coverage distribution are shown in Table S5 and Figure S3. These results indicate the whole-genome sequencing is high-quality and reliable.

The SNP and InDel variations were examined using the GATK tool (version: 4.1.0.0), with the *Nipponbare* genome as a reference [38]. The total number of SNPs was 3,384,568, 3,438,287, 3,728,757, and 3,475,370, and the total number of InDels was 548,409, 561,615, 608,871, and 569,020 for HT2, HT4, HTRM2, and HTRM12, respectively (Tables S6 and S7). The total number of InDels was much less than the total number of SNPs; HTRM2 had the most SNPs and InDels, followed by HTRM12 (Tables S6 and S7). The SNPs and InDels were unevenly distributed across chromosomes. Chromosome (chr) 1 had the most SNPs and InDels, and the SNPs and InDels on chr11 and chr12 significantly differed among the four materials (Figures 3A–D and S4). Although HTRM2 was a tetraploid-revertant diploid from HT4, which was derived from HT2 by chromosome doubling, the total SNP and InDel variation in HT4 and HTRM2 were more abundant than that of HT2. This shows that the autotetraploid and its revertant diploid have the potential to produce new variations within new materials.

HTRM12 had the homozygous SNPs with a homozygosity rate of 17.74%, which was slightly lower than the 18.09% homozygosity rate of HT2, and HT4 had the smallest number of homozygous SNPs and the lowest homozygosity rate (Figure 3E,F). Although the total number of InDels was small, the number of homozygous InDels for each material accounted for a large proportion (Figure 3E,G). The homozygous InDel rate was similar between HT2 and HTRM12 (nearly 83%). The homozygous InDel rate of HTRM2 was 75.89% and that of HT4 was only 64.86%. The overall homozygous SNP and InDel variation proportion of tetraploids and revertant diploids was low, and the number of homozygous SNP and

InDel variations increased after revertant crosses, providing guidance for breeding practice (Figure 3F,G). Therefore, these results show that HTRM2 and HTRM12 had more abundant variation sites, which is conducive to the cultivation of new materials for breeding superior hybrid rice varieties.



**Figure 3.** Number statistics and distribution of SNPs and InDels on chromosomes for HT2, HT4, HTRM2, and HTRM12. (A) Distribution of genomic variants in HT2 on chromosomes; (B) distribution of genomic variants in HT4 on chromosomes; (C) distribution of genomic variants in HTRM2 on chromosomes; (D) distribution of genomic variants in HTRM12 on chromosomes. Circle 1: chromosomes; Circle 2: purple loci indicate the distribution of genomic SNP densities; Circle 3: blue loci indicate the distribution of genomic InDel densities. Circle 1: chromosomes; (E) statistics of homozygous SNP, homozygous InDel, total SNP, and total InDel of each sample; (F) proportion of homozygous SNP and heterozygous SNP in each sample; (G) proportion of homozygous InDel and heterozygous InDel in each sample.

#### 2.4. SNP and InDel Site Annotation and Mutation Type Analysis of HT2, HT4, HTRM2, and HTRM12

To explore the potential effects of polymorphic sites on gene function among HT2, HT4, HTRM2, and HTRM12, we performed polymorphic site annotation and mutation type statistical analysis. SNP annotation showed that, after removing the SNP mutations in the upstream and downstream regions, the SNP mutations of the four rice materials occurred in the protein coding region (CDS), accounting for 62%. The intergenic region reached 16% (Figure S5A; Table S8), and missense mutations (34%) accounted for the largest proportion of SNP mutations in the CDS. The proportions of synonymous and stop-gain mutations were slightly different among the four materials. Synonymous mutations reached 27% in HT2 and HTRM2 and 26% in HT4 and HTRM12. Both HT2 and HTRM2 had 1% stop-gain mutations, and both HT4 and HTRM12 had 2% stop-gain mutations (Figure S5A; Table S8). The high proportion of stop-gain mutations, one of the most severe mutation types, in HT4 and HTRM12 indicates that polyploid rice has high mutational ability. In addition, we determined the frequency of SNP mutation types in each sample. Among all samples, the highest mutation frequency was C > T, followed by G > A, T > C, and A > G. The mutation frequency of the other types decreased significantly (Figure S5C; Table S9). Notably, regardless of the mutation type, the mutation frequency was the highest in HTRM2, followed by that in HTRM12.

The InDel annotation results showed that the mutations of HT2, HTRM2, and HTRM12, except for the upstream and downstream regions, accounted for 31% of the mutations in the CDS, while the mutations in HT4 accounted for 30%. Among the mutation types in the CDS of HT2, HT4, HTRM2, and HTRM12, infra-shift mutations accounted for 10%, and frameshift mutations accounted for 21%. The frameshift mutations in HT4 accounted for 20%. The InDel mutations in the untranslated region (UTR) and the stop-loss type in the CDS of HT4 accounted for 22% and 49.0%, respectively, which were the highest among the four samples. This indicates that polyploid rice HT4 has high potential for creating new rice germplasm (Figure S5B; Table S10).

The effects of SNP and InDel variants on gene function were predicted using the SnpEff tool (version: 4.4) [39], and four types with high, moderate, low, and modifier effects were clustered (Table S11). The number of effects of SNP and InDel mutations on genes followed the order Modifier > Moderate > Low > High and Modifier > High > Moderate > Low, respectively. In the four materials, HTRM2 and HTRM12 had a higher number of high and moderate SNP and InDel variation types, indicating a greater influence on gene function and the corresponding phenotype than low and modifier types.

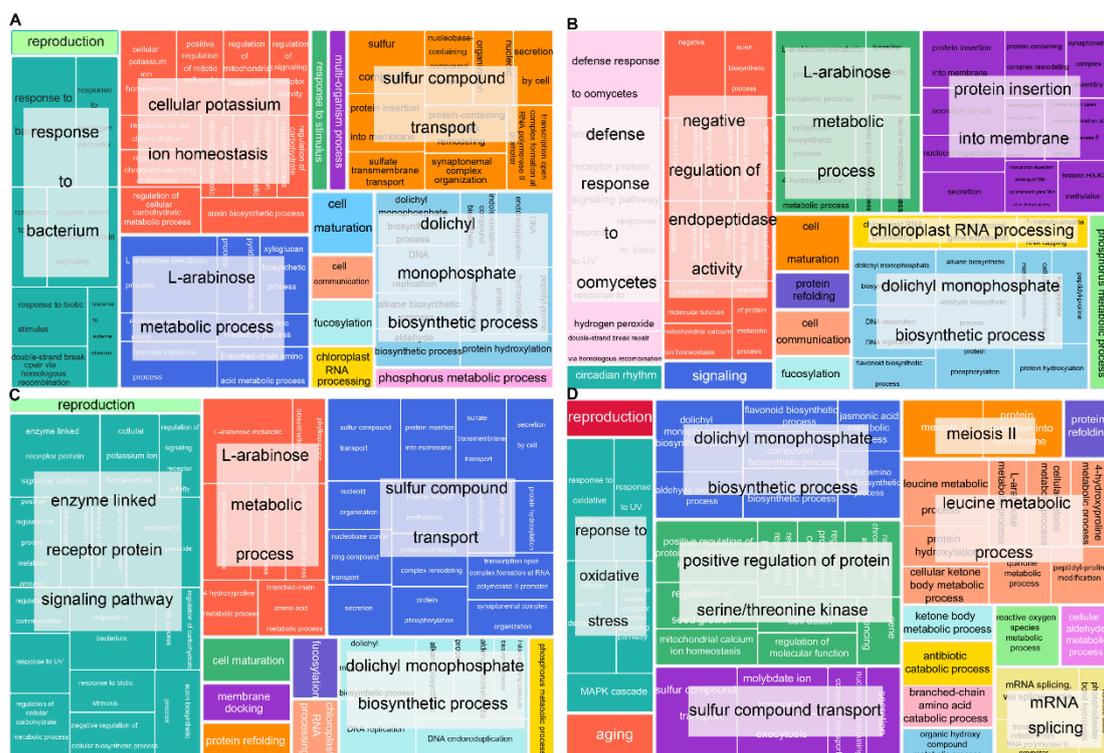
In summary, tetraploid HT4 had more variations (SNPs and InDels) in important genome regions (such as CDS), and tetraploid-revertant diploid HTRM2 as well as hybrid descendant HTRM12 harbored more high and moderate variation types. This indicates that tetraploid can be used as a mutation carrier for creating novel diploid materials by ploidy reverting mutation, and the revertant diploid with rich variations has the potential to breed excellent hybrid rice.

#### 2.5. GO Terms and KEGG Enrichment Analysis of SNP Mutant Genes of HT2, HT4, HTRM2, and HTRM12

To further investigate the causes of the differences in HT2, HT4, HTRM2, and HTRM12, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for the genes in which homozygous SNP and homozygous InDel mutations occurred.

Items enriched in HT2 and HTRM2 were mostly the same, such as "L-arabinose metabolic process", "reproduction", "sulfur compound transport", and "dolichyl monophosphate biosynthesis process" (Figure 4A,C). Items enriched in HT4 were similar to those enriched in HT2 and HTRM2, but there were many unique entries, such as "defense response to oomycetes", "negative regulation of endopeptidase activity", and "protein insertion into membrane" (Figure 4B). Endopeptidases play a key role in the mobilization of storage

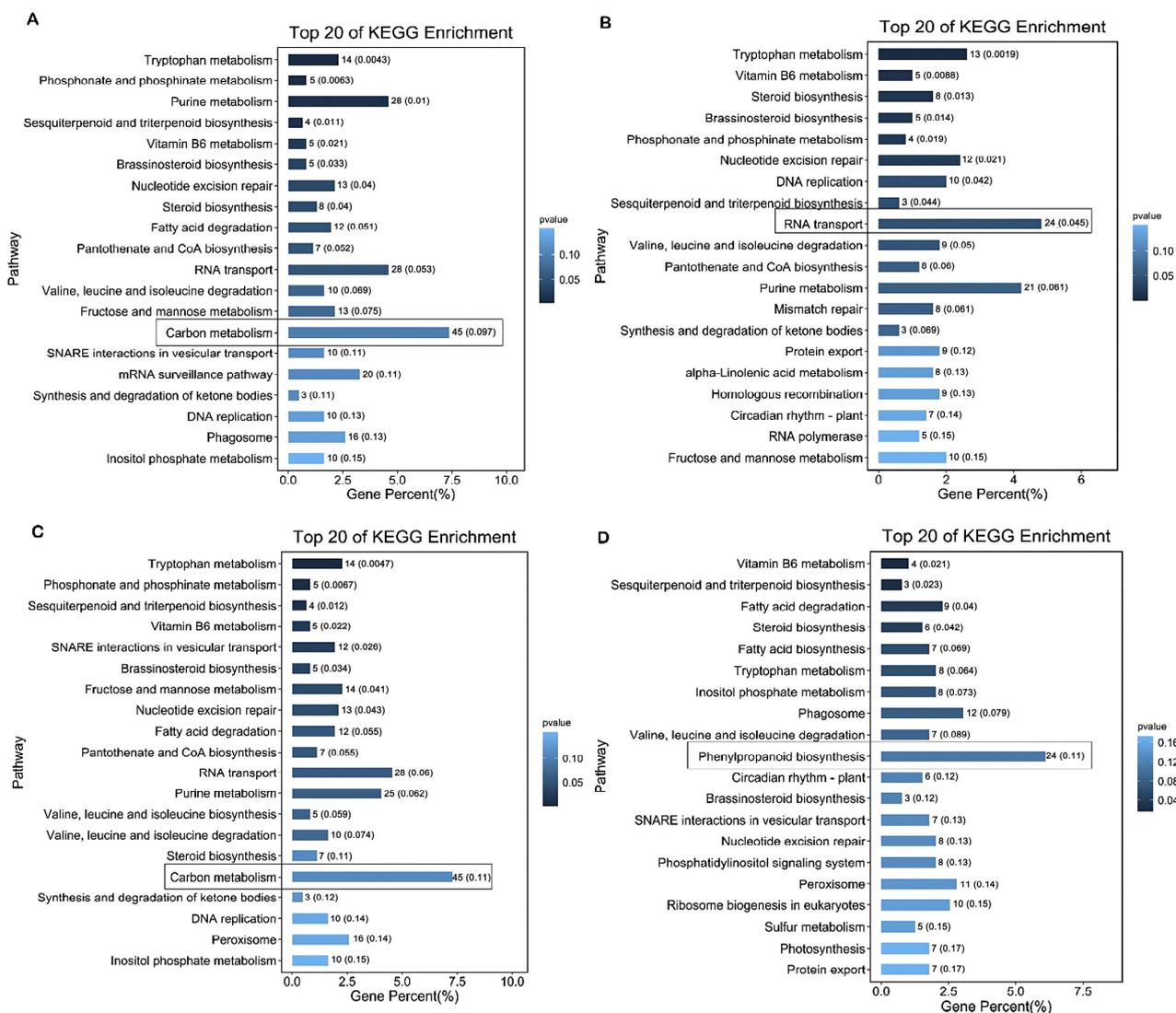
proteins during seed germination [40]. Endopeptidases can be involved in biotic stress and pathogen defense [41]. Polyploid rice is more resistant to stress, and endopeptidases likely play a role in the stress response of polyploid rice. The enriched GO terms in HTRM12, such as “meiosis II”, “mRNA splicing”, “leucine metabolic processes”, “aging”, and “positive regulation of protein serine/threonine kinase”, were significantly different from those in the other three samples; these entries existed only in HTRM12 (Figure 4D). These items all play an important role in plant development. For example, mRNA splicing is a very important biological process in eukaryotic gene expression. Through RNA splicing, many functional mRNAs with coding information can be produced, which is vital to the development and evolution of organisms [42]. Meiosis is the cytological basis of heredity and variation, and it plays a key role in the life process of plants [43].



**Figure 4.** GO terms enrichment analysis results of genes with homozygous SNP mutations in HT2, HT4, HTRM2, and HTRM12. (A) GO terms enrichment analysis results of genes with homozygous SNP mutations in HT2; (B) GO terms enrichment analysis results of genes with homozygous SNP mutations in HT4; (C) GO terms enrichment analysis results of genes with homozygous SNP mutations in HTRM2; (D) GO terms enrichment analysis results of genes with homozygous SNP mutations in HTRM12. The results were visualized by the “TreeMap” view of REVIGO. Each rectangle is a single cluster representative. The representatives are joined into “superclusters” of loosely related terms, visualized with different colors. The size of the rectangles is adjusted to reflect the  $p$ -value.

The KEGG enrichment results showed the same change trend as the GO term enrichment analysis among the four materials. In the results, HT2 and HTRM2 were similar. The pathway with the highest gene percentages in HT2 and HTRM2 was “carbon metabolism”, followed by “RNA transport” and “purine metabolism”. Interestingly, there were a small number of unique pathways between HT2 and HTRM2. For example, “mRNA surveillance pathway” was unique to HT2, and “valine, leucine, and isoleucine biosynthesis” and “peroxisome” were unique to HTRM2. These KEGG pathways may cause differences between HT2 and HTRM2 in RNA, protein metabolism, and oxidative stress response, although HTRM2 was a tetraploid-revertant diploid derived from HT4 by chromosome doubling of HT2 (Figure 5A,C). The enrichment results of HT4, HT2, and HTRM2 were similar, and

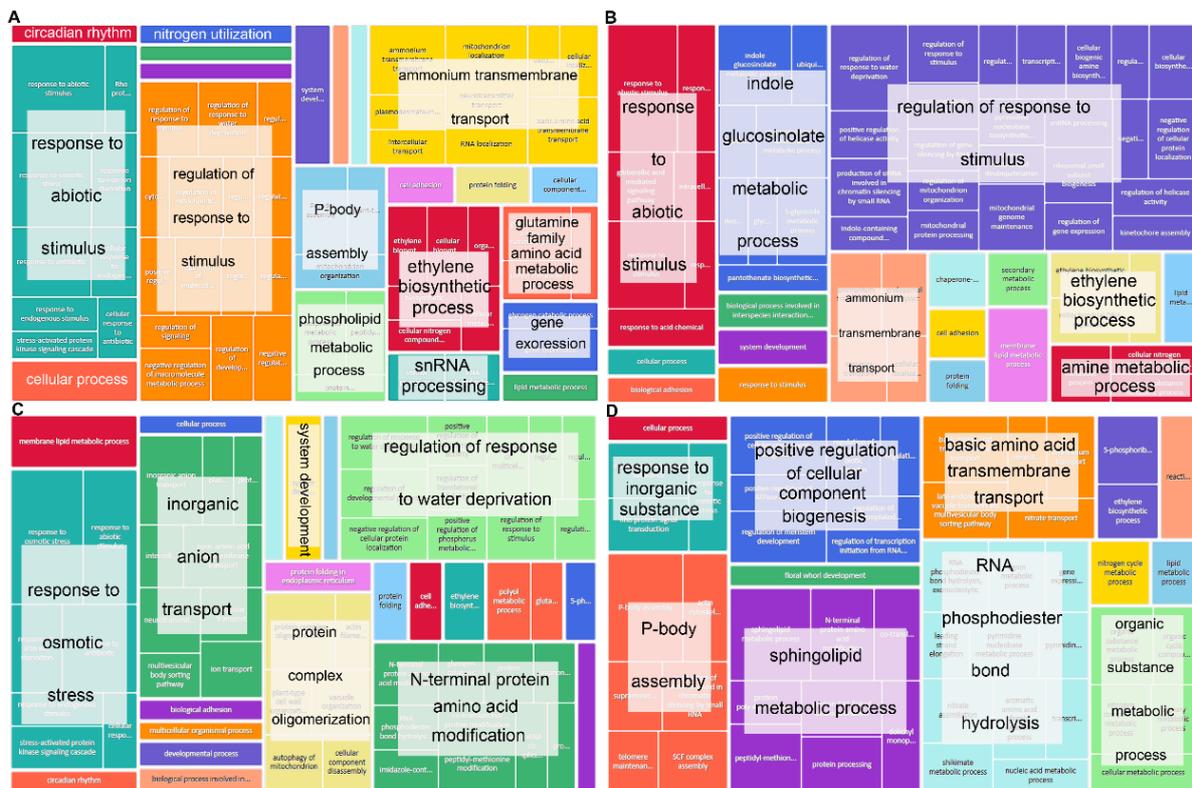
the pathway with the largest number of genes was “RNA transport”, followed by “purine metabolism” and “tryptophan metabolism” (Figure 5B). The above-mentioned pathways are all primary metabolic pathways [44]. In contrast, the pathway with the most enriched genes for HTRM12 was “phenylpropane biosynthesis”, followed by “peroxisome” and “phagosome” (Figure 5D). Phenylpropane metabolism is one of the most important plant secondary metabolic pathways, producing more than 8000 metabolites, which play an important role in plant growth and development and plant–environment interactions [45]. The phenylpropanoid metabolites play a vital role in many aspects, which can regulate plant growth and development and play a role in disease resistance and anti-aging [46]. Unique KEGG pathways in HTRM12, such as “phenylpropanoid biosynthesis”, “ribosome biogenesis in eukaryotes”, “sulfur metabolism”, and “photosynthesis”, may confer the hybrid rice variety HTRM12 with the highest grain number per panicle and weight per panicle among the four materials.



**Figure 5.** KEGG enrichment analysis results of genes with homozygous SNP mutations in HT2, HT4, HTRM2, and HTRM12. (A) Top 20 of KEGG enrichment analysis results of genes with homozygous SNP mutations in HT2; (B) top 20 of KEGG enrichment analysis results of genes with homozygous SNP mutations in HT4; (C) top 20 of KEGG enrichment analysis results of genes with homozygous SNP mutations in HTRM2; (D) top 20 of KEGG enrichment analysis results of genes with homozygous SNP mutations in HTRM12. The darker the blue, the smaller the *p*-value. The labeled boxes indicate the entries with the highest number of genes enriched in the KEGG metabolic pathway for each of these four materials.

## 2.6. GO Terms and KEGG Enrichment Analysis of InDel Mutant Genes of HT2, HT4, HTRM2, and HTRM12

The enrichment results of GO terms of InDel mutant genes showed that the items with the most genes in HT2, HT4, and HTRM2 were related to stress and environmental responses. For example, genes related to “response to biological stimulation” and “regulation of response to stimulation” accounted for the largest proportion in HT2 and HT4 (Figure 6A,B). The items with the most genes in HTRM2 were “response to osmotic stress” and “regulation of response to water deprivation” (Figure 6C). There were no stress response items enriched in HTRM12, and it differed from the other three samples, except for the enrichment of “P-body assembly”, which was also found in HT2 (Figure 6D). Other high-volume items such as “response to inorganic substances”, “positive regulation of cellular component biosynthesis”, “RNA phosphodiester bond hydrolysis”, “sphingolipid metabolic process”, and “basic amino acid transmembrane transport” were only present in HTRM12 (Figure 6D). These results showed that the gene similarity of the InDel homozygous mutation between HT2 and HT4 was high, while that in HTRM12 was relatively unique. The gene function of its InDel homozygous mutations differed from those of the other three materials.

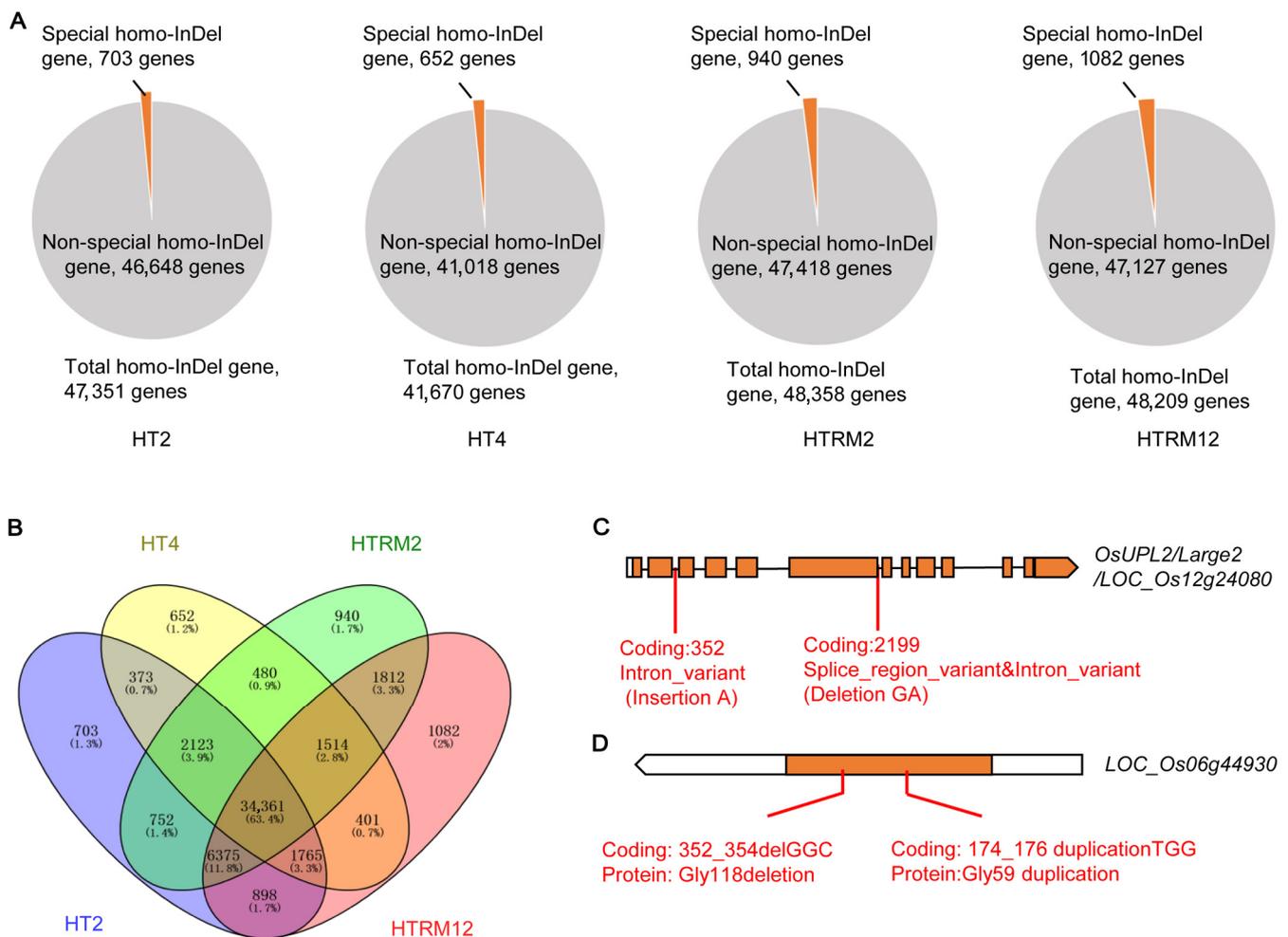


**Figure 6.** GO terms enrichment analysis results of genes with homozygous InDel mutations in HT2, HT4, HTRM2, and HTRM12. (A) GO terms enrichment analysis results of genes with homozygous InDel mutations in HT2; (B) GO terms enrichment analysis results of genes with homozygous InDel mutations in HT4; (C) GO terms enrichment analysis results of genes with homozygous InDel mutations in HTRM2; (D) GO terms enrichment analysis results of genes with homozygous InDel mutations in HTRM12. The results were visualized by the “TreeMap” view of REVIGO. Each rectangle is a single cluster representative. The representatives are joined into “superclusters” of loosely related terms, visualized with different colors. The size of the rectangles is adjusted to reflect the  $p$ -value.

The KEGG enrichment analysis results of the four samples were almost the same, except for a slight difference in the number of genes. For example, the most enriched genes in metabolic processes were related to “global and overview maps”, followed by

some primary metabolic processes, such as “carbohydrate metabolism”, “amino acid metabolism”, and “lipid metabolism”. Most genes enriched in the gene information process were related to “translation”, followed by “folding, sorting, and degradation” and “transcription”. The same was true for entries in other metabolic processes (Figure S6). The InDel mutant genes caused differences in GO term enrichment results among the four samples, especially between HTRM12 and the other three samples, while the KEGG enrichment results were the same among HT2, HT4, HTRM2, and HTRM12.

There were 47,351, 41,670, 48,358, and 48,209 genes harboring at least one homozygous InDel (homo-InDel gene) in HT2, HT4, HTRM2, and HTRM12, respectively (Figure 7A). There were 703, 652, 940, and 1082 unique homo-InDel genes (only present in one material) in HT2, HT4, HTRM2, and HTRM12, respectively (Figure 7A,B). A genome-wide association study (GWAS) phenotype enrichment analysis [47] was performed to identify the candidate genes for special phenotypes in each material. Two terms, “alkali spreading value” and “flag leaf length”, were only found in HT2. “Awn presence” was found only in HT4. Four terms, i.e., “amylose content”, “awn presence”, “pericarp color”, and “flowering time at Arberdeen”, were found in HTRM2, and three terms, i.e., “heading date”, “shattering degree”, and “tiller number”, were found in HTRM12. Interestingly, “awn presence” was present in both HT4 and HTRM2, which agreed with their awn phenotypes (Table S12).

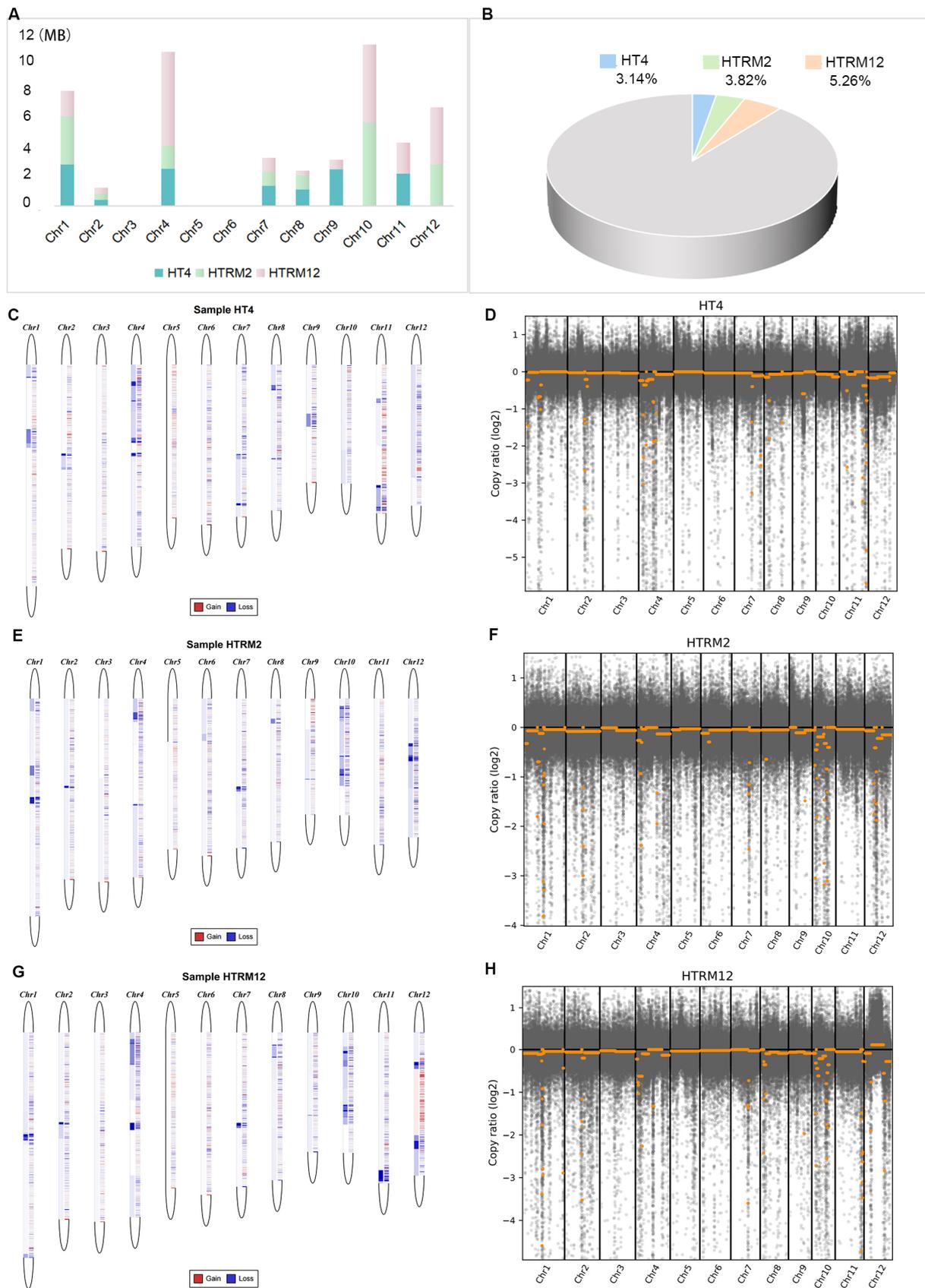


**Figure 7.** The awn genes were identified by GWAS phenotype enrichment analysis for special homo-InDel gene sets of four materials. (A) Statistics of homo-InDel genes and special homo-InDel genes for four materials; (B) the Venn diagram of homo-InDel genes of four materials; (C,D) the InDel variations in coding regions of awn genes identified by GWAS phenotype enrichment analysis (see Table S12). Rectangles represent exons, orange rectangles represent protein-coding regions, and lines represent introns.

We focused on the genes responsible for the grain and awn phenotypes. The GWAS phenotype enrichment analysis showed that LOC\_Os12g24080 (*large2*) was enriched in HT4 (Table S12). The *Large2* gene has been reported to regulate the rice grain type, and the loss of function of this gene leads to a larger grain type [48]. Moreover, GWAS association analysis has previously shown that this gene is associated with the awn phenotype [49]. The phenotypes of HT4, such as the large grain and awn phenotypes, were obviously different from those of the other three materials, which is likely related to the variations in the *Large2* gene. Analysis of homo-InDel variation showed that an adenine base was inserted in the 2nd intron of this gene, and G and A bases were absent at the splicing of the 6th exon and 6th intron (Figure 7C). These variations may lead to a change in the pre-mRNA splicing of the *Large2* gene, thus causing a change in the function of the translated protein. These variations are likely to be important to the grain and awn phenotype differences in HT4 compared with the other three materials. Interestingly, HTRM2 also showed an awn phenotype, but the GWAS phenotype enrichment analysis of its homo-InDel genes highlighted a new gene, LOC\_Os06g44930 (Table S12), which has also been associated with the awn phenotype of rice using GWAS [47]. Analysis of the homo-InDel variation in the coding region of this gene showed that there were two homo-InDel variations in the coding region, leading to changes in the coding protein sequence (Figure 7D). Collectively, the identified genes and corresponding InDel variations, for example, the awn genes, may be important for special phenotype variations in the four materials, which may have potential in polyploidy rice breeding.

### 2.7. Identification and Comparison of Copy Number Variation (CNV) for HT4, HTRM2, and HTRM12

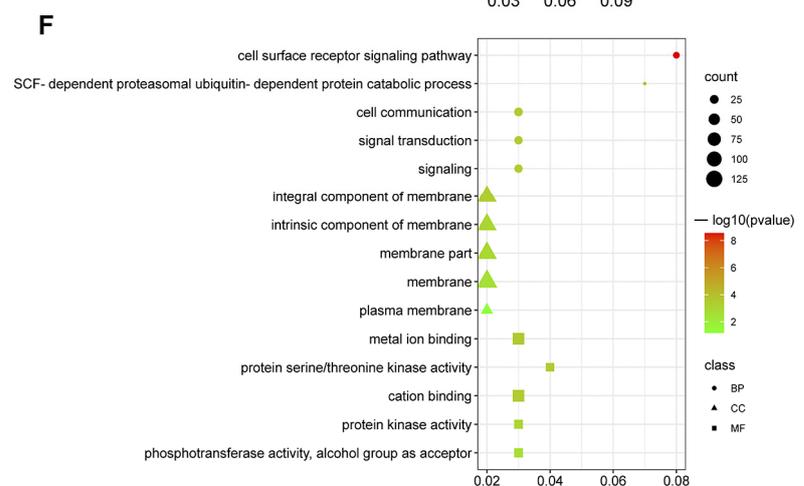
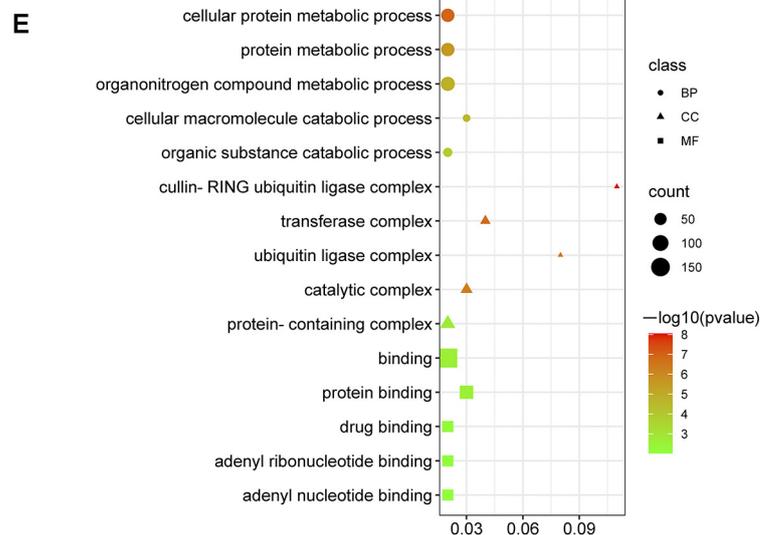
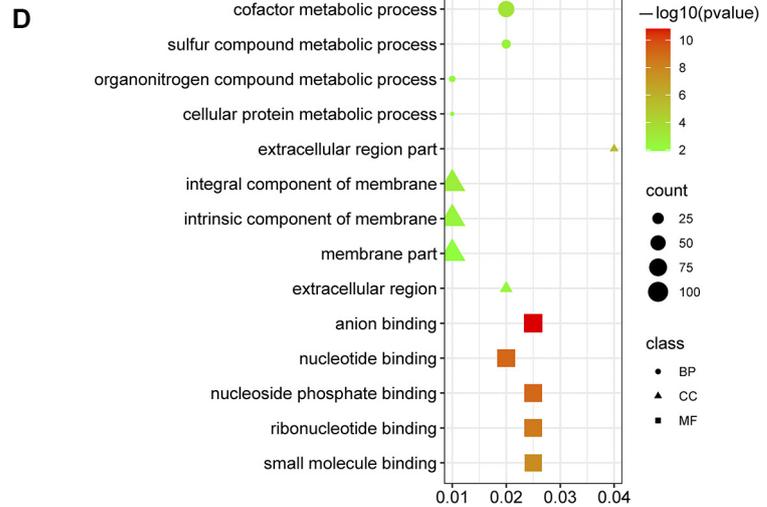
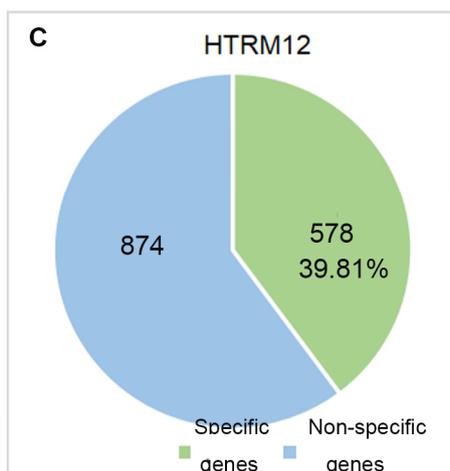
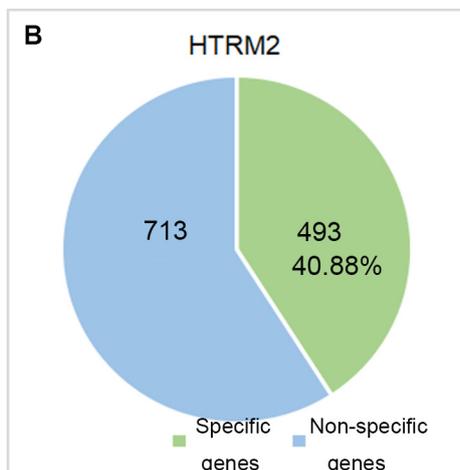
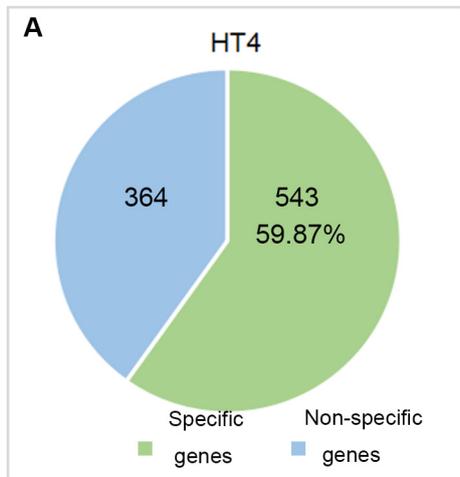
Copy number variation (CNV), which is expressed as an increase (gain) or decrease (loss) in the copy number of genomic segments, is widely distributed in the human genome and is an important genetic basis for individual differences [50]. CNVs may be biologically important for species-specific genome composition, species evolution, phylogeny, and the expression and regulation of genes in specific regions of the genome [51]. Therefore, CNVs have become a topic of interest in genomics research in recent years. To study CNVs in rice with ploidy changes, the HT4, HTRM2, and HTRM12 CNV segments were identified using the HT2 genome as a control. There were no CNV segments identified on chr3, chr5, and chr6, and CNV segments were identified for the three samples on chr1, chr2, chr4, chr7, and chr8 (Figure 8A). However, only HT4 and HTRM12 contained CNVs on chr9 and chr11, and only HTRM2 and HTRM12 contained CNVs on chr10 and chr12 (Figure 8A). The HT4 genome had a CNV interval of 13.20 Mb, accounting for 3.14% of the whole genome relative to the HT2 genome. The HTRM2 genome had a CNV interval of 16.06 Mb, accounting for 3.82% of the whole genome relative to the HT2 genome. Compared with the HT2 genome, there was a 22.10 Mb CNV interval in the HTRM12 genome, accounting for 5.26% of the whole genome (Figure 8B). Analysis of the CNV distribution on different chromosomes of the samples revealed that most of the deletions of the copy number regions (CNRs) and copy number segments (CNSs) occurred in three samples, especially HT4 and HTRM2, where only CNSs were lost (Figure 8C,E). However, there were CNR gains on chr12 of HTRM12 (Figure 8G). According to the principle of CNV calling, most lost CNVs have a small copy ratio [52]. There were more CNSs with extremely small copy ratios of loss CNVs on chr4, chr7, and chr11 in HT4 (Figure 8D). In HTRM2, the lost CNVs with extremely small copy ratios were distributed on chr1 and chr10 (Figure 8F). In HTRM12, they were distributed on chr1, chr10, and chr11 (Figure 8H). These different chromosome distributions of lost CNVs with extremely small copy ratios in each material may contribute different characteristics to each rice variety.



**Figure 8.** Identification and comparison of genomic CNV intervals of HT4, HTRM2, and HTRM12. (A) Distribution and size of the CNVs of HT4, HTRM2, and HTRM12 on the chromosome; the

horizontal coordinates indicate the different chromosomes, and the vertical coordinates indicate the size of the CNV. Different colors indicate different samples; (B) the proportion of the specific CNV interval occupied by each sample; (C,D) distribution of differential CNVs on different chromosomes in HT4 relative to HT2; (E,F) distribution of differential CNVs on different chromosomes in HTRM2 relative to HT2; (G,H) distribution of differential CNVs on different chromosomes in HTRM12 relative to HT2; in (C,E,G) Red, increased copy number; blue, decreased copy number; left, increase or decrease in CNS (Copy Number Segment); right, increase and decrease in CNR (Copy Number Region); in (D,F,H) gray dot, CNR; yellow dot, CNS.

Analysis of the genes presented in the CNV segments of the three samples revealed that fewer genes were present in the CNV segments of HT4 than in the CNV segments of HTRM2 and HTRM12, but the percentage of CNV-specific genes presented only in HT4 was the highest (59.87%) (Figure 9A–C). The CNV segments of HTRM12 contained the most genes (1452 genes) but the lowest percentage of CNV-specific genes (39.81%) (Figure 9C). We also explored the GO term enrichment analysis of specific genes contained in the CNSs of HT4, HTRM2, and HTRM12. The top five items in the biological processes (BPs) in HT4 were all metabolic processes, including primary and secondary metabolism. The most frequent items in the cellular components (CCs) were extracellular regions and cell membranes, and the top five entries in the molecular functions (MFs) were all related to binding, including anion binding, nucleotide binding, and small molecule binding (Figure 9D). The genes in the specific CNV compartment in HTRM2 were enriched in more BP items. In addition to metabolic processes, there were catabolic processes. In contrast to HT4, the most enriched items in the CCs of HTRM2 were enzyme complexes, catalytic complexes, and protein complexes. However, the top five items in the MFs of HTRM2 were similar to those of HT4 in that they were all binding-related items (Figure 9E). The genes in the specific CNV compartment in HTRM12 were enriched for more entries in BPs and were different from those in HT4 and HTRM2, which were mainly enriched in signaling-related items (Figure 9F). The entries enriched in the CCs of HTRM12 were more similar to HT4 in that they were membrane-related items. The top five enriched items in the MFs of HTRM12 were related to kinase activity and binding (Figure 9F). These specific items of gene enrichment for specific CNV intervals may be candidates for phenotypic differences among the three materials.



**Figure 9.** GO Terms enrichment analysis of genes within the specific CNV intervals of HT4, HTRM2, and HTRM12. (A–C) The proportion of specific genes for HT4, HTRM2, and HTRM12; (D–F) top 5 GO terms in each BP, CC, and MF enrichment analysis of genes within the specific CNV intervals of HT4, HTRM2, and HTRM12. Circles represent biological processes, triangles represent cellular components, and squares represent molecular functions; the size of the graphs indicates the number of genes.

### 3. Discussion

At present, breeders are promoting the application of tetraploid rice by cultivating highly fertile tetraploid materials and constructing tetraploid heterosis utilization systems and have made continued progress (see Section 1) [23–31]. However, as far as we know, there is no commercial variety directly bred with tetraploid. In this study, we used tetraploid rice HT4 as a mutant carrier to breed a new excellent diploid material HTRM2, and then bred the hybrid rice variety HTRM12. This hybrid rice variety has passed the national variety approval. It has excellent agronomic traits, with a high yield of 9.67 tons/ha, high quality, and multi-resistance (Figure 1; Table S1). This hybrid rice is suitable for planting in the middle and lower reaches of the Yangtze River, China. Furthermore, the agronomic traits analysis of four materials with ploidy change shows the tetraploid self-reverted diploid HTRM2 can inherit some excellent traits of tetraploid HT4, and some traits are even better than original diploid HT2 (Figure 2).

Diploid can be generated with low probability from tetraploid self-bred progeny in rice. The reverted diploid rice may have more phenotypic variation than the original diploid. Xing et al. found the diploid can be produced in tetraploid self-bred progeny in rice [53]. Xiong et al. generated the self-reverted diploids from self-bred progenies from japonica tetraploids, indica tetraploids, and indica-japonica hybrid tetraploids, respectively. The selfing-reverted diploids showed great phenotypic variations and strong heterosis [33]. This reverted diploid can also be generated by *in vitro* gametocyte culture of tetraploid. For example, Li et al. generated reverted diploid rice by anther culture of PMeS, 9311-4x, and other tetraploid rice varieties, and these reverted diploid rice varieties showed high quality [54].

In summary, this study verified that tetraploid rice can be used as a mutation carrier to breed excellent parent rice materials and can cultivate excellent hybrid rice varieties with high yield, high quality, and high resistance. Thus, it provides excellent new germplasm for rice breeding and a new way for the application of tetraploid rice in breeding.

Polyploidization can cause genetic redundancy when the homologous or heterologous genomes are combined in the course of the evolution of plants, which will lead to the “genome impact”. In order to cope with the “genome impact” caused by the genome combination, the new polyploid will establish a new balance of the genome through chromosome rearrangement, sequence elimination, gene variation, and so on. On the other hand, plants will also undergo polyploidization to diploidization during the process of evolution. This will induce genome loss, variation, and gene expression changes. This “polyploid cycle” is considered to promote plant diversity and adaptive evolution [55,56]. In this study, we generated the tetraploid rice HT4 by artificial chromosome duplication of diploid rice HT2 and selected the excellent diploid HTRM2 from tetraploid self-bred progeny diploid. Hence, we constructed an artificial “polyploid cycle” in rice.

In order to study the relationship between genomic variation and ploidy change, the genomic variation of four materials was analyzed. Through genome-wide SNP and InDel variation analysis, we found that the number of variations in four materials conforms to the following rules: HTRM2 > HTRM12 > HT4 > HT2 (Figure 3E; Tables S6 and S7). The tetraploid self-reverted diploid HTRM2 has the largest number of SNP and InDel variants, and the variation in HTRM2 and HTRM12 is more serious because of the more high, moderate SNP and InDel variation types discovered in the two materials. Specially, the stop-gain variation type of SNP is higher in HT4 and HTRM12, and the percent of InDel variation in the untranslated region (UTR) and the stop-loss type in the CDS of HT4 is highest among four materials. These results suggest that the genomes of rice subjected to the artificial “polyploid cycle” and hybridization are very different from those of the original diploid. At present, a small number of studies have reported the genomic variation between the original diploid and the corresponding tetraploid. Fozia et al. sequenced the genomes of diploid and corresponding tetraploid in rice for studying the embryonic sac fertility in autotetraploid. A large number of SNPs, InDels, SVs, and CNVs were detected between diploids and corresponding tetraploids. It was speculated that changes

in the number of chromosomes may induce changes in the genome sequence and produce specific variations in tetraploids [22]. In our study, many variations were detected after chromosome ploidy changes in rice, which further confirms this speculation. Thus, our work provides a good method by using an artificial “polyploid cycle” for creating variation in rice. Moreover, it provides genomic resources for genetic research in the “polyploid cycle” process and for molecular breeding of polyploid and corresponding diploid rice.

GO and KEGG enrichment entries of homozygous SNP were similar among HT2, HT4, and HTRM2, and HTRM12 mostly differed from them. Although the GO entries of HT4 are similar to those of HT2 and HTRM2, there are some stress-related entries specifically enriched in HT4, such as “defense response to oomycetes”, “negative regulation of endopeptidase activity”, “protein insertion into membrane”, etc. These involved genes may contribute to the HT4 higher resistance. Interestingly, there were many different entries related to plant development, such as “meiosis II”, “mRNA splicing”, “leucine metabolic processes”, etc., that were significantly enriched in hybrid HTRM12. These involved genes may contribute to HTRM12’s excellent agronomic properties.

GO enrichment analysis of InDel homologous variation genes found that the stress and environmental response-related entries were enriched in HT2, HT4, and HTRM2, such as “response to biological stimulation”, “regulation of response to stimulation”, and so on. This indicates that stress-related genes were inherited in different ploidy materials. Interestingly, the above entries were not enriched in HTRM12. But some special entries, such as “response to inorganic substances” and “positive regulation of cellular component biosynthesis”, were only present in HTRM12. It may be because HTRM12 is a hybrid rice and the genes embraced in the above entries are heterozygous. However, KEGG enrichment entries of genes with homozygous InDel variants were similar in the four materials. Interestingly, the analysis of the homo-InDel gene found that most of the homo-InDel genes were shared by the four materials, indicating that this gene was inherited in the four materials. The specific homo-InDel gene analysis found that two different awn genes were enriched in HT4 and HT2, respectively, indicating that the InDel variations occurred in the two different awn genes during ploidy changes.

CNVs are reported to impact the characteristics and environmental adaptation in plants and animals [57,58]. The identification and analysis of CNV intervals for HT4, HTRM2, and HTRM12 were consistent with the SNP and InDel results. The size of CNV intervals followed the trend HTRM12 > HTRM2 > HT4, and the size ranking of specific CNV intervals was the same (Figure 8). Compared to the whole genome, HT4, HTRM2, and HTRM12 had 3.14%, 3.82%, and 5.62% CNVs (Figure 8B), respectively, with some specific CNV intervals containing 543, 493, and 578 genes, respectively (Figure 9A–C). GO term enrichment analysis of the genes at specific intervals of the three materials showed that HTRM12 was significantly different from HT4 and HTRM2 (Figure 9D–F). These genes with specific intervals may have resulted in the outstanding morphological characteristics of HTRM12 in the variety test trial.

In summary, our results indicate that genomic variation can occur in the process of chromosome ploidy change in rice, and some variations can be inherited to the reverted diploid from tetraploid rice. The reverted diploid rice has a large variation relative to the original diploid rice, so it can be used to breed high-yield and strong heterosis rice materials. Therefore, this study provides genomic information, new breeding materials, and a new way of application for tetraploid rice breeding.

## 4. Materials and Methods

### 4.1. Rice Material and Phenotype Examination

The variety *Haitian* (HT2) is a superior diploid restorer line for three-line hybrid rice breeding. HT4 is an autotetraploid rice developed from HT2 by chromosome doubling with colchicine treatment [59], and HTRM2 is a high yield, high quality, and high combining ability line that has been selected from the self-bred progeny of tetraploid HT4 for many

years. We crossed the excellent two-line sterile rice material Y58S and HTRM2 to breed the two-line hybrid rice variety *Y Liang You Duo Hui 14* (HTRM12) in the 2014 year.

For Figure 1 and Table S1, the data were collected from results of the variety test in 18 regional trial sites distributed in Anhui, Fujian, Henan, Hubei, Hunan, Jiangsu, Jiangxi, and Zhejiang provinces of the area of the lower and middle reaches of the Yangtze River, China. The experiment was carried out in accordance with the Agricultural Industry Standard of the People's Republic of China [60]. The varieties were sown and transplanted at the same time, the fertilization level was medium to upper, and cultivation management measures were the same as local field production. The values in bar graphs were shown as means  $\pm$  SD ( $n = 18$  trial sites). For each trial site, the materials were planted on an area of 13 square meters in three replicates (three plots), and the average value of the three repeats of each agronomic trait represents the observed value of corresponding traits. The morphological trait measurement methods are described in detail as follows: Yield: Harvest according to the varieties after maturity in separate plots, drying, weighing after cleaning, calculating the yield per hectare. Number of Effective Panicles: At each trial site, the average panicle number of continuous 10 plants in the middle region of two repeat plots was investigated (a panicle with less than 5 grains is not considered an effective panicle) at the rice maturation stage and then converted to ten thousand/ha, which was used as the effective panicle number of the trial site. Total Grain Number Per Panicle: The sum of the number of grains and the number of falling grains with more than 1/3 of the filling degree of 5 plants divided by the total number of panicles of 5 plants. 1000-Grains Weight: The weight of 1000 dry solid grains; each variety is randomly selected as two 1000 grains, respectively weighed; the difference is not greater than 3% of its average value; take two repeated average values. Panicle Length: The length from the node to the top of the panicle (without awns) is taken as the average of all the rice panicles in 5 plants after rice maturity. Plant Height: The length from the top of the panicle (without awns) to the base of the stem was measured at 10 plants in the middle of the plot at the mature stage of rice, and the average value was taken as the plant height of the trial site. Total Spikelet Number Per Panicle: Total spikelet number of 5 plants divided by total panicle number of 5 plants. Seed Setting Rate: Total grain number per panicle divided by total spikelet number per panicle. Heading Date: The heading period is from the sowing time to the day when 80% of the panicles in the plot are exposed to the blade sheath; take the average value of 3 plots as the heading date of each variety in each trial site. Whole Growth Period: The number of days from the sowing date to the maturity date; take the average value of 3 plots as the whole growth period of each variety in each trial site.

To compare the agronomic properties in Figure 2, the four rice materials HT2, HT4, HTRM2, and HTRM12 were grown under normal paddy conditions from November in the winter of 2021 to April in the next year in Sanya City, Hainan, China. After seedling cultivation, each material was transplanted into a plot, respectively, and 30 plants were planted in each plot at a distance of 16 cm  $\times$  26 cm. The field management measures were consistent with the local field production methods. The morphological trait measurements were as follows: Plant height: At the rice mature stage, take about 30 individual plants in the middle of the plot, measure the length from the top of the panicle (without awn) to the stem base, and take the average value as the plant height of the material. Tillering number: At the rice maturity stage, take about 30 individual plants in the middle of each plot and count the tillering number of each plant, and then take the average value as the number of tillerings of the material. Grain number per panicle: For each material, at the maturity stage, harvest about 30 main panicles (the first heading panicles) of a single plant in the middle of the plot, count the number of solid grains per panicle of each material (more than 1/3 filling degree), and take the average as the number of grains per panicle of the material. And a typical panicle of each material was photographed (Figure 2A). Grain weight per panicle: After the main panicle was dried at 50 °C for 5 days, the grain was separated from the panicle, and the weight of each single panicle was weighed after the blighted grain was removed, and the average value was taken as the grain weight per panicle of

the material. Grain length and grain width: After mixing all single panicle grains of each material, respectively, take about 300–500 grains and measure the grain length and grain width using an automatic seed analyzer (Wan Sheng SC-G, Hangzhou Wanshen testing Technology Co., Ltd., Hangzhou, China), and 10 representative grains of each material were photographed.

#### 4.2. The Disease and Pest Resistance Test

For Table 1, the rice blast resistance test data were obtained from the results of 6 trial sites (Zhejiang, Hunan, Hubei, Anhui, Fujian, and Jiangxi province, China) of the rice variety regional test in the 2019 year, and the data in the table represent the average values of the 6 test sites. The identification of each test site was carried out in accordance with the Agricultural Industry Standard of the People's Republic of China [61]. The susceptibility index of rice blast = Disease grade of leaf blast of seedling  $\times$  25% + Disease grade of panicle blast  $\times$  25% + Disease grade of the spike loss  $\times$  50%; Here, the evaluation standard of disease grade of leaf blast of seedling is: 0 grade (incidence rate = 0), 1 grade (incidence rate  $\leq$  5.0%), 3 grade (incidence rate = 5.1~10.0%), 5 grade (incidence rate = 10.1~25.0%), 7 grade (incidence rate = 25.1~50.0%), 9 grade (incidence rate  $\geq$  50.1%); the standard of disease grade of the spike loss is: 0 grade (loss rate = 0), 1 grade (loss rate  $\leq$  5.0%), 3 grade (loss rate = 5.1~15.0%), 5 grade (loss rate = 15.1~30.0%), 7 grade (loss rate = 30.1~50.0%), 9 grade (loss rate  $\geq$  50.1%); the standard of disease grade of the panicle blast is: 0 grade (rate of infected panicle = 0), 1 grade (rate of infected panicle  $\leq$  5.0%), 3 grade (rate of infected panicle  $\geq$  5.1% and  $\leq$  10.0%), 5 grade (rate of infected panicle  $\geq$  10.1% and  $\leq$  25.0%), 7 grade (rate of infected panicle  $\geq$  25.1% and  $\leq$  50.0%), 9 grade (rate of infected panicle  $\geq$  50.1%).

The test of resistance to white leaf blight was carried out in Anren, Hunan Province, in 2019 and 2020, respectively. The test method was carried out according to the standard of technical regulations for bacterial blight resistance of rice varieties in Hunan Province [62], and the data in the table were the average values of two years' data. The evaluation standard of disease grade is: 0 grade (lesion length = 0 cm), 1 grade (some small lesion around incision and lesion length  $\leq$  2 cm), 3 grade (lesion length  $>$  2 cm and  $\leq$  1/4 of leaf length), 5 grade (lesion length  $>$  1/4 and  $<$  1/2 of leaf length), 7 grade (lesion length  $>$  1/2 and  $<$  3/4 of leaf length), 9 grade (lesion length  $>$  3/4 of leaf length).

The test of resistance to rice planthopper was performed at the seedling stage, according to the international standard seedling group screening method (SSST method), with 3 replicates. The data in Table 1 are the average of the results of 2019 and 2020. For the brown planthopper resistance test, the evaluation standard is: 0 grade (no suffering from pests), 1 grade (suffering from pests very slightly), 3 grade (partial yellowing of the first and second leaves of most plants), 5 grade (obviously yellowing, the plants wither or die of about half the plants), 7 grade (more than half the plants died), and 9 grade (all plants died).

#### 4.3. The High Temperature Resistance Test

For Table S2, the high temperature resistance test in field condition was carried out under the natural high temperature conditions from mid-July to mid-August in Wuhan, Hubei Province. HTRM12 and control FLY4 were sown at the same time in three periods. The tillering that continuously lasted more than 5 days under average daily temperature above 32 °C and the highest temperature above 35 °C was taken as the tillering that had been subjected to high temperature stress after heading period, and the average seed setting rate was taken as the seed setting rate under high temperature stress in field conditions. The tillering with the average seed setting rate, which did not experience the above high temperature condition after heading period, was taken as the seed setting rate under normal temperature conditions. Relative heat resistance coefficient = Seed setting rate at high temperature condition of testing variety / Seed setting rate at high temperature condition of control variety.

The high-temperature test of variety in pot culture was carried out in a glass greenhouse. The two rice varieties were planted in pots, setting 4 replicates. When the first

spikelet of each material was flowering, the continuous 5 days of high temperature treatment in a glass greenhouse was applied (the air temperature in a glass greenhouse is about 40 °C), and then the materials were transferred to a natural condition (average daily air temperature of about 32 °C during the flowering period) for growth to maturity. At the same time, normal temperature control was set under natural conditions (the average daily temperature of the flowering period was about 32 °C), and then the seed setting rate of high temperature treatment and normal temperature control was tested. The relative heat resistance coefficient is calculated in the same manner as above. The combined heat resistance coefficient is the average value of the relative heat resistance coefficient of the two conditions.

#### 4.4. Genome Sequencing and Data Quality Assessment

The construction of the library for genome sequencing and sequencing was performed by the Shanghai Ouyi Biomedical Technology Company (Shanghai, China). Briefly, DNA was extracted from the young roots of HT2, HT4, HTRM2, and HTRM12 using the hexadecyl trimethyl ammonium bromide (CTAB) method. Moreover, 1000 ng of DNA in 52.5 µL was sheared to 350 bp–500 bp using the Covaris S220 (Covaris, LLC, Woburn, MA, USA) at 175 W, 5% duty, 200 cycles for a duration of 20 s. The TruSeq Nano DNA LT Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA) was used in library construction following the manufacturer's instructions. After the library was qualified, whole-genome sequencing was performed by double-end sequencing using the Illumina HiSeq2500 sequencing platform (Illumina, Inc., San Diego, CA, USA). After the raw sequencing data (raw data) were obtained from sequencing, the adaptor sequences, reads with a number of  $N \geq 5$ , and sequences with an average base quality value of less than 20 were removed to obtain clean reads using the software FASTP (version: 0.23.4) [63]. The results of sequencing data quality control can be found in Table S4, Figures S1 and S2.

After removing the reads with a length less than 75 bp and an average base quality value less than 15, the clean reads were aligned to the reference genome (*Nipponbare*, MSU-release7) using BWA (version: 0.7.12) [64], with the "bwa mem" alignment algorithm and default parameter. Then, the format conversion and ranking of alignment results were performed using SAMtools (version: 1.4) [65], and the PCR repeat reads were removed using Picard (version: 4.1.0.0) (<https://broadinstitute.github.io/picard/>, accessed on 10 October 2020). The final results were analyzed by Qualimap 2 software (version: 2.2.2-dev) [66] for quality control. The results of mapping clean reads to the reference can be found in Figure S3 and Table S5.

#### 4.5. Analysis of Variation Information

The results of alignment to the reference genome (*Nipponbare*, MSU-release7) (BAM files) were analyzed for SNP and InDel variation information. Briefly, SNP and InDel detection were performed using the Haplotypecaller module of GATK4 software (version: 4.1.0.0) [38]. The criterion of variation quality value (QD)  $\geq 2.0$  was chosen for filtering the real SNP and InDel variations, in which the QD is equal to the variant quality value (Quality) divided by the depth of coverage (Depth). In fact, the QD reflects the variant quality value per unit depth, as most of the false-positive variants had a QD less than 2. The detected SNP and InDel were annotated with SnpEff software (version: 4.4) [39]. The detected genomic variation information was visualized using Circos plots [67] (Figure 3A–D).

#### 4.6. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichment Analyses

Homozygous SNP and InDel variant genes within the GO database were annotated, and the number of differential genes in each GO entry was calculated. Statistical analysis was performed using hypergeometric tests. The annotation results were compared with the genomic background to screen GO entries that were significantly enriched in the variant genes. The calculated *p*-value was corrected through multiple hypothesis tests and then

used as a threshold for GO entries with significantly enriched differentially expressed genes with a  $q$ -value  $< 0.05$ . GO annotation results were visualized using Revigo [68]. Statistical analysis of genes with homozygous SNP and InDel specificity was conducted to determine the enriched KEGG pathways. The statistical principles and methods were similar to GO analysis, and significantly enriched pathways were identified through analysis. The KEGG annotation results were visualized using R language [69].

#### 4.7. GWAS Phenotype Enrichment Analysis

The GWAS phenotype enrichment analysis in comprehensive annotation of rice multi-omics data (CARMO), an integrated annotation platform for functional exploration of rice multi-omics data, is a method of gene functional enrichment analysis like GO enrichment and KEGG enrichment analysis, but the raw gene set (reference gene set) of GWAS phenotype enrichment analysis was collected from genes affected by a set of single nucleotide polymorphisms (SNPs) associated with a certain trait identified from genome-wide association studies (GWAS) and integrated in CARMO. For any input study or gene list, it will be apparent which genes participate in regulation of key agronomic traits [47]. In this study, the homo-InDel genes of HT2, HT4, HTRM2, and HTRM12 were used for Venn diagram analysis to identify the special homo-InDel genes in the four materials (Figure 7A,B). The special homo-InDel genes of each material were used for GWAS phenotype enrichment analysis by the web-based tool CARMO, and the output-enriched gene lists of each material were compiled into Table S12.

#### 4.8. Identification and Comparative Analysis of Copy Number Variation Intervals

Copy number variation is one of the major types of polymorphisms in rice genome research. In this study, to identify the CNV intervals of the test material relative to the HT2 reference genome, CNVkit software (version: 0.9.7) [52] was used to detect CNVs in our samples. In general, the whole genome of the test sample was divided into small windows of 1 Mb in the analysis (CNV region), and the ratio of read coverage of the test sample to the reference sample and the  $\log_2^{\text{ratio}}$  values were calculated to identify chromosomal segment gain or loss using the sliding window algorithm. The  $\log_2^{\text{ratio}} < -0.4$  or  $\log_2^{\text{ratio}} > 0.3$  was used as an indication of chromosomal segment loss or gain, respectively.

#### 4.9. Statistical Methods

The data in Figure 1 were analyzed and plotted using GraphPad Prism (version: 9.0.0) software (Boston, MA, USA). The normality of data was tested (Anderson–Darling test and D’Agostino and Pearson test), and Student’s  $t$ -test (two-tailed) was used for difference significance analysis. Specially for Figure 1K, the Welch’s  $t$ -test (two-tailed) was used for difference significance analysis because the variances (SDs) between two samples are not equal. For Figure 1H–J, as the data of FLY4 does not fit a normality, the Mann–Whitney U test (two-tailed) was used for difference significance analysis. ns presents no significant difference; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ .

In Figure 2, data analysis was performed using R language, and Student’s  $t$ -test (two-tailed) was used for difference significance analysis after normality of the data was tested (Anderson–Darling test and D’Agostino and Pearson test). Moreover, ns presents no significance; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ .

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25169012/s1>.

**Author Contributions:** Y.Z.: formal analysis, writing—original draft, visualization, investigation, methodology, and validation. A.D.: formal analysis, writing—original draft, writing—review and editing, visualization, investigation., validation, and funding acquisition. L.T.: formal analysis, writing—original draft, and methodology. G.Y.: formal analysis, writing—original draft, and visualization. L.L.: data curation, methodology, investigation, and visualization. Y.Y.: data curation and methodology. X.F.: investigation and data curation. H.Y.: investigation and data curation. H.L.:

resources, writing—review and editing, and funding acquisition. W.H.: resources and writing—review and editing. D.C.: conceptualization and resources. Z.S.: methodology. X.Z.: methodology. Y.H.: resources, conceptualization, supervision, funding acquisition, project administration, and writing—review and editing. S.T.: resources, conceptualization, supervision, funding acquisition, project administration, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Alliance of International Science Organizations, (Grant No. ANSO-PA-2023-09), the Sichuan Science and Technology Program (2023YFH0019, 2021YFH0052), the “One Belt and One Road” Project of the International Partners Program from the Bureau of International Cooperation, Chinese Academy of Science (151751KYSB20180006), the Chinese National Natural Science Foundation (Grant No. 31960068), the Wuhan Science and Technology Major Project “Development and application of polyploid rice” (2021–2025), the Technology Innovation Guidance Project of Science and Technology Department of Jiangxi Province (20212BDH81022), the Open Project Program of Crop Characteristic Resources Creation and Utilization Key Laboratory of Sichuan Province (No. myzdsys23-01), the Open Project Program (Grant No. SKL-KF202312) of State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, and the 2022 Sanya Science and Technology innovation project (2022KJCX05).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data generated during the current study are available in online repositories and the Supplementary Information Files. The genome sequencing data can be found in the National Center for Biotechnology Information (NCBI) Short Read Archive under accession number PRJNA1017988.

**Acknowledgments:** We thank Chuanmin Zeng and Zaishu Zhu (Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China) for their help with the field experiments. We thank Yuhua Liu (Wuhan Polyploid Biotechnology Limited Company, Wuhan, China) for his cooperation in cultivating “Y Liang You Duo Hui 14” (HTRM12) varieties, promoting the acquisition of certification of variety, and contributing in part to the field trial. We thank the Hubei Provincial Key Research and Development (R&D) Plan “Breeding of polyploid high protein and nutrient rich characteristic rice varieties” (2022032BBA). We thank Lanying Wang, Ting Zhang, and Sheng Pan (Shanghai OE Biotech Co., Ltd., Shanghai, China) for providing technical support for bioinformatics analysis. We thank Cui Yu, Yuhuan Geng, and Can Chen (Chengdu Institute of Biology, Chinese Academy of Sciences) for providing support and assistance to this work. We thank the cooperative groups for their close teamwork.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

## References

1. Mishra, P.; Devi, M.; Fatih, C.; Devi, M.; Williams, A.J. Forecasting of Rice Production using the Meteorological Factor in Major States in India and its Role in Food Security. *Int. J. Agric. Environ. Biotechnol.* **2021**, *14*, 51–62.
2. Khush, G.S. Green revolution: The way forward. *Nat. Rev. Genet.* **2001**, *2*, 815–822. [[CrossRef](#)]
3. Masterson, J. Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. *Science* **1994**, *264*, 421–424. [[CrossRef](#)]
4. Dubcovsky, J.; Dvorak, J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **2007**, *316*, 1862–1866. [[CrossRef](#)]
5. Doebley, J.; Bacigalupo, A.; Stec, A. Inheritance of Kernel Weight in Two Maize-Teosinte Hybrid Populations: Implications for Crop Evolution. *J. Hered.* **1994**, *85*, 191–195.
6. Begna, T. Effects of crop evolution under domestication and narrowing genetic bases of crop species. *Open J. Plant Sci.* **2021**, *6*, 49–54.
7. Yoo, M.J.; Liu, X.X.; Pires, J.C.; Soltis, P.S.; Soltis, D.E. Nonadditive Gene Expression in Polyploids. *Annu. Rev. Genet.* **2014**, *48*, 485–517. [[CrossRef](#)]
8. Fridman, E. Consequences of hybridization and heterozygosity on plant vigor and phenotypic stability. *Plant Sci.* **2015**, *232*, 35–40. [[CrossRef](#)]

9. Cai, D.T.; Yuan, L.P.; Lu, X.G. A new strategy of rice breeding in the 21st century: II. searching a new pathway of rice breeding by utilization of double heterosis of wide cross and polyploidization. *Acta Agron. Sin.* **2001**, *27*, 110–116.
10. Hattersley, P.W.; Watson, L. Diversification of photosynthesis. In *Chapman GP (ed) Grass Evolution and Domestication*; Cambridge University Press: Cambridge, UK, 1992.
11. International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **2018**, *361*, eaar7191. [[CrossRef](#)]
12. Nieto Feliner, G.; Casacuberta, J.; Wendel, J.F. Genomics of evolutionary novelty in hybrids and polyploids. *Front. Genet.* **2020**, *11*, 792. [[CrossRef](#)]
13. Koide, Y.; Kuniyoshi, D.; Kishima, Y. Fertile Tetraploids: New Resources for Future Rice Breeding? *Front. Plant Sci.* **2020**, *11*, 1231. [[CrossRef](#)]
14. Chen, R.R.; Feng, Z.Y.; Zhang, X.H.; Song, Z.J.; Cai, D.T. A New Way of Rice Breeding: Polyploid Rice Breeding. *Plants* **2021**, *10*, 422. [[CrossRef](#)]
15. Wang, N.N.; Fan, X.H.; Lin, Y.J.; Li, Z.; Wang, Y.K.; Zhou, Y.M.; Meng, W.L.; Peng, Z.W.; Zhang, C.Y.; Ma, J. Alkaline Stress Induces Different Physiological, Hormonal and Gene Expression Responses in Diploid and Autotetraploid Rice. *Int. J. Mol. Sci.* **2022**, *23*, 5561. [[CrossRef](#)]
16. Wu, J.W.; Hu, C.Y.; Shahid, M.Q.; Guo, H.B.; Zeng, Y.X.; Liu, X.D.; Lu, Y.G. Analysis on genetic diversification and heterosis in autotetraploid rice. *Springerplus* **2013**, *2*, 439. [[CrossRef](#)]
17. Tu, S.B.; Kong, F.L.; Xu, Q.F.; He, T. Study on new system of heterosis utilization in autotetraploid rice. *Bull. Chin. Acad. Sci.* **2003**, *6*, 426–428. (In Chinese)
18. Shahid, M.Q.; Sun, J.F.; Wei, C.M.; Zhang, P.; Liu, X.D. Studies on the Abnormality of Embryo Sac and Pollen Fertility in Autotetraploid Rice during Different Growing Seasons. *Pak. J. Bot.* **2010**, *42*, 7–19.
19. He, Y.C.; Wei, Q.O.; Ge, J.; Jiang, A.M.; Gan, L.; Song, Z.J.; Cai, D.T. Genome duplication effects on pollen development and the interrelated physiological substances in tetraploid rice with polyploid meiosis stability. *Planta* **2010**, *232*, 1219–1228. [[CrossRef](#)] [[PubMed](#)]
20. Wu, J.W.; Shahid, M.Q.; Chen, L.; Chen, Z.X.; Wang, L.; Liu, X.D.; Lu, Y.G. Polyploidy Enhances F1 Pollen Sterility Loci Interactions That Increase Meiosis Abnormalities and Pollen Sterility in Autotetraploid Rice. *Plant Physiol.* **2015**, *169*, 2700–2717. [[CrossRef](#)]
21. Li, X.; Yu, H.; Jiao, Y.M.; Shahid, M.Q.; Wu, J.W.; Liu, X.D. Genome-wide analysis of DNA polymorphisms, the methylome and transcriptome revealed that multiple factors are associated with low pollen fertility in autotetraploid rice. *PLoS ONE* **2018**, *13*, e0201854. [[CrossRef](#)]
22. Ghouri, F.; Zhu, J.N.; Yu, H.; Wu, J.W.; Baloch, F.S.; Liu, X.D.; Shahid, M.Q. Deciphering global DNA variations and embryo sac fertility in autotetraploid rice line. *Turk. J. Agric. For.* **2019**, *43*, 554–568. [[CrossRef](#)]
23. Tu, S.B.; Luan, L.; Liu, Y.H.; Long, W.B.; Kong, F.L.; He, T.; Xu, Q.F.; Yan, W.N.; Yu, M. Production and heterosis analysis of rice autotetraploid hybrids. *Crop Sci.* **2007**, *47*, 2356–2363. [[CrossRef](#)]
24. Cai, D.T.; Chen, J.G.; Chen, D.L.; Dai, B.C.; Zhang, W.; Song, Z.J.; Yang, Z.F.; Du, C.Q.; Tang, Z.Q.; He, Y.C.; et al. The breeding of two polyploid rice lines with the characteristic of polyploid meiosis stability. *Sci. China Ser. C Life Sci.* **2007**, *50*, 356–366. [[CrossRef](#)]
25. Guo, H.B.; Shahid, M.Q.; Zhao, J.; Li, Y.J.; Wang, L.; Liu, X.D. Agronomic Traits and Cytogenetic Evaluation of Newly Developed Autotetraploid Rice Line. *Pak. J. Agric. Sci.* **2016**, *53*, 291–301. [[CrossRef](#)]
26. Guo, H.B.; Mendrikahy, J.N.; Xie, L.; Deng, J.F.; Lu, Z.J.; Wu, J.W.; Li, X.; Shahid, M.Q.; Liu, X.D. Transcriptome analysis of neo-tetraploid rice reveals specific differential gene expressions associated with fertility and heterosis. *Sci. Rep.* **2017**, *7*, 40139. [[CrossRef](#)]
27. Bei, X.; Shahid, M.Q.; Wu, J.; Chen, Z.; Liu, X. Re-sequencing and transcriptome analysis reveal rich DNA variations and differential expressions of fertility-related genes in neo-tetraploid rice. *PLoS ONE* **2019**, *14*, e0214953.
28. Yu, H.; Shahid, M.Q.; Li, Q.H.; Li, Y.D.; Li, C.; Lu, Z.J.; Wu, J.W.; Zhang, Z.M.; Liu, X.D. Production Assessment and Genome Comparison Revealed High Yield Potential and Novel Specific Alleles Associated with Fertility and Yield in Neo-Tetraploid Rice. *Rice* **2020**, *13*, 32. [[CrossRef](#)]
29. Ghaleb, M.A.A.; Li, C.; Shahid, M.Q.; Yu, H.; Liang, J.H.; Chen, R.X.; Wu, J.W.; Liu, X.D. Heterosis analysis and underlying molecular regulatory mechanism in a wide-compatible neo-tetraploid rice line with long panicles. *BMC Plant Biol.* **2020**, *20*, 83. [[CrossRef](#)]
30. Zhang, X.H.; Zuo, B.; Song, Z.J.; Wang, W.; He, Y.C.; Liu, Y.H.; Cai, D.T. Breeding and study of two new photoperiod- and thermo-sensitive genic male sterile lines of polyploid rice (*Oryza sativa* L.). *Sci Rep.* **2017**, *7*, 14744. [[CrossRef](#)]
31. Chen, Y.; Shahid, M.Q.; Wu, J.W.; Deng, R.L.; Chen, Z.X.; Wang, L.; Liu, G.Q.; Zhou, H.; Liu, X.D. Thermo-Sensitive Genic Male Sterile Lines of Neo-Tetraploid Rice Developed through Gene Editing Technology Revealed High Levels of Hybrid Vigor. *Plants* **2022**, *11*, 1390. [[CrossRef](#)]
32. Bassil, N.V.; McCallum, S.; Graham, J.; Olmstead, J.W.; Itle, R.A.; Brown, A.; Buck, E.J.; Wiedow, C.; Finn, C.E.; Hancock, J.F.; et al. Progress toward a Universal Linkage Map in Tetraploid Highbush Blueberry. In Proceedings of the 2011 ASHS Annual Conference, Kohala, HI, USA, 25–28 September 2011.
33. Xiong, X.Y.; Xiao, X.C.; Lei, X.F. Breeding Utilization of the Diploid Rice in Autopolyploid Self Returning Progeny. *J. Anhui Agric. Sci.* **2016**, *44*, 28+51. (In Chinese)

34. Kamara, N.; Jiao, Y.M.; Lu, Z.J.; Aloryi, K.D.; Wu, J.W.; Liu, X.D.; Shahid, M.Q. Cytological Observations and Bulk-Segregant Analysis Coupled Global Genome Sequencing Reveal Two Genes Associated with Pollen Fertility in Tetraploid Rice. *Int. J. Mol. Sci.* **2021**, *22*, 841. [[CrossRef](#)]
35. Yu, H.; Li, Q.H.; Li, Y.D.; Yang, H.J.; Lu, Z.J.; Wu, J.W.; Zhang, Z.M.; Shahid, M.Q.; Liu, X.D. Genomics Analyses Reveal Unique Classification, Population Structure and Novel Allele of Neo-Tetraploid Rice. *Rice* **2021**, *14*, 16. [[CrossRef](#)]
36. Song, Z.J.; Du, C.Q.; Zhang, X.H.; Chen, D.L.; He, Y.C.; Cai, D.T. Studies on awns in polyploid rice (*Oryza sativa* L.) and preliminary cross experiments of a special awnless tetraploid rice. *Genet. Resour. Crop Evol.* **2014**, *61*, 797–807. [[CrossRef](#)]
37. Zhao, J.; Liu, X.; Wang, M.; Xie, L.J.; Wu, Z.X.; Yu, J.M.; Wang, Y.C.; Zhang, Z.Q.; Jia, Y.F.; Liu, Q.P. The miR528-D3 Module Regulates Plant Height in Rice by Modulating the Gibberellin and Abscisic Acid Metabolisms. *Rice* **2022**, *15*, 27. [[CrossRef](#)] [[PubMed](#)]
38. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[PubMed](#)]
39. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.Y.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)] [[PubMed](#)]
40. Shi, C.; Xu, L.L. Characters of Cysteine Endopeptidases in Wheat Endosperm during Seed Germination and Subsequent Seedling Growth. *J. Integr. Plant Biol.* **2009**, *51*, 52–57. [[CrossRef](#)] [[PubMed](#)]
41. Höwing, T.; Huesmann, C.; Hoefle, C.; Nagel, M.K.; Isono, E.; Hüchelhoven, R.; Gietl, C. Endoplasmic reticulum KDEL-tailed cysteine endopeptidase 1 of Arabidopsis (*AtCEP1*) is involved in pathogen defense. *Front. Plant Sci.* **2014**, *5*, 58. [[CrossRef](#)]
42. Yap, K.; Makeyev, E.V. Regulation of gene expression in mammalian nervous system through alternative pre-mRNA splicing coupled with RNA quality control mechanisms. *Mol. Cell. Neurosci.* **2013**, *56*, 420–428. [[CrossRef](#)]
43. Hunter, N.; Verma, V. Cell division | Meiosis. In *Encyclopedia of Biological Chemistry*, 3rd ed.; Elsevier: Amsterdam, the Netherlands, 2021; pp. 49–56.
44. Deamer, D.W. Bioenergetics and Primitive Metabolic Pathways. In *Assembling Life*; Oxford Academic: Oxford, UK, 2019.
45. Dong, N.Q.; Lin, H.X. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. *J. Integr. Plant Biol.* **2021**, *63*, 180–209. [[CrossRef](#)]
46. Kao, Y.T.; Gonzalez, K.L.; Bartel, B. Peroxisome Function, Biogenesis, and Dynamics in Plants. *Plant Physiol.* **2018**, *176*, 162–177. [[CrossRef](#)]
47. Wang, J.; Qi, M.; Liu, J.; Zhang, Y. CARMO: A Comprehensive Annotation Platform for Functional Exploration of Rice Multi-Omics Data. *Plant J.* **2015**, *83*, 359–374.
48. Huang, L.; Hua, K.; Xu, R.; Zeng, D.; Wang, R.; Dong, G.; Zhang, G.; Lu, X.; Fang, N.; Wang, D.; et al. The LARGE2-APO1/APO2 regulatory module controls panicle size and grain number in rice. *Plant Cell* **2021**, *33*, 1212–1228.
49. Zhao, K.; Tung, C.W.; Eizenga, G.C.; Wright, M.H.; Ali, M.L.; Price, A.H.; Norton, G.J.; Islam, M.R.; Reynolds, A.; Mezey, J. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2011**, *2*, 467.
50. Palacios, R.; Jáuregui, C.G.; Flores, M.; Palacios-Flores, K. Copy Number Variation. In *Reference Module in Life Sciences*; Elsevier: Amsterdam, The Netherlands, 2022. [[CrossRef](#)]
51. Cacabelos, R.; Tellado, I.; Cacabelos, P. The Epigenetic Machinery in the Life Cycle and Pharmacopigenetics. In *Pharmacopigenetics*; Cacabelos, R., Ed.; Academic Press: Cambridge, MA, USA, 2019; Volume 10, pp. 1–100.
52. Talevich, E.; Shain, A.H.; Botton, T.; Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **2016**, *12*, e1004873. [[CrossRef](#)]
53. Xing, S.; Zhou, K. Genetic study on the self-bred progenies of special autopolyploid rice. *J. Sichuan Agric. Univ.* **2000**, *18*, 308–310+314. (In Chinese)
54. Li, L.; Zhou, X.; Wang, J.; Mei, X.; Zhang, X.; Song, Z.; Cai, D. Analysis of the Grain Quality of Elite Diploid Rice Lines Reverted from Polyploid Rice. *Hans J. Agric. Sci.* **2021**, *11*, 184–190. (In Chinese)
55. Wendel, J.F. The wondrous cycles of polyploidy in plants. *Am. J. Bot.* **2015**, *102*, 1753–1756. [[CrossRef](#)]
56. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **2005**, *6*, 836–846. [[CrossRef](#)]
57. Zmienko, A.; Samelak, A.; Kozłowski, P.; Figlerowicz, M. Copy number polymorphism in plant genomes. *Theor. Appl. Genet.* **2014**, *127*, 1–18. [[CrossRef](#)]
58. Lye, Z.N.; Purugganan, M.D. Copy Number Variation in Domestication. *Trends Plant Sci* **2019**, *24*, 352–365. [[CrossRef](#)]
59. Luan, L.; Long, W.-B.; Wang, X.; Chen, Y.; Qiu, L.; Tu, S.-B.; Kong, F.-L.; Xiao, X.-Y. Research on creating auto-tetraploid rice using colchicum to induce the panicle callus of diploid rice. *J. Sichuan Univ. (Nat. Sci. Ed.)* **2009**, *46*, 829–837. (In Chinese)
60. *NY/T 1300 2007*; Technical Procedures for Rice Variety Trials. The Ministry of Agriculture of the People’s Republic of China: Beijing, China, 2007.
61. *NY/T 2646-2014*; Technical Specification for the Identification and Evaluation of Blast Resistance in Rice Variety Testing. The Ministry of Agriculture of the People’s Republic of China: Beijing, China, 2014.
62. *HNZ070-2015*; Technical Regulations of Bacterial Blight Resistance of Rice Varieties. Department of Agriculture and Rural Affairs of Hunan: Changsha, China, 2015.

63. Chen, S.F.; Zhou, Y.Q.; Chen, Y.R.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, 884–890. [[CrossRef](#)]
64. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
65. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
66. Okonechnikov, K.; Conesa, A.; García-Alcalde, F. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **2016**, *32*, 292–294. [[CrossRef](#)]
67. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)]
68. Supek, F.; Bonjak, M.; Kunca, N.; Muc, T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE* **2011**, *6*, e21800.
69. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS J. Integr. Biol.* **2012**, *16*, 284–287.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.