



Article

Introgression of a Complex Genomic Structural Variation Causes Hybrid Male Sterility in *GJ* Rice (*Oryza sativa* L.) Subspecies

Na Xu [†], Hai Xu [†], Zhengjin Xu, Fengcheng Li ^{*} and Quan Xu ^{*}

Rice Research Institute, Shenyang Agricultural University, Shenyang 110866, China

^{*} Correspondence: fengchengli@syau.edu.cn (F.L.); kobexu34@syau.edu.cn (Q.X.)

[†] These authors contributed equally to this work.

Abstract: Hybrids between different subspecies of rice *Oryza sativa* L. commonly show hybrid sterility. Here we show that a widely planted commercial *japonica/GJ* variety, DHX2, exhibited hybrid sterility when crossing with other *GJ* varieties. Using the high-quality genome assembly, we identified three copies of the *Sc* gene in DHX2, whereas Nipponbare (Nip) had only one copy of *Sc*. Knocking out the extra copies of *Sc* in DHX2 significantly improved the pollen fertility of the F₁ plant of DHX2/Nip cross. The population structure analysis revealed that a slight introgression from Basmati1 might occur in the genome of DHX2. We demonstrated that both DHX2 and Basmati1 harbored three copies of *Sc*. Moreover, the introgression of *GS3* and *BADH2/fgf* from Basmati1 confers the slender and fragrance grain of DHX2. These results add to our understanding of the hybrid sterility of inter-subspecies and intra-subspecies and may provide a novel strategy for hybrid breeding.

Keywords: *Oryza sativa* L.; hybrid sterility; copy number variation; DUF1618 protein; genome introgression



Citation: Xu, N.; Xu, H.; Xu, Z.; Li, F.; Xu, Q. Introgression of a Complex Genomic Structural Variation Causes Hybrid Male Sterility in *GJ* Rice (*Oryza sativa* L.) Subspecies. *Int. J. Mol. Sci.* **2022**, *23*, 12804. <https://doi.org/10.3390/ijms232112804>

Academic Editor: Deyong Ren

Received: 1 October 2022

Accepted: 22 October 2022

Published: 24 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

For over 2000 years, two major types of *O. sativa*—*O. sativa indica/Xian* (*XI*) and *O. sativa japonica/Geng* (*GJ*) Group—have historically been recognized [1]. There is great breeding potential in *XI/GJ* hybridization. Its hybrid F₁ has strong heterosis in yield, quality, and stress resistance [2]. However, due to the semi-sterile characteristics of the *XI/GJ* hybrid, *XI/GJ* heterosis has encountered great difficulties in practical application, including a series of problems, such as high plant height, difference in flowering time, and low seed setting rate. The most important one is the low seed setting rate (pollen fertility) of the F₁ generation of the *XI/GJ* hybrid [3]. Kato et al. (1928) showed that the average fertility of hybrids among cultivated rice varieties was more than 50%, while the fertility of *XI/GJ* F₁ Hybrid was between 0% and 33% [4]. A series of subsequent studies showed that different *XI/GJ* hybrid combinations contained different sterile loci. The fertility of F₁ generation of *XI/GJ* inter subspecies hybrid varied greatly from complete sterility to complete fertility, and the seed setting rate of the inter-subspecies hybrid was significantly lower than that of the intra-subspecies hybrid [5].

As early as 1962, Yang et al. proposed that *XI/GJ* hybrid rice can be used as a common strategy for rice breeding, and further explored the breeding methods regarding the improvement of seed setting rate of *XI/GJ* offspring, combination of the advantages of *XI* and *GJ* rice and heterosis utilization [6]. In 1987, Yuan et al. divided hybrid rice breeding into three development stages from the level of heterosis: intra-subspecies, inter-subspecies and utilization of distant heterosis, and strategically put forward the idea of hybrid rice breeding [7]. Because of the close relationship of intra-subspecies, the heterosis is limited in yield. The yield potential of an *XI/GJ* F₁ hybrid is expected to be an effective way to further improve the yield of hybrid rice. The main limitation of direct utilization of strong *XI/GJ* heterosis is hybrid sterility [8]. Tremendous efforts have been made to overcome

inter-subspecific hybrid sterility in rice, and genetic studies have identified approximately 50 loci involved in hybrid sterility [9]. The *XI/GJ* hybrid sterility is mainly affected by the genes at *Sc*, *S5*, *SA*, *hsa1*, *S7*, *dpl1/dpl2*, and *S27/S28* loci [10–15]. The isolation and cloning of these hybrid sterile genes have deepened the understanding of the molecular mechanism of plant reproductive isolation. The discovery of its genetic and molecular mechanism provides a new opportunity for overcoming hybrid sterility and utilizing heterosis.

Daohuaxiang2 (DHX2) is a large-scale variety planted in northern China. DHX2 is favored by the market because of its slender and fragrant grain. Thus, DHX2 was used as a backbone parent to breed new varieties in northern China (Figure 1a). During the breeding process, partial sterility due to the abortion of pollen was observed in the cross between DHX2 and other *GJ* varieties, whereas the F_1 pollen derived from the cross between DHX2 and *XI* varieties showed normal fertility. In this study, we demonstrated that the semi-sterility between DHX2 and *GJ* was caused by introgression of a copy number variation at the *Sc* locus from Basmati1. The F_1 hybrid sterility was rescued when knocking out an extra copy of *Sc* in DHX2 using CRSIPR gene editing technology. Our results identified a novel allele of *Sc* and will provide an effective approach to conduct crossbreeding using DHX2 as a parent line.

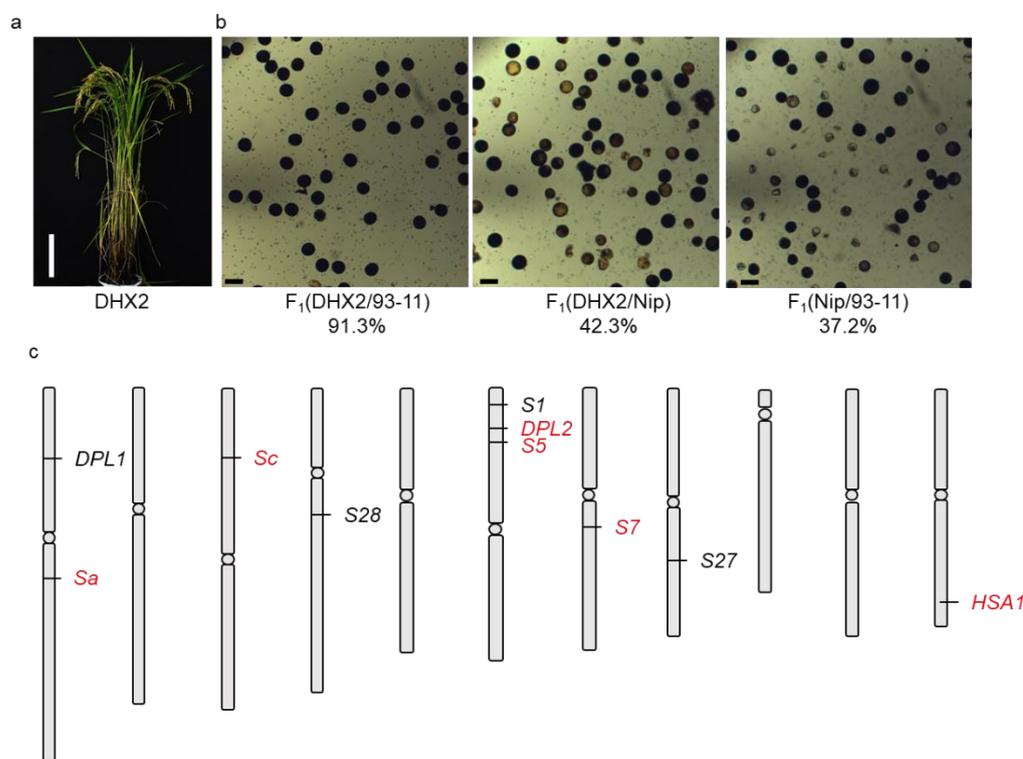


Figure 1. DHX2 exhibited hybrid sterility when crossing with other *GJ* varieties (a) The plant architecture of *GJ* variety DHX2. Scale bar: 20 cm; (b) The pollen phenotypes of F_1 plants derived from the cross of DHX2/93-11, DHX2/Nip, and Nip/93-11. Scale bar: 50 μ m; (c) The loci for hybrid sterility. The loci that have been cloned from *XI/GJ* cross were highlighted in red color.

2. Results

2.1. F_1 Hybrid Sterility between *Geng* Varieties

Rice hybrid sterility occurs extensively in the hybrid between *XI* and *GJ* [16]. For instance, the pollen fertility of the F_1 hybrid between *GJ* variety Nipponbare (Nip) and *XI* variety 93–11 was 37.2% (Figure 1b). DHX2 is a *GJ* variety with slender and fragrant grain and was widely planted in northern China. Interestingly, F_1 hybrid sterility was observed in the cross between DHX2 and other *GJ* varieties, such as Nip (Figure 1b). Moreover, the pollen fertility of the F_1 hybrid derived from the cross between DHX2 and an *XI* variety 93–11 was significantly improved compared to the cross between Nip and 93–11 (Figure 1b).

As a recent study has published the de novo genome assembly of DHX2 [17], we compared the sequence of published hybrid causal genes in an *XI/GJ* hybrid (Figure 1c). The result showed that the sequences of *Sa*, *DPL2*, *S5*, *S7*, and *HSA1* in DHX2 were identical to those of Nip. However, we found a complex structural variation at the *Sc* locus in DHX2 compared to Nip (Figure 2a).

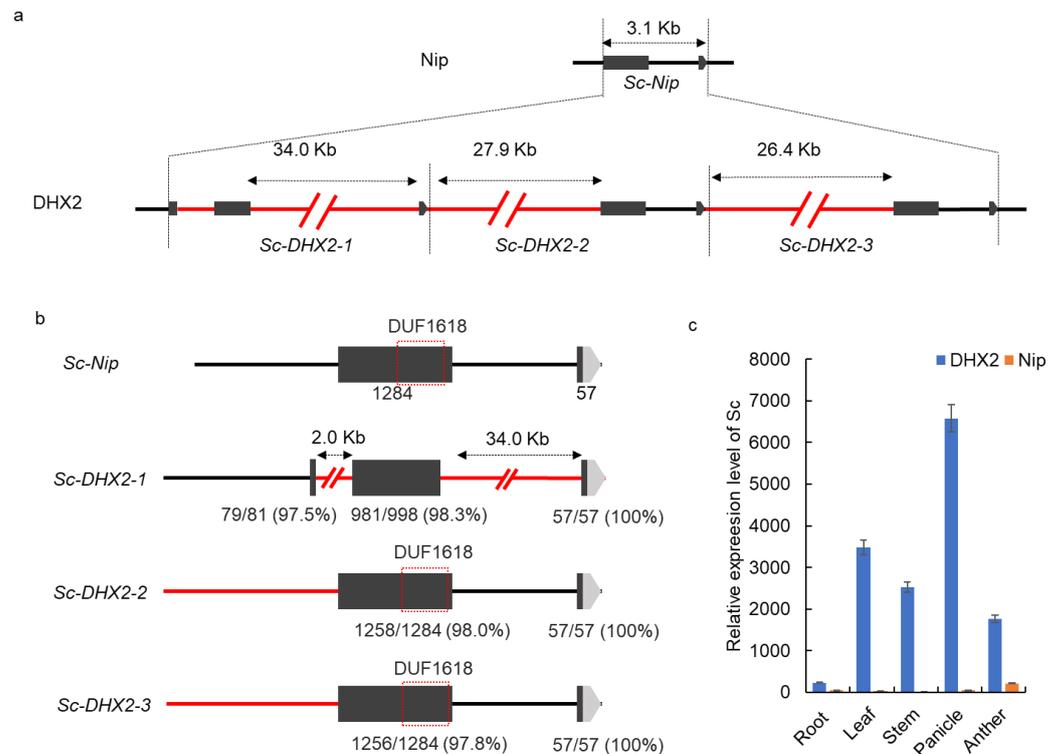


Figure 2. The structure variation of DHX2 at the *Sc* locus. (a) The copy number variation of the *Sc* locus in DHX2 and Nip; (b) the sequence variation of three copies of *Sc* in DHX2. The red line indicates the sequence of DHX2 is different from Nip; (c) the expression level of *Sc* on various tissues of DHX2 and Nip.

2.2. Complex Genomic Structural Variation in the *Sc*-DHX2 Alleles

The copy number variation at the *Sc* locus was reported to confer the *XI/GJ* hybrid male sterility [15]. Normally, the *GJ* allele (*Sc*-Nip) contains a pollen-essential gene encoding a DUF1618 domain protein, whereas the *XI* allele harbors at least two extra copies besides *Sc*-Nip. Here, we found that there were three tandem duplicated segments, and each of them contained a copy of *Sc* homolog protein (*Sc*-DHX2-1, *Sc*-DHX2-2

, and *Sc*-DHX2-3) in DHX2 (Figure 3a). The *Sc*-DHX2-1 shared a similar promoter sequence to *Sc*-Nip, but lost the DUF1618 domain due to the 2.0 kb insertion in the first exon and a 34.0 kb insertion at the intron of *Sc*-DHX2-1. Thus, we concluded that the *Sc*-DHX2-1 is a pseudogene. The other two copies, *Sc*-DHX2-2 and *Sc*-DHX2-3, each contain the entire sequence of *Sc*-Nip. However, there are several SNPs in the exon of *Sc*-DHX2-2 and *Sc*-DHX2-3 compared to that of *Sc*-Nip. The first exon of *Sc*-Nip contains 1284 bp, the first exon of *Sc*-DHX2-2 was 98.0% (1258/1284), identical to *Sc*-Nip, and the first exon of *Sc*-DHX2-3 was 97.8% (1256/1284), identical to *Sc*-Nip (Figure 3b). We then checked whether the extra copies of *Sc* were altering the expression pattern and expression level of *Sc*. The RT-PCR showed that *Sc* was specifically expressed in anther with a low level in Nip, whereas, they were broadly in leaf, stem, panicle, and anther, with a significantly higher level in DHX2 compared to Nip (Figure 3c). Thus, we hypothesized that the copy number variation at the *Sc* locus in DHX2 conferred the *F*₁ hybrid sterility when DHX2 was crossed to other typical *GJ* varieties.

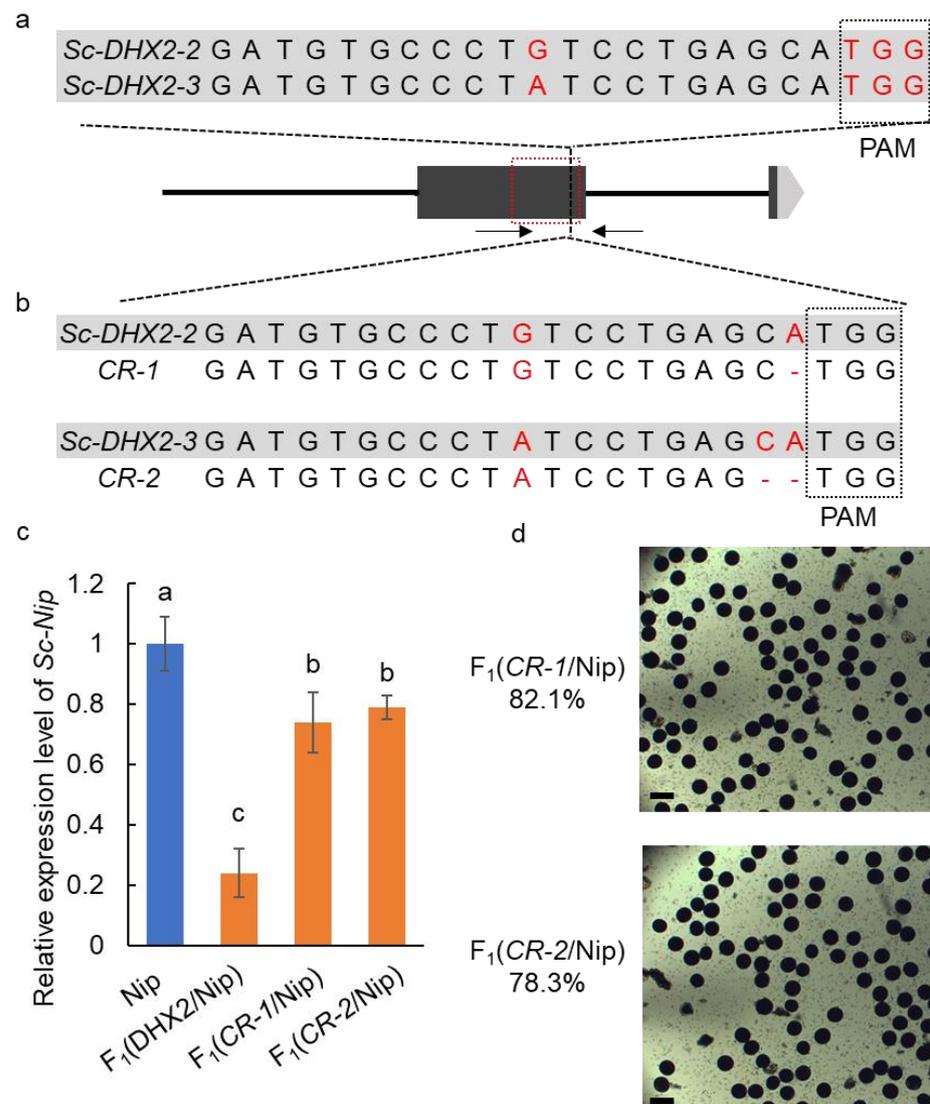


Figure 3. The CRISPR/Cas9 gene editing of the extra copy of *Sc* in *DHX2*. (a) The sgRNA sequence of *Sc-DHX2-2* and *Sc-DHX2-3*. The red dash line box indicates the DUF1618 domain; (b) the sequence of two mutant lines; (c) the expression level of *Sc-Nip* in the anthers of the F₁ plants. Different letters denote significant differences ($p < 0.05$) from Duncan's multiple range tests; (d) the pollen phenotypes of F₁ plants derived from the cross of *CR-1/Nip* and *CR-2/Nip*. Scale bar: 50 μ m.

2.3. Knockout of *Sc-DHX2-2* or *Sc-DHX2-3* Rescues the F₁ Hybrid Sterility

To confirm whether the extra copies of *Sc-DHX2-2* and *Sc-DHX2-3* caused the F₁ hybrid sterility of the cross between *DHX2* and other *GJ* varieties, we used the CRISPR/cas9 plant genome editing system to knockout the *Sc-DHX2-2* and *Sc-DHX2-3* in *DHX2*. As there are two SNPs between *Sc-DHX2-2* and *Sc-DHX2-3*, we designed the different sgRNA for *Sc-DHX2-2* and *Sc-DHX2-3* based on the SNPs (Figure 3a). By transforming *DHX2* with these two constructs, we successfully identified a plant *CR-1* with 1bp deletion at the first exon of *Sc-DHX2-2*, and a plant *CR-2* with 2 bp at the at first exon of *Sc-DHX2-3*. The homozygous mutants of *Sc-DHX2-2* and *Sc-DHX2-3* did not exhibit a significant difference of agronomic traits compared to *DHX2*. We then crossed *CR-1* and *CR-2* with *Nip* to generate F₁ plants (Figure 3b). The *Sc-Nip* expression level of F₁ (*CR-1/Nip*) and F₁ (*CR-2/Nip*) was significantly higher than that of F₁ (*DHX2/Nip*) (Figure 3c), and the pollen fertility of these F₁ plants was significantly improved (Figure 3d).

2.4. Introgression from Basmati Variety

As most GJ varieties only have a single copy of *Sc* [15], we hypothesized that the copy number variation of the *Sc* locus in DHX2 was caused by introgression from XI or other subspecies. We first compared the genome of DHX2 to Nip and identified 16,873 SVs (Figure 4a). Then we collected 58 long-read *de novo* assemblies to conduct a population structure analysis [17–19]. The result showed that the slight introgression from the circum-Basmati group (cB) might occur in the genome of DHX2 (Figure 4b). The cB group comprises the famous Basmati and Sadri aromatic varieties which we term the circum-Basmati group (cB) [1]. Basmati1 is a high-yield Basmati variety, that possesses extra-long slender grains, a pleasant aroma, appealing taste, good mouthfeel, and easy digestibility [20]. Considering that DHX2 exhibited similar characteristics to Basmati1, such as slender and fragrance grain, we speculated that DHX2 inherited these characteristics from Basmati1. As the truncated GS3 protein contributed to the slender grain of Basmati1 [21], and the fragrance of Basmati1 was regulated by gene *BADH2/fgr* [22,23], we conducted the haplotype network analysis of *Sc*, *GS3*, and *BADH2/fgr* using sequence data of the 3000 Rice Genomes Project [24]. The results indicated that *Sc*, *GS3*, and *BADH2/fgr* differentiated among cB, XI, and GJ groups (Figure 4c). We subsequently compared the sequence of the *Sc* locus between DHX2 and Basmati1. The result showed that both DHX2 and Basmati1 had three copies of *Sc*, although the interval of the three copies was different between DHX2 and Basmati1 (Figure 5a). The *BADH2/fgr* of DHX2 was identical to that of Basmati1, which was different to other GJ variety, such as Nip (Figure 5b). We subsequently compared the sequencing of *GS3* between DHX2 and Basmati1. The result exhibited that both DHX2 and Basmati1 shared a C/A SNP compared to Nip. The C/A SNP generated a premature stop codon, which caused a frameshift mutation in the C terminus that yielded a truncated protein of *GS3* (Figure 5c).

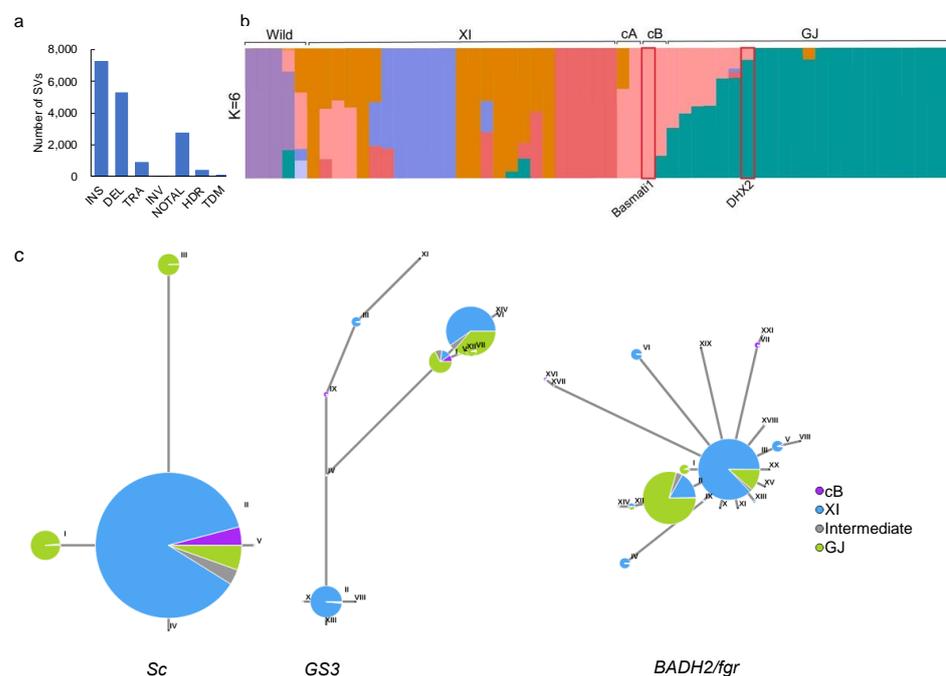


Figure 4. The population structure analysis of 58 assemblies. (a) The different types of SVs between DHX2 and Nip. INS: insertion, DEL: deletion, TRA: translocation, INV: inversion, NOTAL: not aligned region, TDM: tandem repeat, HDR: highly diverged regions; (b) Structure analysis of 58 accessions with different numbers of clusters, $K = 6$. The different colors represent different groups of rice accessions. (c) Haplotype network of *Sc*, *GS3*, and *BADH2/fgr* using sequence data of the 3000 Rice Genomes Project. Circle size is proportional to the sample number for a given haplotype. Different colors represent ecotypes, as shown in the illustration. Different Roman numerals indicate various haplotypes of *Sc*, *GS3*, and *BADH2/fgr*.

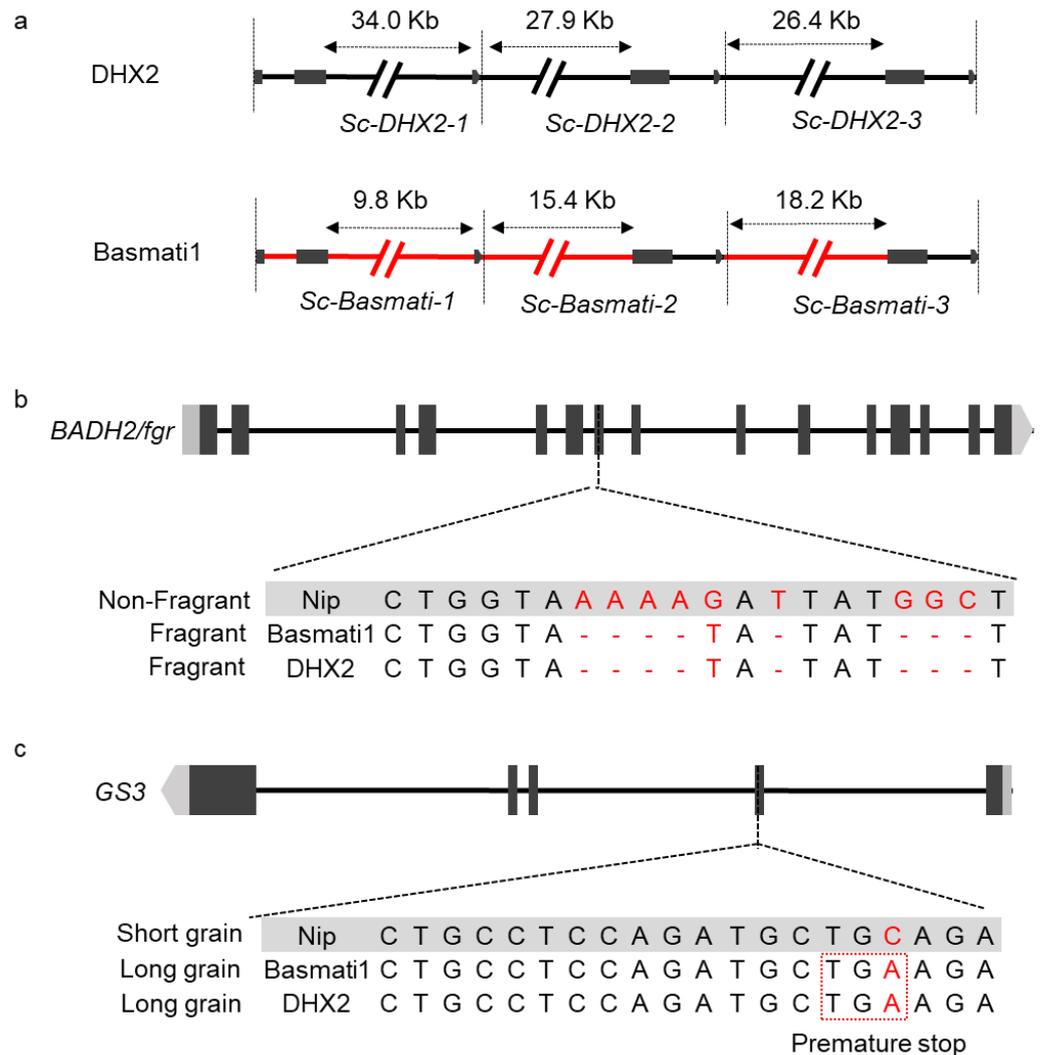


Figure 5. The introgression loci in DHX2 from Basmati1. (a) The structural variation of the *Sc* locus in DHX2 and Basmati1. The red line indicates that the sequence of Basmati1 is different from DHX2; (b) the sequence variation of *BADH2/fgr* in DHX2, Nip, and Basmati1; (c) the sequence variation of the *GS3* locus in DHX2, Nip, and Basmati1.

3. Discussion

A combination of the advantages of *XI* and *GJ* rice through the *XI/GJ* crossbreeding was the basic breeding strategy in northern China. Hybridization between *XI* and *GJ* rice combined with the utilization of the ideal plant type has led to the development of high-yielding *GJ* rice in northern China. Crossbreeding causes genome introgression from *XI* and *cB* varieties. Our previous study revealed that the *XI* pedigree introgression frequencies were significantly increased in cultivars bred after 1990 and the *XI* pedigree introgression frequencies were significantly positively correlated with grain number per panicle [25]. In an investigation of 1200 Chinese accessions, an average of 6.8 Mb *XI* genome introgression was found in *GJ* accessions [26]. Notably, the important gene haplotypes controlling plant architecture, yield components, and pest and disease resistance, including *IPA1*, *SMG1*, *DEP3*, *Pib*, *Pi-d2*, and *Bph3*, were introduced from *XI* rice to *GJ* by introgression [27]. In this study, we found that there is *cB* introgression in some Chinese *GJ* varieties, such as DHX2. The introgression of *Sc*, *GS3*, and *BADH2/fgr* from *cB* variety Basmati1 caused the special characteristics of DHX2 in *GJ/GJ* hybrid sterility, slender grain shape, and fragrance.

Copy number variations are widely distributed in plant genomes [28]. Recently reported pan-genomes have revealed hidden copy number variations and demonstrated

that copy number variations regulate important agronomic traits [17,29,30]. The copy number variation at the *GL7* locus contributes to the grain size diversity in rice [31], the extra copy number at the *GNP1* locus significantly increased the grain number per panicle [30], the dual copies of *OsMADS18* are likely a causal candidate accounting for the early flowering phenotype Koshihikari [17]. However, only the copy number variation of *Sc* differentiated between *XI* and *GJ* subspecies [15]. Our study detected a *GJ* variety DHX2 harboring three tandem copies at the *Sc* locus, which might be caused by the introgression of Basmati1. Moreover, the sequence of the extra copies of *Sc* showed a slight difference compared to the *XI* allele [15]. These findings suggested that the presence of rich copy number variation is a treasure of the total genetic diversity of *O.sativa*, revealing that the copy number variation originated from the intricate breeding history.

DUF1618 is a new gene family that originated after the dicot–monocot divergence. DUF1618 family members in plants possess a 56–199-amino acid conserved domain, and there are 121 DUF1618 genes identified in the rice genome [32]. Recent studies have demonstrated that the DUF1618 gene family is involved in the regulation of hybrid sterility. The *GJ* type allele of *HSA1a* encodes a highly conserved plant-specific domain of DUF1618, whereas the *XI* type allele has two deletion mutations that cause disruption of domain structure. The recombinant haplotype of *HSA1a* and *HSA1b* caused semi-sterility [14]. Shen et al. (2017) genotyped 14 *GJ* cultivars and 21 *XI* cultivars and found that all tested *GJ* type alleles of *Sc* contain a pollen-essential gene encoding a DUF1618 domain, and all tested *XI* type alleles contain two or three tandem duplicated *Sc* segments. The genomic structural variation at the *Sc* locus causes hybrid male sterility between *XI* and *GJ* [15]. Our study found that there are copy number variations at the *Sc* locus in *GJ* variety DHX2 and the extra two copies of *Sc* cause hybrid male sterility when DHX2 is crossed to other *GJ* varieties. However, only these two genes have been functionally studied among 121 DUF1618 gene families. The function of the other 119 DUF1618 genes needs to be further studied.

4. Materials and Methods

4.1. Plant Materials

In this study, the *GJ* variety DHX2, *GJ* variety Nipponbare (Nip), and *XI* variety 93-11, and the F₁ plants were employed. Shenyang Agricultural University's Rice Research Institute (41° N, 123° E) was used to conduct field tests. The seeds were sowed on April 16 and transplanted to the field on May 22 in 2020. Each line was planted in three rows, with 10 plants per row and a 30 cm × 13.3 cm plant spacing.

4.2. Pollen Fertility Test

Five young panicles of each F₁ plant were sampled. Pollen fertility was evaluated using pollen grains stained with I₂KI solution. We counted 500 pollen to distinguish the stained/unstained pollen grains for each young panicle under a microscope. The average data of five young panicles was presented in the Figures 1 and 3.

4.3. Expression Analysis

The pollen was sampled according to the method described previously [15]. The primer for the RT-PCR of *Sc* and *Actin* was listed in Table S1. The investigation was conducted by three biological replicates, and the significance was analyzed by Duncan's multiple range tests.

4.4. Vector Construction and Plant Transformation

To conduct the CRISPR/Cas9 gene editing, the vector construction was performed as described by Li et al. [33]. We designed the specific single-guide RNA (sgRNA) sequences targeting the *Sc-DHX2-2* and *Sc-DHX2-3* copies. The specificity of the targeting sequence was confirmed by BLAST searching against the Nip genome [34]. The rice transformation was conducted as described elsewhere [35]. We extracted the genomic DNA from transformants, and the genomic DNA was sequenced for mutant identification. The PCR

products (200–500 bp) were sequenced and identified using the degenerate sequence decoding method [36]. We examined 20 independent transgenic plants for each construction at T₀ generation. The sequence analysis detected 12 and 9 plants with mutations of *Sc-DHX2-2* and *Sc-DHX2-3*, respectively. Among these mutants, 3 and 4 plants were putative homozygous mutants, of *Sc-DHX2-2* and *Sc-DHX2-3*, respectively.

4.5. Population Structure Analysis

We collected 55 de novo assembled genomes of wild type, *XI, GJ*, circum-Aus group, (cA) which encompasses the Aus, Boro, and Rayada ecotypes from Bangladesh and India, and cB, based on the long-read sequencing techniques [17–19]. After discarding duplicated assemblies, a total of 57 assembled genomes were compared to the genome of Os-Nipponbare-Reference-IRGSP-1.0 (<https://rapdb.dna.affrc.go.jp/download/irgsp1.html>) (Kawahara et al., 2013) using MUMmer (v 4.0) [37]. After filtering the low-quality structure variations using metrics of minor allele frequency (MAF) > 0.05' from the raw SV dataset, we retained 156,319 high-confidence SVs for population structure analysis. Detailed information on SVs was described in our previous study [30]. The population genetic structure was examined using the program ADMIXTURE (v1.23) [38] with K values (the putative number of populations) from 2 to 10. The K = 6 values were chosen to display the genetic admixtures of rice populations. A haplotype network of *Sc, GS3*, and *BADH2/fgf* was conducted using the tools of Haplotype Network Analysis on RiceVarMap v2.0 (http://ricevarmap.ncpgr.cn/hap_net/ (accessed on 22 September 2022)) [39].

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232112804/s1>.

Author Contributions: Conceptualization, Q.X. and Z.X.; investigation, N.X., H.X. and F.L.; data curation, F.L.; writing—original draft preparation, Q.X. and F.L.; writing—review and editing, Q.X.; supervision, Q.X.; project administration, Q.X.; funding acquisition, Q.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 32071982.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study did not report any data.

Acknowledgments: We thank Zhibo Cui at Rice Research Institute of Shenyang Agricultural University for the technical support of CRISPR/Cas9 gene editing. We thank Jian Sun at Rice Research Institute of Shenyang Agricultural University for the technical support of haplotype analysis.

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

References

1. Wang, W.; Mauleon, R.; Hu, Z.; Chebotarov, D.; Tai, S.; Wu, Z.; Li, M.; Zheng, T.; Fuentes, R.R.; Zhang, F.; et al. Genomic variation in 3010 diverse accessions of Asian cultivated rice. *Nature* **2018**, *557*, 43–49. [[CrossRef](#)] [[PubMed](#)]
2. Huang, X.; Yang, S.; Gong, J.; Zhao, Q.; Feng, Q.; Zhan, Q.; Zhao, Y.; Li, W.; Cheng, B.; Xia, J. Genomic architecture of heterosis for yield traits in rice. *Nature* **2016**, *537*, 629–633. [[CrossRef](#)] [[PubMed](#)]
3. Ouyang, Y.; Liu, Y.G.; Zhang, Q. Hybrid sterility in plant: Stories from rice. *Curr. Opin. Plant Biol.* **2010**, *13*, 186–192. [[CrossRef](#)] [[PubMed](#)]
4. Kato, A. On the affinity of rice varieties as shown by the fertility of rice plants. *Centr. Agric. Inst. Kyushu Imp. Univ.* **1928**, *2*, 241–279.
5. Liu, K.D.; Zhou, Z.Q.; Xu, C.G.; Zhang, Q.; Maroof, M.A.S. An analysis of hybrid sterility in rice using a diallel cross of 21 parents involving indica, japonica and wide compatibility varieties. *Euphytica* **1996**, *90*, 275–280. [[CrossRef](#)]
6. Yang, S.; Shen, X.; Gu, W.; Cao, D. The breeding research of cross breeding between indica and japonica. *Acta Agron. Sin.* **1962**, *1*, 97–102.
7. Yuan, L. The strategy of hybrid rice breeding. *Hybrid Rice* **1987**, *1*, 1–3.

8. Zhang, G.Q. Prospects of utilization of inter-subspecific heterosis between indica and japonica rice—ScienceDirect. *J. Integr. Agric.* **2020**, *19*, 1–10. [[CrossRef](#)]
9. Guo, J.; Xu, X.; Li, W.; Zhu, W.; Zhu, H.; Liu, Z.; Luan, X.; Dai, Z.; Liu, G.; Zhang, Z. Overcoming inter-subspecific hybrid sterility in rice by developing indica-compatible japonica lines. *Sci. Rep.* **2016**, *6*, 26878. [[CrossRef](#)]
10. 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](#)]
11. 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](#)] [[PubMed](#)]
12. Yamagata, Y.; Yamamoto, E.; Aya, K.; Win, K.T.; Doi, K.; Sobrizal; Ito, T.; Kanamori, H.; Wu, J.; Matsumoto, T.; et al. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1494–1499. [[CrossRef](#)] [[PubMed](#)]
13. 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](#)]
14. Kubo, T.; Takashi, T.; Ashikari, M.; Yoshimura, A.; Kurata, N. Two Tightly Linked Genes at the hsa1 Locus Cause Both F1 and F2 Hybrid Sterility in Rice. *Mol. Plant* **2016**, *9*, 221–232. [[CrossRef](#)]
15. 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](#)]
16. Xie, Y.; Shen, R.; Chen, L.; Liu, Y.G. Molecular mechanisms of hybrid sterility in rice. *Sci. China. Life Sci.* **2019**, *62*, 737–743. [[CrossRef](#)]
17. Qin, P.; Lu, H.; Du, H.; Wang, H.; Chen, W.; Chen, Z.; He, Q.; Ou, S.; Zhang, H.; Li, X.; et al. Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell* **2021**, *184*, 3542–3558.e16. [[CrossRef](#)]
18. Stein, J.C.; Yu, Y.; Copetti, D.; Zwickl, D.J.; Zhang, L.; Zhang, C.; Chougule, K.; Gao, D.; Iwata, A.; Goicoechea, J.L.; et al. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nat. Genet.* **2018**, *50*, 285–296. [[CrossRef](#)]
19. Zhang, F.; Xue, H.; Dong, X.; Li, M.; Zheng, X.; Li, Z.; Xu, J.; Wang, W.; Wei, C. Long-read sequencing of 111 rice genomes reveals significantly larger pan-genomes. *Genome Res.* **2022**, *32*, 853–863. [[CrossRef](#)]
20. Singh, V.; Singh, A.K.; Mohapatra, T.; Krishnan, S.G.; Ellur, R.K. Pusa Basmati 1121—a rice variety with exceptional kernel elongation and volume expansion after cooking. *Rice* **2018**, *11*, 19. [[CrossRef](#)]
21. Mao, H.; Sun, S.; Yao, J.; Wang, C.; Yu, S.; Xu, C.; Li, X.; Zhang, Q. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19579–19584. [[CrossRef](#)]
22. Chen, S.; Yang, Y.; Shi, W.; Ji, Q.; He, F.; Zhang, Z.; Cheng, Z.; Liu, X.; Xu, M. Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. *Plant Cell* **2008**, *20*, 1850–1861. [[CrossRef](#)] [[PubMed](#)]
23. Kovach, M.J.; Calingacion, M.N.; Fitzgerald, M.A.; McCouch, S.R. The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14444–14449. [[CrossRef](#)] [[PubMed](#)]
24. Li, J.Y.; Wang, J.; Zeigler, R.S. The 3000 rice genomes project. *GigaScience* **2014**, *3*, 7. [[CrossRef](#)]
25. Sun, J.; Liu, D.; Wang, J.; Ma, D.; Tang, L.; Gao, H.; Xu, Z.; Chen, W. The contribution of intersubspecific hybridization to the breeding of super-high-yielding japonica rice in northeast China. *Theor. Appl. Genet.* **2012**, *125*, 1149–1157. [[CrossRef](#)]
26. Chen, Z.; Li, X.; Lu, H.; Gao, Q.; Du, H.; Peng, H.; Qin, P.; Liang, C. Genomic atlases of introgression and differentiation reveal breeding footprints in Chinese cultivated rice. *J. Genet. Genom. = Yi Chuan Xue Bao* **2020**, *47*, 637–649. [[CrossRef](#)]
27. Cui, D.; Zhou, H.; Ma, X.; Lin, Z.; Sun, L.; Han, B.; Li, M.; Sun, J.; Liu, J.; Jin, G.; et al. Genomic insights on the contribution of introgressions from Xian/Indica to the genetic improvement of Geng/Japonica rice cultivars. *Plant Commun.* **2022**, *3*, 100325. [[CrossRef](#)]
28. Lye, Z.N.; Purugganan, M.D. Copy Number Variation in Domestication. *Trends Plant Sci.* **2019**, *24*, 352–365. [[CrossRef](#)]
29. Shang, L.; Li, X.; He, H.; Yuan, Q.; Song, Y.; Wei, Z.; Lin, H.; Hu, M.; Zhao, F.; Zhang, C.; et al. A super pan-genomic landscape of rice. *Cell Res.* **2022**, *32*, 878–889. [[CrossRef](#)]
30. Wang, Y.; Li, F.; Zhang, F.; Wu, L.; Xu, N.; Sun, Q.; Chen, H.; Yu, Z.; Lu, J.; Jiang, K.; et al. Time-ordering japonica/geng genomes analysis indicates the importance of large structural variants in rice breeding. *Plant Biotechnol. J.* **2022**. *early view*. [[CrossRef](#)]
31. Wang, Y.; Xiong, G.; Hu, J.; Jiang, L.; Yu, H.; Xu, J.; Fang, Y.; Zeng, L.; Xu, E.; Xu, J. Copy number variation at the GL7 locus contributes to grain size diversity in rice. *Nat. Genet.* **2015**, *47*, 944–948. [[CrossRef](#)] [[PubMed](#)]
32. Wang, L.; Shen, R.; Chen, L.T.; Liu, Y.G. Characterization of a novel DUF1618 gene family in rice. *J. Integr. Plant Biol.* **2014**, *56*, 151–158. [[CrossRef](#)] [[PubMed](#)]
33. Li, M.; Chen, L.; Tian, S.; Lin, Y.; Tang, Q.; Zhou, X.; Li, D.; Yeung, C.K.L.; Che, T.; Jin, L.; et al. Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies. *Genome Res.* **2017**, *27*, 865–874. [[CrossRef](#)] [[PubMed](#)]
34. Hsu, P.D.; Scott, D.A.; Weinstein, J.A.; Ran, F.A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E.J.; Wu, X.; Shalem, O. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **2013**, *31*, 827–832. [[CrossRef](#)] [[PubMed](#)]

35. Nishimura, A.; Aichi, I.; Matsuoka, M. A protocol for *Agrobacterium*-mediated transformation in rice. *Nat. Protoc.* **2006**, *1*, 2796. [[CrossRef](#)]
36. Ma, X.; Zhang, Q.; Zhu, Q.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.; Li, H.; Lin, Y. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* **2015**, *8*, 1274–1284. [[CrossRef](#)]
37. Marçais, G.; Delcher, A.L.; Phillippy, A.M.; Coston, R.; Salzberg, S.L.; Zimin, A. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **2018**, *14*, e1005944. [[CrossRef](#)]
38. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)]
39. Zhao, H.; Li, J.; Yang, L.; Qin, G.; Xia, C.; Xu, X.; Su, Y.; Liu, Y.; Ming, L.; Chen, L.L.; et al. An inferred functional impact map of genetic variants in rice. *Mol. Plant* **2021**, *14*, 1584–1599. [[CrossRef](#)]