

## RESEARCH ARTICLE

Allelic variations of the *Wx* locus in cultivated rice and their use in the development of hybrid rice in ChinaYe Shao<sup>1,2</sup>\*, Yan Peng<sup>2</sup>\*, Bigang Mao<sup>2</sup>, Qiming Lv<sup>2</sup>, Dingyang Yuan<sup>2</sup>, Xionglun Liu<sup>1\*</sup>, Bingran Zhao<sup>1,2\*</sup>**1** College of Agronomy, Hunan Agricultural University, Changsha, China, **2** State Key Laboratory of Hybrid Rice, Hunan Hybrid Rice Research Centre, Changsha, China

\* These authors contributed equally to this work.

\* [xionglun@hunau.edu.cn](mailto:xionglun@hunau.edu.cn) (XL); [brzhao123@163.com](mailto:brzhao123@163.com) (BZ)

## Abstract

To make better use of global germplasm resources for improving the eating quality of hybrid rice, using the resequencing data from the 3,000 rice genomes project (3K RGP), the allelic variations of the rice *Wx* locus were analysed. With the exception of five rare alleles discovered for the first time in our study, most of these alleles were known alleles of *Wx*. Furthermore, a set of Kompetitive allele-specific PCR (KASP) markers based on these *Wx* alleles have been developed, and thirty-six main parents of hybrid rice from 1976 to 2018 were selected for *Wx* genotyping. The results showed that only three *Wx* alleles existed in the main parents of hybrids, and the allelic combination of the hybrids changed from  $Wx^a/Wx^b$  and  $Wx^v/Wx^b$  to  $Wx^b/Wx^b$  with the development of hybrid rice.  $Wx^b$  is widely used in the male parents of hybrid rice.  $Wx^a$  and  $Wx^v$  were used in the female parents of early hybrid rice, and they were gradually replaced by  $Wx^b$ . In the future, more favourable *Wx* alleles from cultivated rice should be identified, introduced, and effectively used to improve hybrid rice quality.

## OPEN ACCESS

**Citation:** Shao Y, Peng Y, Mao B, Lv Q, Yuan D, Liu X, et al. (2020) Allelic variations of the *Wx* locus in cultivated rice and their use in the development of hybrid rice in China. PLoS ONE 15(5): e0232279. <https://doi.org/10.1371/journal.pone.0232279>

**Editor:** Prasanta K. Subudhi, Louisiana State University College of Agriculture, UNITED STATES

**Received:** January 27, 2020

**Accepted:** April 10, 2020

**Published:** May 5, 2020

**Copyright:** © 2020 Shao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was financially supported by funds from the Pre-cultivation Project of National Biological Seed Industry Technology Innovation Center (2018XK2005), the Key Project of Hunan Province (2018NK1020-1), National Natural Science Foundation of China U19A2032, and the National Natural Science Foundation of China 31171526. The funders had no role in study

## Introduction

Rice is one of the most important crops with the great success in heterosis applications. For a long time, considerable attention has been given to improve yield in hybrid rice. But research on rice quality started fairly late in China, especially in southern areas. With the improvement in the standard of living, consumers are more concerned about rice quality. Over the past 15 years, the quality traits of *indica* hybrid rice have improved to a certain extent. However, further improvements are still required to make hybrid rice comparable to high-quality conventional rice, especially in terms of eating quality [1,2]. As many widely used hybrid parents are derived from the same basic materials, the genetic polymorphisms in hybrid rice may be relatively low [3]. Thus, an improved understanding the diversity of genes associated with eating quality in cultivated rice and hybrid rice is needed for improving the quality of hybrids.

Rice quality is a complex characteristic, with amylose content (AC) serving as a key determinant of eating and cooking quality [4,5]. The *Wx* gene, which encodes a granule-bound starch synthase (OsGBSS1), is a critical gene controlling AC. The transcription and OsGBSS1

design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** AC, amylose content; cA, circum-Aus; CAM, Central America and Caribbean; cB, circum-Basmati; CNCGB, China National Crop Genebank; CT, Cytosine and thymidine; dCAPS, derived cleaved amplified polymorphic sequence; EAF, East Africa; EAR, East Asia Islands; EAS, East Asia; EUR, Europe; GJ, *Geng/Japonica*; IOC, Indian Ocean; IRGC, International Rice Genebank Collection; KASP, Kompetitive allele-specific PCR; NAF, North Africa; NAM, North America; NTC, non-template control; OCE, Oceania; OsGBSS1, granule-bound starch synthase; RFGB, Rice Functional Genomics-based Breeding; SAC, South Asia–Central; SAE, South Asia–East; SAM, South America; SAW, South Asia–West; SEA, Southeast Asia; SER, SEA islands; WAF, West Africa; WAS, West Asia; XI, *Xian/Indica*.

protein activity of this gene are positively correlated with AC [6]. To date, at least eight *Wx* alleles have been identified in rice— $Wx^a$ ,  $Wx^b$ ,  $Wx^{in}$ ,  $Wx^{op}$ ,  $Wx^{mq}$ ,  $Wx^{mw}$ ,  $Wx^{lv}$ , and  $wx$ . The AC of these alleles decreases in the following order:  $Wx^{lv}$  (>25%),  $Wx^a$  (20–25%),  $Wx^{in}$  (18–22%),  $Wx^b$  (15–18%),  $Wx^{mw}$  (10–14%),  $Wx^{mq}$  (8–12%),  $Wx^{op}$  (5–10%), and  $wx$  (AC<2%) [7–13]. However, whether there exist some new functional alleles in cultivated rice resources worldwide has not been elucidated to date. Moreover, we hope to develop a set of molecular markers for high-throughput genotyping in the *Wx* locus, but whether the known alleles can cover most of the *Wx* allelic variations and explain variation in AC has not been determined.

In 2015, 3,000 germplasm accessions from 89 different countries/regions were selected for genome-wide resequencing, and a comprehensive SNP and InDel sub-database was established for the Rice Functional Genomics-based Breeding (RFGB) Database [14]. Of these accessions, 2,466 were retrieved from a core collection of over 101,000 rice accessions in the International Rice Genebank Collection (IRGC), while 534 accessions were selected from a core collection of 61,470 rice accessions preserved in the China National Crop Genebank (CNCGB) [15]. Thus, the 3,000 germplasm accessions represent a panel with abundant genetic diversity worldwide. Using the currently available genome information from the 3,000 rice genomes project (3K RGP) sequencing data, allelic variations at the *Wx* locus can be comprehensively analysed, and new allelic variants responsible for different AC classes can be discovered to improve the quality of hybrid rice.

A prior study revealed a significant correlation between cytosine and thymidine (CT) repeat and AC, and microsatellite markers have been developed [16]. However, some of the AC variations cannot be accurately explained by these markers [17]. With a series of *Wx* alleles identified in rice [18], some SNP genotyping methods have been used for polymorphism analysis of the *Wx* locus. Kompetitive allele-specific PCR (KASP) is a high-throughput method that can be employed for SNP and InDel genotyping at specific sites. Different genotypes at a single site can be detected by two-colour fluorescence based on terminal fluorescence reading [19]. Therefore, owing to advantages such as high efficiency, accuracy, and throughput, the KASP marker may help accelerate rice quality breeding.

In this study, to make better use of the *Wx* allele resources and breed hybrid rice with improved eating quality, we investigated the diversity of *Wx* alleles in cultivated rice and their development in hybrid rice. Using the 3K RGP resequencing data, we identified not only eight known alleles that existed widely across the globe but also five allelic variations that had not been previously reported. Furthermore, a set of KASP markers based on these *Wx* alleles was developed, and thirty-six main parents of hybrid rice were selected for *Wx* genotyping. We found only three *Wx* alleles existed in these main parents, and the allelic combination of hybrid rice changed from  $Wx^a/Wx^b$  and  $Wx^{lv}/Wx^b$  to  $Wx^b/Wx^b$  as the quality improved. Because the allelic variation of the *Wx* locus in hybrid rice is relatively low, more favourable *Wx* alleles from cultivated rice should be introduced for further quality breeding of hybrid rice.

## Materials and methods

### Plant materials

By referring to previously described data [20], the Hybrid Rice Variety Resources Database (<http://www.hybridrice.com.cn/>), and the China Rice Data Centre (<http://www.ricedata.cn/>), thirty-six main parents of hybrid rice from 1976 to 2018 were selected for *Wx* genotyping. All the combinations of these parents are the hybrid rice with the largest planting area over the years. The hybrid parents included eighteen female parents of Erjiunan 1A, Zhenshan 97A, V20A, Gang 46A, Longtefu A, Bo A, II-32A, Xieqingzao A, Jin 23A, Zhong 9A, Tianfeng A,

Annong-S1, Peiai 64S, Guangzhan 64-4S, Y58S, Zhu 1S, C815S, Longke 638S and eighteen male parents of IR24, IR26, Minghui 63, Ce 64–7, Duoxi 1, Fuhui 838, Xianhui 207, Shuhui 527, CDR22, Miyang 46, R402, Minghui 86, Mianhui 725, Gui 99, Guanghui 3550, Yangdao 6, Bing 4114, Huazhan. The control group included Zhenshan 97A ( $Wx^a$ ), Nipponbare ( $Wx^b$ ), Kasalath ( $Wx^l$ ), Basmati ( $Wx^m$ ), Haopi ( $Wx^{op}$ ), and Nanjing 46 ( $Wx^{mq}$ ). All hybrid parental lines and control materials for genotyping were stored in our laboratory. Four germplasm resources—IRIS\_313–9445, IRIS\_313–10892, IRIS\_313–10866, and IRIS\_313–8956—were obtained from the Crops Research Institute, Chinese Academy of Agricultural Sciences. All varieties were cultivated under normal growing conditions in the experimental field of the Hunan Hybrid Rice Research Centre in Changsha.

### Resequencing data of the *Wx* alleles

The region for sequence analysis is located in a region (1765622–1770574) on chromosome 6 according to the Nipponbare genome (IRGSP V.1). Data for the SNPs and InDels at the *Wx* locus in 3,000 rice accessions were downloaded from the Rice SNP-Seek Database (<http://oryzasnp.org/iric-portal/index.zul>).

### Sequence analysis

Sequences were aligned using CLUSTAL X version 2.0 and adjusted manually with Microsoft Office Excel 2010 [21]. Haplotype diversity was calculated using DNASP v5.0 [22], and the haplotype network was constructed using PopART 1.7 [23]. A geographical distribution map of the *Wx* alleles was generated using rworldmap V1.36 [24].

### Crystal structure analysis

The crystal structure of the rice OsGBSSI catalytic domain in a complex with ADP was downloaded from the Protein Data Bank (<http://www.rcsb.org/pdb/>). The PyMOL Molecular Graphics System (Schrödinger LLC) was used to display the structural features of the OsGBSSI protein (PDB: 3VUF) with a focus on the novel mutations identified.

### Detection of the five allelic variations by Sanger sequencing

Genomic DNA was extracted from fresh leaves of the IRIS\_313–9445, IRIS\_313–10892, IRIS\_313–10866, and IRIS\_313–8956 varieties using a modified CTAB method. Briefly, a 266-bp sequence containing the mutation site of Ex2+160 was amplified using the primer pair 5′-ATGTCGGCTCTCACCACG-3′ and 5′-CCGACGAACACGACGTTTCATG-3′; a 297-bp sequence containing the mutation site of Ex4+73 was amplified using the primer pair 5′-GATACCAGCGTTGTGGCTGAG-3′ and 5′-CAGTCCAACCTGCTAAATGCACTG-3′; a 194-bp sequence containing the mutation site of Ex14+2 was amplified using the primer pair 5′-GAGTGACAAATTCAGGCAATCGAG-3′ and 5′-CCAGAAGAACGATCTGGACGTC-3′; and a 174-bp sequence containing the mutation sites of Ex14+2 and Ex14+28 was amplified using the primer pair 5′-CAGAGATTACCTGTCTGATGCTG-3′ and 5′-TCAAGGAGCAGC CACGTTCTC-3′. PCR amplification was carried out as follows: initial DNA denaturation at 95 °C for 4 min; 30 cycles of denaturation at 95 °C for 30 s; annealing at 58 °C for 30 s; extension at 72 °C for 30 s; and final extension at 72 °C for 5 min. After gel purification, the PCR products were sequenced by TsingKe Biology Technology, Bei Jing, China.

## KASP genotyping

The allele-specific primers were designed to carry the standard FAM (5'GAAGGTGACCAA GTTCATGCT3') and HEX (5' GAAGGTCGGAGTCAACGGATT 3') tails and the targeted SNP at the 3' end. Assays were carried out in 384-well formats and 10- $\mu$ l reactions (20–30 ng/ $\mu$ l DNA, 5  $\mu$ l of 1 $\times$  KASP master mixture, 0.14  $\mu$ l of KASP assay mix, and 4.86  $\mu$ l of water). PCR was conducted using the following protocol: hot start at 94 °C for 15 min, ten touchdown cycles (94 °C for 20 s; initial touchdown at 61 °C and then a decrease by  $-0.6$  °C per cycle for 60 s), and 26 additional cycles of annealing (94 °C for 20 s; 55 °C for 60 s). Finally, the PCR product with fluorescent labelling was scanned using a Roche Light Cycler 480.

## Amylose content measurement

The AC of rice seeds was measured using the iodine colorimetry assay described previously [25].

## Results

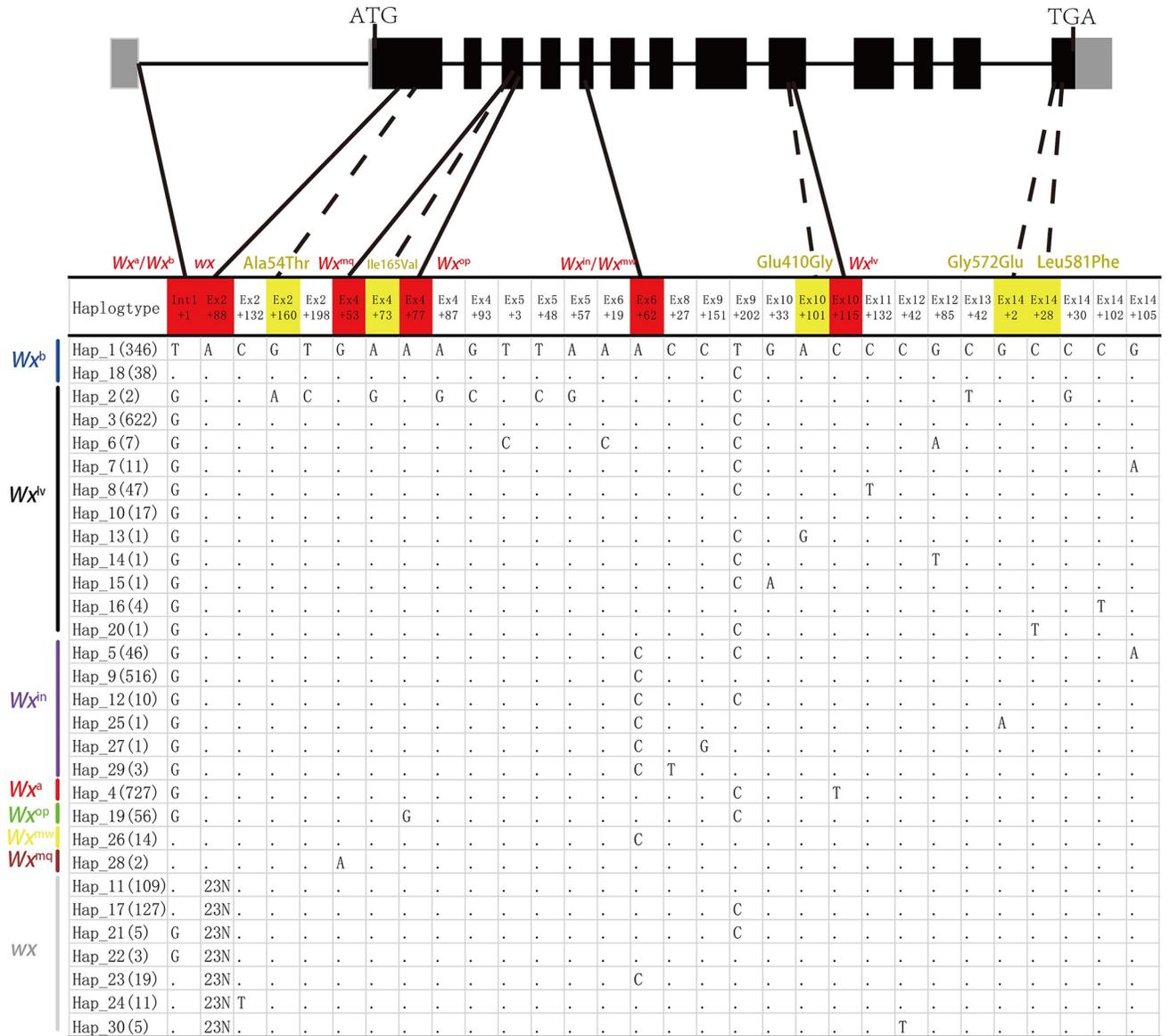
### Analysis of *Wx* allelic variations in 3K RGP

Using the coding sequence (CDS) of Nipponbare as a reference, allelic variations in the coding region of the *Wx* gene were analysed according to the 3K RGP sequencing data. After eliminating the alleles with heterozygous sequences in the coding region and the possible deletion caused by insufficient sequencing coverage, a total of 2,752 lines were obtained for variation analysis. Based on the G/T SNP at the splice site of the first intron (In1G/T SNP), two important alleles, namely,  $Wx^a$  and  $Wx^b$ , have been identified previously [26]. Although it does not belong to the coding region, this functional SNP should be taken into account for analyzing variation.

Based on 1 insertion and 29 nucleotide polymorphic sites (SNPs), a total of 30 haplotypes of *Wx* alleles were identified (Fig 1). Among these haplotypes, Int1+1, Ex2+88, Ex4+53, Ex4+77, Ex6+62, and Ex10+115 formed the eight known alleles, representing 99.9% of the allelic variation in the 2,752 lines. In addition, five non-synonymous mutations—Ex2+160, Ex4+73, Ex10+101, Ex14+2, and Ex14+28—were identified for the first time, and they resulted in amino acid changes from Ala to Thr (54 cases), Ile to Val (165), Glu to Gly (410), Gly to Glu (572) and Leu to Phe (581). Ex2+160 and Ex4+73 coexisted in IRIS\_313–9445, Ex10+101, Ex14+2 and Ex14+28 existed in IRIS\_313–10892, IRIS\_313–10866 and IRIS\_313–8956, respectively. Except for these novel mutations, IRIS\_313–9445, IRIS\_313–8956 and IRIS\_313–10892 contained the  $Wx^{lv}$  allele (AC $\approx$ 25%), and IRIS\_313–10866 contained the  $Wx^{in}$  allele (AC $\approx$ 18%). By determining the AC in the mature seeds of these four IRIS lines, we found that except for IRIS\_313–10892 (Ex10+101), the AC of the other three materials remained generally unchanged compared with that of the wild type (Fig 2). By analysing the crystal structure of OsGBSS1, we found that Glu410Gly (Ex10+101) is located in the active centre of the enzyme, next to the ligand ADP of OsGBSS1, while the remaining four mutation sites were located at a distance from the active centre (Fig 3B). In addition, Ala54Thr did not even exist in the truncated body of OsGBSS1.

### Construction of the *Wx* haplotype flowchart

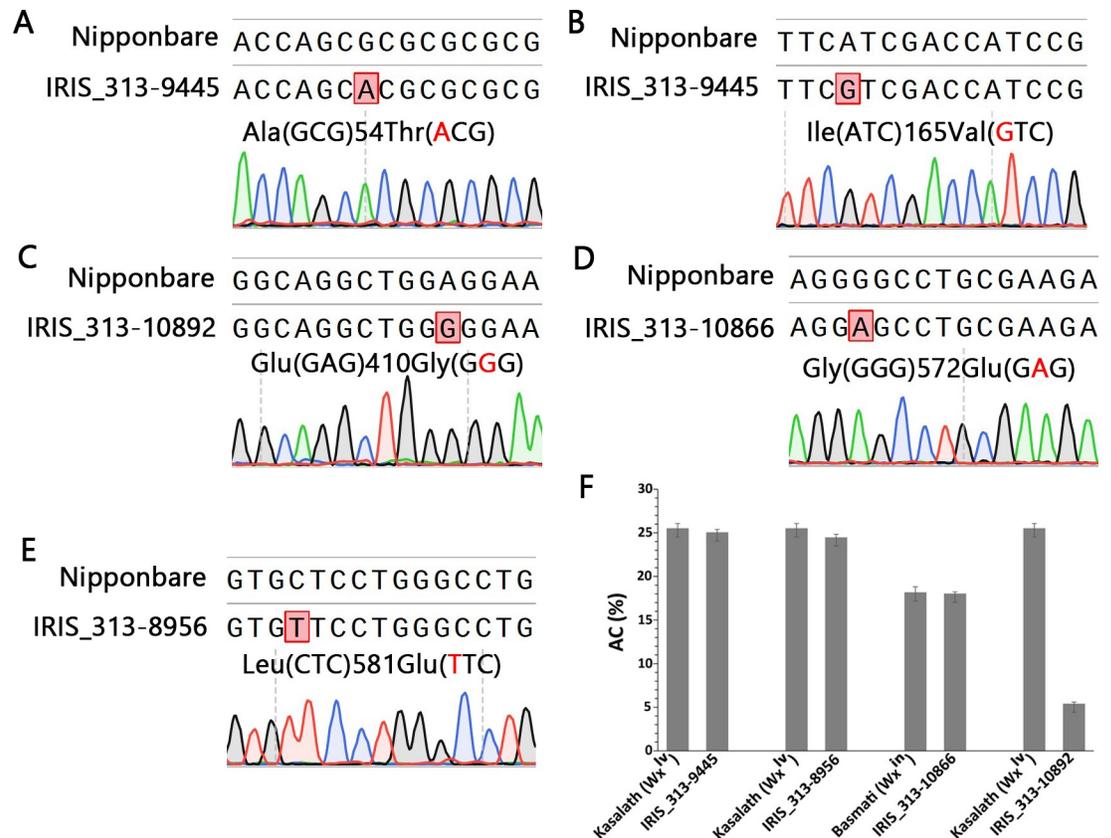
To understand the genetic background of each haplotype and the relationships among them, a haplotype flowchart was constructed to describe the mutational steps of 30 haplotypes. Hap1, Hap3, Hap4, and Hap9, which represent  $Wx^b$ ,  $Wx^{lv}$ ,  $Wx^a$ , and  $Wx^{in}$ , respectively (Fig 4), demonstrated absolute predominance compared to the other haplotypes in cultivated rice. With



**Fig 1. Analysis of the allelic variation among the 2,752 lines.** Each vertical line with a unique colour represents the same type of allele. Coloured boxes represent non-synonymous mutations; red box: mutation sites of each known allele, yellow box: mutation sites of each unreported allele. Boxes without colour represent synonymous mutations. Numbers in parentheses represent the total number for each specific haplotype. Hap1 (346) including 345 lines from 3 K RGP and one control material (Nipponbare).

<https://doi.org/10.1371/journal.pone.0232279.g001>

the exception of *wx*, *indica-japonica* background differences existed among the remaining seven alleles. Notably, the *Wx<sup>a</sup>*, *Wx<sup>lv</sup>*, and *Wx<sup>op</sup>* alleles belonged to the *indica* background, while *Wx<sup>b</sup>*, *Wx<sup>in</sup>*, *Wx<sup>mw</sup>*, and *Wx<sup>mq</sup>* represented the *japonica* background. *Wx<sup>lv</sup>* was recognized as the predominant allele in the cA subgroup, while *Wx<sup>in</sup>* is the predominant allele of cB. These findings indicated that *Wx* alleles have regional preference and a corresponding environmental adaptability.



**Fig 2. DNA sequence validation and AC measurement.** (A-E) Sequencing validation of five non-synonymous mutations. (F) Amylose content in mature seeds of four IRIS lines and WT ( $Wx^{lv}$  and  $Wx^{in}$ ). Kasalath ( $Wx^{lv}$ ) and Basmati ( $Wx^{in}$ ) were used as the control group.

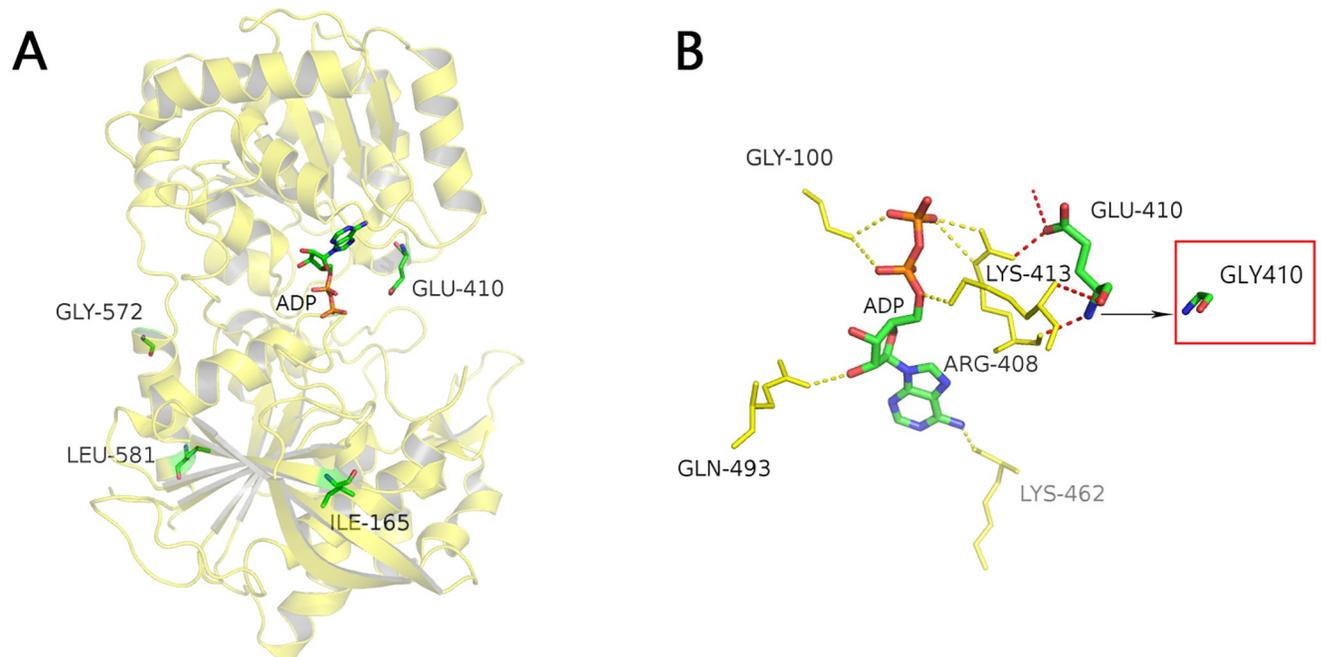
<https://doi.org/10.1371/journal.pone.0232279.g002>

## Worldwide distribution of different *Wx* alleles

A global map showing the distributions of the eight known alleles based on the 3k RGP information (Fig 5) indicates that  $Wx^a$ ,  $Wx^b$ ,  $Wx^{in}$ , and  $Wx^{lv}$  are four predominant alleles with a wide distribution in most rice-producing regions of the world. However, regional differences existed among the alleles. Compared with other alleles,  $Wx^b$  is widely distributed at high latitudes,  $Wx^{in}$  is widely distributed in tropical areas at low latitudes, and  $Wx^a$  is located in the mid-latitude region. In addition,  $Wx^{lv}$  was mainly distributed in the South Asia-Central (SAC) and South Asia-East (SAE) areas, and  $wx$  was commonly distributed in Southeast Asia (SEA) and East Asia (EAS). These findings are consistent with the eating habits of the regions, as high-amylose varieties are popular in Myanmar, Sri Lanka, provinces of Indonesia, and many states of India [28].

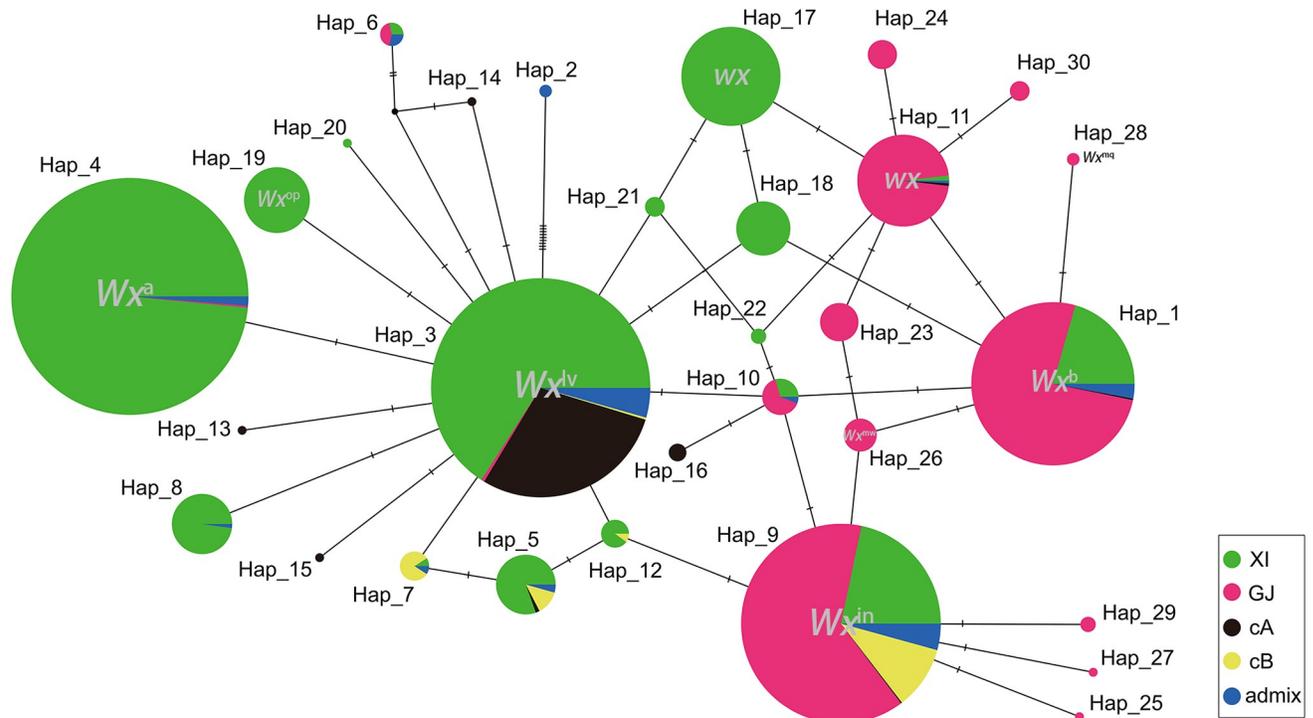
## Correlation analysis between the *Wx* alleles and CT repeats

In most prior studies, AC was primarily defined according to the CT repeat. Generally, varieties with 17 or 18 CT repeats are considered to be “low-amylose” types, while those with 10 or 11 CT repeats are generally classified as “high-amylose” types [17]. Although previous studies found a significant correlation between the number of CT repeats and AC, some related



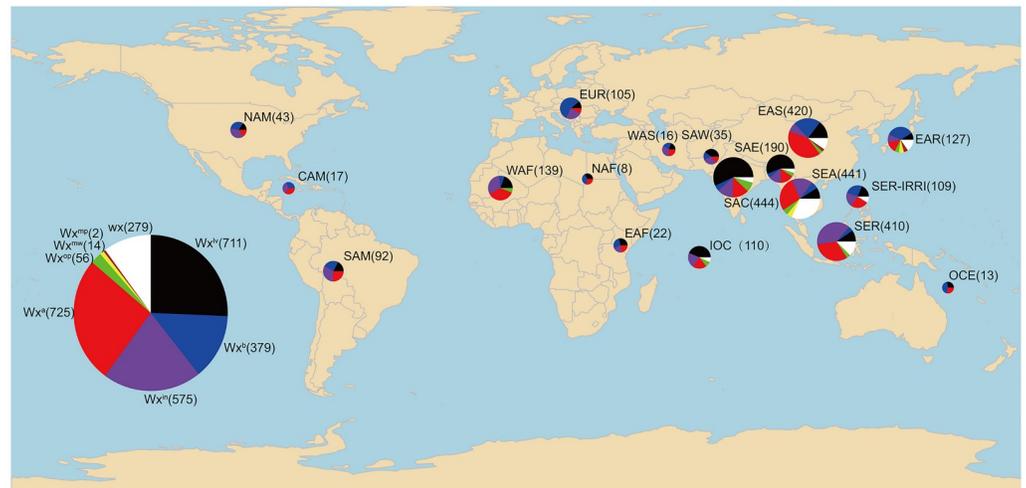
**Fig 3. Crystal structure of rice OsGBSS1.** (A) Crystal structure of the OsGBSS1 catalytic domain in complex with ADP. ADP, Glu410, Gly572, Leu581, and Ile165 are each represented by the highlighted stick models. (B) The enlarged active centre of OsGBSS1. ADP is located in the middle area, the five amino acids linked to ADP by hydrogen bonds are indicated by yellow sticks, and the dotted yellow line represents the hydrogen bonds between them. The dotted red line represents the hydrogen bonds between Glu410 and two important amino acids. Gly410 in the red box represents mutated amino acids.

<https://doi.org/10.1371/journal.pone.0232279.g003>



**Fig 4. Flowchart of the 30 *Wx* haplotypes.** Different colours represent different varietal groups. XI, GJ, cA, cB, and admix represent the *Xian/Indica*, *Geng/Japonica*, *circum-Aus*, *circum-Basmati*, and *intermediate* groups, respectively [27]. Each pie chart indicates a unique haplotype and the proportions of different varietal groups. The size of each pie chart represents the number of lines in each haplotype. Hap1, Hap3, Hap4, and Hap9 represent  $Wx^b$ ,  $Wx^{iv}$ ,  $Wx^a$ , and  $Wx^{in}$ , respectively. Each short line segment on a solid line represents one SNP between two haplotypes.

<https://doi.org/10.1371/journal.pone.0232279.g004>



**Fig 5. Worldwide distribution of different *Wx* alleles.** Each colour represents a unique *Wx* allele. Red:  $Wx^a$ , black:  $Wx^{lv}$ , purple:  $Wx^{in}$ , blue:  $Wx^b$ , white:  $wx$ , green:  $Wx^{op}$ , yellow:  $Wx^{mw}$ , and red-brown:  $Wx^{mq}$ . Each pie chart indicates the proportions of different *Wx* alleles in a certain area, and the numbers in parentheses indicate the sample size for each pie chart. The pie charts at the bottom left indicate the proportions of different *Wx* alleles in 3K RGP, and the numbers in parentheses indicate the numbers of a certain *Wx* allele.

<https://doi.org/10.1371/journal.pone.0232279.g005>

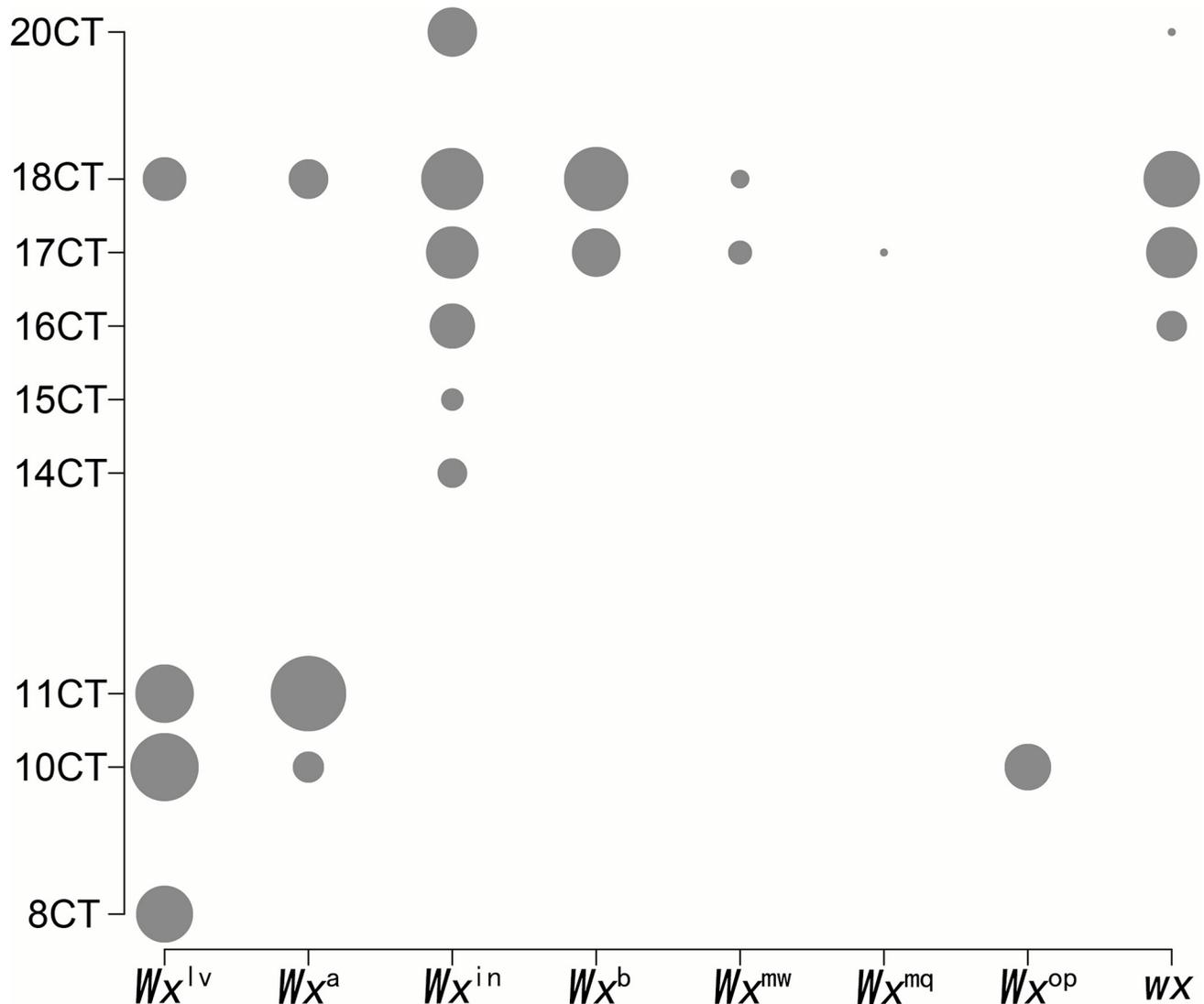
molecular markers have been developed. Nonetheless, a considerable amount of AC variation could not be explained by the CT repeats. By analysing the number of CT repeats corresponding to different *Wx* alleles in the 3K data, we found that most low-amylose varieties had long CT repeats, while some high-amylose varieties had short CT repeats. These results are consistent with the findings of previous studies [17,29]. However, some high-amylose alleles, such as  $Wx^a$  and  $Wx^{lv}$ , also contained long CT repeats, and some low-amylose alleles, such as  $Wx^{op}$ , had only short CT repeats (Fig 6). These exceptions help to explain the AC variation, which could not be explained by microsatellite markers.

### KASP genotyping of the *Wx* alleles in hybrid parents

Considering that these eight known *Wx* alleles covered 99.9% of the allelic variation, a set of KASP markers was developed based on these *Wx* alleles (Fig 7). Thirty-six main parent lines, including 18 female parents and 18 male parents of hybrid rice from 1976 to 2018 [20], were selected for genotyping (Fig 8). Fig 9 indicates that  $Wx^b$  is the major allele for nearly all restorer lines.  $Wx^a$  is the major allele involved in the male sterility of three-line hybrid rice, such as Zhenshan97A.  $Wx^{lv}$  is the major allele involved in the male sterility of early two-line hybrid rice, such as Annong S-1 and Peiai 64S, and with the development of hybrid rice,  $Wx^{lv}$  was gradually replaced by  $Wx^b$ .

### Discussion

This study identified five new allelic variations for the first time. By determining the AC in four lines that contained these five variation sites, we found that only the AC of IRIS\_313–10892 (Ex10+101) changed significantly compared with the AC of the wild type. Analysis of the crystal structure of OsGBSS1 showed that Lys413 and Arg408 are two important amino acids that directly interact with ADP, and Glu410 was found to be linked with these two amino acids via hydrogen bonds. Previously, many artificial point mutations in OsGBSS1 have been identified, including E410 [6]. In their study, Glu410Asp and Glu410Gln resulted in a

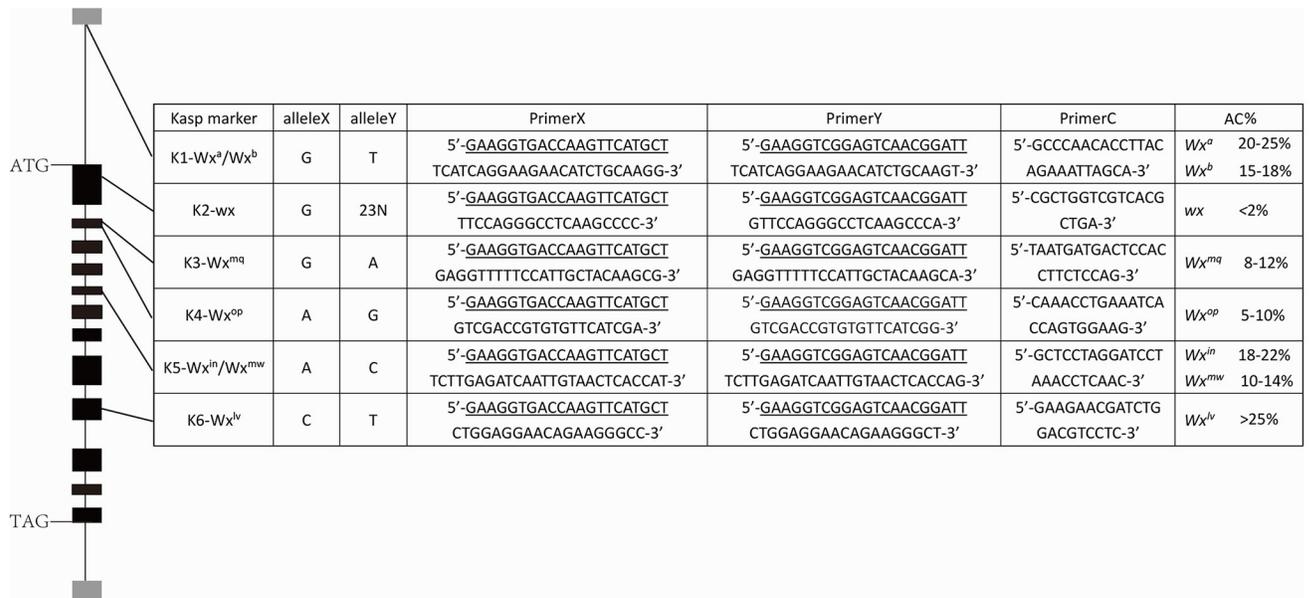


**Fig 6. Relationship between the *Wx* alleles and number of CT repeats.** The size of each grey dot represents the amount of each allele that corresponds to different CT repeats.

<https://doi.org/10.1371/journal.pone.0232279.g006>

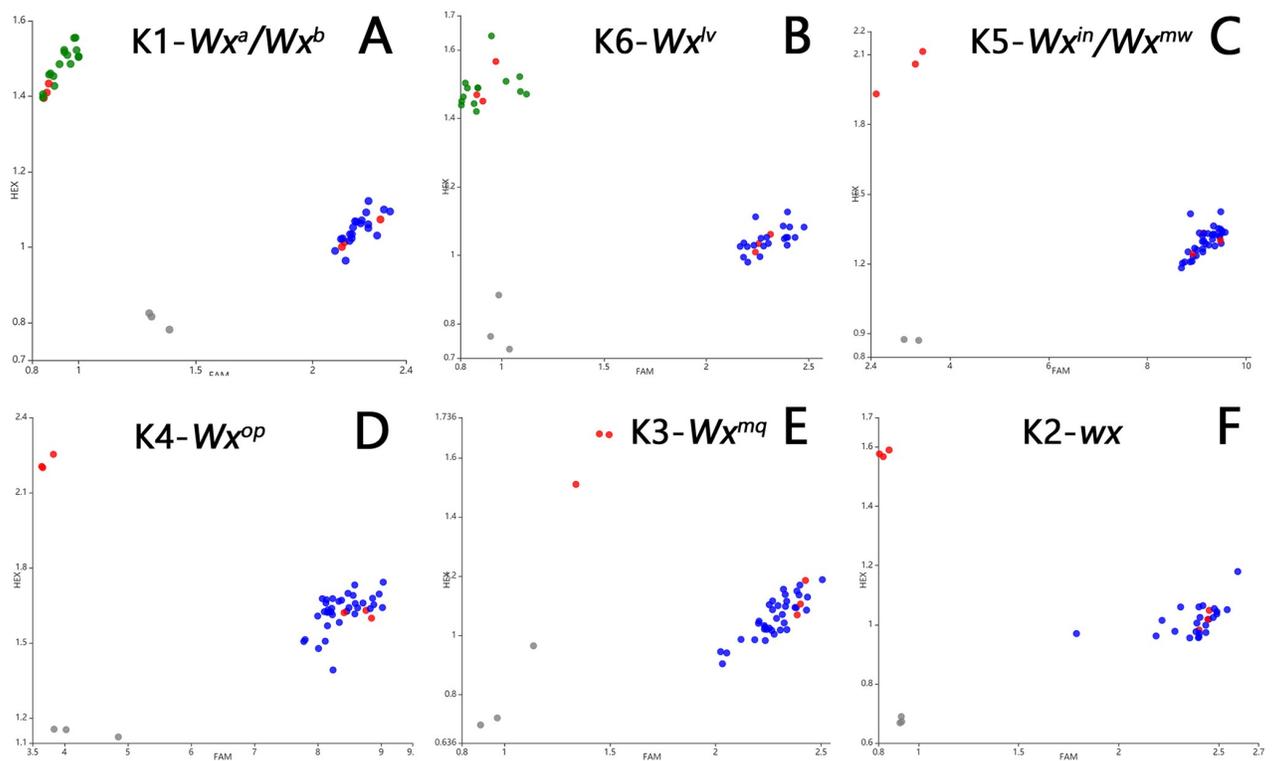
reduction in AC and OsGBSS1-specific activity. Glu410Gly was a natural allelic variation, where the acidic amino acid was replaced by a non-polar amino acid without any R group; this type of change would probably have a strong effect on the conformation of Lys413 and Arg408. Except for *wx* in glutinous rice, *Wx<sup>op</sup>* is currently the allelic variation with the lowest AC in non-glutinous rice. In this study, we found a novel allele with AC between *wx* and *Wx<sup>op</sup>*, which increased the diversity of the variation in AC in cultivated rice.

By analysis of the *Wx* allelic variation in thirty-six main parents of hybrid rice from 1976 to 2018, we found that only three *Wx* alleles existed, and the allelic combinations in hybrid rice changed from *Wx<sup>a</sup>/Wx<sup>b</sup>* and *Wx<sup>lv</sup>/Wx<sup>b</sup>* to *Wx<sup>b</sup>/Wx<sup>b</sup>* as quality improved. *Wx<sup>a</sup>* and *Wx<sup>lv</sup>* are two alleles with high amylose content and hard texture of cooked rice, and *Wx<sup>b</sup>* is a low-amylose allele with higher taste value [7]. Therefore, removing these two alleles and introducing *Wx<sup>b</sup>* might be the main reason for the quality improvement of hybrid rice. In addition, With



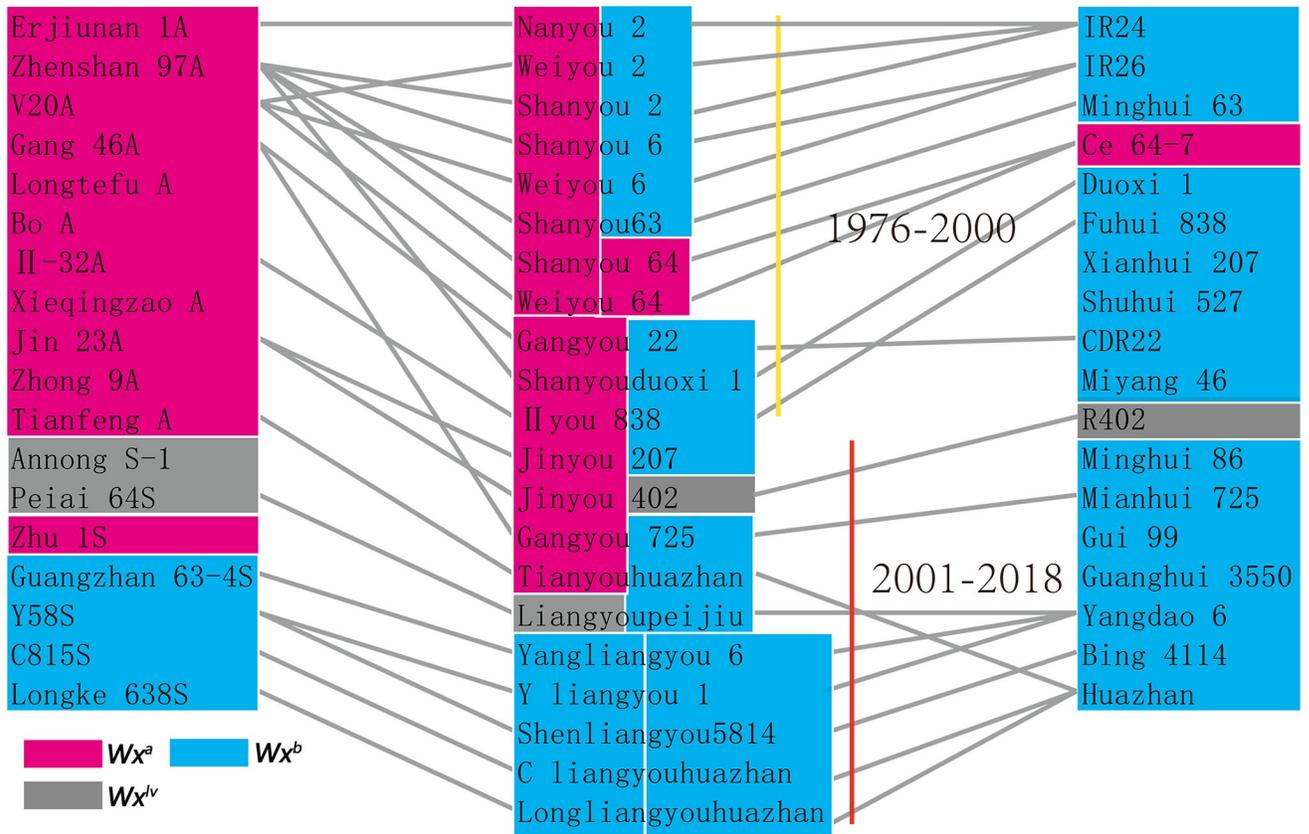
**Fig 7. Primer sequences for KASP genotyping of different *Wx* alleles.** Primers X and Y are two allele-specific primers (one for each SNP allele), and each primer contains 5' fluoro-labelled oligos (underlined). Primer C represents a common (reverse) primer. Alleles X and Y represent SNPs or InDels at the 3' end of primers X and Y, respectively. AC% represent amylose content of each allele.

<https://doi.org/10.1371/journal.pone.0232279.g007>



**Fig 8. KASP genotyping of the different *Wx* alleles in main parent lines.** Scatter plots of the selected KASP assays reveal the clustering of varieties on the X- (FAM) and Y- (HEX) axes. Varieties coloured blue represent the FAM-type allele, while those coloured green represent the HEX-type allele. Red dots represent the control groups, and black dots represent the non-template control (NTC). (A) KASP genotyping for the In1+1 SNP (G/T), Zhenshan 97A (*Wx<sup>a</sup>*) and Nipponbare (*Wx<sup>b</sup>*) were used as the control group. (B) KASP genotyping for the Ex10+115 SNP (C/T), Kasalath (*Wx<sup>lv</sup>*) and Zhenshan 97A (*Wx<sup>a</sup>*) were used as the control group. (C) KASP genotyping for the Ex6+62 SNP (A/C); Basmati (*Wx<sup>in</sup>*) and Nipponbare (*Wx<sup>b</sup>*) were used as the control group [13]. (D) KASP genotyping for the Ex4+77 SNP (A/G); Haopi (*Wx<sup>op</sup>*) and Nipponbare (*Wx<sup>b</sup>*) were used as the control group [9]. (E) KASP genotyping for the Ex4+53 SNP (G/A); Nanjing 46 (*Wx<sup>mq</sup>*) and Nipponbare (*Wx<sup>b</sup>*) were used as the control group [30]. (F) KASP genotyping for the 23-bp insertion at the Ex2+88 site; Suyunuo (*wx*) and Nipponbare (*Wx<sup>b</sup>*) were used as the control group [31].

<https://doi.org/10.1371/journal.pone.0232279.g008>



**Fig 9. Genotyping of *Wx* in the hybrid parents.** The left column represents the sterile line, right column represents restorer lines, and middle column represents hybrid combinations. Each vertical line with a unique colour represents different time periods in the development of hybrid rice.

<https://doi.org/10.1371/journal.pone.0232279.g009>

homozygosis of the  $Wx^b$  allele in hybrid rice, the AC of each grain in the hybrid combination tends to be more homogeneous. However, as high temperatures occur frequently in the southern regions of China, the AC of hybrids in these regions decreased significantly, which have strongly affected their appearance and eating quality. Many studies revealed that the splicing efficiency of the first intron of *Wx* is affected by temperature [32–36]. This might be the important reason why  $Wx^b$  is sensitive to high temperature.  $Wx^{in}$  is an intermediate amylose allele with the G base at the leader intron splice site. In addition, our analysis found that  $Wx^{in}$  is the predominant allele in the circum-Basmati group (cB), which includes many high-quality rice varieties, such as Basmati and Sadri aromatic varieties. Thus, we suggest that  $Wx^{in}$  is the preferred choice for improving the quality of *indica*-type rice in southern China.

As only three allelic variations were present in the hybrid rice, more favourable *Wx* alleles from cultivated rice should be identified, introduced, and effectively used to improve hybrid rice quality in the future. By evaluating the genetic effects of different allelic combinations, appropriate genotypic combination to achieve improved rice quality can be identified.

## Supporting information

**S1 Appendix.**  
(XLS)

**S2 Appendix.**  
(XLSX)

## Author Contributions

**Conceptualization:** Xionglun Liu, Bingran Zhao.

**Data curation:** Ye Shao.

**Formal analysis:** Ye Shao, Yan Peng.

**Methodology:** Ye Shao, Yan Peng.

**Resources:** Bigang Mao, Qiming Lv, Dingyang Yuan.

**Writing – original draft:** Ye Shao, Yan Peng.

**Writing – review & editing:** Ye Shao, Yan Peng.

## References

1. Chen LY, Lei DY, Tang WB, Deng HB, Xiao YH, Zhang GL. Challenges and strategies of hybrid rice development in China. *Hybrid Rice*. 2015; 30: 1–4.
2. Tang WB, Zhang GL, Xiong YD, Ming XQ, Ding XC, Yi GL. Reflection on rice breeding strategy in transitional period of China-developing high quality hybrid rice. *Hybrid Rice*. 2016; 31: 1–5.
3. Li YH, Qian Q, Zeng DL, Sun ZX. Identification by RAPD analysis and studies on genetic relationship of main parents of hybrid rice in China. *Zuo Wu Xue Bao*. 2000; 26: 171–176.
4. Tian Z, Qian Q, Liu Q, Yan M, Liu X, Yan C, et al. Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc Natl Acad Sci U S A*. 2009; 106: 21760–21765. <https://doi.org/10.1073/pnas.0912396106> PMID: 20018713
5. Qian Q, Guo L, Smith SM, Li J. Breeding high-yield superior quality hybrid super rice by rational design. *Natl Sci Rev*. 2016; 3: 283–294.
6. Liu D, Wang W, Cai X. Modulation of amylose content by structure-based modification of OsGBSS1 activity in rice (*Oryza sativa* L.). *Plant Biotechnol J*. 2014; 12: 1297–1307. <https://doi.org/10.1111/pbi.12228> PMID: 25052102
7. Zhang C, Zhu J, Chen S, Fan X, Li Q, Lu Y, et al. *Wx*(lv), the ancestral allele of rice waxy gene. *Mol Plant*. 2019; 12: 1157–1166. <https://doi.org/10.1016/j.molp.2019.05.011> PMID: 31181338
8. Yang J, Wang J, Fan F-J, Zhu J-Y, Chen T, Wang C-L, et al. Development of AS-PCR marker based on a key mutation confirmed by resequencing of *Wx*-mp in milky princess and its application in japonica soft rice (*Oryza sativa* L.) breeding. *Plant Breed*. 2013; 132: 595–603.
9. Liu L, Ma X, Liu S, Zhu C, Jiang L, Wang Y, et al. Identification and characterization of a novel Waxy allele from a Yunnan rice landrace. *Plant Mol Biol*. 2009; 71: 609–626. <https://doi.org/10.1007/s11103-009-9544-4> PMID: 19760367
10. Suto M, Ando I, Numaguchi K, Horisue N. Breeding of low amylose content paddy rice variety "milky queen" with good eating quality. *Jpn J Breed*. 1996; 46: 221.
11. Mikami I, Aikawa M, Hirano H-Y, Sano Y. Altered tissue-specific expression at the *Wx* gene of the opaque mutants in rice. *Euphytica*. 1999; 105: 91–97.
12. Mikami I, Dung LV, Hirano HY, Sano Y. Effects of the two most common *Wx* alleles on different genetic backgrounds in rice. *Plant Breed*. 2000; 119: 505–508.
13. Mikami I, Uwatoko N, Ikeda Y, Yamaguchi J, Hirano HY, Suzuki Y, et al. Allelic diversification at the *Wx* locus in landraces of Asian rice. *Theor Appl Genet*. 2008; 116: 979–989. <https://doi.org/10.1007/s00122-008-0729-z> PMID: 18305920
14. Li JY, Wang J, Zeigler RS. The 3,000 rice genomes project: new opportunities and challenges for future rice research. *Gigascience*. 2014; 3: 8. <https://doi.org/10.1186/2047-217X-3-8> PMID: 24872878
15. Zhang H, Zhang D, Wang M, Sun J, Qi Y, Li J, et al. A core collection and mini core collection of *Oryza sativa* L. in China. *Theor Appl Genet*. 2011; 122: 49–61. <https://doi.org/10.1007/s00122-010-1421-7> PMID: 20717799
16. Larkin PD, Park WD. Association of waxy gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). *Mol Breed*. 2003; 12: 335–339.

17. Dobo M, Ayres N, Walker G, Park WD. Polymorphism in the GBSS gene affects amylose content in US and European rice germplasm. *J Cereal Sci.* 2010; 52: 450–456.
18. Zhu JH, Zhang CQ, Gu MH, Liu QQ. Progress in the allelic variation of *Wx* gene and its application in rice breeding. *Chin J Rice Sci.* 2015; 29: 431–438.
19. He C, Holme J, Anthony J. SNP genotyping: the KASP assay. *Methods Mol Biol.* 2014; 1145: 75–86. [https://doi.org/10.1007/978-1-4939-0446-4\\_7](https://doi.org/10.1007/978-1-4939-0446-4_7) PMID: 24816661
20. Hu ZX, Tian Y, Xu QS. Review of extension and analysis on current status of hybrid rice in China. *Hybrid Rice.* 2016; 31: 1–8.
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and clustal X version 2.0. *Bioinformatics.* 2007; 23: 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
22. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics.* 2003; 19: 2496–2497. <https://doi.org/10.1093/bioinformatics/btg359> PMID: 14668244
23. Leigh JW, Bryant D. popart: full-feature software for haplotype network construction. *Methods Ecol Evol.* 2015; 6: 1110–1116.
24. South A. rworldmap: a new R package for mapping global data. *R J.* 2011; 3: 35–43.
25. Williams P, Kuzina F, Hlynka I. Rapid colorimetric procedure for estimating the amylose content of starches and flours. *Cereal Chem.* 1970; 47: 411–421.
26. Isshiki M, Morino K, Nakajima M, Okagaki RJ, Wessler SR, Izawa T, et al. A naturally occurring functional allele of the rice waxy locus has a GT to TT mutation at the 5' splice site of the first intron. *Plant J.* 1998; 15: 133–138. <https://doi.org/10.1046/j.1365-313x.1998.00189.x> PMID: 9744101
27. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature.* 2018; 557: 43–49. <https://doi.org/10.1038/s41586-018-0063-9> PMID: 29695866
28. Calingacion M, Laborte A, Nelson A, Resurreccion A, Concepcion JC, Daygon VD, et al. Diversity of global rice markets and the science required for consumer-targeted rice breeding. *PLoS ONE.* 2014; 9: e85106. <https://doi.org/10.1371/journal.pone.0085106> PMID: 24454799
29. Cheng A, Ismail I, Osman M, Hashim H. Simple and rapid molecular techniques for identification of amylose levels in rice varieties. *Int J Mol Sci.* 2012; 13(5):6156–6166. <https://doi.org/10.3390/ijms13056156> PMID: 22754356
30. Wang CL, Zhang YD, Zhu Z, Chen T, Zhao L, Lin J, et al. Breeding of a new rice variety with good eating quality by marker assisted selection. *Chin J Rice Sci.* 2009; 23: 25–30.
31. Tian Z, Yan C, Qian Q, Yan S, Xie H, Wang F, et al. Development of gene-tagged molecular markers for starch synthesis-related genes in rice. *Chin Sci Bull.* 2010; 55: 3768–3777.
32. Yamakawa H, Hirose T, Kuroda M, Yamaguchi T. Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiol.* 2007; 144: 258–277. <https://doi.org/10.1104/pp.107.098665> PMID: 17384160
33. Umemoto T, Nakamura Y, Ishikura N. Activity of starch synthase and the amylose content in rice endosperm. *Phytochemistry.* 1995; 40(6): 1613–1616.
34. Sano Y, Hirano H-Y, Nishimura M. Evolutionary significance of differential regulation at the *wx* locus of rice. *Rice Genetics II: (In 2 Parts): World Scientific;* 1991. p. 11–20.
35. Larkin PD, Park WD. Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. *Plant molecular biology.* 1999; 40(4): 719–727. <https://doi.org/10.1023/a:1006298608408> PMID: 10480395
36. Hirano H-Y, Sano Y. Enhancement of *Wx* gene expression and the accumulation of amylose in response to cool temperatures during seed development in rice. *Plant and Cell physiology.* 1998; 39(8): 807–712.