

RESEARCH

Open Access



Genome wide association study on development and evolution of glutinous rice

Conghui Jiang³, Muhammad Abdul Rehman Rashid^{4,5}, Yanhong Zhang⁶, Yan Zhao^{2*} and Yinghua Pan^{1*}

Abstract

Background: Glutinous rice as a special endosperm type is consumed as a staple food in East Asian countries by consumers' preference. Genetic studies on glutinous rice could be conducive to improve rice quality and understand its development and evolution. Therefore, we sought to explore more genes related to glutinous by genome wide association study and research the formation history for glutinous.

Results: Here, genome-wide association study was performed to explore the associated loci/genes underlying glutinous rice by using 2108 rice accessions. Combining the expression patterns analysis, 127, 81, and 48 candidate genes were identified to be associated with endosperm type in whole rice panel, *indica*, and *japonica* sub-populations. There were 32 genes, including three starch synthesis-related genes *Wx*, *SSG6*, and *OsSSIIa*, detected simultaneously in the whole rice panel and subpopulations, playing important role in determining glutinous rice. The combined haplotype analyses revealed that the waxy haplotypes combination of three genes mainly distributed in Southeast Asia (SEA), SEA islands (SER) and East Asia islands (EAR). Through population structure and genetic differentiation, we suggest that waxy haplotypes of the three genes firstly evolved or were directly inherited from wild rice in *japonica*, and then introgressed into *indica* in SER, SEA and EAR.

Conclusions: The cloning and natural variation analysis of waxy-related genes are of great significance for the genetic improvement of quality breeding and comprehend the history in glutinous rice. This work provides valuable information for further gene discovery and understanding the evolution and formation for glutinous rice in SEA, SER and EAR.

Keywords: Glutinous rice, Genome-wide association study, *Oryza Sativa* L., Haplotype analysis, Evolution

Background

Rice (*Oryza Sativa* L.) is one of the most important crop grain feeding more than half the world's population [1]. High yield and good quality are two important goals of rice production [2]. Since the green revolution, new farming methods and breeding techniques have greatly

increased food production in many countries [3–6]. but the improvement of quality breeding have lagged. The demand of rice with good quality is more urgent for consumers and producers with the improvement of people's living standards. The yield and quality of rice largely determined by the starch content, the ratio of amylose to amylopectin, and the fine structure of amylopectin.

There are two unique subpopulations of rice, *japonica* and *indica*. But whether in *indica* or *japonica*, it can be divided into glutinous and non-glutinous rice. Rice endosperm type are routinely classified according to their amylose content: high (> 25%), intermediate (20–25%), low (10–19%), very low or soft (3–9%), and waxy or glutinous (< 2%) [7]. Compared to non-glutinous rice,

*Correspondence: zhaoyan1216@163.com; panyinghua2008@163.com

¹ Rice Research Institute, Guangxi Academy of Agricultural Sciences/ Guangxi Key Laboratory of Rice Genetics and Breeding, Nanning 530007, China

² State Key Laboratory of Crop Biology, Shandong Key Laboratory of Crop Biology, College of Agronomy, Shandong Agricultural University, Tai'an, Shandong 271018, PR China

Full list of author information is available at the end of the article



the texture of glutinous rice is very sticky. Glutinous rice is a major type of cultivated rice with long-standing cultural importance in Asia, and glutinous rice is also eaten as a staple food of East Asian countries, including Laos and northern Thailand, known as the “center of the glutinous rice region” [8–11]. Therefore, the formation of waxiness in rice is not only affected by natural selection but also human preference.

Starch biosynthesis is a complex system composed of synthesis of substrate adenosine diphosphate glucose, direct starch, and amylopectin [12]. It involves 18 starch synthase enzymes related genes, and each gene plays a different role in various stages of starch synthesis [13–16]. In rice grains, the *Waxy* (*Wx*) gene encodes granule-bound starch synthase (GBSS), is a major gene controlling amylose synthesis, and directly affects the amylose content (AC). The rice type (glutinous or non-glutinous) is mainly governed by two alleles (*wx* and *Wx*) of same gene [17, 18]. Recently, a study was reported to enrich the range of breeding materials by using a base editing system at the third, fourth, and fifth exon of *Wx^b* to create a series of mutants with AC of 1.4–11.9% [19]. Rice starch content is a comprehensive trait contributed by a series of starch synthesis genes. A fine regulatory network that regulates the eating and cooking qualities (ECQs) in edible rice has been clarified by association analysis and transgenic verification experiments in the starch biosynthesis pathway [20]. *Wx* and *SSII-3* are two major genes that determine ECQs by affecting AC, gel consistency (GC), and gelatinization temperature (GT). *Wx* is the only gene that has a major effect on AC and GC, and has a minimal effect on GT. Interaction of multiple pairs of genes has a significant effect on rice apparent amylose contents (AAC) [21]. Owing to the interaction among starch-synthesis genes, mutation in a single gene will cause changes in the effects of multiple other genes.

Due to the limitations of traditional parental mapping and the special characteristics of waxiness, only *wx* for waxiness had been cloned and there is molecular evidence of a strong selective sweep in a starch of about ~250kb flanking the *Wx* locus among landraces cultivated in Asian countries [17, 18, 22]. More genes related to waxiness are required to be urgently discovered to enrich the natural variation of waxiness and the information about evolutionary origin, domestication and adaptation of key genes of glutinous is conducive to study the formation and evolution of glutinous rice. With the development of sequencing technology, genome wide association study (GWAS) has become an effective mean to discover genes and QTLs for grain qualities [23–25]. The sequencing of 3K core germplasm rice provides strong guarantee for the discovery of waxy genes [26].

In this study, 2108 rice germplasm were used in GWAS to identify the significant loci and candidate genes controlling the development of glutinous rice. Haplotype analysis was performed to identify the corresponding glutinous haplotypes of three key waxiness related genes. Combined haplotype analyses were carried out to reveal the genetic characteristics of glutinous rice in Southeast Asia (SEA), SEA islands (SER), and East Asia islands (EAR). Phylogenetic tree and population structure analysis for the origin and evolution of three key waxy genes. Our findings provide important information for further gene discovery and, to gain insight into the evolution and formation of glutinous rice in SEA, SER and EAR.

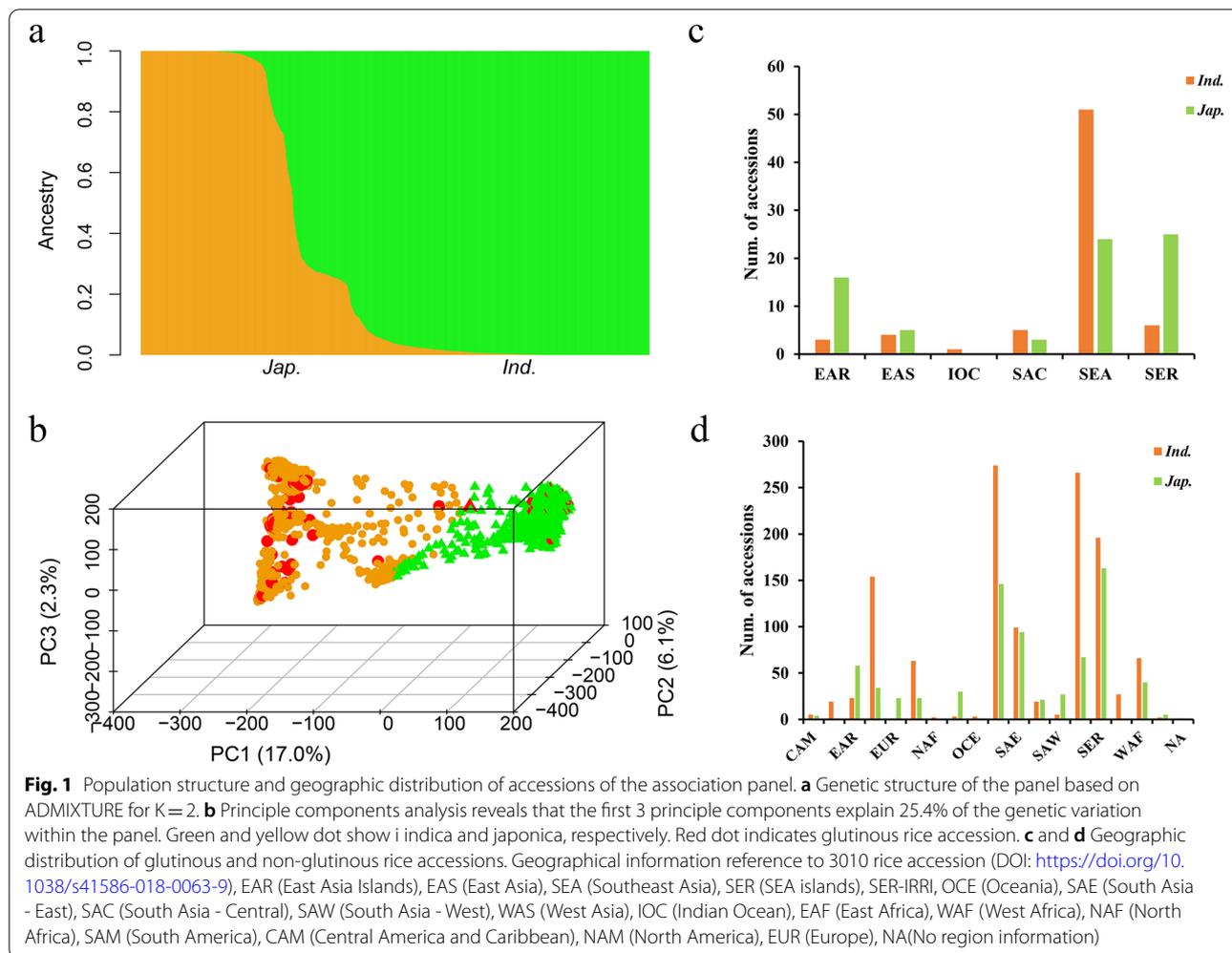
Results

Endosperm types analysis within different subgroups

A total of 2108 rice accessions, including 1965 non-glutinous (or non-waxy) and 143 glutinous (or waxy) rice accessions (<http://snp-seek.irri.org/>), were used to identify waxiness-related genetic loci and analyze the differentiation for development of both endosperm types (Table S1). Meanwhile, 17,132,232 SNPs of the rice panel were obtained from 3KRGF [26]. Subsets of these data were further filtered and used in the subsequent analyses.

Reasonable assessment of population structure is conducive to detect the phenotypic differences and subsequent GWAS of natural population. Using Admixture software [27], we calculated varying levels of K means within the rice population. The *indica* and *japonica* subpopulations appear clearly at $K=2$ (Fig. 1a). The principal component (PC) analysis indicated that top three PCs explained 17.0, 6.1 and 2.3% of the genetic variation within the rice panel, which supported that there were two main subpopulations (Fig. 1b). Referring to the recent results of 3010 rice accessions [26], we classified the panel into two major subpopulations, 1298 *indica* and 810 *japonica*, although there were several atypical *indica* and *japonica* accessions (Table S1). Hence, the endosperm type of each rice subgroups was compared, and top three PCs were used as covariates to control for subgroup structure in GWAS.

Among the 143 glutinous rice accessions, there were 70 *indica* and 73 *japonica* (Table S1, Fig. 1b), suggesting broad genetic variation of trait occurred in *indica* and *japonica*. To study the underlying external factors affecting glutinous differentiation, the geographic distribution of accessions with different glutinous traits was investigated. The vast majority of glutinous rice accessions are distributed in SEA, SER and EAR with 75, 31 and 19 accessions, respectively (Fig. 1c). In contrast, non-glutinous rice, as a major endosperm type, was widely distributed in the whole rice growing area (Fig. 1d). This geographic distribution was in consistence with previous

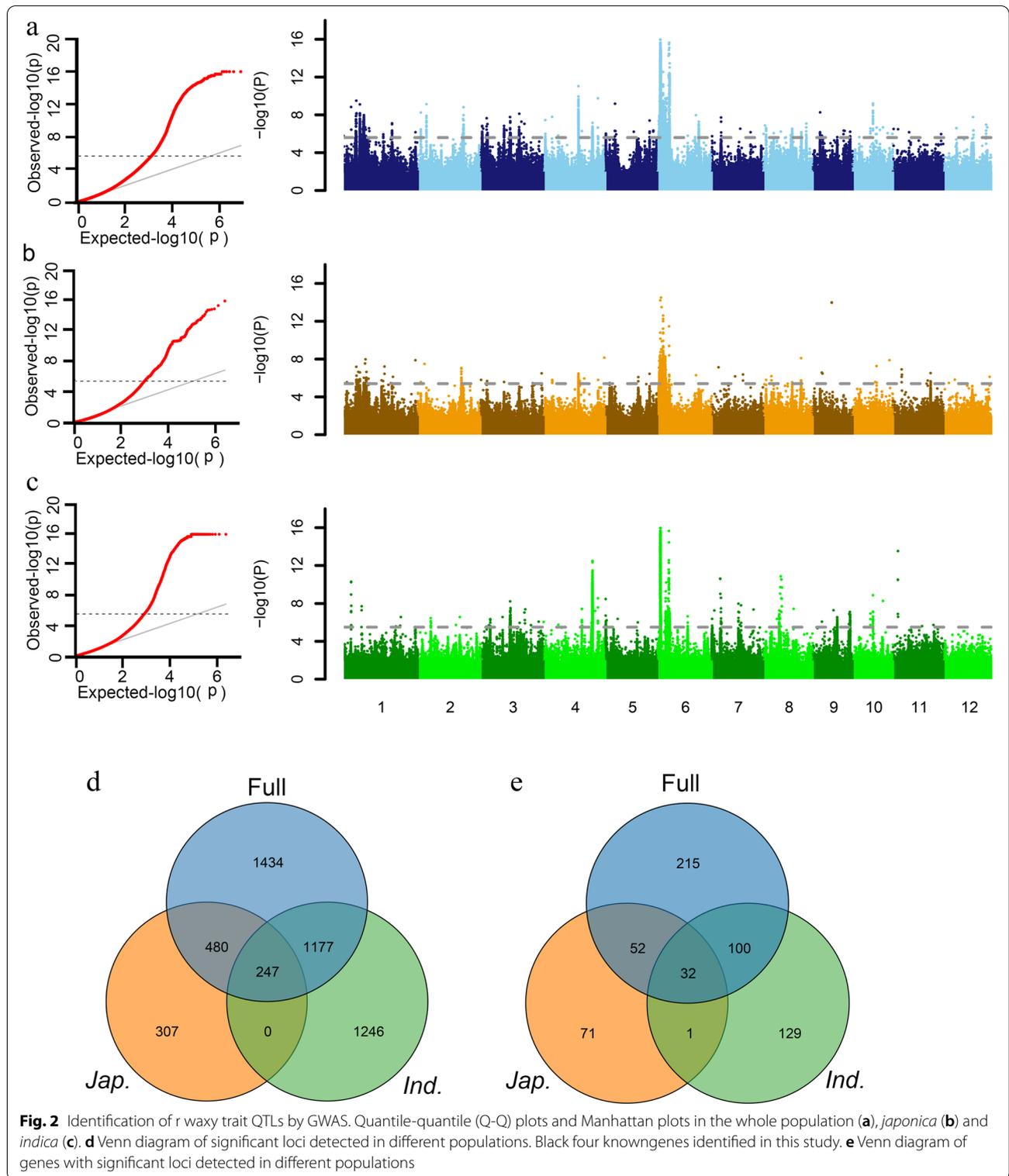


research that reported the artificial selection of glutinous rice in Southeast Asia [10, 28]. Taken together, these results suggested that there were large genetic differentiation among glutinous rice accessions, although they were relatively geographically concentrated.

Identification of waxy trait QTLs by GWAS

Under linear mixed model (LMM) with kinship matrix (K) and top three PCs (Q), GWAS was performed to study the genetic basis of endosperm types. Quantile-quantile (Q-Q) plot showed that LMM efficiently controlled population structure and relationships as there was no inflated P values and a majority (95%) of markers exhibited P value equal to or lower than the expected with accordance to null hypothesis (Fig. 2a, b and c). Finally, a total 3338 SNPs located in 399 annotated genes (including gene region and 2kb promoter region) were identified to associate with endosperm type with threshold of $-\log(P)=5.6$ (Table S2). Taking into account the large genetic differences between

the glutinous accessions of *japonica* and *indica* (Fig. 1a and b), we further conducted GWAS of *indica* and *japonica* to explore subpopulation-specific waxy genes. According to the above criteria, a total of 2670 and 1034 associated SNPs were identified in *indica* and *japonica*, located in 262 and 156 annotated genes, respectively (Fig. 2b and c, Table S2). The GWAS detection efficiency of the whole panel was higher with the most associated sites identified (Table S2). By comparing GWAS results of three populations, 1424 significant SNPs (53.3%) of *indica* and 727 significant SNPs (70.3%) of *japonica* could be detected in the whole panel (Table S2). Interestingly, a certain degree of significant loci was detected simultaneously among the whole panel and subpopulations, including 244 common SNPs located in 32 annotated genes (Fig. 2d and e, Table S3), indicating that these genes were important and have conserved gene response for endosperm type between subpopulation.



Exploration of candidates for endosperm type in rice

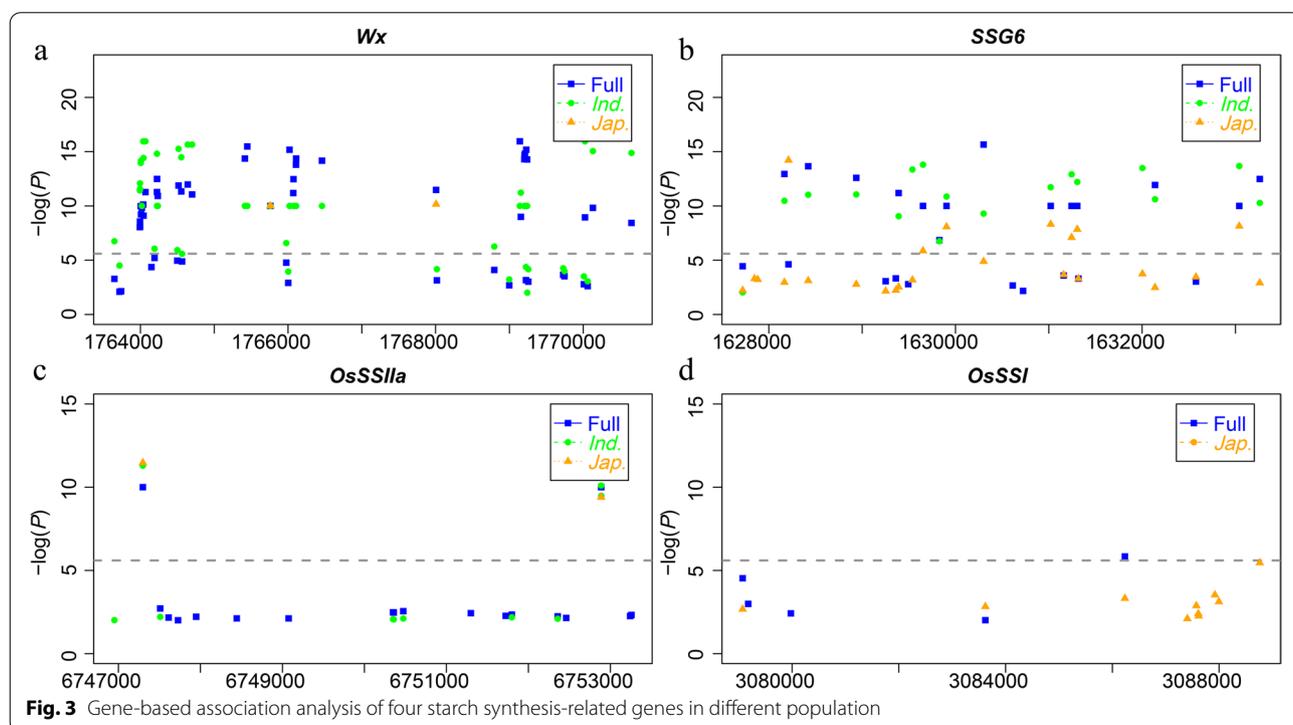
The real genes related to rice endosperm type were required to be adequately expressed in seeds at

grain filling stage such as *OsAGPL2*, *Wx* and *OsSSI-Ila* (Fig. S1). To further screen candidate genes for endosperm types in QTL regions, we firstly analyzed

the expression level of candidate genes in rice seeds at two periods (7–8 and 10–14 days after flowering) of seeds development in rice. Among them, 127 of 399 candidates in whole rice panel, 81 of 262 candidates in *indica*, and 48 of 156 candidates in *japonica* showed moderate expression at least one period (FPKM and RPKM > 10), including the 32 annotated genes detected simultaneously in different population (Table S4). To further verify the reliability of combined analysis of GWAS, and expression level, the comparison between GWAS detected candidate genes and the known waxy genes was performed. Three starch synthesis-related genes, *Wx* (*LOC_Os06g04200*), *SSG6* (*LOC_Os06g03990*) and *OsSSIa* (*LOC_Os06g12450*) were detected among three populations, respectively. As Manhattan plots showed, these known genes showed top signals in whole rice panel and subpopulations (Fig. 3a, b and c). Meanwhile, *OsSSI* (*LOC_Os06g06560*) showed association with endosperm types in the GWAS results of whole rice panel (Fig. 3d). The comparison of the GWAS results with known starch synthesis-related genes indicated that the GWAS results for endosperm type were credible. The four known genes were key loci for natural variations of rice endosperm type, and the other 28 genes are important gene for endosperm type needed to be further verified by transgenic experiment.

Natural variation in three key genes responsible for rice endosperm type

The exploration of natural variation of key endosperm type is beneficial in breeding for high-quality rice. We performed haplotype analysis to identify their elite alleles of the three key genes (*Wx*, *SSG6* and *OsSSIa*) for rice endosperm types. Firstly, association analysis of candidate genes was performed between endosperm types and 537 SNPs with MAF > 0.01 located in three known genes. Of these, 100 SNPs were significant associated with rice endosperm type ($-\log(p) > 2$). Here, we focused on non-synonymous SNPs, SNPs at splice site and SNPs in promoter (Table S5), as these SNPs could be responsible for functional variation through changes in expression and protein sequence [29–31]. A total of 37 significant SNPs were identified within *Wx* gene, including two non-synonymous SNP, one SNP at split site, and 34 in promote or UTR regions. Twenty-six haplotypes, named *Wx-1* to *Wx-26*, were identified in whole panel (Fig. 4a). Twenty-four of 26 haplotypes were detected in *indica*, eight of which showed moderate frequencies ranging from 5 to 23.3%. By comparison, 47.8 and 24.1% *japonica* carried *Wx-8* and *Wx-9*, suggesting there were large genetic variation of *Wx* in *indica* than *japonica* (Fig. 4a). Previous studies showed that Chr6_1765761 was a key functional SNP for post-transcriptional modification of *Wx* [32]. The mutant of fifth exons of *Wx^b* induced to lower AC than that of glutinous rice. In our study, we did not



a

HAP	Promoter and 5'UTR																	Nonsys	Num of accessions																							
	1763647	1763716	1763739	1763992	1763995	1764004	1764019	1764038	1764068	1764149	1764178	1764224	1764235	1764499	1764516	1764533	1764564		1764700	1765415	1765448	1765761	1765976	1766003	1766020	1766071	1766074	1766108	1766110	1766460	1768006	1768998	Ind./Glu ind	Jap./Glu jap								
Wx-1	A	A	A	A	T	T	G	T	C	A	G	T	C	G	A	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	6/0	20/0			
Wx-2	A	A	A	A	T	T	G	T	C	A	G	T	C	G	C	C	G	T	T	G	C	C	A	T	A	G	T	G	A	C	G	T	G	T	C	T	C	C	A	C	7/0	20/0
Wx-3	A	A	A	A	T	T	G	T	C	A	G	T	T	G	A	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	10/1	28/0			
Wx-4	A	A	A	A	T	T	G	T	C	A	G	T	T	G	C	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	4/0	7/0			
Wx-5	A	A	A	A	T	T	G	T	C	A	G	T	T	G	C	C	G	T	T	G	C	C	A	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	9/0	19/0		
Wx-6	A	A	A	C	G	G	T	A	C	G	A	C	C	G	C	T	A	T	T	T	A	C	G	A	C	G	A	C	A	C	A	C	A	C	T	T	C	C	0/0	10/0		
Wx-7	A	A	A	C	G	G	T	A	T	G	A	C	C	G	A	C	T	A	C	T	T	A	C	G	A	C	T	A	C	A	C	A	C	A	C	T	T	A	C	5/0	37/0	
Wx-8	A	A	A	C	G	G	T	A	T	G	A	C	C	G	C	T	A	C	T	T	T	A	C	G	A	C	G	A	C	A	C	A	C	A	C	T	T	C	C	65/0	343/5	
Wx-9	A	A	A	C	G	G	T	A	T	G	A	C	C	G	C	T	A	C	T	T	T	A	C	G	A	C	T	A	C	A	C	A	C	A	C	T	T	A	C	108/37	173/51	
Wx-10	A	A	A	C	G	G	T	A	T	G	A	C	C	G	C	T	A	C	T	T	T	A	C	G	A	C	T	A	C	A	C	A	C	A	C	T	T	C	C	0/0	21/7	
Wx-11	G	A	A	C	G	G	T	A	T	G	A	C	C	G	C	T	A	C	T	T	T	A	C	G	A	C	T	A	C	A	C	A	C	A	C	T	T	A	C	23/9	0/0	
Wx-12	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	12/0	1/0			
Wx-13	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	G	T	C	G	C	G	T	A	G	T	G	G	T	G	T	C	T	C	C	A	C	38/0	1/0			
Wx-14	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	G	T	C	G	C	G	T	A	G	T	G	G	T	G	T	C	T	C	C	A	T	90/0	0/0			
Wx-15	G	G	A	A	T	T	G	T	C	A	G	T	C	G	A	C	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	65/1	4/0		
Wx-16	G	G	A	A	T	T	G	T	C	A	G	T	C	G	A	C	C	G	T	C	G	C	G	T	A	G	T	G	G	T	G	T	C	T	C	C	A	C	20/0	0/0		
Wx-17	G	G	A	A	T	T	G	T	C	A	G	T	C	G	A	C	C	G	T	C	G	C	G	T	A	G	T	G	G	T	G	T	C	T	C	C	A	T	142/3	1/0		
Wx-18	G	G	A	A	T	T	G	T	C	A	G	T	C	G	C	C	G	T	T	G	C	A	T	A	G	T	G	G	C	G	T	C	T	C	C	A	C	9/0	3/0			
Wx-19	G	G	A	A	T	T	G	T	C	A	G	T	C	G	C	C	G	T	T	G	C	A	T	A	G	T	G	G	T	G	T	C	T	C	C	A	C	13/0	1/0			
Wx-20	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	C	G	C	G	T	A	A	C	T	A	C	A	C	A	C	T	T	A	C	11/5	0/0		
Wx-21	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	C	G	C	G	T	A	G	T	G	A	C	G	T	C	T	C	C	A	C	113/2	13/0		
Wx-22	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	T	G	C	A	T	A	G	T	G	G	T	G	T	C	T	C	C	A	C	64/0	1/0		
Wx-23	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	C	G	C	G	T	A	G	T	G	G	T	G	T	C	T	C	C	A	T	273/3	4/0		
Wx-24	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	T	G	C	A	T	A	G	T	G	G	C	G	T	C	T	C	C	A	C	45/0	1/0		
Wx-25	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	T	G	C	A	T	A	G	T	G	G	C	G	T	C	T	C	C	C	C	28/1	6/0		
Wx-26	G	G	A	A	T	T	G	T	C	A	G	T	T	G	A	C	C	G	T	T	G	C	A	T	A	G	T	G	G	T	G	T	C	T	C	C	A	C	11/0	4/0		

b

HAP	Promoter										5'UTR	Nonsys	3'UTR	Num of accessions					
	1627720	1627847	1627881	1628167	1628211	1628422	1628937	1629252	1629361	1629391	1629494	1629652	1629828	1629904	1630301	1633040	1633260	Ind./Glu ind	Jap./Glu jap
SSG6-1	A	C	T	C	T	C	T	G	T	A	T	A	T	G	C	G	C	1/0	82/0
SSG6-2	A	T	G	C	A	C	T	G	T	A	T	A	T	G	C	G	C	197/6	20/0
SSG6-3	A	T	G	C	T	C	T	G	T	A	T	A	T	G	C	G	C	563/18	91/0
SSG6-4	A	T	T	C	A	C	T	G	T	A	T	A	T	G	C	G	C	13/0	0/0
SSG6-5	G	T	G	A	A	G	C	G	T	T	T	T	G	A	T	A	T	3/0	117/47
SSG6-6	G	T	G	A	T	G	C	A	A	T	A	A	G	G	C	G	T	10/0	154/5
SSG6-7	G	T	G	A	T	G	C	G	T	T	T	T	G	A	T	A	T	131/41	236/20
SSG6-8	G	T	G	C	A	C	T	G	T	A	T	A	T	G	C	G	C	289/4	11/0
SSG6-9	G	T	G	C	T	C	T	G	T	A	T	A	T	G	C	G	C	57/0	75/1

c

HAP	Promoter					5'UTR	CDS	3'UTR		Num of accessions		
	6747298	6747510	6747612	6747727	6747948	6748443	6752357	6752888	6753241	6753257	Ind./Glu ind	Jap./Glu jap
OsSSIIa-1	A	A	T	C	G	A	A	T	A	A	296/58	212/54
OsSSIIa-2	G	A	T	C	G	A	A	C	A	A	384/7	375/18
OsSSIIa-3	G	A	T	C	G	A	A	C	G	G	12/0	0/0
OsSSIIa-4	G	G	C	A	C	T	G	C	G	G	550/4	215/1
OsSSIIa-5	G	G	C	A	G	A	G	C	G	G	35/0	4/0

Fig. 4 Haplotype analyses for three key cloned genes conferring waxy traits. Haplotypes filled in color are waxy haplotypes. Number on the left and right represent the number of total accessions and glutinous accessions in *indica* and *jaпонica*, respectively

detect a unique waxy haplotype of *Wx*. *Wx*-9 (allele T at Chr6_1765761) considered as the waxy haplotype, with 37 of 108 in *indica* and 51 of 173 in *japonica* were glutinous rice (Fig. 4a and Fig. S2). The results suggested that *Wx* was not the only key gene accounted for natural variation in proportion of amylopectin and amylose in rice. Waxiness of rice, as a physiological trait, is often the result of the continuous joint change of multiple biochemical processes of starch biosynthesis.

Based on 17 significant SNPs within *SSG6* (twelve in the promoter, one in the 5'UTR, 3 non-synonymous SNPs and one in the 3'UTR), nine major haplotypes, named *SSG6*-1 to *SSG6*-9, were identified in whole panel. *SSG6*-2, *SSG6*-3, *SSG6*-7 and *SSG6*-8 were predominantly represented *indica* varieties, accounting for 15.2, 43.4, 10.1 and 22.3% of the total, respectively. Moreover, *SSG6*-1, *SSG6*-3, *SSG6*-5, *SSG6*-6 and *SSG6*-7 were predominant within *japonica*, accounting for 10.1, 11.2, 14.4, 19.0 and 29.1% (Fig. 4b). The results indicated the existence of a certain degree of genetic differentiation of *SSG6* between *indica* and *japonica*, although there were two shared haplotypes between *indica* and *japonica*. Further study showed that *SSG6*-7 could be considered as main waxy haplotype, due to that 41 of 131 *indica* and 20 of 236 *japonica* carrying *SSG6*-7 were glutinous rice. Additionally, a *japonica*-special glutinous haplotype *SSG6*-5 was detected with 47 of 117 *japonica* carrying *SSG6*-5 were glutinous rice (Fig. 4b and Fig. S2). Meanwhile, we detected 5 haplotypes (named *OsSSIIa*-1 to *OsSSIIa*-5) of *OsSSIIa* gene, based on 10 significant SNPs (five in promoter, one in 5'UTR, two non-synonymous SNPs and two in 3'UTR). *OsSSIIa*-1, *OsSSIIa*-2 and *OsSSIIa*-3 were predominant in whole panel (Fig. 4c). There was no obvious genetic differentiation of *OsSSIIa* between *indica* and *japonica*. *OsSSIIa*-1 could be considered as waxy haplotype, as 58 of 296 *indica* and 54 of 212 *japonica* accessions carrying *OsSSIIa*-1 were glutinous rice (Fig. 4c and Fig. S2).

Taken together, we identified the key glutinous rice haplotype of each gene (Fig. S2), although none of them completely determined the waxiness of rice. Furthermore, it provides an important message that waxiness of rice, as a physiological trait, is also determined by a complex network, rather than simple genes in the biochemical synthesis pathway in the traditional sense. To prove the above hypothesis, we first examined the geographical distribution of different haplotype combinations of the three genes. Totally, there were 27 haplotype combinations in 124 glutinous rice accessions, haplotype combinations with more than three accessions were listed (Fig. 5a). Among 75 glutinous rice of SEA, 33 accessions carried the haplotype combination of *Wx*-9, *SSG6*-7 and *OsSSIIa*-1, 12 accessions carried the haplotype

combination of *Wx*-9, *SSG6*-5 and *OsSSIIa*-1. The haplotype combination of *Wx*-9, *SSG6*-5 and *OsSSIIa*-1 was also the predominant in SER, while most glutinous accession of EAR carried the haplotype combinations of *Wx*-10, *SSG6*-7 and *OsSSIIa*-1 or *Wx*-10, *SSG6*-5 and *OsSSIIa*-1 (Fig. 5b), indicating the combination of the waxy haplotypes of three genes were conducive to the formation of glutinous rice in SEA, SER and EAR.

Population structure and genetic differentiation of three key genes between both endosperm types

The sequence alignment of three key waxy genes and geographical distribution of their different haplotype combinations suggested that the genetic differences underlying waxiness trait among regions was greater than that between subpopulations in rice. To confirm the above hypothesis, we investigated the population structure and admixture patterns of each gene in the whole rice panel. We first estimated ancestry proportions of *Wx*, *SSG6* and *OsSSIIa* for individuals by Admixture. Population structure based on each of three genes showed different genetic structures from the whole genome. Admixture model using 202 SNPs within *Wx* gene indicated that 53 of 70 glutinous *indica* accessions clustered with glutinous *japonica* accessions (Fig. 6a). Meanwhile, admixture model using 123 SNPs within *SSG6* gene indicated that 42 of 70 glutinous *indica* accessions clustered with glutinous *japonica* accessions, and one glutinous *japonica* accession clustered with other 28 glutinous *indica* accessions (Fig. 6b). Additionally, admixture model using 194 SNPs within *OsSSIIa* gene showed that 66 of 70 glutinous *indica* accessions clustered with glutinous *japonica* accessions, and one glutinous *japonica* accession clustered with other 4 glutinous *indica* accessions (Fig. 6c). The results confirmed that there was no obvious genetic differentiation of the three key waxy genes between *japonica* and *indica* distributed in SEA, SER and EAR, which was supported by further PC analysis (Fig. 6d, e and f).

The exceptional genetic similarity among glutinous rice revealed by PC and admixture analyses could be caused by a unique domestication process. The origin of waxy haplotypes of the three genes and how they spread in *japonica* and *indica* rice are two key issues to reveal the formation of glutinous rice. Here, we firstly examined haplotypes of three known genes in wild rice. There were 72, 64 and 52 haplotypes in *Wx*, *SSG6* and *SSIIa* of wild rice. The waxy haplotypes *Wx*-9 of *Wx* gene could be detected in 3 wild rice accessions, which were from Thailand and China. The results indicated that the waxy haplotype *Wx*-9 could be inherited from wild rice, but it is a very unlikely scenario that all waxy haplotype in both rice subpopulations originate directly

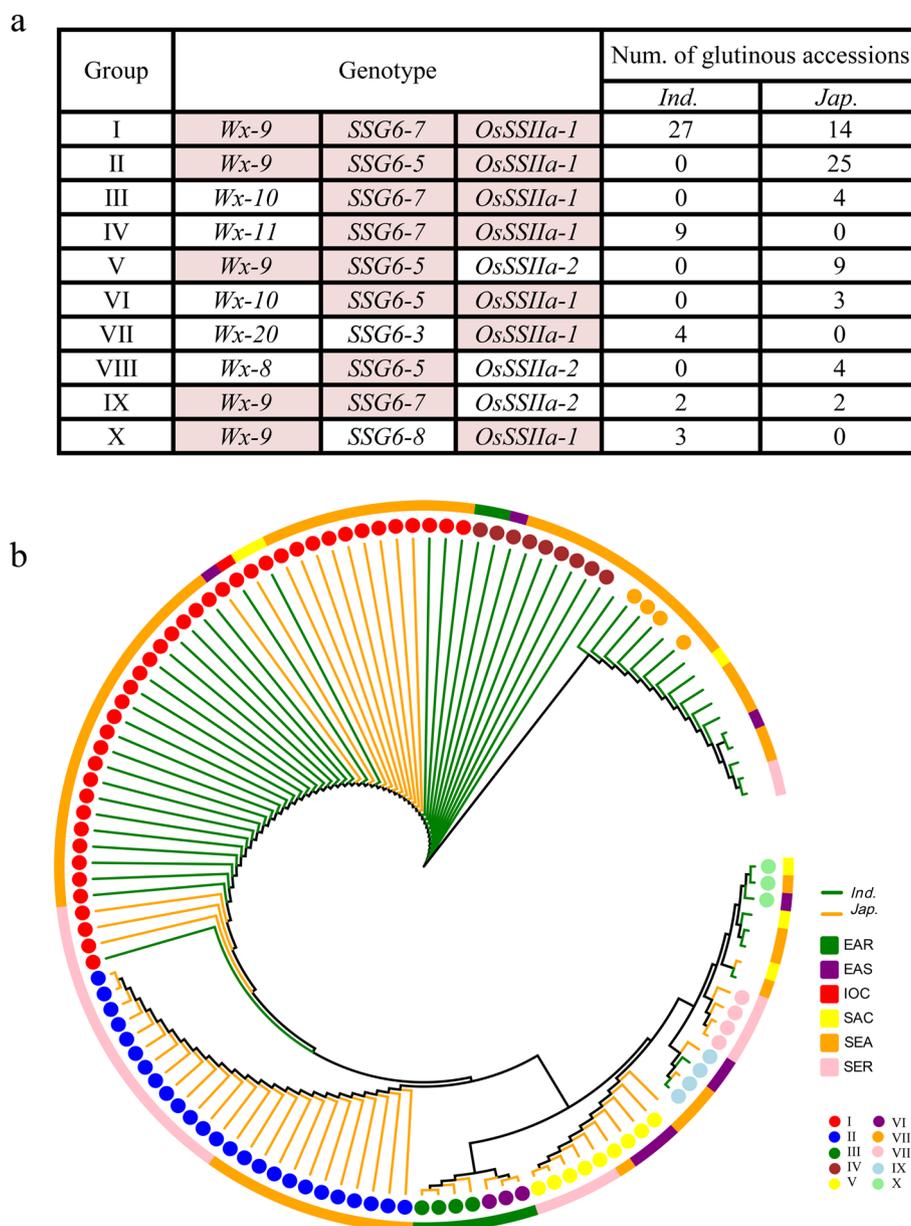
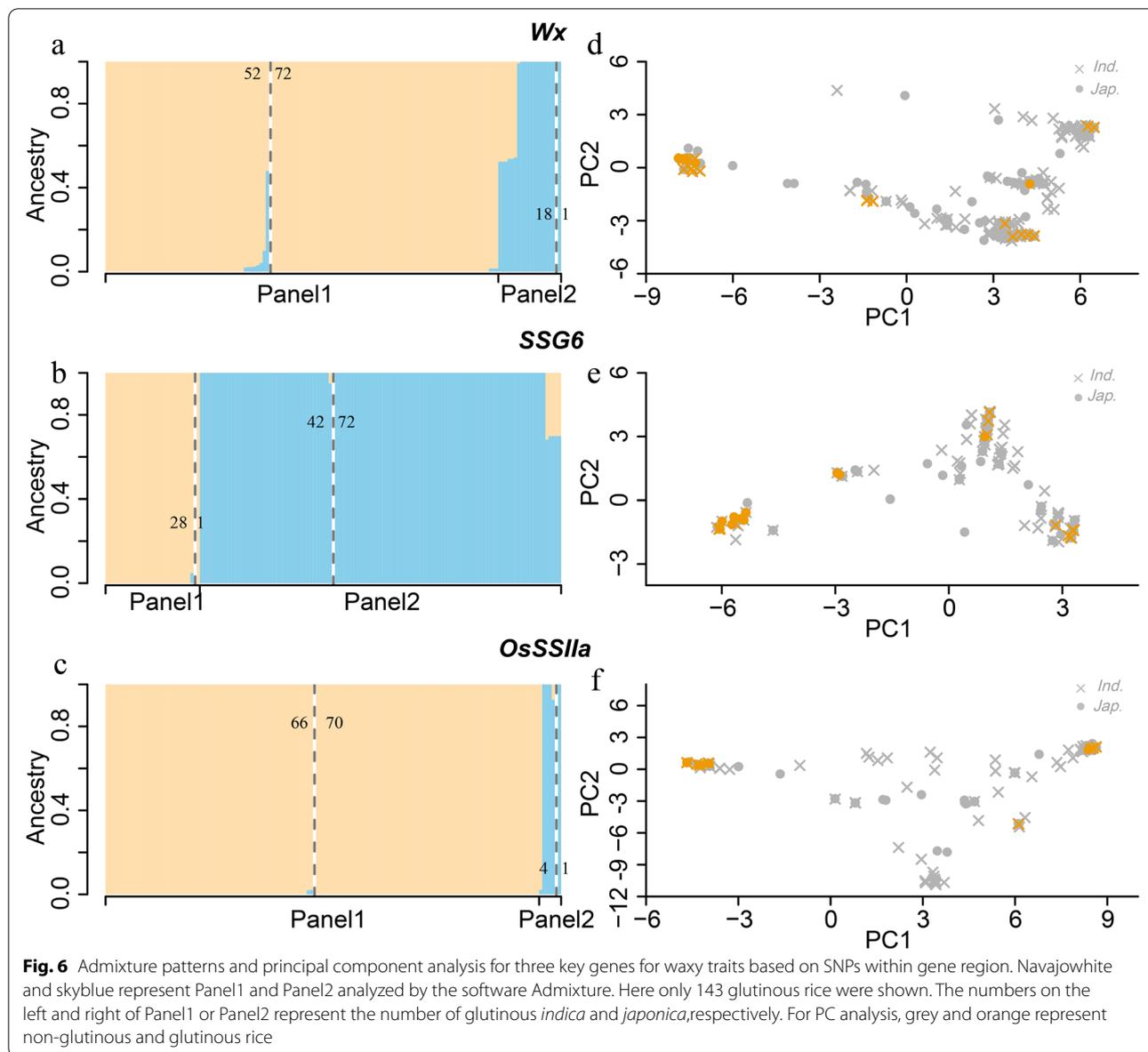


Fig. 5 Combined haplotypes and phylogenetic tree of haplotype combinations based on the haplotype of *Wx*, *SSG6* and *OsSSIIa*. **a** Combined haplotypes of *Wx*, *SSG6* and *OsSSIIa* in 124 glutinous accessions. The haplotype combinations with more than 3 accessions are listed. Haplotypes filled in color are waxy haplotypes. **b** The phylogenetic tree based on combined haplotypes of *Wx*, *SSG6* and *OsSSIIa* in 124 glutinous accessions. Linear represent subpopulation. Colored squares represent different regions. Colored circles represent different combined haplotypes

from a small amount of wild rice (Fig. 7a). Additionally, none of wild rice carried waxy haplotypes *SSG6-5* and *SSG6-7* of *SSG6* and waxy haplotype *OsSSIIa-1* of *OsSSIIa* (Fig. 7b and c), suggesting that the waxy haplotypes of *SSG6* and *OsSSIIa* newly generated during rice domestication. Taken together, a more possible hypothesis for the exceptional genetic similarity among glutinous rice is substantial local gene flow of *Wx*, *SSG6* and

SSIIa between *indica* and *japonica* in SEA, EAR, and EAR.

To further determine the hypothesis of gene flow and examine the direction of gene flow, we performed phylogenetic analyses using all haplotype types of each gene. For *Wx* gene, waxy haplotype *Wx-9* clustered with other *japonica* haplotypes and formed a monophyletic group (Fig. 7a). Meanwhile, two waxy haplotypes *SSG6-5* and



SSG6-7 of *SSG6* clustered together with long genetic distance to other haplotypes of cultivated rice (Fig. 7b). Additionally, waxy haplotype *SSIIa-1* clustered with *SSIIa-2* and two wild haplotypes (Fig. 7c). Phylogenetic trees in cultivated rice indicated that the waxy haplotypes of each gene were closer to their corresponding *japonica* haplotypes than *indica* haplotypes, such as *Wx-9* closed to *Wx-7/8/10*, *SSG6-5/7* closed to *SSG6-6*, and *OsSSIIa-1* mainly closed to *japonica* as the haplotype *OsSSIIa-2* account for 46.5% of total in *japonica* accessions (Fig. S3). Further analyses of introgressed regions in 143 glutinous rice showed that the waxy haplotype *Wx-9*, *SSG6-7* and *OsSSIIa-1* were transferred from *japonica*

to the *indica* population (Fig. 8). Above all, we suggested that glutinous haplotypes of the three genes in *japonica* rice firstly evolved or were directly inherited from wild rice, and then introgressed into *indica* rice in SER, SEA and EAR.

Discussion

Waxiness is a complex polygenic trait

The endosperm type is an important characteristic of rice quality. Glutinous rice is one of the traditional grains loved by mankind. There are various reports of research on waxiness in the past decades, but only *Wx* was cloned. In order to analyze the genetic composition more

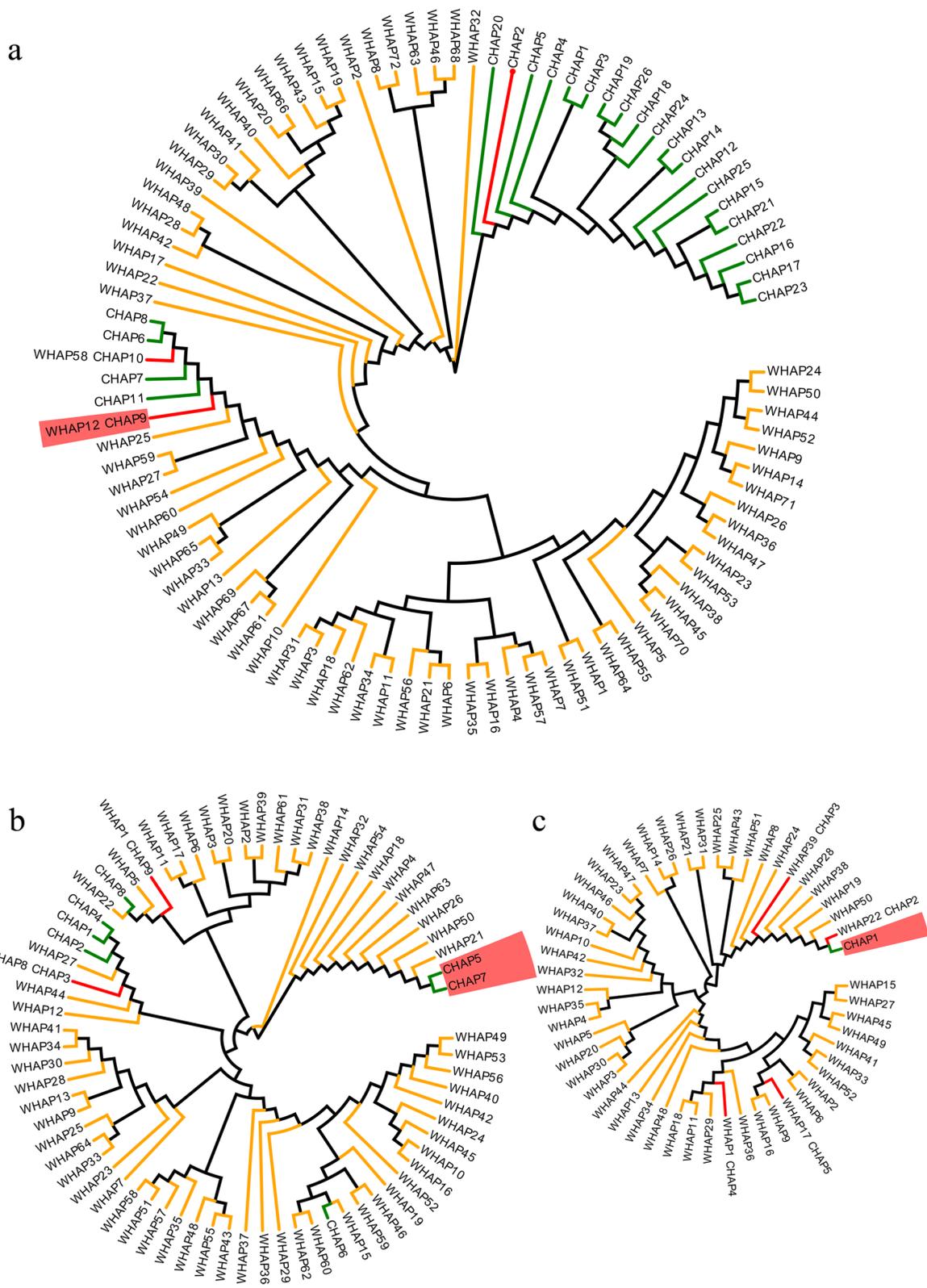


Fig. 7 Phylogenetic tree of three keys gene based on haplotypes on cultivated rice and wild rice. **a**, **b** and **c** represent *Wx*, *SSG6* and *OsSSIIa*. The prefix W stands for wild rice and the prefix C stands for cultivated rice. Green, orange and red linear represent haplotypes in cultivated, wild rice and shared between them, taxa name with red background represents waxy haplotypes of each gene

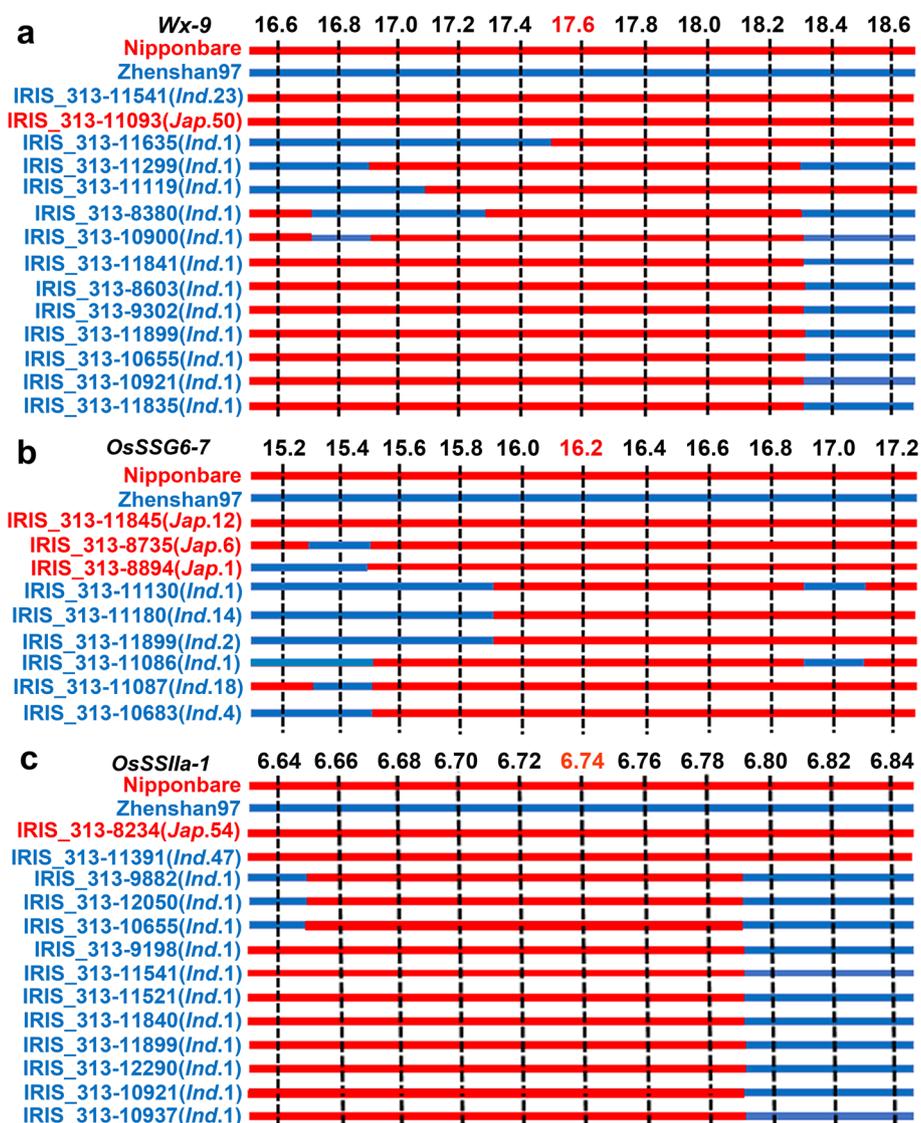


Fig. 8 Analyses of introgressed regions for *Wx*, *OsSSG6* and *OsSSIIa*. Eleven SNP markers in 200 kb upstream and downstream intervals of these genes were used. *Ind.*, *indica*. *Jap.*, *japonica*. Nipponbare and Zhenshan97 were reference sequences. Red and blue bars represent *japonica* and *indica* genotypes respectively. Markers at 17.6, 16.2 and 6.74 Mb were the closest to *Wx*, *OsSSG6* and *OsSSIIa* respectively. The number next to *Ind.* and *Jap.* represent the number of glutinous *indica* and *japonica*, respectively

accurately, it is necessary to identify more genes that regulate this trait. Through genome wide association analysis in this study, a large number of waxiness-related loci were identified, indicating that the waxiness trait is a complex quantitative trait controlled by multiple genes. The *Wx*, *SSG6* and *OsSSIIa* as the starch synthesis controlling genes were identified in full population and subpopulations with strong correlation signals in GWAS result, are proved as key genes controlling waxiness. Sequence analysis showed several alleles of *Wx* for AC [7, 32–36]. Previous studies showed that the base substitutions in coding

sequence of *ALK* may cause the alteration in gelatinization temperature and the mutation in *ssg6* introduced a premature stop codon developed enlarged starch grains in endosperm [37–39]. In this study, natural variations for waxiness in three key genes were investigated on the basis of 2108 rice accessions. The 26, 9, and 5 haplotypes of *Wx*, *SSG6*, and *OsSSIIa* were found, respectively. Further analysis showed that *Wx-9*, *SSG6-5/7* and *OsSSIIa-1* were considered as the waxy haplotypes. The research provides important breeding variations for glutinous rice breeding.

Three key genes cooperate in evolution of glutinous rice

In previous studies, *Wx* was observed as the key gene for waxiness. Other widely conducted evolutionary studies and the manipulation of *Wx* in rice breeding indicated that the *Wx* cooperated with others starch synthesis gene to form a fine regulating network that controls the eating and cooking quality [7, 20, 35, 40]. In our research, through the combination of haplotypes, the interaction between the three genes (*Wx*, *SSG6*, *OsSSIIa*) was studied, which provided the reference for molecular pyramid breeding. The combined haplotype analysis of *Wx*, *SSG6* and *OsSSIIa* showed *Wx-9 / SSG6-5(7)* and *OsSSIIa-1* genotypes may be the best allele combination for waxy rice and quality breeding, because the combination of these three alleles accounted for the largest proportion in the main distribution area of waxy materials. This study revealed the potential gene combination types of glutinous rice that are popular in Southeast Asia. Population structure and genetic differentiation of three waxiness-related genes showed that glutinous haplotypes of the three genes in *japonica* rice firstly generated or were directly inherited from wild rice, and then flowed into *indica* rice in SER, SEA and EAR. Although the mechanism of these genes to control waxiness is still not clear, elucidating their molecular characteristics and evolutionary patterns in rice germplasm will help to promote glutinous rice breeding. These cloned genes and ongoing gene cloning work will provide a comprehensive understanding of mechanism behind the waxiness, which can then be applied to design varieties of the desired quality.

Conclusions

As a globally known staple food, rice is well domesticated in the world. Various regions have different food preferences, which lead to design the goals for breeding programs. Glutinous rice as a special endosperm type is also consumed as a staple food in East Asian countries by consumers' preference. But no genetic study on development and evolution of glutinous rice, specifically in this region has been reported. Here, 2108 rice germplasm were used in GWAS to identify the significant loci and candidate genes controlling the development of glutinous rice. Candidate genes were screened in whole rice panel, *indica*, and *japonica* sub-populations with transcriptome analysis. There were 32 genes, including three starch synthesis-related genes *Wx*, *SSG6*, and *OsSSIIa*, detected simultaneously in the whole rice panel and subpopulations, playing important role in determining glutinous rice. Combined haplotype analysis revealed that the waxy combined haplotype of three genes mainly distributed in Southeast Asia (SEA), SEA islands (SER) and East Asia islands (EAR). This study provides valuable information for further gene discovery and understanding the

evolution and formation for glutinous rice in SEA, SER and EAR.

Methods

Plant material

A total of 2108 cultivated rice varieties were used in the present study, which were obtained from the 3000 Rice Genome Project (3KRGP) [26, 41]. The phenotype of endosperm type for the 3000 rice varieties collected from the International Rice Genebank Collection Information System (IRGCIS) was recorded in phenotype data dictionary. The endosperm type recorded as three grades, non-glutinous (non waxy) glutinous (waxy) and intermediate. In our study, 1965 non-glutinous (or non-waxy) and 143 glutinous (or waxy) rice were used to study glutinous (<http://snp-seek.irri.org/>). Additionally, 446 wild rice accessions from a previous report were used to study evolutionary aspects [28].

Population structure and genetic differentiation

In the whole accessions panel, 5,039,852 independent SNPs across the whole genome determined by PLINK (window size 50, step size 50, $R^2 \geq 0.3$) [42] were used to population structure and admixture patterns analysis by Admixture and GAPIT [43, 44]. For the population structure against the three known genes, SNPs located in gene region and 2kb promoter region were used. Meanwhile, these SNPs were used to construct neighbor-joining tree in accessions panel with and/or without 446 wild rice accessions. Neighbor-joining tree were developed in MEGA version 7 with the bootstrap method and 1000 replicates [45].

Genome-wide association study

Total of 5,039,852 SNPs with missing rates $\leq 50\%$ and minor allele frequencies $\geq 5\%$ obtained from 3000 Rice Genome Project (3KRGP) using an in-house Perl script. GWAS were performed using GAPIT under the LMM model [46, 47]. Here, the top three principal components (PCs) were used to estimate population structure. Given that it was too stringent for significant association detection when the threshold was derived from the total number of markers [47], the threshold to control the type I error rate was defined at $-\log(p) = 5.6$ after Bonferroni-adjusted correction [48].

Candidate genes expression analysis

According to the results of the association analysis, genes with significant loci were screened, including gene region and 2kb promoter region. To study the expression pattern of each gene, two sets of transcriptome data of rice seeds at two periods (GSE98924 for 7–8 days and GSE132303 for 10–14 days after flowering) were obtained

from NCBI (<https://www.ncbi.nlm.nih.gov/guide/genes-expression>), Gene expression with the value of FPKM or RPKM > 10 in at least one set deemed as stable expression gene, then genes with stable expression in rice seeds were selected as candidate genes.

Haplotype analysis

Based on information on coding sequence (CDS) coordinates and the transcript from MSU RGAP 7, we separated non-synonymous SNPs, SNPs at splice site and SNPs in promoter from all SNPs across the 2108 accessions using an in-house Perl script. Non-synonymous SNPs, SNPs at splice site and SNPs in promoter significant associated with rice endosperm type ($-\log(p) > 2$) were used for haplotype analysis.

Introgressed regions analyses

Use the genotypes of the 3K database to perform introgressed analysis, mainly for materials containing waxy haplotypes with Wx -9 for Wx , $OsSSIIa$ -1 for $OsSSIIa$ and $OsSSG6$ -7 for $OsSSG6$ in 143 glutinous. Eleven SNP markers evenly distributed in 200 kb upstream and downstream intervals of these genes were used. Nipponbare and Zhenshan97 with better sequencing quality were the reference sequences as *japonica* and *indica*, respectively.

Abbreviations

GWAS: Genome-wide association study; AC: Amylose content; ECQs: Eating and cooking qualities; GC: Gel consistency; GT: Gelatinization temperature; AAC: Apparent amylose contents; EAR: East Asia Islands; EAS: East Asia; SEA: Southeast Asia; SER: SEA Islands; IRRI: Philippines; OCE: Oceania; SAE: South Asia - East; SAC: South Asia - Central; SAW: South Asia - West; WAS: West Asia; IOC: Indian Ocean; EAF: East Africa; WAF: West Africa; NAF: North Africa; SAM: South America; CAM: Central America and Caribbean; NAM: North America; EUR: Europe; NA: No region information.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12863-022-01033-1>.

Additional file 1: Supplementary Table S1: List and brief introduction of 2108 rice accessions used in this study.

Additional file 2: Supplementary Table S2: Summary of SNPs associated significantly with glutinous detected in full population, *indica* and *japonica* by GWAS mapping.

Additional file 3: Supplementary Table S3: Thirty two candidate genes overlapped in full population, *indica*, and *japonica*

Additional file 4: Supplementary Table S4: Expression analysis of candidate genes in full population, *indica* and *japonica*.

Additional file 5: Supplementary Table S5: SNPs used for haplotype analysis in three key genes.

Additional file 6: Figure S1. Expression patterns of *OsAGPL2*, Wx and *OsSSIIa* by Rice eFPBrowser (<http://bar.utoronto.ca/efprice/cgi-bin/efpWeb.cgi>). **Figure S2.** The proportion of glutinous rice among different haplotypes of three key genes. **Figure S3.** Phylogenetic tree based on haplotypes of three key genes in cultivated rice.

Acknowledgments

Not applicable.

Authors' contributions

C.J. and Y.Z. contributed for conception and design of the study, and performed most of the experiments. C.J. analyzed data and wrote the initial draft. M.A.R.R. and Yh.Z. reviewed the study, edited and constructed the final manuscript. Y.Z. and Y.P. supervised and helped in each step of the study. The final manuscript was read and approved by all authors.

Funding

The present study was supported by the Project of Shandong Natural Science Foundation (Grant No. ZR201910280138), the National Natural Science Foundation of China (Grant Nos. NSFC32060452, NSFC32060454, U20A2032), the key project of Guangxi Natural Science Foundation (Grant Nos. 2018GXNSFDA281053, 2019GXNSFBA245006), the project of the central government guides the development of local science and technology (GuikeZY20198015, GuikeZY19183020, GuikeZY18057004), the project of the base and talent (Guike AD17129064 and Guike AD20297093, GuiKe2021JM50, GuiKe AD18281069, 2018-15-Z06-KF16), the Base Business Project of Guangxi Academy of Agricultural Sciences (Grant Nos. 2018 QN21), and the Project of Guangxi Rice High-quality Breeding Research (Grant No. Talent small highland for rice 160-380-16-2), Special Cultivation Program for Scientific and Technological Innovation of Xinjiang Academy of Agricultural Sciences (xjkcpy-005) and the funding bodies had no role in the design of the study; collection, analysis, and interpretation of data; and in writing the manuscript.

Availability of data and materials

The endosperm type of 2108 cultivated rice varieties used in our study can be obtained from the 3K-RG dataset: <http://snp-seek.irri.org>. The 3K-RG sequencing data used for our analyses can be obtained via project accession PRJEB6180 from NCBI (<https://www.ncbi.nlm.nih.gov/sra/?term=PRJEB6180>). The transcriptome data of rice seeds at 7-8 days and 10-14 days after flowering were obtained from NCBI (<https://www.ncbi.nlm.nih.gov/gds>), the series accession ID: GSE98924 and GSE132303.

Declarations

Ethics approval and consent to participate

The research was conducted in accordance with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora were performed.

Consent for publication

Not applicable.

Competing interests

Authors declare that they have no competing interests.

Author details

¹Rice Research Institute, Guangxi Academy of Agricultural Sciences/Guangxi Key Laboratory of Rice Genetics and Breeding, Nanning 530007, China. ²State Key Laboratory of Crop Biology, Shandong Key Laboratory of Crop Biology, College of Agronomy, Shandong Agricultural University, Tai'an, Shandong 271018, PR China. ³Shandong Rice Engineering Technology Research Center, Shandong Rice Research Institute, Shandong Academy of Agricultural Sciences, Jinan 250100, China. ⁴Department of Bioinformatics and Biotechnology, Government College University, Faisalabad 38000, Pakistan. ⁵State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Research Center of Perennial Rice Engineering and Technology in Yunnan, School of Agriculture, Yunnan University, Kunming 650500, China. ⁶Institute of Nuclear and Biological Technologies, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China.

Received: 23 June 2021 Accepted: 2 March 2022

Published online: 04 May 2022

References

- Roy SC, Shil P. Assessment of genetic heritability in Rice breeding lines based on morphological traits and caryopsis ultrastructure. *Sci Rep-UK*. 2020;10(1):1–17.
- Fitzgerald MA, McCouch SR, Hall RD. Not just a grain of rice: the quest for quality. *Trends Plant Sci*. 2009;14(3):133–9.
- Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, et al. Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature*. 2002;416(6882):701–2.
- Spielmeier W, Ellis MH, Chandler PM. Semidwarf (sd-1), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proc Natl Acad Sci U S A*. 2002;99(13):9043–8.
- Xing Y, Zhang Q. Genetic and molecular bases of rice yield. *Annu Rev Plant Biol*. 2010;61:421–42.
- Cui Y, Li R, Li G, Zhang F, Zhu T, Zhang Q, et al. Hybrid breeding of rice via genomic selection. *Plant Biotechnol J*. 2020;18(1):57–67.
- Zhang C, Yang Y, Chen S, Liu X, Zhu J, Zhou L, et al. A rare *Waxy* allele coordinately improves rice eating and cooking quality and grain transparency. *J Integr Plant Biol*. 2021;63(5):889–901.
- Clarkson JD, Watabe T, Gluck SH. Glutinous Rice in northern Thailand. *Rep Res Southeast Asia Nat Sci*. 1968;27(4):922.
- Juliano BO, Villareal CP. Grain quality evaluation of world rices. 1993.
- 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(1):e85106.
- Olsen KM, Purugganan MD. Molecular evidence on the origin and evolution of glutinous Rice. *Genetics (Austin)*. 2002;162(2):941–50.
- Bao JS, Sun M, Corke H. Analysis of the genetic behavior of some starch properties in indica rice (*Oryza sativa* L.): thermal properties, gel texture, swelling volume. *Theor Appl Genet*. 2002;104(2–3):408–13.
- Nakamura Y. Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. *Plant Cell Physiol*. 2002;43(7):718–25.
- James MG, Denyer K, Myers AM. Starch synthesis in the cereal endosperm. *Curr Opin Plant Biol*. 2003.
- Vandeputte GE, Vermeylen R, Geeroms J, Delcour JA. Rice starches. I. Structural aspects provide insight into crystallinity characteristics and gelatinisation behaviour of granular starch. *J Cereal Sci*. 2003;38(1):43–52.
- Hannah LC, James M. The complexities of starch biosynthesis in cereal endosperms. *Curr opin biotechnol*. 2008;19(2):160–65.
- Yamanaka S, Nakamura I, Watanabe KN, Sato Y. Identification of SNPs in the *waxy* gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice. *Theor Appl Genet*. 2004;108(7):1200–4.
- Jeng TL, Wang CS, Tseng TH, Wu MT, Sung JM. Nucleotide polymorphisms in the *waxy* gene of NaN-induced *waxy* rice mutants. *J Cereal Sci*. 2009;49(1):112–6.
- Xu Y, Lin Q, Li X, Wang F, Chen Z, Wang J, et al. Fine-tuning the amylose content of rice by precise base editing of the *Wx* gene. *Plant Biotechnol J*. 2021;19(1):11–3.
- 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*. 2009;106(51):21760–5.
- He Y, Han Y, Jiang L, Xu C, Lu J, Xu M. Functional analysis of starch-synthesis genes in determining rice eating and cooking qualities. *Mol Breeding*. 2006;18(4):277–90.
- Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD. Selection under domestication: evidence for a sweep in the rice *waxy* genomic region. *Genetics*. 2006;173(2):975–83.
- Zhao C, Zhao L, Zhao Q, Chen T, Yao S, Zhu Z, et al. Genetic dissection of eating and cooking qualities in different subpopulations of cultivated rice (*Oryza sativa* L.) through association mapping. *Bmc Genet*. 2020;21(1).
- Sun D, Cen H, Weng H, Wan L, Abdalla A, El-Manawy AI, et al. Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. *Plant Methods*. 2019;15(1).
- Wang X, Pang Y, Wang C, Chen K, Zhu Y, Shen C, et al. New candidate genes affecting rice grain appearance and milling quality detected by genome-wide and gene-based association analyses. *Front Plant Sci*. 2017;7.
- Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*. 2018;557(7703):43–9.
- Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *Bmc Bio*. 2011;12:246.
- Huang X, Kurata N, Wei X, Wang Z, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature*. 2012;490(7421):497–501.
- Yano K, Yamamoto E, Aya K, Takeuchi H, Lo PC, Hu L, et al. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet*. 2016;48(8):927–34.
- Zhao Y, Zhang H, Xu J, Jiang C, Yin Z, Xiong H, et al. Loci and natural alleles underlying robust roots and adaptive domestication of upland ecotype rice in aerobic conditions. *Plos Genet*. 2018;14(8):e1007521.
- Yu J, Miao J, Zhang Z, Xiong H, Zhu X, Sun X, et al. Alternative splicing of *OsLG3b* controls grain length and yield in japonica rice. *Plant Biotechnol J*. 2018; 16(9):1667–78.
- Cai XL, Wang ZY, Xing YY, Zhang JL, Hong MM. Aberrant splicing of intron 1 leads to the heterogeneous 5' UTR and decreased expression of *waxy* gene in rice cultivars of intermediate amylose content. *Plant J*. 1998;14(4):459.
- Isshiki M, Morino K, Nakajima Okagaki JR, Wessler RS, Izawa, 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.
- Wanchana S, Toojinda T, Tragoonrungrong S, Vanavichit A. Duplicated coding sequence in the *waxy* allele of tropical glutinous rice (*Oryza sativa* L.). *Plant Sci*. 2003;165(6):1199.
- Mikami I, Uwatoko N, Ikeda Y, Yamaguchi J, Hirano HY, Suzuki Y, Sano Y. Allelic diversification at the *wx* locus in landraces of Asian rice. *Theor Appl Genet*. 2008;116(7):979–89.
- Yang J, Wang J, Fan F, Zhu J, Chen T, Wang C, et al. Development of AS-PCR marker based on a key mutation confirmed by resequencing of *Wx*-mpin Milky Princess and its application in japonicasoft rice (*Oryza sativa* L.) breeding. *Plant Breeding*. 2013.
- Gao Z, Zeng D, Cui X, Zhou Y, Yan M, Huang D, et al. Map-based cloning of the *ALK* gene, which controls the gelatinization temperature of rice. *Sci China C Life Sci*. 2003;46(6):661–8.
- Fujita N, Yoshida M, Asakura N, Ohdan T, Miyao A, Hirochika H, Nakamura Y. Function and characterization of starch synthase I using mutants in Rice. *Plant Physiol*. 2006;140(3):1070–84.
- Matsushima R, Maekawa M, Kusano M, Tomita K, Kondo H, Nishimura H, Crofts N, Fujita N, Sakamoto W. Amyloplast membrane protein SUBSTANDARD STARCH GRAIN6 controls starch GRAIN size in Rice endosperm. *Plant Physiol*. 2016;170(3):1445–59.
- 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(8):1157–66.
- Lenihan J, Kvist S, Fernández R, Giribet G, Ziegler A. The 3,000 rice genomes project. *Gigascience*. 2014;3(1):7.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*. 2007;81(3):559–75.
- Martínez-Cortés G, Salazar-Flores J, Fernández-Rodríguez LG, Rubi-Castellanos R, Rodríguez-Loya C, Velarde-Félix JS, et al. Admixture and population structure in Mexican-Mestizos based on paternal lineages. *J Hum Genet*. 2012;57(9):568.
- Tang Y, Liu X, Wang J, Li M, Wang Q, Tian F, et al. GAPIT version 2: an enhanced integrated tool for genomic association and prediction. *The Plant Genome*. 2016;9(2). <https://doi.org/10.3835>.
- Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33(7):1870–74.
- Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. FaST linear mixed models for genome-wide association studies. *Nat Methods*. 2011;8(10):833–5.
- Yang W, Guo Z, Huang C, Duan L, Chen G, Jiang N, et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat com*. 2014;5(1).

48. Li M, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet.* 2012; 131(5):747–56.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

