Available online at www.sciencedirect.com

ScienceDirect

Rice Science, 2021, 28(1): 58–68

Research Paper

Genetic and Geographic Patterns of Duplicate *DPL* Genes Causing Genetic Incompatibility Within Rice: Implications for Multiple Domestication Events in Rice

XU Xun[1, 2], GE Song[1, 2], ZHANG Fumin[1, 2]

([1]*State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China;* [2]*University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Strong genetic incompatibilities exist between two primary rice subspecies, *indica* and *japonica*. However, the wild ancestors of rice, *O. nivara* Sharma et Shastry and *O. rufipogon* Griff., are genetically compatible. How this genetic incompatibility became established has not been clearly elucidated. To provide insights into the process, we analyzed a pair of hybrid sterility genes in rice, *DOPPELGANGER 1* (*DPL1*) and *DOPPELGANGER 2* (*DPL2*). Either of the two loci can have one defective allele (*DPL1-* and *DPL2-*). Hybrid pollen carrying both *DPL1-* and *DPL2-* alleles is sterile. To explore the origination of *DPL1-* and *DPL2-*, we sequenced the *DPL1* and *DPL2* genes of 811 individual plants, including *Oryza sativa* (132), *O. nivara* (296) and *O. rufipogon* (383). We then obtained 20 *DPL1* and 34 *DPL2* sequences of *O. sativa* from online databases. Using these sequences, we analyzed the genetic and geographic distribution patterns of *DPL* genes in modern rice and its wild ancestors. Compared with the ancestral populations, *DPL1-* and *DPL2-* showed reduced diversity but increased frequency in modern rice. We speculated that the diversity reduction was due to a historic genetic bottleneck, and the frequency had likely increased because the defective alleles were preferred following this artificial selection. Such results indicated that standing variances in ancestral lines can lead to severe incompatibilities among descendants. Haplotype analysis indicated that the *DPL1-* haplotype of rice emerged from an *O. nivara* population in India, whereas the *DPL2-* haplotype emerged from *O. rufipogon* in South China. Hence, the evolutionary history of DPLs conforms to the presumed multiple domestication events of modern rice.

Key words: rice; *DPL* gene; domestication; genetic incompatibility; phylogeography

It is well known that strong genetic incompatibility exists between two rice subspecies (*Oryza sativa* L.), *japonica* and *indica*. In contrast, two wild rice ancestors, *O. nivara* Sharma et Shastry and *O. rufipogon* Griff., are genetically compatible (Cai et al, 2019). Modern rice and these two wild ancestors constitute an excellent system for studying the evolutionary development of incompatible genes. The crossbreeding barriers between the two modern subspecies represent a major problem for rice cultivation. Deeper knowledge of their evolutionary history therefore has substantial practical significance for improving rice cultivation.

The genetic incompatibility between *japonica* and *indica* is a complex biological event with many loci involved (Ouyang et al, 2010; Wang et al, 2014; Xie et al, 2019), such as *S5* (Chen et al, 2008; Yang J Y et al, 2012) leading to endoplasmic reticulum stress; *Sa* (Long et al, 2008), *DOPPELGANGERs* (*DPLs*) (Mizuta et al, 2010) and *Sc* (Shen et al, 2017), resulting in hybrid pollen sterility. Additional loci have been identified by genetic maps (Li et al, 2017). Although many loci have been identified, the evolutionary

Received: 7 April 2020; **Accepted:** 20 June 2020

Corresponding author: ZHANG Fumin (zhangfm@ibcas.ac.cn)

Peer review under responsibility of China National Rice Research Institute

http://dx.doi.org/10.1016/j.rsci.2020.11.007

history of the development of overall incompatibility remains poorly understood.

In the present study, we utilized the *DPL* system as a suitable framework for exploring this challenge. This system contains only two genes: *DOPPELGANGER 1* (*DPL1*) and *DOPPELGANGER 2* (*DPL2*), both of which cause pollen sterility via reciprocal gene loss (Mizuta et al, 2010). Both genes of the *DPL* system have two alleles: *indica*-like *DPL1* (*DPL1*-, nonfunctional) and *japonica*-like *DPL1* (*DPL1*+, functional), *indica*-like *DPL2* (*DPL2*+, functional) and *japonica*-like *DPL2* (*DPL2*-, nonfunctional). Theoretically, a quarter of all pollen in an F_1 hybrid cross between *indica* and *japonica* would have *DPL1*- and *DPL2*- alleles and therefore fail to germinate.

The genetic distribution pattern of *DPLs* in modern *O. sativa* remains unknown, let alone the pattern in its wild ancestors. Modern rice includes six groups (Wang et al, 2014). *Japonica* can be divided into *temperate japonica*, *tropical japonica*, *aromatic* and *rayada*, whereas *indica* consists of *indica* and *aus*. Previous reports have indicated that *DPL1*- is fixed in *aus* but relatively rare in *indica*, while *DPL2*- is common in *japonica* (Mizuta et al, 2010; Craig et al, 2014). Among wild rice strains, *DPL1*- has been reported but *DPL2*- has not. The absence of *DPL2*- in wild rice strains indicates that the DPL system is established during domestication and not inherits from wild ancestors. Nevertheless, previous studies have lacked population sampling of wild rice strains. In addition, previous studies focused only on the major groups of *japonica* (*temperate japonica* and *tropical japonica*), and *aromatic* and *rayada* were excluded.

In this study, we used the *DPL* system to elucidate the establishment process of genetic incompatibility between *indica* and *japonica* and attempted to answer the following questions: 1) Does the *DPL* system exist in wild rice strains? and 2) How does the *DPL* system develop and what role has domestication played in this process?

RESULTS

Allele distribution

A total of 695 *DPL1* and 787 *DPL2* sequences were obtained. In addition, we collected 20 *DPL1* sequences and 34 *DPL2* sequences from online databases (Mizuta et al, 2010; Craig et al, 2014) (Table S1). Alignments with gaps produced lengths of *DPL1* and *DPL2* are 1407 bp and 439 bp, respectively.

The frequency of *DPL1*- in *O. sativa* (10.2%) was between those of *O. nivara* (37.0%) and *O. rufipogon* (5.0%) (Table 1). *DPL2*- was present only in one wild rice strain (*O. rufipogon*), with a much lower frequency (4.9%) overall than that in *O. sativa* (30.4%). In contrast to previous studies (Mizuta et al, 2010; Craig et al, 2014), we identified both *DPL1*- and *DPL2*- in wild rice strains, indicating that the *DPL* system was established prior to the domestication of rice.

DPL1- and *DPL2*- exhibited clear frequency differences among rice groups. Only *aus*, *indica* and *tropical japonica* carried *DPL1*-, with frequencies of 70.6%, 3.0% and 4.8%, respectively. *DPL2*- was detected in all the six groups of *O. sativa*, with the highest frequency in *temperate japonica* (87.8%) and the lowest frequency in *aus* rice (4.8%) (Table 1). We speculated that the *DPL* system more likely causes reproductive barriers between the two groups, *aus* and *temperate japonica*, rather than the two subspecies, *indica* and *japonica*.

Phylogeography

The distribution of nonfunctional alleles among different populations confirmed the strong geographic nature of *DPL1* and *DPL2* (Table 1). *DPL1*- was found in both *O. nivara* and *O. rufipogon* but clustered largely in *O. nivara* of South Asia (India and Nepal) and Southeast Asia (Thailand, Cambodia and Laos). *DPL2*- was only present in two *O. rufipogon* populations in southern China (GDGZ and GXHZ). Considering the geographic isolation between these strains, we speculated that the incompatible alleles failed to trigger hybrid fertility under natural conditions.

The *DPL1* and *DPL2* networks (Figs. 1 and 2) contained 20 and 5 haplotypes, respectively. The *DPL1* network (Fig. 1-A) established that the *DPL1*- allele in *O. sativa* only clustered in H_1, and the haplotype in rice had only one original source, the wild ancestors, which also contained H_1. These results stand in contrast to the multiple emergence hypothesis for *DPL1*- previously suggested (Mizuta et al, 2010). The two most frequent haplotypes of *DPL1*+ in rice (H_5 and H_36) were closely related with each other but separated from H_1, suggesting independent originations of *DPL1*- and *DPL1*+ in rice. The *DPL1*- haplotype in wild rice strains showed discontinuous geographic distribution (Fig. 1-B). H_1 was present mostly in *O. nivara* from India, and also in *O. rufipogon* from

Table 1. Nucleotide polymorphism and results of neutral test in *O. rufipogon*, *O. nivara* and *O. sativa*.

Species/Population/	Number		Frequency (%)		*S*		*π*		*h*		*Hd*		Tajima's *D*	
Group	*DPL1-*	*DPL2-*	*DPL1-*	*DPL2-*	*DPL1*	*DPL2*	*DPL1*	*DPL2*	*DPL1*	*DPL2*	*DPL1*	*DPL2*	*DPL1*	*DPL2*
O. nivara	91	0	37.0	0.0	41	12	0.00336	0.00090	23	9	0.819	0.267	-1.73831	-1.85460*
O. rufipogon	16	18	5.0	4.9	79	34	0.00421	0.00290	53	25	0.914	0.429	-2.19350**	-2.11384**
O. sativa	15	49	10.2	30.4	11	6	0.00118	0.00363	7	4	0.223	0.571	-1.21896	1.04607
nIND	18	0	45.0	0.0	12	5	0.00210	0.00140	6	5	0.638	0.410	-1.18432	-1.17921
rIND	2	0	11.8	0.0	17	8	0.00422	0.00292	9	7	0.955	0.584	-1.31101	-1.22445
nTHA	5	0	62.5	0.0	6	3	0.00370	0.00105	3	3	0.762	0.295	1.59761	-1.65231
rTHA	2	0	13.3	0.0	12	3	0.00345	0.00110	5	3	0.813	0.362	-1.08464	-1.34917
nKHM	29	0	93.5	0.0	11	1	0.00202	0.00014	5	2	0.743	0.125	-1.19393	-1.14244
rKHM	6	0	22.2	0.0	6	0	0.00182	0.00000	5	4	0.667	0.200	-0.48253	NA
nMMR	0	0	0.0	0.0	0	3	0.00000	0.00050	1	2	0.000	0.063	NA	-1.56135
rMMR	0	0	0.0	0.0	22	10	0.00429	0.00361	9	1	0.714	0.000	-1.52558	-1.00632
nNEP	0	0	0.0	0.0	4	1	0.00083	0.00028	3	3	0.530	0.170	-0.76705	-0.77374
rNEP	2	0	7.1	0.0	5	3	0.00123	0.00071	3	7	0.574	0.696	-0.51862	-1.37016
nLAO1	20	0	100.0	0.0	13	1	0.00341	0.00015	4	2	0.591	0.067	1.061730	-1.14700
rLAO1	0	0	0.0	0.0	12	13	0.00278	0.00240	8	5	0.837	0.261	-1.04354	-2.24274**
nLAO2	19	0	90.5	0.0	8	0	0.00142	0.00000	4	1	0.348	0.000	-1.54473	NA
rLAO2	1	0	3.4	0.0	13	2	0.00297	0.00102	7	2	0.817	0.228	-0.78196	-0.24788
nLKA04	0	0	0.0	0.0	1	0	0.00023	0.00000	2	1	0.200	0.000	-1.11173	NA
nLKA05	0	0	0.0	0.0	3	0	0.00187	0.00000	2	1	0.533	0.000	1.83053*	NA
nLKA07	0	0	0.0	0.0	2	0	0.00047	0.00000	3	1	0.378	0.000	-1.40085	NA
nLKA08	0	0	0.0	0.0	1	0	0.00026	0.00000	2	1	0.222	0.000	-1.88230	NA
nLKA09	0	0	0.0	0.0	0	0	0.00000	0.00000	1	1	0.000	0.000	NA	NA
nLKA10	0	0	0.0	0.0	0	3	0.00000	0.00198	1	3	0.000	0.600	NA	-0.65748
nLKA11	0	0	0.0	0.0	3	0	0.00098	0.00000	3	1	0.417	0.000	-0.93613	NA
nLKA12	0	0	0.0	0.0	4	2	0.00201	0.00091	2	1	0.429	0.000	0.48523	-1.40085
rLKA01	0	0	0.0	0.0	3	1	0.00195	0.00076	4	2	1.000	0.333	0.16766	-0.93302
rLKA06	0	0	0.0	0.0	1	0	0.00033	0.00000	1	1	0.000	0.000	-1.00623	NA
rLKA13	0	0	0.0	0.0	0	1	0.00000	0.00046	1	1	0.000	0.000	NA	-1.11173
rGDGZ	0	10	0.0	47.6	8	4	0.00152	0.00354	7	3	0.569	0.648	-1.48037	1.12924
rGXBH	0	0	0.0	0.0	6	2	0.00285	0.00198	2	2	1.000	0.400	-0.49605	-0.05002
rGXHZ	0	8	0.0	34.8	8	4	0.00213	0.00236	6	4	0.748	0.679	-0.60490	-0.12320
rGXTD	0	0	0.0	0.0	18	2	0.00440	0.00230	11	2	0.912	0.500	-1.13215	1.81115
rHNCL	0	0	0.0	0.0	7	2	0.00235	0.00249	4	3	0.810	0.607	-0.85010	1.82766
rHNJY	0	0	0.0	0.0	11	3	0.00407	0.00294	3	3	0.600	0.679	-0.38658	0.77501
rHNDZ	0	0	0.0	0.0	8	4	0.00306	0.00370	2	3	0.476	0.733	-0.51253	0.56555
rHNWC	0	0	0.0	0.0	11	5	0.00332	0.00332	7	4	0.944	0.750	-0.23157	-0.56682
rJXDX	0	0	0.0	0.0	8	2	0.00331	0.00213	6	2	0.952	0.533	-0.25553	1.03299
rYNJH	0	0	0.0	0.0	5	1	0.00251	0.00046	5	2	0.806	0.200	0.28702	-1.11173
rIDN	0	0	0.0	0.0	5	4	0.00304	0.00365	4	3	1.000	0.700	0.56199	-1.09380
rPNG	3	0	60.0	0.0	5	2	0.00281	0.00178	3	2	0.700	0.400	0.00000	0.19590
sARO	0	1	0.0	25.0	0	4	0.00000	0.00456	1	2	0.000	0.500	NA	-0.78012
sAUS	12	1	70.6	4.8	7	3	0.00286	0.00068	4	2	0.596	0.100	0.61839	-1.72331
sINDI	2	5	3.0	7.5	5	5	0.00035	0.00135	2	3	0.060	0.252	-1.64155	-0.99668
sRAY	0	2	0.0	66.7	1	3	0.00078	0.00456	2	2	0.667	0.667	NA	NA
sTEJ	0	36	0.0	87.8	1	4	0.00007	0.00172	2	3	0.056	0.292	-1.13321	-0.46012
sTRJ	1	4	4.8	16.0	7	5	0.00078	0.00417	3	4	0.186	0.684	-2.13123*	1.11236

S, Number of segregating sites; *π*, Nucleotide diversity; *h*, Number of haplotypes; *Hd*, Haplotype diversity. Sample names are composed of information of the species and their origins. Initial letters of sample names: n, *O. nivara*; r, *O. rufipogon*; s, *O. sativa*. Origins of samples: IND, India; THA, Thailand; KHM, Cambodia; MMR, Myanmar; NEP, Nepal; LAO, Laos; LKA, Sri Lanka; GDGZ, Gaozhou, Guangdong Province, China; GXBH, Beihai, Guangxi Province, China; GXHZ, Hezhou, Guangxi Province, China; GXTD, Tiandong, Guangxi Province, China; HNCL, Chaling, Hunan Province, China; HNJY, Jiangyong, Hunan Province, China; HNDZ, Danzhou, Hainan Province, China; HNWC, Wenchang, Hainan Province, China; JXDX, Dongxiang, Jiangxi Province, China; YNJH, Jinghong, Yunnan Province, China; IDN, Indonesia; PNG, Papua New Guinea; ARO, *aromatic* rice; AUS, *aus* rice; INDI, *indica* rice; RAY, *rayada* rice; TEJ, *temperate japonica* rice; TRJ, *tropical japonica* rice.
*, $P < 0.05$; **, $P < 0.02$.

Oceania. Previous studies have shown that South Asian countries play an important role in the domestication of rice (Kovach et al, 2007; Sweeney and McCouch, 2007; Vaughan et al, 2008). Therefore, we presumed that the *DPL1-* of *O. sativa* emerged in India. The reason underlying the discontinuous distribution of H_1 in wild rice strains remained unclear. H_5 and H_36 were present in wild rice strains from almost

every region, indicating that *DPL1+* was established before *O. sativa* and the two wild species diverged.

The *DPL2* network (Fig. 2-A) indicated that all *DPL2-* sequences in rice were concentrated in H_13 and the *DPL2+* appeared in H_1 and H_5. Similar to *DPL1*, the *DPL2+* and *DPL2-* alleles were separated. We speculated that *DPL2-* and *DPL2+* also originated separately in rice. H_13 included haplotypes from two adjoining *O. rufipogon* populations: rGDGZ and rGXHZ (Fig. 2-B). H_1 and H_5 included samples from all habitats of wild ancestor strains, indicating that *DPL2+* has likely existed in wild rice strains for a long time. Therefore, we speculated that *DPL2-* originated from rGDGZ and rGXHZ was subsequently inherited throughout rice.

In summary, *DPL1-* and *DPL2-* of *O. sativa* are both immediate successor strains from wild ancestor strains that originated in geographically isolated areas. *DPL1-* originated from *O. nivara* in India (nIND) whereas *DPL2-* originated from *O. rufipogon* in South

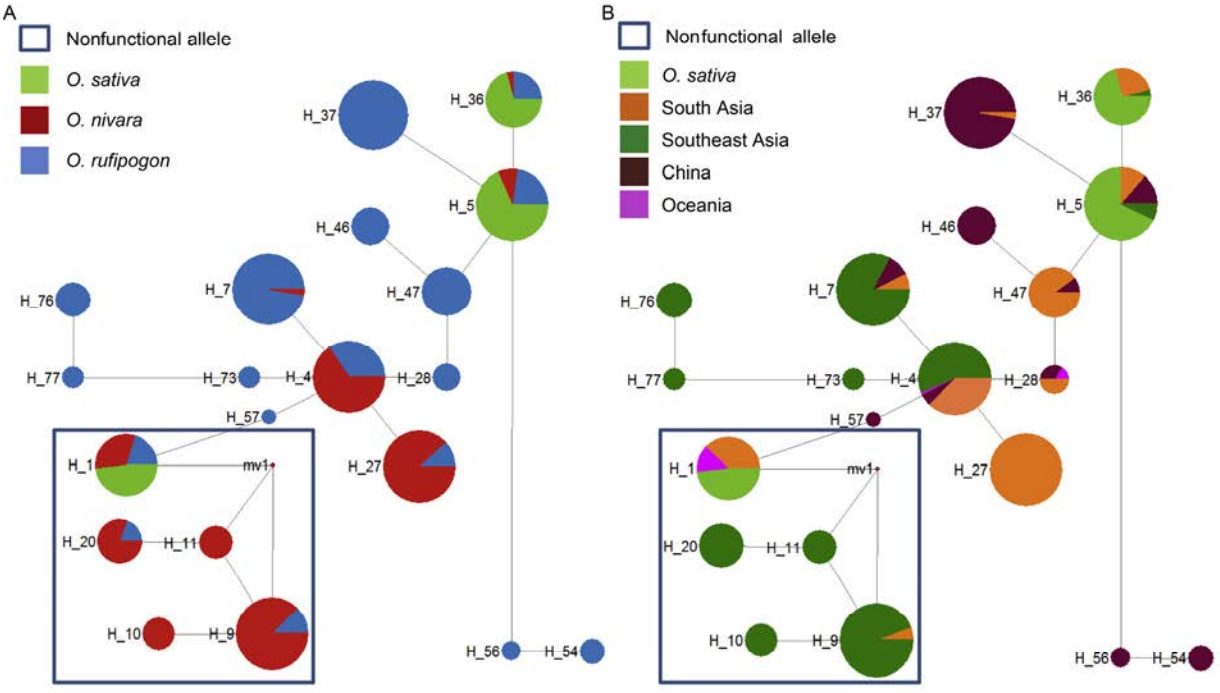

Fig. 1. Median-joining haplotype networks of *DPL1*.
Sections of each circle represent the proportion species (**A**) or locations (**B**) present in each haplotype. The size of each circle corresponds to the number of sequences of each haplotype. The lengths of the gray lines correspond to the number of mutations. Nonfunctional haplotypes are indicated by hollow squares.

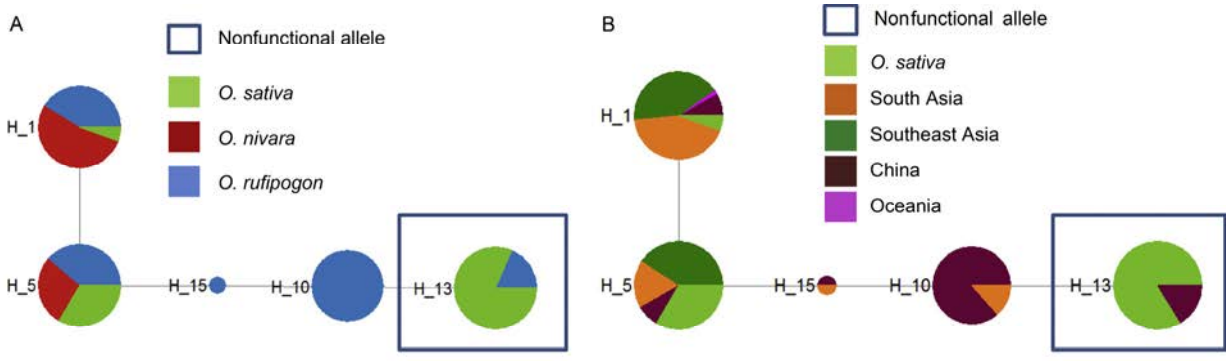

Fig. 2. Median-joining haplotype networks of *DPL2*.
Sections of each circle represent the proportion species (**A**) or locations (**B**) present in each haplotype. The size of each circle corresponds to the number of sequences of each haplotype. The lengths of the gray lines correspond to the number of mutations. Nonfunctional haplotypes are indicated by hollow squares.

China (rGDGZ and rGXHZ). The definitive origination areas for *DPL1*+ and *DPL2*+ are still unclear because of their extensive geographic distributions.

Tests for selection

Frequency variations of DPL1- and DPL2-

In addition to modern rice, *DPL2*- was identified only in the two adjoining *O. rufipogon* populations (rGDGZ and rGXHZ), we therefore merged them into one ancestral population, rGDGX. We compared the frequencies of *DPL1*- and *DPL2*- with their ancestral populations (Table 1). The frequency of *DPL1*- in *aus* was 70.6%, higher than that of nIND (45.0%). The frequency of *DPL2*- in *temperate japonica* was 87.8%, which was also higher than that of rGDGX (41.2%). These results indicated that the frequencies of *DPL1*- and *DPL2*- increased after domestication, suggesting they might be preferred in the process of artificial selection.

Neutral test

We used the Tajima's *D* statistic to assess whether *DPL1* and/or *DPL2* underwent selection, and if so, whether the selection differed among species, populations or groups (Table 1). In *O. rufipogon*, significant and negative Tajima's *D* values were detected in both *DPL1* and *DPL2*. In *O. nivara*, only the Tajima's *D* of *DPL2* was significant negative. No significant result was tested in *O. sativa*. At the population level, we found significant Tajima's *D* on *DPL2* only in the population rLAO1, and on *DPL1* only in nLKA05. Although not significant, the values of Tajima's *D* for most populations/groups were negative. One of the basic hypotheses of neutral tests is that the sample must be a population with internal gene flow (Tajima, 1989). However, the two wild species encompass many geographically isolated populations, and modern rice includes two subspecies that are severely incompatible. We speculated that the significance and lack of significant results of Tajima's *D* testing at the species level may be caused to this intraspecies structure. Methods other than neutral testing are likely required to understand the effects of artificial selection on rice *DPLs*.

An absence of polymorphic sites (*S*) was identified in several populations/groups (Table 1). This observation indicated possibly purifying selection at the population/group level, as well as a divergence of selection within species.

Nucleotide and haplotype diversity analysis

We next estimated the nucleotide diversity (π) of *DPL1* and *DPL2* in all the three species and each population/group (Table 1). In wild rice, we found that the π of *DPL1* was greater than that of *DPL2* (*O. nivara*: 0.00336 > 0.00090; *O. rufipogon*: 0.00421 > 0.00290). In modern rice, we found the opposite: the π of *DPL1* (0.00118) was lower than that of *DPL2* (0.00363). After excluding recombined sequences, the number of recombined sequences and haplotype diversity (*Hd*) are also listed in Table 1. The *Hd* of *DPL1* in *O. sativa* was 0.223, lower than that of *DPL2* (0.571). In *O. nivara* and *O. rufipogon*, the *Hd* values of *DPL1* were higher than those of *DPL2*. Therefore, we determined that artificial selection likely affected the genetic pattern of *DPLs* in rice.

To further understand the effects of selection on *DPL1*- and *DPL2*-, we compared the nucleotide and haplotype diversities of *DPL1* between *aus* and the wild ancestral population nIND, then compared *DPL1*- between them. The same comparison was used for *DPL2* and *DPL2*- between *temperate japonica* and rGDGX. The *DPL1*- of *aus* contained 12.37% nucleotide diversity and 72.43% haplotype diversity, but the number of haplotypes was equal to nIND (Fig. 3). The *DPL2*- of *temperate japonica* showed a more severe reduction in diversity, containing only one haplotype and no diversity (indicated via blanks in Fig. 3). *DPL1*- and *DPL2*- showed marked reductions in diversity compared with *DPL1* and *DPL2*. In light of the high frequencies and low diversities, we speculated that the haplotypes of *DPL1*- and *DPL2*- in *O. sativa*

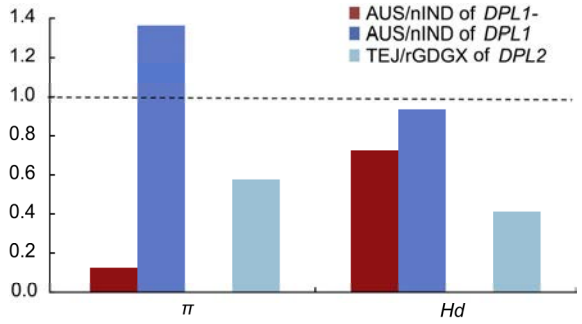

Fig. 3. Ratios of nucleotide diversity (π) and haplotype diversity (*Hd*) between groups of *O. sativa* and ancestral populations.
AUS and TEJ indicate *aus* and *temperate japonica* rice, respectively. nIND and rGDGX indicate the ancestral populations: *O. nivara* from India and the two *O. rufipogon* populations carrying *DPL2*- from South China. The ratio of *DPL2*- nucleotide diversity between *temperate japonica* and the ancestral populations is represented by a blank shape, since *DPL2*- of *temperate japonica* has no diversity.

were randomly selected due to genetic bottleneck and the nonfunctional alleles were subsequently preferred by artificial selection, which led to their increase in frequencies in specific rice groups.

DISCUSSION

Domestication events have shaped *DPL* system derived from wild ancestor strains

The origin of the genetic incompatibility between *japonica* and *indica* has been unclear since the ancestor strains *O. nivara* and *O. rufipogon* are compatible (Cai et al, 2019). In the present study, we wondered whether the incompatibility appeared during the process of domestication or whether the deficiency already existed in the ancestral strains.

To date, several major incompatible systems have been identified within rice: *Sc* (Shen et al, 2017), *S5* (Chen et al, 2008; Yang J Y et al, 2012), *DPLs* (Mizuta et al, 2010) and *Sa* (Long et al, 2008). The *Sc* system includes two alleles, *japonica* type *Sc-j* and *indica* type *Sc-i*. High-expressed *Sc-i* in sporophytic cells of hybrids leads to abortion of pollen carrying *Sc-j*. Both *Sc-j* and *Sc-i* are inherited respectively from diverged wild rice populations (Shen et al, 2017). However, the origin areas of *Sc-j* and *Sc-i* are still unclear. The *S5* system can be roughly classified into three types: *indica*-like, *japonica*-like and wide compatibility. *Indica*-like and *japonica*-like are incompatible, and the wide compatibility shows compatibility with both of the others (Du et al, 2011). All the three types are present in wild rice, which provides evidence of potential reproductive isolation (Du et al, 2011; Craig et al, 2014). *Sa* causes hybrid pollen sterility and includes two adjacent genes, *SaM* and *SaF*. *Indica* and *japonica* carry *SaM+SaF+* and *SaM-SaF-*, respectively. *SaM+*, *SaF+* and *SaM-* are necessary for pollen sterility and are present in wild rice (Long et al, 2008; Craig et al, 2014). These studies have established that several potential reproductive barriers between *japonica* and *indica* are present in their ancestors. However, it is still unclear how the incompatible alleles have been maintained under natural conditions and the influence of domestication. Ouyang and Zhang (2018) proposed three models to explain the establishment of incompatibilities within rice. The models focused on how the incompatible genes accumulated in each divergent lineage and may not be accurate for several loci due to the lack of clear evolutionary background information. For example, Ouyang and Zhang (2018) indicated that *DPL2-* did not exist in wild rice strains and suggested that the incompatibility occurred during the divergence of the two subspecies.

Our study identified that both *DPL1-* and *DPL2-* are present in wild rice. Therefore, the three major incompatibility systems are all inherited rather than *de novo* mutations that arose during domestication. Larger-scale studies involving more groups of rice are warranted. In a study including a typical *japonica* and two typical *indica*, 43 loci and 223 interactions involved in the fertility of embryo-sac, pollen and spikelet were identified (Li et al, 2017). The study failed to identify *DPLs* and *Sa* but is consistent with our findings that the *DPL* system was more likely found between *aus* and *temperate japonica* rice. We therefore hypothesized that the incompatibility between *indica* and *japonica* may contain a more complex genetic background than previously believed.

We found that domestication had affected the genetic background of *DPL1-* and *DPL2-* in rice. After determining that the *Hd* of *DPL2-* was severely reduced, we surprisingly found that the frequency of *DPL2-* was much higher in *temperate japonica* (87.8%) than the ancestral population rGDGX (41.2%). Similarly, the frequency of *DPL1-* was 45.0% in nIND, which then increased to 70.6% in *aus*, although the *Hd* of *DPL1-* was reduced in *aus* (Fig. 3). These results indicated that *DPL1-* and *DPL2-* were preferred by artificial selection.

Significant and negative Tajima's *D* values are a sign of extremely low polymorphisms and are considered a consequence of purifying selection, but population expansions or unknown structures can also lead to significant negative changes (Tajima, 1989). Neutral tests are generally designed for one population with fluid gene flow and are generally not suitable for widespread species. In the wild ancestors, we found significant and negative results of neutral testing. Considering the complex structures of *O. rufipogon* and *O. nivara* (Londo et al, 2006; Liu et al, 2015), the significance was more likely caused by the intraspecies structure rather than selection. A similar phenomenon has been reported for *S5*: the neutral test results are also significantly negative in one wild ancestor (*O. rufipogon*) but not in *O. sativa* (Du et al, 2011). We speculated this may also be due to the interspecies structure of the wild species, and detailed studies

among the groups of *O. sativa* are warranted to clearly elucidate the role of artificial selection, as well as how this selection has formed the *DPL* system.

We compared the frequencies and diversities of *DPL1*- and *DPL2*- between groups of modern rice and populations of wild rice strains. At both loci, we found a decrease in diversity but an increase in frequency (Table 1 and Fig. 3). Therefore, we believed the effects of domestication can be divided into two parts: a sampling effect caused by genetic bottleneck and subsequent artificial selection. A few haplotypes of *DPL1*- and *DPL2*- were randomly preserved, then the frequencies of nonfunctional alleles arose during long-term artificial selection.

Ancestral structures contributed to establishment of *DPL* system

Mizuta et al (2010) suggested that *DPLs* are a barrier between the two rice subspecies, *india* and *japonica*. However, their study did not consider *aus* separately from *indica*, and their results actually indicated that *DPL1*- clustered in *aus*, not in *indica*. Craig et al (2014) and our results further established that *DPL1*- has a high frequency in *aus* but a low frequency in *indica*. Therefore, we presumed that *aus* inherits *DPL1*- directly from a wild ancestor and then spreads it into the other groups. Our study indicated that *DPL2*- had the highest frequency in *temperate japonica* (87.8%). The groups belonging to *japonica* (*aromatic*, *rayada*, *temperate japonica* and *tropical japonica*) had higher frequencies than the *indica* groups (*aus* and *indica*) (Table 1). We therefore speculated that the *japonica* groups (most likely *temperate japonica*) inherit *DPL2*- first and then spread it into other groups.

DPL1 network analysis showed that the *DPL1*- haplotype of modern rice is related to nIND and *O. rufipogon* in Oceania (rPNG). India is thought to be the primary potential domestication center of *aus* (Khush, 1997; Vaughan et al, 2008; Civan et al, 2015) and is closely connected with the domestication of *O. sativa* (Morishima et al, 1992; Kovach et al, 2007; Sweeney and McCouch, 2007; Vaughan et al, 2008). *DPL2* network analysis indicated that the haplotype H_13 of rice *DPL*2- was from *O. rufipogon* in South China (rGDGZ and rGXHZ), which is generally considered the origin center of *japonica* (Khush, 1997; Londo et al, 2006; Vaughan et al, 2008; Civan et al, 2015). We speculated that the ancestral population of *DPL1*- was nIND and those of *DPL2*- were rGDGZ and rGXHZ. It appears that at least two domestication events are required for *O. sativa* to obtain both *DPL1*- and *DPL2*-. Our study is therefore consistent with the 'multiple domestications' hypothesis, in which *O. sativa* is thought to have arisen as a result of more than one domestication event (Sang and Ge, 2007; Yang C C et al, 2012).

These results raise an interesting question regarding the discontinuous distribution of H_1, which is present in wild rice strains not only from India but also from New Guinea. Similar to *DPL1*-, the study of another incompatible system in *indica* and *japonica*, *S5*, has also indicated the presence of an *indica*-like haplotype (H9) in New Guinea and South Asia but nowhere else (Du et al, 2011). Repeated extinction and colonization throughout the history of *O. rufipogon* has previously been reported (Liu et al, 2015), therefore, we suspected the discontinuous distribution of *DPL1*- is a consequence of historical population dynamics.

The *japonica* haplotype (*SaM-SaF*-) in the *Sa* system has been proved to originate from an *O. rufipogon* population in southern China, but there is no evidence for the origination area of the *indica* haplotypes (Long et al, 2008). Du et al (2011) provided the evolutionary history and geographic distribution of *S5*. Nevertheless, they did not definitively locate the origin areas of the incompatible alleles. Our study confirmed the origination centers of both parts and established that the genetic incompatibility between *indica* and *japonica* is formed from an ancestral geographic pattern.

Not solely limited to modern rice, the effects of ancestral geographic patterns on incompatible genes have been widely reported. The same mechanism occurs in other *Oryza* species. *S27*/*S28* causes hybrid pollen sterility between *O. sativa* and *O. glumaepatula*. This system is influenced by an ancestral geographic pattern and rapid sampling effects during the divergence of the two species (Yamagata et al, 2010).

DPLs are maintained in wild ancestors because of the absence of selection pressure and geographic isolation

Reproductive isolation genes lead to hybrid incompatibility, which means a certain number of hybrid descendants fail to survive or breed. Therefore, incompatible genes are generally deleterious for populations and are expected to result in loss, especially in gene flow conditions (Gavrilets, 1997; Bank et al, 2012). In contrast, incompatible genes are commonly found to

be polymorphic, especially in plants (Rieseberg and Blackman, 2010; Bank et al, 2012; Cutter, 2012; Corbett-Detig et al, 2013; Lindtke and Buerkle, 2015). An interesting question is how incompatible alleles are maintained in nature.

Our results indicated that geographic isolation is a possible explanation. *DPL1*- mainly clustered in Southeast and South Asia, whereas *DPL2*- was only found in a small area of South China (Table 1). The geographic distance prevents incompatible pairing of *DPL1*- and *DPL2*-. Furthermore, neither *DPL1*- nor *DPL2*- is deleterious individually. In *O. glaberrima* and its wild ancestor *O. barthii*, *DPL1* losses are due to the lack of the longest exon (Mizuta et al, 2010). In two *O. nivara* populations from Laos in our study, *DPL1*- had high frequencies (100.0% and 90.5%). *DPL2*- also showed 34.8% and 47.6% frequencies in two *O. rufipogon* populations in South China. However, none of the species/populations mentioned above showed any limitations from the lack of one gene copy. Thus, it is possible that *DPL1*- or *DPL2*- alone exists in wild rice.

Incompatibility caused by ancestral geographic structure can be found in many species. Genes leading to male sterility are variable within grasshopper populations, establishing that selection of incompatible genes can be weak or even absent (Shuker et al, 2005). Furthermore, the incompatible genes of *Capsella* species are stable under balanced selection (Sicard et al, 2015). In the ancestor *C. grandiflora*, balanced selection maintains the polymorphisms of two potentially incompatible genes, *NPR1* and *RPP5*, facilitating the establishment of genetic incompatibility between two selfing descendants, *C. rubella* and *C. grandiflora*. A pair of closely linked genes causes a globally distributed incompatibility in *Caenorhabditis elegans* that is maintained by balanced selection (Seidel et al, 2008). It is generally thought that purifying selection may be helpful for polymorphisms. *Hl13* and *Hl14* cause hybrid lethality between *Mimulus guttatus* and *M. nasutus*, and both genes are polymorphic in the two species (Zuellig and Sweigart, 2018). *Hl13* and *Hl14* have been maintained for decades, presumably because continual gene flow continues purging lethal combinations and reestablishing polymorphisms at both loci.

Our results further suggested that the establishment of the *DPL* system relies on the increase in frequency of nonfunctional alleles. In this hypothesis, an ancestral polymorphism is a crucial precondition. In addition to the nonfunctional alleles reported by Mizuta et al (2010), we have identified other alleles that may possibly result in a loss of function. For example, a 7 bp insertion in the longest exon occurs only in wild rice strains from Sri Lanka. It is plausible that the hybrid sterility caused by *DPLs* may be established in a more complex way under natural conditions. We speculated that the frequency of any other nonfunctional alleles increased after domestication, and genetic incompatibility would still have been established between *indica* and *japonica*. We therefore hypothesized that incompatible systems of rice, or even other species, may arise in similar ways as *DPLs*. To clarify the evolutionary process of incompatible systems, large-scale analyses at the population level are warranted.

METHODS

Rice materials

We sampled 296 *O. nivara* and 383 *O. rufipogon* individual plants, represented as 'n' and 'r', respectively. The entire natural distribution range of the wild ancestors was covered (Fig. 4). Detailed material notes are listed in Table S2. These individuals were collected from different populations. Twenty-one natural populations were from Southeast Asia and South Asia, including Sri Lanka, Nepal, Myanmar, Cambodia and Laos. The number of sampled individuals in these populations ranged from 29 to 36, except for Sri Lanka (9–10). We further collected a total of 115 individual samples from the International Rice Research Institute (IRRI), including 6 populations from India, Thailand, Indonesia and New Guinea. *O. rufipogon* of New Guinea is a distinct line and is generally regarded as a special type (Oka, 1988). Ten Chinese *O. rufipogon* populations were sampled, each of which included 10–11 individuals. Of the Chinese populations, one was from Jiangxi Province (JXDX), the northern-most frontier of the natural range of *O. rufipogon* (Morishima et al, 1992); two from Hunan Province (HNCL and HNJY); one from Guangdong Province (GDGZ); three from Guangxi Province (GXHZ, GXBH and GXTD); one from Yunnan Province (YNJH); and two from Hainan Province (HNDZ and HNWC). China and South Asia countries are commonly regarded as areas that are closely related with the domestication of rice (Morishima et al; 1992; Garris et al, 2005; Sang and Ge, 2007; Huang et al, 2012). In addition to these wild samples, 132 *O. sativa* (represented by 's') landraces were sampled from all 6 groups: *aromatic*, *aus*, *indica*, *rayada*, *temperate japonica* and *tropical japonica*.

Sequencing and genotype analysis

We extracted genomic DNA from fresh or silica gel-desiccated

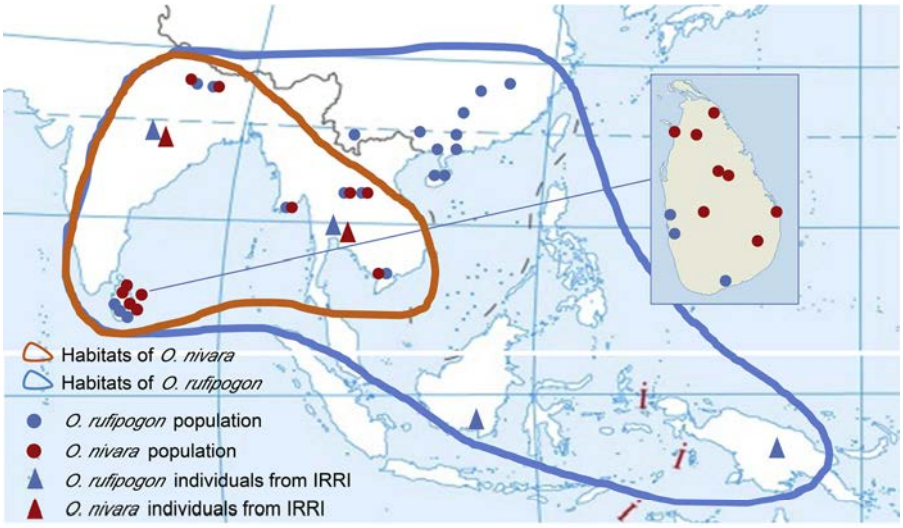

Fig. 4. Distribution of wild rice samplings, as well as functional and nonfunctional alleles of *DPLs*.
Populations of *O. nivara* and *O. rufipogon* are indicated with spots. Triangles indicate individual samples from the International Rice Research Institute (IRRI). Red and blue lines indicate *O. nivara* and *O. rufipogon*, respectively. Circles indicate the frequencies of functional and nonfunctional alleles.

leaves using a DNAsecure Plant Kit (TIANGEN, Beijing, China). The primers and thermal cycling procedure used was previously described by Craig et al (2014). The genomic DNA and PCR products were analyzed via gel electrophoresis. PCR products were Sanger sequenced (Majorbio, Shanghai, China). When any double peaks in a sequencing string occurred, indicating heterozygous individuals. The PCR products were cloned into pEASY-T1 vectors (TransGen Biotech, Beijing, China). We consulted previous studies to confirm singletons (Zheng and Ge, 2010). All sequences are available in GenBank (http://www.ncbi. nlm.nih.gov). We also downloaded 20 *DPL1* and 34 *DPL2* sequences for *O. sativa* from the National Centre for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih. gov) (Table S1).

For simplicity, *DPL1*- and *DPL2*- were defined according to Mizuta et al (2010). *DPL1*- was defined by a 518 bp transposable element insertion in the first coding sequence that failed to translate. *DPL2*- was defined by a single nucleotide transition from 'A' to 'G' in the intron, leading to a premature stop codon but producing a read-through protein (Fig. S1).

Genetic diversity and phylogeographic analysis

All obtained sequences were aligned using Clustal X 1.83 (Thompson et al, 1997) and revised manually with BioEdit 7.0.9.0 (Hall, 1999). To detect potential traces of selection, we evaluated the number of polymorphic sites (S) and nucleotide diversity (π) and subsequently performed neutrality tests (Tajima's D) with DnaSP 5.1 (Librado and Rozas, 2009). The number of haplotypes (h) and haplotype diversity (Hd) were also calculated with DnaSP 5.1 after excluding recombined sequences via the online tool IMgc (Woerner et al, 2007). Insertions and deletions were excluded during analysis of diversities. Comparisons of π between each group of *O. sativa* and the wild ancestors were performed using the ratio of π of each group to that of wild ancestor populations. The comparison of Hd was performed in the same way.

We used networks of haplotypes to investigate the origination of *DPL1* and *DPL2* in rice. The median-joining method (Bandelt et al, 1999) was used to construct networks of haplotypes via Network 5011 (Fluxus Technology Ltd., England). Recombined sequences were excluded by using IMgc (Woerner et al, 2007). Maximum Parsimony Calculation (Polzin and Daneshmand, 2003) was used to clarify the skeleton.

Data archiving

All sequence data used are available in GenBank (http://www.ncbi.nlm.nih.gov) under the accessions MN446022–MN446138, MN446139–MN446234, MN446235–MN446531, MN446532–MN446716, MN476103–MN476405, MN476406–MN476830, MN476831–MN476868 and MN476869–MN476888.

ACKNOWLEDGEMENTS

This work was supported by the Chinese Academy of Sciences (Grant Nos. XDA08020103 and XDB31000000) and the National Natural Science Foundation of China (Grant Nos. 31470332, 91731301 and 91231201). We thank CHEN Chengbin of Guangxi Academy of Agricultural Sciences and ZHENG Xiaoming of Institute of Crop Science of Chinese Academy of Agricultural Science for providing samples. We also thank the IRRI for providing samples and the members of GE Song's laboratory for field collections of wild rice.

SUPPLEMENTAL DATA

The following materials are available in the online version of this article at http://www.sciencedirect.com/science/journal/16726308; http://www.ricescience.org.

Fig. S1. Schematic representations of *DPLs* and positions of study primers.

Table S1. List of downloaded *DPL* sequences.

Table S2. Material list.

REFERENCES

Bandelt H J, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, **16**(1): 37–48.

Bank C, Burger R, Hermisson J. 2012. The limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a Continent-Island Model. *Genetics*, **191**(3): 845–345.

Cai Z, Zhou L, Ren N N, Xu X, Liu R, Huang L, Zheng X M, Meng Q L, Du Y S, Wang M X, Geng M F, Chen W L, Jing C Y, Zou X H, Guo J, Chen C B, Zeng H Z, Liang Y T, Wei X H, Guo Y L, Zhou H F, Zhang F M, Ge S. 2019. Parallel speciation of wild rice associated with habitat shifts. *Mol Biol Evol*, **36**(5): 875–889.

Chen J J, Ding J H, Ouyang Y D, Du H Y, Yang J Y, Cheng K, Zhao J, Qiu S Q, Zhang X L, Yao J L, Liu K D, Wang L, Xu C G, Li X H, Xue Y B, Xia M, Ji Q, Lu J F, Xu M L, Zhang Q F. 2008. A triallelic system of *S5* is a major regulator of the reproductive barrier and compatibility of *indica-japonica* hybrids in rice. *Proc Natl Acad Sci USA*, **105**(32): 11436–11441.

Civan P, Craig H, Cox C J, Brown T A. 2015. Three geographically separate domestications of Asian rice. *Nat Plants*, **1**(11): 1–5.

Corbett-Detig R B, Zhou J, Clark A G, Hartl D L, Ayroles J F. 2013. Genetic incompatibilities are widespread within species. *Nature*, **504**: 135–137.

Craig S M, Reagon M, Resnick L E, Caicedo A L. 2014. Allele distributions at hybrid incompatibility loci facilitate the potential for gene flow between cultivated and weedy rice in the US. *PLoS One*, **9**(1): e86647.

Cutter A D. 2012. The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. *Trends Ecol Evol*, **27**(4): 209–218.

Du H, Ouyang Y, Zhang C, Zhang Q. 2011. Complex evolution of *S5*, a major reproductive barrier regulator, in the cultivated rice *Oryza sativa* and its wild relatives. *New Phytol*, **191**(1): 275–287.

Garris A J, Tai T H, Coburn J, Kresovich S, McCouch S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics*, **169**(3): 1631–1638.

Gavrilets S. 1997. Hybrid zones with Dobzhansky-type epistatic selection. *Evolution*, **51**(4): 1027–1035.

Hall T A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*, **41**: 95–98.

Huang X H, Kurata N, Wei X H, Wang Z X, Wang A, Zhao Q, Zhao Y, Liu K Y, Lu H Y, Li W J, Guo Y L, Lu Y Q, Zhou C C, Fan D L, Weng Q J, Zhu C R, Huang T, Zhang L, Wang Y C, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X P, Xu Q, Dong G J, Zhan Q L, Li C Y, Fujiyama A, Toyoda A, Lu T, Feng Q, Qian Q, Li J Y, Han B. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature*, **490**: 497–501.

Khush G S. 1997. Origin, dispersal, cultivation and variation of rice. *Plant Mol Biol*, **35**(1): 25–34.

Kovach M J, Sweeney M T, McCouch S R. 2007. New insights into the history of rice domestication. *Trends Genet*, **23**(11): 578–587.

Li G W, Li X T, Wang Y, Mi J M, Xing F, Zhang D H, Dong Q Y, Li X H, Xiao J H, Zhang Q F, Ouyang Y D. 2017. Three representative inter and intra-subspecific crosses reveal the genetic architecture of reproductive isolation in rice. *Plant J*, **92**(3): 349–362.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**(11): 1451–1452.

Lindtke D, Buerkle C A. 2015. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. *Evolution*, **69**(8): 1987–2004.

Liu R, Zheng X M, Zhou L, Zhou H F, Ge S. 2015. Population genetic structure of *Oryza rufipogon* and *Oryza nivara*: Implications for the origin of *O. nivara*. *Mol Ecol*, **24**(20): 5211–5228.

Londo J P, Chiang Y C, Hung K H, Chiang T Y, Schaal B A. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc Natl Acad Sci USA*, **103**(25): 9578–9583.

Long Y M, Zhao L F, Niu B X, Su J, Wu H, Chen Y L, Zhang Q Y, Guo J X, Zhuang C X, Mei M T, Xia J X, Wang L, Wu H B, Liu Y G. 2008. Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc Natl Acad Sci USA*, **105**(48): 18871–18876.

Mizuta Y, Harushima Y, Kurata N. 2010. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proc Natl Acad Sci USA*, **107**(47): 20417–20422.

Morishima H, Sano Y, Oka H. 1992. Evolutionary Studies in Cultivated Rice and Its Wild Relatives. Oxford: Oxford University Press.

Oka H I. 1988. Origin of Cultivated Rice. Amsterdam, Holland: Elsevier.

Ouyang Y D, Liu Y G, Zhang Q F. 2010. Hybrid sterility in plant: Stories from rice. *Curr Opin Plant Biol*, **13**(2): 186–192.

Ouyang Y D, Zhang Q F. 2013. Understanding reproductive isolation based on the rice model. *Annu Rev Plant Biol*, **64**(1): 111–135.

Ouyang Y D, Zhang Q F. 2018. The molecular and evolutionary basis of reproductive isolation in plants. *J Genet Genom*, **45**(11): 613–620.

Polzin T, Daneshmand S V. 2003. On Steiner trees and minimum spanning trees in hypergraphs. *Oper Res Lett*, **31**(1): 12–20.

Rieseberg L H, Blackman B K. 2010. Speciation genes in plants. *Ann Bot*, **106**(3): 439–455.

Sang T, Ge S. 2007. The puzzle of rice domestication. *J Integr Plant Biol*, **49**(6): 760–768.

Seidel H S, Rockman M V, Kruglyak L. 2008. Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science*, **319**: 589–594.

Shen R X, Wang L, Liu X P, Wu J, Jin W W, Zhao X C, Xie X R, Zhu Q L, Tang H W, Li Q, Chen L T, Liu Y G. 2017. Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat Commun*, **8**: 1310.

Shuker D M, Underwood K, King T M, Butlin R K. 2005. Patterns of male sterility in a grasshopper hybrid zone imply accumulation of hybrid incompatibilities without selection. *Proc R Soc Lond Ser B-Biol Sci*, **272**: 2491–2497.

Sicard A, Kappel C, Josephs E B, Lee Y W, Marona C, Stinchcombe J R, Wright S I, Lenhard M. 2015. Divergent sorting of a balanced ancestral polymorphism underlies the establishment of gene-flow barriers in *Capsella*. *Nat Commun*, **6**: 7960.

Sweeney M, McCouch S. 2007. The complex history of the domestication of rice. *Ann Bot*, **100**(5): 951–957.

Tajima F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**(3): 585–595.

Thompson J D, Gibson T J, Plewniak F, Jeanmougin F, Higgins D G. 1997. The CLUSTAL_X windows interface: Flexible strategies

for multiple sequence alignment aided by quality analysis tools. *Nucl Acids Res*, **25**(24): 4876–4882.

Vaughan D A, Lu B R, Tomooka N. 2008. The evolving story of rice evolution. *Plant Sci*, **174**(4): 394–408.

Wang C H, Zheng X M, Xu Q, Yuan X P, Huang L, Zhou H F, Wei X H, Ge S. 2014. Genetic diversity and classification of *Oryza sativa* with emphasis on Chinese rice germplasm. *Heredity*, **112**: 489–496.

Woerner A E, Cox M P, Hammer M F. 2007. Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**(14): 1851–1853.

Xie Y Y, Shen R X, Chen L T, Liu Y G. 2019. Molecular mechanisms of hybrid sterility in rice. *Sci China Life Sci*, **62**(6): 737–743.

Yamagata Y, Yamamoto E, Aya K, Win K T, Doi K, Sobrizal, Ito T, Kanamori H, Wu J, Matsumoto T, Matsuoka M, Ashikari M, Yoshimura A. 2010. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proc Natl Acad Sci USA*, **107**(4): 1494–1499.

Yang C C, Kawahara Y, Mizuno H, Wu J Z, Matsumoto T, Itoh T. 2012. Independent domestication of Asian rice followed by gene flow from *japonica* to *indica*. *Mol Biol Evol*, **29**(5): 1471–1479.

Yang J Y, Zhao X B, Cheng K, Du H Y, Ouyang Y D, Chen J J, Qiu S Q, Huang J Y, Jiang Y H, Jiang L W, Ding J H, Wang J, Xu C G, Li X H, Zhang Q F. 2012. A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science*, **337**: 1336–1340.

Zheng X M, Ge S. 2010. Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Mol Ecol*, **19**(12): 2439–2454.

Zuellig M P, Sweigart A L. 2018. A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, **72**(11): 2394–2405.

(Managing Editor: WU Yawen)