Received: 21 September 2023 | Revised: 4 January 2024 | Accepted: 8 January 2024

DOI: 10.1111/mec.17268

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genetic architecture of ecological divergence between *Oryza rufipogon* and *Oryza nivara*

Qing-Lin Meng[1,2] | Cheng-Gen Qiang[1,2] | Ji-Long Li[1,2] | Mu-Fan Geng[1,2] | Ning-Ning Ren[1] | Zhe Cai[1] | Mei-Xia Wang[1,2] | Zi-Hui Jiao[1,2] | Fu-Min Zhang[1,2] | Xian-Jun Song[3] | Song Ge[1,2]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Key Laboratory of Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China

Correspondence
Song Ge, State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.
Email: gesong@ibcas.ac.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 32130008, 31900198 and 31800186; Strategic Priority Research Program of Chinese Academy of Sciences, Grant/Award Number: XDB31000000; Ministry of Science and Technology, Grant/Award Number: 2021YFD1200101-02

Handling Editor: Suhua Shi

Abstract

Ecological divergence due to habitat difference plays a prominent role in the formation of new species, but the genetic architecture during ecological speciation and the mechanism underlying phenotypic divergence remain less understood. Two wild ancestors of rice (*Oryza rufipogon* and *Oryza nivara*) are a progenitor-derivative species pair with ecological divergence and provide a unique system for studying ecological adaptation/speciation. Here, we constructed a high-resolution linkage map and conducted a quantitative trait locus (QTL) analysis of 19 phenotypic traits using an F_2 population generated from a cross between the two *Oryza* species. We identified 113 QTLs associated with interspecific divergence of 16 quantitative traits, with effect sizes ranging from 1.61% to 34.1% in terms of the percentage of variation explained (PVE). The distribution of effect sizes of QTLs followed a negative exponential, suggesting that a few genes of large effect and many genes of small effect were responsible for the phenotypic divergence. We observed 18 clusters of QTLs (QTL hotspots) on 11 chromosomes, significantly more than that expected by chance, demonstrating the importance of coinheritance of loci/genes in ecological adaptation/speciation. Analysis of effect direction and *v*-test statistics revealed that interspecific differentiation of most traits was driven by divergent natural selection, supporting the argument that ecological adaptation/speciation would proceed rapidly under coordinated selection on multiple traits. Our findings provide new insights into the understanding of genetic architecture of ecological adaptation and speciation in plants and help effective manipulation of specific genes or gene cluster in rice breeding.

KEYWORDS
ecological speciation, genetic architecture, phenotypic divergence, QTL mapping, wild ancestors of rice

Qing-Lin Meng, Cheng-Gen Qiang and Ji-Long Li contributed equally to this work.

1 | INTRODUCTION

Ecological divergence between populations, often arising from local adaptation, is driven by divergent natural selection between contrasting environments, which in turn results in ecological speciation through the evolution of reproductive isolation (Nosil, 2012; Schluter & Rieseberg, 2022; Seehausen et al., 2014). Multiple lines of evidence from plant and animal studies have demonstrated that ecological divergence due to habitat difference plays a prominent role in the formation of new species (Erickson et al., 2004; Peichel & Marques, 2017; Schluter & Rieseberg, 2022). During the process of ecological speciation, phenotypic differentiation occurs and reproductive isolation evolves as a consequence of divergent natural selection (Erickson et al., 2004; Faria et al., 2014; Nosil, 2012; Schluter & Rieseberg, 2022). Despite substantial studies, many fundamental questions regarding the genetic basis underlying ecological divergence and speciation remain debated or largely elusive, including the relative contributions of loci/genes with large versus small effects to phenotypic divergence, the randomly distributed versus clustered genetic architecture and their evolutionary implications, mechanisms underlying the link between evolution of divergent phenotypes and the emergence of reproductive isolation, and the role of divergent selection in the process of ecological divergence and speciation (Bomblies & Peichel, 2022; Faria et al., 2014; Kitano et al., 2022; Nosil et al., 2021; Schluter & Rieseberg, 2022; Seehausen et al., 2014).

Two *Oryza* species, *O. rufipogon* Griff. and *O. nivara* Sharma et Shastry, are most closely related and collectively regarded as the ancestors of cultivated rice (*O. sativa* L.) (Cai et al., 2019; Khush, 1997; Sang & Ge, 2007; Vaughan et al., 2008). The perennial *O. rufipogon*, characterized by photoperiod sensitivity and predominate cross-fertilization, is widely distributed throughout southern China, South and Southeast Asia, Papua New Guinea, and northern Australia. In contrast, the annual *O. nivara*, characterized by photoperiod insensitivity and predominant self-fertilization, has a more restricted distribution in South and Southeast Asia (Sang & Ge, 2007; Vaughan, 1994; Vaughan et al., 2008). In addition, interspecific differences in a few dozens of traits have been documented by experimental and field investigations (Banaticla-Hilario et al., 2013; Barbier, 1989; Cai et al., 2019; Eizenga et al., 2022; Guo et al., 2016; Jing et al., 2023; Morishima et al., 1984; Ren, 2019; Sano et al., 1980). Moreover, studies show that the annual *O. nivara* evolved from the perennial *O. rufipogon* to associate with a habitat shift from a persistently wet to a seasonally dry habitat, in which flowering time change in the derived *O. nivara* was the major component contributing to the reproductive isolation between *O. rufipogon* and *O. nivara* (Barbier, 1989; Cai et al., 2019; Liu et al. 2015; Morishima et al., 1984; Xu et al., 2020). Therefore, the progenitor-derivative species pair with distinct differences in morphology, life history traits, and habitat preference represents a feasible system for the study of ecological adaptation and speciation (Cai et al., 2019; Grillo et al., 2009; Zheng & Ge, 2010).

QTL analysis is powerful approach to uncover the genetic architecture of ecologically important traits and to determine the targets of natural selection and thus has been used successfully for studies of evolutionary process and mechanisms in various plants and animals (Barton & Keightley, 2002; Connallon & Hodgins, 2021; Erickson et al., 2004; Jakobson & Jarosz, 2020; Saltz et al., 2017; Tanksley, 1993). In this study, we present a quantitative traits locus (QTL) analysis of an F_2 population derived from a cross between *O. rufipogon* and *O. nivara*, using SNPs generated from specific-locus amplified fragment-sequencing (SLAF-seq) technology (Sun et al., 2013). First, we examine the number, effect size, and distribution pattern of QTLs controlling phenotypic divergence between species. Specifically, we ask: (1) How many genomic regions contribute to the phenotypic divergence between *O. rufipogon* and *O. nivara* given substantial phenotypic differentiation between species? (2) Are the traits differentiating two species controlled by a large number of loci with small effects or a small number of loci with large effects? Although substantial studies involving morphological variation have been undertaken on *O. nivara* and *O. rufipogon* (e.g., Banaticla-Hilario et al., 2013; Barbier, 1989; Cai et al., 2004; Cai et al., 2019; Eizenga et al., 2022; Guo et al., 2016; Kim et al., 2016; Morishima et al., 1961), no effort has been attempted to explore the genetic basis of phenotypic divergence between the two species until Grillo et al. (2009) who performed a QTL analysis to investigate the genetic architecture for phenotypic divergence between *O. rufipogon* and *O. nivara*. Nevertheless, Grillo et al. (2009)'s study provided limited knowledge of the genetic basis underlying phenotypic divergence because of the low marker density (116 SSRs) and relatively small mapping populations (less than 200). Here, based on the high-resolution markers, we were able to identify QTLs for phenotypic traits that diverge between species and explore the full genetic architecture of ecological speciation.

Second, we address how the identified loci are distributed across the genome and whether they cluster (colocalize) in particular chromosomal regions. Multiple lines of evidence indicate that adaptation to multiple different aspects of new environments can be facilitated by coinheritance of adaptive phenotypes, embodied as enriched QTLs in some genomic regions (Bomblies & Peichel, 2022; Hoffmann & Rieseberg, 2008; Jacobs et al., 2017; Nosil et al., 2021). Indeed, accumulating studies in both plants (e.g., Ferris et al., 2017; Lowry et al., 2015; Nakazato et al., 2013; Roda et al., 2017) and animals (e.g., Archambeault et al., 2020; Jacobs et al., 2017; Linnen et al., 2013) revealed that many QTLs were not distributed randomly across the genome but rather in hotspots involving a variety of adaptive traits. Clustering of QTLs responsible for domestication-related traits is also common in crop species (e.g., Burke et al., 2002; Cai & Morishima, 2002; Geng et al., 2021; Wang et al., 2011; Yang et al., 2019). These studies suggested that QTL clustering, due to either pleiotropy or tight linkage, might be a mechanism for preventing

unfit combinations of genotypes and thus facilitates rapid adaptation and speciation (Bomblies & Peichel, 2022; Nosil et al., 2021; Peichel & Marques, 2017). Despite these, the prevalence of the clustered genetic architecture and the mechanisms that facilitate the coinheritance of adaptive phenotypes during ecological speciation are less understood (Archambeault et al., 2020; Bomblies & Peichel, 2022; Yang et al., 2019). The two *Oryza* species provide a unique opportunity to gain further insights into the distribution pattern of QTLs and the underlying mechanisms during ecological speciation.

Finally, we investigated the potential roles of natural selection in trait divergence between species. Evidence shows that the origin of *O. nivara* from *O. rufipogon* was associated with a suite of phenotypic changes in a pattern consistent with ecological speciation (Barbier, 1989; Cai et al., 2019; Guo et al., 2016; Morishima et al., 1984; Ren, 2019). Although previous studies demonstrated the roles of natural selection rather than random genetic drift in the phenotypic divergence between *O. rufipogon* and *O. nivara* (Cai et al., 2019; Guo et al., 2016), these studies were unable to distinguish between direct and indirect selection acting on the traits because selection on one trait could have caused substantial divergence in other traits due to genetic correlations (Feng et al., 2019; Muir et al., 2014; Via & Hawthorne, 2005). Recently developed *v*-test (Fraser, 2020) provides a feasible and powerful approach to determine whether traits evolved under directional selection based on phenotype divergence of parental and phenotype distribution of the crossing population. Therefore, we were interested in whether divergent natural selection is responsible for the coordinated differentiation of a suite of traits as expected during ecological divergence between *O. rufipogon* and *O. nivara*. Addressing these questions not only provides additional insights into the process and mechanisms of ecological adaptation and speciation in plants but also facilitates rice genetic improvements given abundant unique genetic resources maintained in wild *Oryza* species.

2 | MATERIALS AND METHODS

2.1 | Development of F_2 mapping population

To explore the genetic basis of divergence between the two species, we created an F_2 mapping population between *O. rufipogon* (Ruf-I, IRGC 81881) and *O. nivara* (Niv-I, IRGC 101508) inbred lines (Figure 1, Table S1) that were self-pollinated for five generations. The *O. rufipogon* and *O. nivara* individuals were sampled from India and showed morphologies typical of the two species that diverge significantly in numerous traits, including the taxonomically diagnostic characters such as flowering time, anther length, culm length, panicle exsertion, and shape (Cai et al., 2019; Jing et al., 2023). The construction of inbred lines of two parents, crossing between parental lines and subsequent development of F_1 and F_2 populations were conducted from 2014 to 2017 at Lingshui Station (18°30.6′ N, 110°2.4′ E) in Hainan Province, China (Meng, 2021).

Because our previous studies showed that *O. nivara* usually flowered over 60 days earlier than *O. rufipogon* in the wild (Cai et al., 2019; Xu et al., 2020), we germinated the seeds of *O. nivara* in three batches at an interval of 10 days to ensure the concurrence of the flowering time of the parental lines. The Niv-I was designated as female parent, while Ruf-I was selected as male parent. To avoid self-pollination in crossing, we emasculated the panicles of the female parents before blossoming of the spikelets and then removed the immature anthers, and finally sprayed water on the emasculated panicles in case of residual pollen (Xu et al., 2020). A single F_1 individual from the crosses between Ruf-I and Niv-I was chosen randomly to produce the F_2 population ($N = 1174$) by selfing (Meng, 2021).

The F_2 population was grown together with self-fertilizing seeds from two mapping parents (Ruf-I, $N = 28$ and Niv-I, $N = 23$) in Lingshui Station in November 2016. Seeds were processed at 50°C for 5 days to break dormancy and then germinated in a growth chamber under long-day condition (day: 14 h, 36°C; night: 10 h,

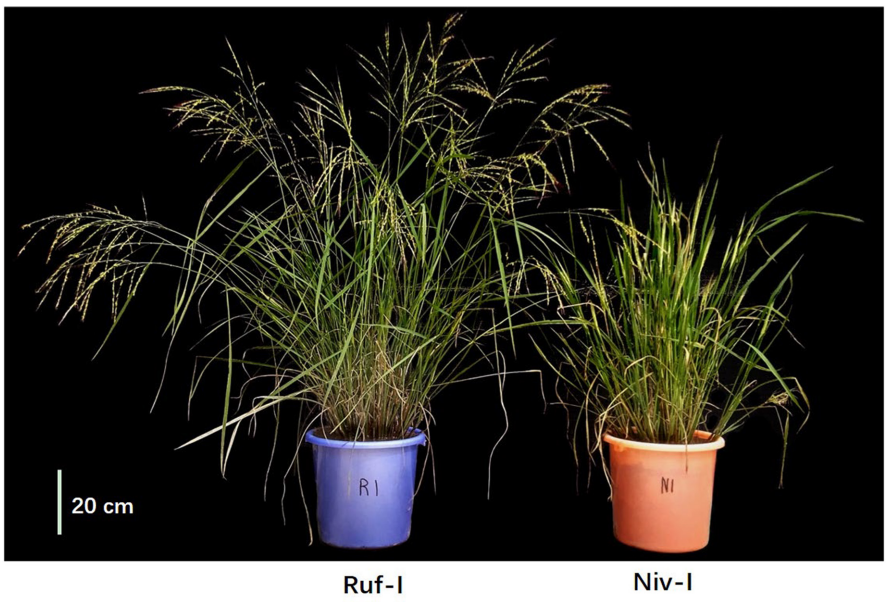

FIGURE 1 Gross morphology of two parental lines used to generate F_2 mapping population, *Oryza rufipogon* (Ruf-I, IRGC 81881) and *Oryza nivara* (Niv-I, IRGC 101508). [Colour figure can be viewed at wileyonlinelibrary.com]

33°C). One week later, the young seedlings were planted in greenhouse at natural daylength condition. After additional 3 weeks, the seedlings with more than three tillers were transplanted into paddy field randomly with a spacing of 1.5 × 1.5 meters. Finally, 862 F_2 plants survived to the flowering period and were recorded phenotypically.

2.2 | Phenotypic analyses

We measured 19 phenotypic traits (Table S2) that were either adaptive or taxonomically and agronomically important according to previous studies (Banaticla-Hilario et al., 2013; Cai et al., 2019; Eizenga et al., 2022; Guo et al., 2016; Jing et al., 2023; Ren, 2019). Given premating reproductive isolation (flowering time difference) and habitat divergence between *O. rufipogon* and *O. nivara* are two main factors associated with the process of speciation, we classified all traits into three categories: reproduction-related traits (RR traits) (three quantitative traits), habitat-related traits (HR traits) (13 quantitative traits) and colour traits (three qualitative traits) (Table S2). The RR traits were related to mating or reproductive isolation and the HR traits involved habitat preference of the derived *O. nivara*. Like many other studies on phenotypic variation in which different classifications of traits were used (e.g., Grillo et al., 2009; Hall et al., 2006; Lexer et al., 2005; Peichel & Marques, 2017), trait division in our case seems a bit arbitrary but feasible to help our analyses by relating phenotypic variation to ecological speciation.

Trait measurements were taken on all plants that flowered following the methods detailed in Biodiversity-International et al. (2007). Measurements were taken for three tillers/culms and averaged for each trait, except for first heading, culm habit, grain, and colour traits (Table S2). First heading (FH) and culm habit (CH) were recorded for the primary culm. Grain length (GL) and width (GWI) were calculated for 10 full seeds, and grain weight (GWE) was measured for 30 full seeds. Three qualitative colour traits that might be related with biotic and abiotic stress in plants (Dabravolski & Isayenkov, 2023; Qin et al., 2021), that is, awn colour (AWC), basal leaf sheath colour (BLSC), and stigma colour (SC), were scored as binary traits, with 1 and 0 indicating the presence and absence, respectively (Table S2).

To test for trait divergence between parental populations, *t*-test and *chi-square* (χ^2) test were conducted for the 16 quantitative and three qualitative traits, respectively. We used the method of Pearson correlation to calculate the correlations among traits. All the calculations and plotting were performed in R (R Core Team, 2020).

2.3 | Sequencing and genotyping

Fresh leaves of two parental lines and 600 F_2 individuals randomly chosen from the 862 F_2 population were collected and dried with silica gel. Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (Murray & Thompson, 1980). The libraries of mapping parents were constructed following the manufacturer's recommendations (Illumina) for 500 bp insert size and sequenced by BGIseq500 platform (BGI; Shenzhen, China) with 150 bp paired-end reads. The F_2 individuals were genotyped by a specific-locus amplified fragment-sequencing (SALF-seq) method (Sun et al., 2013). In brief, two restriction enzymes (*RsaI* and *HaeIII*) were selected to digest the genomic DNA. The digested fragments (SLAF tags) were ligated to the adapters with T4 DNA ligase. After PCR amplification, purification, sample mixing, and electrophoresis with agarose gels, the size of fragments ranging from 264 to 314 bp were obtained and purified. Subsequently, the products were sequenced using Illumina HiSeq 2500 platform (Illumina; San Diego, USA) with 125 paired-end reads according to the manufacturer's instruction.

The short reads of parental lines and F_2 individuals were filtered by removing low-quality reads with more than 10% of bases missing. Then short reads were aligned to the reference sequence of Nipponbare genome (IRGSP-1.0) (Kawahara et al., 2013) using BWA (Li & Durbin, 2010) with the MEM algorithm. Furthermore, Samtools (Li et al., 2009) were applied to sort the mapping results and built index for each BAM file. Variant calling was conducted using the Genome Analysis Toolkit (GATK, version 4.0.2.1) (Van der Auwera et al., 2013). SNPs were filtered with VariantFiltration of GATK "AC < 2 || QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < −12.5 || ReadPosRankSum < −8.0" as suggested by the manual.

The filtered SNPs were genotyped based on the following criteria: (1) homozygous for each mapping parent and different between two parents; (2) bi-allelic polymorphism among the F_2 individuals; (3) the missing rates were no more than 5%; (4) the extreme segregation distorted SNPs were excluded at the cut-off ($P \leq 1 \times 10^{-7}$) in the separation of Mendelian law (Wang et al., 2022; Zhang et al., 2010; Zuo et al., 2019). Finally, we obtained 38,144 SNPs in total.

2.4 | Linkage map construction and QTL analyses

SNPs markers were converted into bins using SNPbinner (Gonda et al., 2019; Oren et al., 2019). Cross points were calculated with minimum ratio (r) set as 0.01, and bins were generated with minimum bin length (–m) set as 5 kb. Based on the position, 6579 bin markers were divided into linkage groups corresponding to 12 rice chromosomes (Figure S1). We applied *est.rf* and *est.map* in R/qtl (Broman et al., 2003) to demonstrate marker order and the genetic position of bins.

We performed QTL mapping in R/qtl (Broman et al., 2003). The composite interval mapping (CIM) was conducted with *cim* function, the window size was set as 10 cM, and the cofactor was stepwise increased until the detected QTLs were stable. The genome-wide significance threshold of each trait was determined separately by 1000 permutations with $\alpha = 0.05$. Two-way ANOVA was used to test for epistatic interactions among QTLs (Grillo

et al., 2009; Westerbergh & Doebley, 2004), and QTL interaction pairs that reached statistical significance ($p < 0.05$) were kept. We further used *q* value (Storey & Tibshirani, 2003) to evaluate the significance of *p*-values for each trait at false discovery rate (FDR) of 5% to obtain more robust QTL interactions. We estimated the percent variance explained (PVE), additive effect, and dominance effect with *fitqtl* function. The confidence intervals were determined as 1.5-LOD confidence intervals.

The direction of QTL effect was scored as a positive if the effect of a parental allele was consistent with species divergence and otherwise was scored as negative if the effect was opposite of species divergence. We noted that the threshold level for classification of QTLs in terms of effect size was variable in previous studies in which the criteria of 10%–25% of the total phenotypic variation have been used for a criterion of major QTLs (e.g., Bradshaw et al., 1998; Davey et al., 2006; Hall et al., 2006; Kumar et al., 2017; Tanksley, 1993). For simplicity, we accept a criterion of 10% for distinguishing between major-effect and minor-effect QTLs and define a major-effect QTL as a large-effect QTL if it explains >25% of the total phenotypic variance in the mapping population.

To explore the extent to which the QTLs overlapped across the genome and in their correlations with phenotypic traits, we first compared the distribution of the QTL peaks on each chromosome with uniform probability distribution using the Kolmogorov–Smirnov test (Massey, 1951), performed by using *ks.test* function in R. If the QTL positions significantly deviated from uniform distribution, the QTLs were not distributed randomly (Arnegard et al., 2014). Then, following the methods in previous studies (Frary et al., 2014; Nakazato et al., 2013; Oakley et al., 2018), we marked all QTLs on the 12 chromosomes and identified the QTLs that exhibited overlapping 1.5-LOD confidence intervals with at least two other QTLs. Those regions with more than three QTLs overlapping across traits were defined as QTL hotspots (Frary et al., 2014; Nakazato et al., 2013).

2.5 | Test for directional selection

To determine whether traits evolved under directional selection, we performed *v*-test of Fraser (2020) based on phenotype divergence of parental and phenotype distribution of the crossing population. This method is a generalization of QTL sign test (Orr, 1998b) and applicable to phenotypic data for almost any genetic cross, thus providing a feasible and powerful approach to detect selection. The *v*-test was performed as described in equation 2 of Fraser (2020). To calculate *v*, we estimated the phenotypic variances within and between parents of the cross, and the variance among F_2 population. As the broad-sense heritability (H^2) was needed to correct the random noise, we calculated as V_g/V_p by using sommer package in R. The V_g is the genetic variance estimated from the kinship matrix, and V_p is the phenotypic variance (Covarrubias-Pazaran, 2016). The constant *c* was equal to 2.0 for F_2 (Fraser, 2020). Significance of *v* was estimated based on the cumulative *F* distribution with (1, *k*-1) degree of freedom; the *k* is the individual number of F_2.

3 | RESULTS

3.1 | Phenotypic variation and correlations

To evaluate variation patterns of phenotypic traits that diverged between *O. rufipogon* and *O. nivara*, we calculated the mean and ranges of 16 quantitative traits for two parental lines. The parental lines differed significantly for all traits (*t*-test, $p < 0.01$) except for spikelet number (SN) and culm diameter (CD) (Table 1). It is noted that all 14 traits that diverged significantly between species except for two flag leaf traits (FLL and FLW) exhibited the same differentiation patterns as those reported in previous studies using population samples (Banaticla-Hilario et al., 2013; Cai et al., 2019; Guo et al., 2016; Ren, 2019). Consistent with previous studies (Cai et al., 2019; Guo et al., 2016), we regarded these 12 traits as adaptive traits (Table 1). Of three qualitative (colour) traits, two (BLSC and SC) exhibited significant interspecific divergence (Table 1, Figure S2). However, divergent patterns for these colour traits might represent the variation within populations/species because no significant differentiation was found between species for them in previous studies of natural populations (Cai et al., 2019; Ren, 2019).

Overall, all 16 quantitative traits, including the two (SN and CD) without significant divergence between two parental lines, showed an increase in variance in F_2 population relative to the parental lines (Figure 2), suggesting the segregation of many genes of small to moderate effect. It is evident that most traits were distributed normally or nearly normally (Figure 2, Table S3), further suggesting that they are under polygenic control. Three exceptions include two panicle traits (PE and PS) that exhibited largely bimodal distribution and the first heading (FH) that showed an extended tail in one direction, implying that these traits may be under the control of genetic loci with major genetic effects. Seven traits (ANL, CD, CH, CL, PL, SN, and FLL) showed obvious transgressive segregation in the F_2 population (Figure 2).

We calculated the pairwise correlation of traits in F_2 population to evaluate the potential roles of single pleiotropic or tightly linked loci in trait divergence between species because a significant correlation between traits suggests the shared genetic basis due to either pleiotropy or linkage of genes (Saltz et al., 2017; Via & Hawthorne, 2005). As shown in Figure 3, 75 (63%) of all 120 pairwise combinations of 16 quantitative traits showed significant correlations, with most of them (89%) being positive. By focusing analyses on 12 putatively adaptive traits, we detected significant correlations for two pairwise combinations of three RR traits (ANL, FH, and PE) and for 20 (55.6%) of all 36 pairwise combinations of nine HR traits, with four traits (ANL, CL, PL, and PS) significantly correlating with almost all other traits (Figure 3). These observations provided an initial indication that single pleiotropic or multiple tightly linked loci may have substantial impacts on the trait divergence between *O. rufipogon* and *O. nivara*.

TABLE 1 Variation of 19 phenotypic traits measured for *Oryza rufipogon* and *Oryza nivara* parental lines in the common garden.

	O. rufipogon		*O. nivara*			
Trait	***N***	**Mean (SD)**	***N***	**Mean (SD)**	***t*-Statistic**	***V*-statistic**
Reproduction-related trait (RR trait)						
Anther length (ANL)	12	**4.23 (0.22)**	21	1.99 (0.05)	34.912***	12.11***
First heading (FH)	23	**109.48 (6.72)**	21	71 (4.89)	21.847***	8.159**
Panicle exsertion (PE)	12	**8.62 (2.41)**	22	0 (0)	21.464***	5.756*
Habitat-related trait (HR trait)						
Awn length (AWL)	12	2.26 (0.67)	22	**6.38 (0.68)**	−17.057***	2.862
Culm diameter (CD)	12	0.55 (0.07)	21	0.56 (0.07)	-0.465^{ns}	NA
Culm habit (CH)	12	**60.83 (5.97)**	21	36.91 (5.59)	11.34***	6.855**
Culm length (CL)	12	**98.83 (10.35)**	20	57.25 (9.74)	13.615***	6.962**
Flag leaf attitude (FLA)	12	**92.67 (10.67)**	21	36.19 (10.45)	14.736***	17.468***
Flag leaf length (FLL)	12	**28.42 (3.68)**	22	24.18 (2.17)	3.654**	NA
Flag leaf width (FLW)	12	0.8 (0.07)	22	**1.23 (0.08)**	−15.518***	NA
Grain length (GL)	10	8.16 (0.23)	10	**9.17 (0.15)**	−11.603***	2.897
Grain weight (GWE)	10	0.461 (0.02)	10	**0.735 (0.034)**	−21.745***	1.692
Grain width (GWI)	10	2.32 (0.09)	10	**2.82 (0.09)**	−12.759***	6.18*
Panicle length (PL)	12	**26.16 (1.55)**	22	17.98 (1.28)	5.027***	3.053
Panicle shape (PS)	12	**44.25 (5.86)**	21	0 (0)	26.14***	4.965*
Spikelet number (SN)	12	49.08 (7.87)	21	49.29 (4.85)	-0.081^{ns}	NA
Colour trait					χ^2	NA
Awn colour (AWC)	23	0.52 (0.51)	21	0.43 (0.51)	0.1^{ns}	NA
Basal leaf sheath colour (BLSC)	12	1	22	0	29.8***	NA
Stigma colour (SC)	12	1	21	0	28.8***	NA

Note: All the quantitative traits except for four in italic (CD, FLL, FLW, and SN) exhibited the differentiation patterns that were same as those found in previous studies of natural populations (Cai et al., 2019; Ren, 2019) and were considered to be adaptive. For CD and SN, no significant differentiation was found between parental lines, while for the two flag leaf traits (FLL and FLW), the opposite patterns of divergence to those reported for natural populations were determined. Figures in boldface represent larger average values in comparison of the *O. rufipogon* and *O. nivara* parental lines.
N, sample size; *t*-statistic was used for differentiation test between parental lines and *v*-test was for selection test.
Abbreviations: NA, not applicable; ns, not significant.
*$p<0.05$; **$p<0.01$; ***$p<0.001$.

3.2 | Linkage map and QTL analysis

Based on 6579 bin markers, we constructed the genetic map spanning 1481.09 cM over 12 LGs corresponding to 12 rice chromosomes, with the length close to cultivated rice (Huang et al., 2009). The genetic distance between adjacent bin markers ranged from 0.08 to 4.62 cM, with the mean distance being 0.23 cM (Figure 4, Table S4). The genetic map was high-resolution and enables us to get a comprehensive and precise mapping result.

To reveal the genetic basis of species divergence between *O. rufipogon* and *O. nivara*, we mapped QTLs involved in all 19 traits. For 16 quantitative traits, we identified a total of 113 QTLs that were located on all 12 chromosomes, with the number of QTLs per trait ranging from 4 (GWE and GWI) to 11 (ANL), and the amount of variation explained by these QTLs ranging from 1.61% (ANL3.b) to 34.1% (AWL4) (Table 2, Table S5). Moreover, we identified 12 major QTLs, that is, the QTLs that explain over 10% of total phenotypic variation, which involved 10 traits (Figure 4, Table 2). Of 12 major QTLs, three (AWL4, FLW1, and SN1.a) exhibit large effect, that is, the QTLs that explain over 25% of total phenotypic variation. For three qualitative (colour) traits, we identified five QTLs (one each for AWC and BLSC, and three for SC) (Table S5).

We detected significant epistatic interactions for 23 pairs of QTLs that affected 11 quantitative traits, with the number ranging from one pair (CL, FLW, GWI, and PS) to five (AWL) (Table 3). Of these combinations, eight involved major-effect QTLs (PVE > 10%) with three remaining significant after FDR correction at 5% level (Table 3). However, all the significant epistatic interactions explained a small amount of variance, implying that the QTL × QTL interactions might not play an important role in divergence of these traits between species (Nakazato et al., 2013).

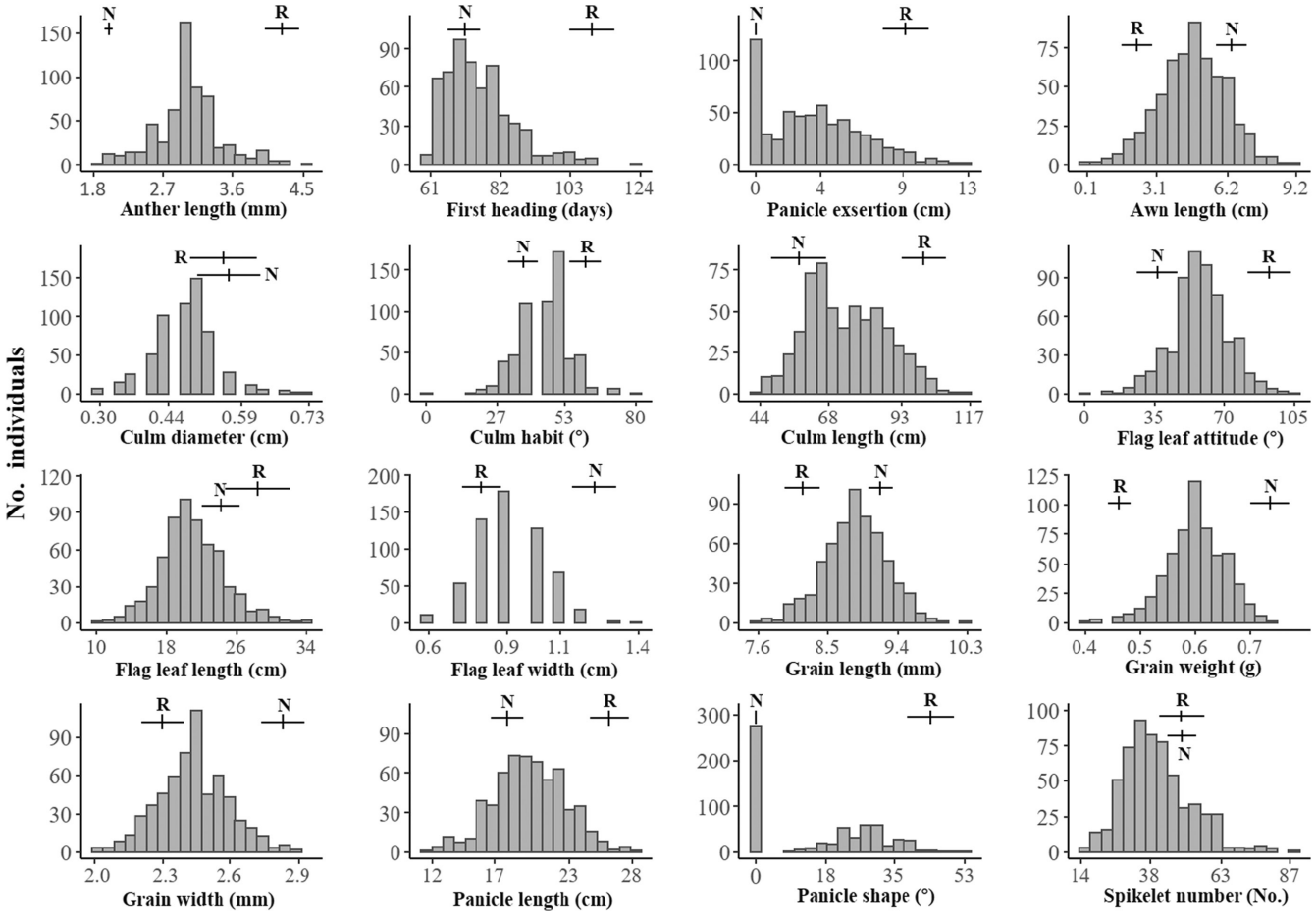

FIGURE 2 Frequency distribution of the parental lines, *Oryza rufipogon* (R) and *Oryza nivara* (N), and 600 F_2 progeny for 16 quantitative traits. Means and standard deviations of two parental lines are indicated by vertical and horizontal lines, respectively.

3.3 | Analysis of directional selection

Of 12 putatively adaptive traits, seven (ANL, FH, AWL, GL, GWI, PL, and PS) included at least one major-effect QTL (Table 2). Moreover, the majority of QTLs of all putatively adaptive traits except for three (GL, PL, and PS) were positive, with over two-thirds of QTLs showing effects in the same directions of the phenotypic divergence between species (positive effects) (Table 2). These results suggested that differentiations of these putatively adaptive traits might be due to directional selection. The three exceptional traits included almost half of QTLs with negative effects (i.e., antagonistic effects) (Table 2), implying that these traits diverged under either weak selection or drift (Ferris et al., 2017; Muir et al., 2014).

To further explore the role of natural selection in trait divergence between *O. nivara* and *O. rufipogon*, we conducted *v*-test for 12 putatively adaptive traits and found that eight traits were significant, including three RR traits, that is, first heading (FH), anther length (ANL), and panicle exsertion (PE) (Table 1), which associate with reproductive isolation between species. These results were in accordance with the above QTL effect analyses in which interspecific differentiation for 9 out of 12 putatively adaptive traits evolved under directional selection.

3.4 | Distribution of effect sizes and clustering of QTLs

To evaluate the relative contribution of the mutations of large and small effects during phenotypic differentiation between the two species, we estimated the distribution of effect sizes for all 113 QTLs of 16 quantitative traits identified in the F_2 population (Figure 5a). It is clear that the effect sizes of these QTLs were typically small to moderate (<10% of PVE), with only three being the large-effect QTLs (>25% of PVE), which is consistent with the Orr's model (Orr, 1998a) in which a few genes of large effect and many genes of small effect underlying the phenotypic divergence. The distribution of effect size for each of the two categories (RR and HR traits as well as putatively adaptive traits) (Figure 5b–d) also followed the Orr's model. These results suggest that the phenotypic evolution during the origin of *O. nivara* involves a dozen of traits through a few mutations of large effect and many mutations of small effect.

We first tested whether the QTL positions significantly deviated from uniform distribution on chromosomes using the Kolmogorov–Smirnov test (Arnegard et al., 2014; Massey, 1951) and found that QTLs on chromosomes 1, 4, 5, 6, 9, and 12 were

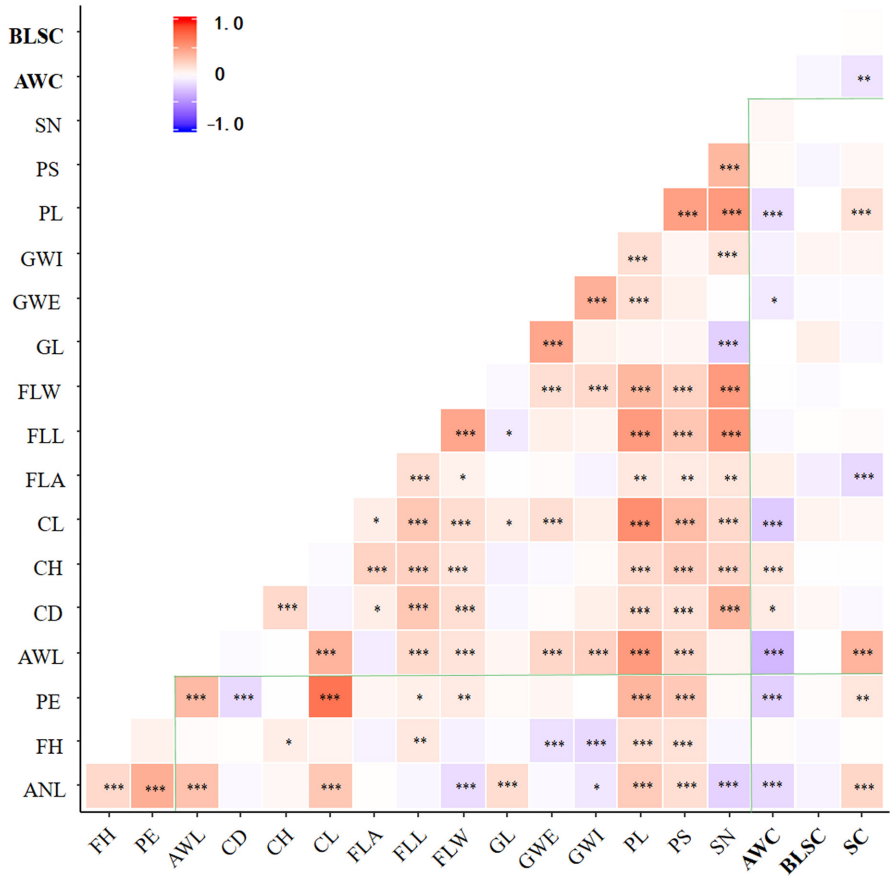

FIGURE 3 Pairwise correlation of 19 traits in the *Oryza rufipogon* × *Oryza nivara* F_2 populations. Three colour (qualitative) traits were in bold. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. The full names for the trait abbreviations are the same as those in Table 1. [Colour figure can be viewed at wileyonlinelibrary.com]

significantly clustered rather than over-dispersed (Table S6). Then, by marking all QTLs on 12 chromosomes, we identified a total of 18 QTL hotspots located on 11 chromosomes, each involving traits from 3 to 13 (Figure 4, Table S7). Interestingly, all major-effect QTLs (PVE > 10%) except for one (GL1.b) were located in QTL hotspots, with four in the hotspots on chromosome 1 (FLW1, GWI1, PL1, and SN1.a) and chromosome 6 (FH6, FLL6, PL6.a, and PS6.a), one each in hotspots on chromosomes 3 (FH3), 4 (AWL4), and 5 (ANL5.b) (Figure 4, Table S7).

It is noteworthy that a hotspot on chromosome 1 (HS1) involved nine species-distinguishing traits, including three RR traits (ANL, FH, and PE) and six HR traits (AWL, GL, GWE, GWI, PL, and PS) (Figure 4, Table S7). Similarly, a hotspot on chromosome 6 (HS11) involved eight putatively adaptive traits (ANL, FH, PE, AWL, CH, CL, PL, and PS). Other hotspots all included the QTLs involving adaptive traits (Figure 4, Table S7). Moreover, significantly positive correlations were detected for pairwise combinations of most adaptive traits with QTLs in the hotspots (Figure 3). For example, in the hotspot on chromosome 6 (HS11) where eight adaptive traits were involved, panicle shape (PS) was significantly correlated with all other adaptive traits, and similarly, anther length (ANL) was significantly correlated with all other adaptive traits except for culm habit (CH) (Figure 3). These results implied that a shared genetic basis (pleiotropy or linkage of multiple genes) might contribute to clustering of the QTLs responsible for interspecific trait divergence.

4 | DISCUSSION

4.1 | Phenotypic divergence between *O. nivara* and *O. rufipogon* and the underlying genetic architecture

Phenotypic variation within and between *O. rufipogon* and *O. nivara* have been extensively investigated (e.g., Banaticla-Hilario et al., 2013; Barbier, 1989; Cai et al., 2004; Cai et al., 2019; Eizenga et al., 2022; Morishima et al., 1961; Morishima et al., 1984; Ren, 2019; Vaughan, 1994) because these two species are direct ancestors of cultivated rice with abundant genetic variation. Our recent studies incorporating common garden experiment, artificial crossing, and population genomics (Cai et al., 2019; Guo et al., 2016; Ren, 2019; Xu et al., 2020) further demonstrated that the significant differentiation between species for a dozen phenotypic traits was associated with habitat differences, as expected for ecological speciation (Cai et al., 2019; Ren, 2019; Zheng & Ge, 2010). In the present study, we showed that most of the phenotypic traits divergent between species exhibited approximately normal distribution, as expected of quantitative or polygenic traits. Of 16 quantitative traits examined, 12 showed significant differentiation as expected for ecological divergence between *O. rufipogon* and *O. nivara* (Table 2), and these traits are associated with either reproductive isolation between species or the fitness of *O. nivara* in dry habitats (Banaticla-Hilario et al., 2013; Cai et al., 2019; Grillo et al., 2009; Xu et al., 2020). Notably, considerable interspecific

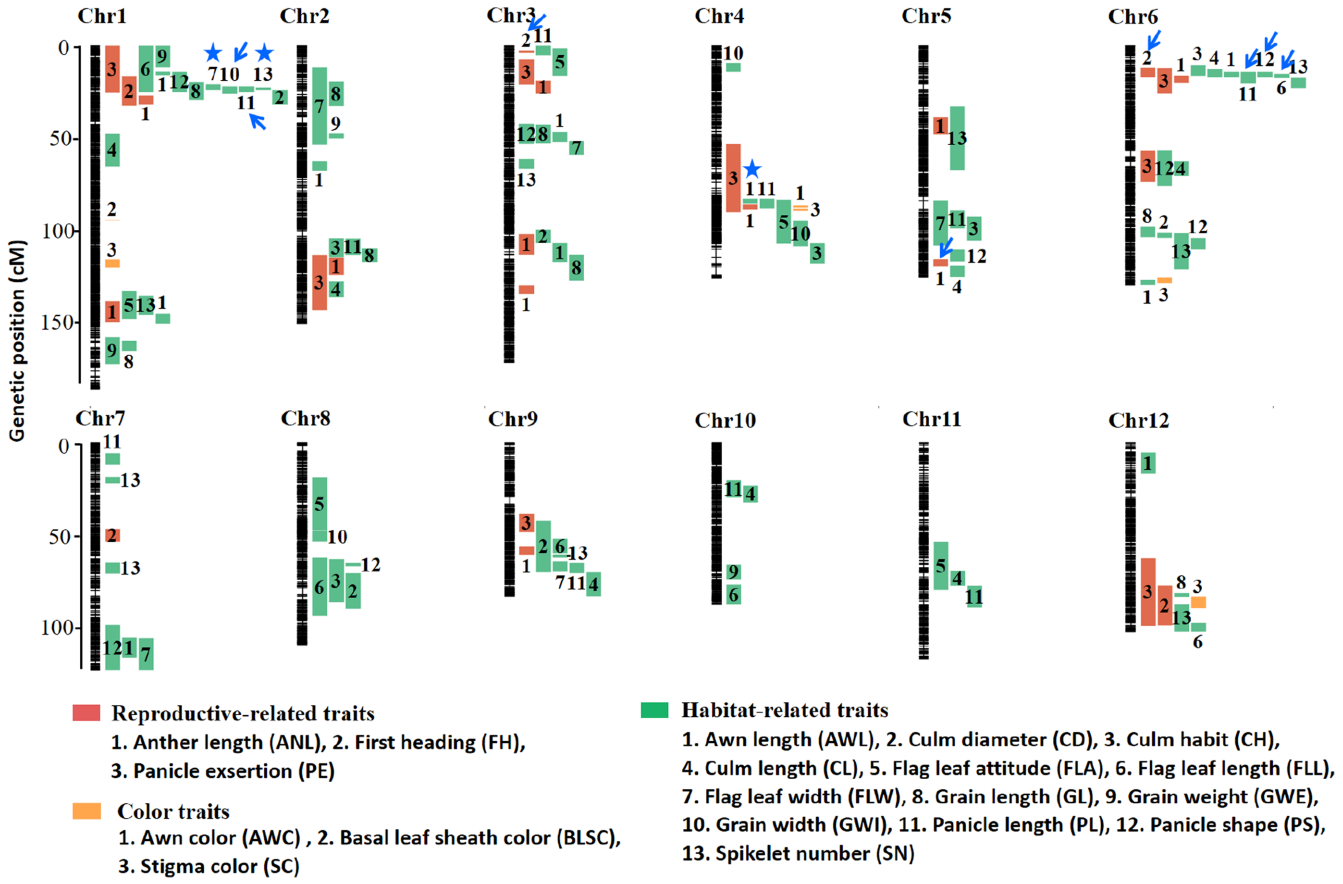

FIGURE 4 Genetic map of the *Oryza nivara* × *Oryza rufipogon* F_2 population, with the QTL locations involving all 19 phenotypic traits presented. Rectangular box indicates 1.5-LOD confidence intervals of each QTL, with width of the boxes corresponding to the range of genomic regions. Number within/beside the boxes represents the QTLs illustrated in the legends: reproduction-related traits (RR traits) (red), habitat-related traits (HR traits) (green) and colour traits (yellow). Arrows and stars above the numbers stand for the major-effect and large-effect QTLs, respectively. QTLs on the map are ordered according their positions on chromosome. [Colour figure can be viewed at wileyonlinelibrary.com]

divergence (>1.5-fold differences) was observed for several traits commonly used for distinguishing species (i.e., diagnostic traits), including anther length (ANL), first heading (FH), panicle exsertion (PE), culm length (CD), and panicle shape (PS) (Banaticla-Hilario et al., 2013; Cai et al., 2019; Eizenga et al., 2022; Jing et al., 2023; Ren, 2019). Overall, compared with the perennial *O. rufipogon*, the annual *O. nivara* flowers earlier with shorter anthers, is shorter with erected flag leaves, has shorter and less exserted panicles with more compactness, and shorter culm length (plant height), consistent with previous studies (Banaticla-Hilario et al., 2013; Cai et al., 2019; Eizenga et al., 2022; Grillo et al., 2009; Guo et al., 2016; Jing et al., 2023).

Relative to numerous studies that investigated genetic basis of trait divergence between cultivated rice and either *O. rufipogon* or *O. nivara* (e.g. Luo et al., 2016; Onishi et al., 2007; Thomson et al., 2003; Uga et al., 2003; Wang et al., 2011; Xiong et al., 1999), only a single study (Grillo et al., 2009) was performed to explore the genetic architecture of divergence between *O. rufipogon* and *O. nivara*. Grillo et al. (2009) identified a total of 30 QTLs involving eight quantitative traits related to life history, mating system, and flowering time and found that the effect sizes of QTLs ranged from 2.9% to 36.5% with an exponential distribution. Nevertheless, the low marker density and small mapping population in Grillo et al. (2009) study results in low mapping resolution and overestimates the percent variance explained for small-effect loci (Connallon & Hodgins, 2021; Slate, 2005; Visscher et al., 1996; Wang et al., 2011), which precludes deeper investigations into the genetic basis of phenotypic divergence between these two species.

In the present study, using a larger mapping population and higher marker density, we identified significantly more QTLs (119) responsible for 19 phenotypic traits with much narrower regions (Figure 4, Table S5). We noted that only seven of 30 QTLs identified in Grillo et al. (2009) were re-located in our study. The differences might arise from (1) the F_2 mapping populations generated from different parent lines, (2) different phenotypic traits studied, and (3) relatively crude estimates of QTL locations and magnitudes in Grillo et al. (2009) due to the low resolution arising from smaller F_2 mapping populations and lower density of SSR markers. In addition to a significant increase in loci underlying interspecific divergence of traits, we detected significant epistatic interactions for 23 pairs of QTLs involving 11 traits (Table 3), suggesting that the trait differences between species were highly polygenic and associated with

TABLE 2 Summary of the 116 QTLs identified for 16 quantitative traits with 600 F_2 progeny genotyped with 6579 bin markers distributed across the 12 rice chromosomes.

Trait (abbreviation)	QTL	Chr	Peak position (cM)	LOD	Mean RR phenotype	Mean RN phenotype	Mean NN phenotype	Effect and direction	PVE (%)
Reproduction-related trait (RR trait)									
Anther length (ANL)	ANL1.a	1	27.61	24.64	0.320	0.302	0.285	0.035	7.68
	ANL1.b	1	144.26	10.23	0.292	0.303	0.307	−0.015	3.01
	ANL2	2	115.71	9.33	0.309	0.303	0.292	0.017	2.74
	ANL3.a	3	20.15	10.93	0.317	0.299	0.291	0.026	3.23
	ANL3.b	3	107.38	5.57	0.299	0.307	0.292	0.007	1.61
	ANL3.c	3	129.58	12.04	0.310	0.302	0.290	0.020	3.57
	ANL4	4	85.27	21.44	0.281	0.302	0.318	−0.037	6.60
	ANL5.a	5	39.03	13.36	0.313	0.302	0.282	0.031	3.98
	ANL5.b	**5**	**114.93**	**45.34**	**0.323**	**0.303**	**0.277**	**0.046**	**15.35**
	ANL6	6	17.33	18.87	0.312	0.302	0.281	0.031	5.75
	ANL9	9	58.19	16.17	0.318	0.302	0.286	0.032	4.87
First heading (FH)	FH1	1	27.27	7.90	79.323	75.226	72.799	6.524	3.37
	FH3	**3**	**3.51**	**47.03**	**82.376**	**74.014**	**71.811**	**10.565**	**23.47**
	FH6	**6**	**13.81**	**36.00**	**79.701**	**74.325**	**72.452**	**7.249**	**17.18**
	FH7	7	51.53	18.03	71.968	75.552	79.196	−7.228	8.01
	FH12	12	79.93	5.40	77.725	74.007	75.911	1.814	2.28
Panicle exsertion (PE)	PE1	1	4.10	5.63	3.147	3.612	4.471	−1.324	3.04
	PE2	2	120.14	4.42	3.983	3.917	3.151	0.832	2.37
	PE3	3	9.03	4.87	4.406	3.600	3.263	1.143	2.62
	PE4	4	69.21	5.42	3.220	3.528	4.574	−1.354	2.92
	PE6.a	6	17.24	7.67	4.778	3.586	2.389	2.389	4.17
	PE6.b	6	67.69	9.37	4.568	3.780	2.336	2.232	5.13
	PE9	9	42.55	11.14	5.229	3.546	2.844	2.385	6.15
	PE12	12	91.56	5.64	4.312	3.702	2.927	1.385	3.04
Habitat-related trait (HR trait)									
Awn length (AWL)	AWL1.a	1	14.14	21.86	4.100	4.649	5.129	1.029	6.32
	AWL1.b	1	145.27	10.31	4.413	4.670	4.841	0.428	2.85
	AWL2	2	63.13	11.61	4.300	4.636	5.019	0.719	3.22
	AWL3.a	3	47.99	21.95	4.195	4.648	5.190	0.995	6.34
	AWL3.b	3	110.83	11.93	4.184	4.783	4.941	0.757	3.32
	AWL4	**4**	**83.60**	**89.48**	**3.177**	**4.846**	**5.528**	**2.351**	**34.14**
	AWL6.a	6	15.32	20.72	5.203	4.593	3.949	−1.254	5.96
	AWL6.b	6	125.08	8.81	4.388	4.623	5.003	0.615	2.42
	AWL7	7	109.29	10.47	4.507	4.648	4.843	0.336	2.89
	AWL12	12	9.62	6.49	4.758	4.702	4.454	−0.304	1.76
Culm diameter (CD)	CD1	1	26.44	7.46	0.473	0.467	0.494	0.021	4.61
	CD3	3	101.78	10.55	0.496	0.474	0.455	−0.041	6.60
	CD6	6	101.14	10.12	0.457	0.475	0.497	0.040	6.32
	CD8	8	84.44	7.73	0.493	0.469	0.464	−0.029	4.78
	CD9	9	64.89	5.18	0.458	0.473	0.493	0.035	3.17

TABLE 2 (Continued)

Trait (abbreviation)	QTL	Chr	Peak position (cM)	LOD	Mean RR phenotype	Mean RN phenotype	Mean NN phenotype	Effect and direction	PVE (%)
Culm habit (CH)	CH2	2	108.69	6.80	42.153	45.795	47.981	−5.828	4.29
	CH4	4	106.77	5.33	47.264	46.473	41.429	5.835	3.34
	CH5	5	96.86	7.45	48.359	46.090	42.134	6.225	4.70
	CH6	6	13.14	5.16	45.848	46.506	41.990	3.858	3.23
	CH8	8	70.40	4.38	47.995	44.184	45.202	2.793	2.73
Culm length (CL)	CL1	1	50.37	7.36	69.522	74.887	76.363	−6.841	3.90
	CL2	2	129.93	8.46	77.358	74.814	70.236	7.122	4.51
	CL5	5	117.86	5.62	75.687	74.646	70.426	5.261	2.96
	CL6.a	6	14.82	15.78	79.327	73.651	65.280	14.047	8.66
	CL6.b	6	67.69	8.68	76.803	74.508	67.797	9.006	4.63
	CL9	9	79.36	7.50	72.574	73.023	76.157	−3.583	3.98
	CL10	10	28.09	5.90	77.651	73.224	71.832	5.819	3.12
	CL11	11	69.08	4.54	70.277	75.432	74.000	−3.723	2.38
Flag leaf attitude (FLA)	FLA1	1	132.81	5.89	61.250	58.659	54.076	7.174	3.82
	FLA3	3	4.43	6.18	63.372	56.823	55.198	8.174	4.01
	FLA4	4	87.86	5.59	62.864	57.433	54.586	8.278	3.62
	FLA8	8	26.31	4.53	59.769	59.094	53.185	6.584	2.92
	FLA11	11	63.97	4.30	54.593	58.069	60.862	−6.269	2.77
Flag leaf length (FLL)	FLL1	1	13.13	5.83	20.081	20.831	21.741	−1.660	3.03
	FLL6	**6**	**16.15**	**26.30**	**22.559**	**20.761**	**18.548**	**4.011**	**14.82**
	FLL8	8	86.20	6.07	21.666	20.846	19.967	1.699	3.16
	FLL9	9	54.52	17.07	19.459	20.763	22.254	−2.795	9.28
	FLL10	10	83.80	4.99	19.915	20.91	21.822	−1.907	2.59
	FLL12	12	100.26	7.22	21.923	20.459	20.624	1.299	3.78
Flag leaf width (FLW)	**FLW1**	**1**	**23.00**	**46.75**	**0.836**	**0.892**	**1.007**	**−0.171**	**25.18**
	FLW2	2	20.83	7.41	0.891	0.898	0.949	−0.058	3.41
	FLW3	3	51.76	8.46	0.94	0.91	0.869	0.071	3.91
	FLW5	5	95.19	8.16	0.893	0.905	0.930	−0.037	3.77
	FLW7	7	111.13	5.32	0.879	0.915	0.934	−0.055	2.43
	FLW9	9	64.22	10.18	0.879	0.895	0.958	−0.079	4.74
Grain length (GL)	GL1.a	1	26.19	11.33	8.995	8.863	8.700	−0.295	5.16
	GL1.b	**1**	**160.28**	**20.38**	**8.661**	**8.839**	**9.037**	**0.376**	**9.61**
	GL2.a	2	27.13	12.99	8.971	8.827	8.765	−0.206	5.95
	GL2.b	2	110.70	10.03	8.912	8.868	8.759	−0.153	4.54
	GL3.a	3	44.48	6.51	8.88	8.878	8.751	−0.129	2.91
	GL3.b	3	121.97	17.79	8.715	8.837	9.016	0.301	8.31
	GL6	6	98.30	11.88	8.738	8.875	8.917	0.179	5.42
	GL12	12	81.02	11.93	8.763	8.909	8.818	0.055	5.44
Grain weight (GWE)	GWE1.a	1	5.10	8.07	0.581	0.599	0.608	0.027	5.15
	GWE1.b	1	159.95	7.04	0.578	0.597	0.613	0.035	4.47
	GWE2	2	48.50	9.41	0.613	0.597	0.582	−0.031	6.04
	GWE10	10	69.92	10.48	0.584	0.6	0.602	0.018	6.76

(Continues)

TABLE 2 (Continued)

Trait (abbreviation)	QTL	Chr	Peak position (cM)	LOD	Mean RR phenotype	Mean RN phenotype	Mean NN phenotype	Effect and direction	PVE (%)
Grain width (GWI)	**GWI1**	**1**	**23.42**	**26.44**	**2.369**	**2.429**	**2.534**	**0.165**	**15.86**
	GWI4.a	4	12.47	13.69	2.496	2.439	2.381	−0.115	7.81
	GWI4.b	4	99.66	5.51	2.405	2.45	2.467	0.062	3.04
	GWI8	8	47.66	7.08	2.441	2.43	2.481	0.040	3.94
Panicle length (PL)	**PL1**	**1**	**22.84**	**36.96**	**18.716**	**19.505**	**21.283**	**−2.567**	**14.22**
	PL2	2	107.52	7.90	20.456	19.791	19.180	1.276	2.71
	PL3	3	2.09	16.06	20.566	19.558	19.495	1.071	5.68
	PL4	4	83.43	12.80	18.503	20.057	20.325	−1.822	4.47
	PL5	5	92.18	11.25	20.829	19.845	18.814	2.015	3.91
	PL6	**6**	**15.74**	**37.92**	**21.294**	**19.645**	**17.747**	**3.547**	**14.64**
	PL7	7	7.61	5.62	19.741	19.572	20.146	−0.405	1.91
	PL9	9	65.31	14.55	19.207	19.576	20.583	−1.376	5.12
	PL10	10	20.65	12.04	20.498	19.834	19.094	1.404	4.20
	PL11	11	84.09	7.55	19.453	19.559	20.655	−1.202	2.59
Panicle shape (PS)	PS1	1	23.00	11.16	11.734	13.669	20.933	−9.199	5.17
	PS3	3	44.56	7.96	19.238	14.548	11.868	7.370	3.64
	PS5	5	112.92	8.99	19.768	14.526	12.038	7.730	4.13
	PS6.a	**6**	**15.15**	**24.4**	**21.16**	**15.367**	**4.121**	**17.039**	**11.91**
	PS6.b	6	71.54	6.40	18.15	15.358	10.551	7.599	2.91
	PS6.c	6	106.83	7.79	10.92	17.016	15.153	−4.233	3.56
	PS7	7	111.80	7.41	9.918	16.734	17.910	−7.992	3.38
	PS8	8	64.63	17.04	19.186	15.324	9.062	10.124	8.08
Spikelet number (SN)	**SN1.a**	**1**	**22.84**	**64.74**	**34.619**	**38.09**	**49.867**	**−15.248**	**25.95**
	SN1.b	1	139.17	10.23	44.343	40.037	37.228	7.115	3.29
	SN3	3	62.72	21.75	44.791	39.529	36.534	8.257	7.33
	SN5	5	36.19	5.43	41.68	40.724	37.530	4.150	1.72
	SN6.a	6	20.09	15.47	43.65	39.558	37.327	6.323	5.08
	SN6.b	6	105.24	8.04	37.875	40.341	43.102	−5.227	2.57
	SN7.a	7	19.57	12.10	38.791	39.343	42.563	−3.772	3.92
	SN7.b	7	66.26	8.82	40.834	41.089	37.595	3.239	2.82
	SN9	9	59.95	26.15	35.908	39.488	45.385	−9.477	8.96
	SN12	12	93.73	5.12	42.158	39.328	40.246	1.912	1.62

Note: All traits except for the four in italics were considered to be adaptive because they showed significant and the same differentiation patterns as those reported in previous studies of large samples/populations. PVE (%), percent phenotypic variance explained by QTL. The QTLs with major effect size (PVE > 10%) are in bold. Effect size was estimated as the difference between homozygous *O. rufipogon* alleles and homozygous *O. nivara* alleles at the QTL, with a positive value if the effect was consistent with species divergence and a negative value if the effect was opposite of species divergence.

epistasis, pleiotropy, and linkage of multiple genes. Interestingly, several important traits that proved to involve reproductive isolation and fitness, including anther length (ANL), first heading (FH), panicle length (PL), and panicle shape (PS) (Cai et al., 2019), were controlled by at least one major-effect QTL (Table 2). This implies that these traits might experience relatively large steps during initial stage of speciation because large-effect loci should be favoured when a population is far from the optimum (Connallon & Hodgins, 2021; Orr, 1998a). It is possible, as theory predicted (Orr, 1998a), that the formation of *O. nivara* proceeded in a way of "adaptive walk", in which several large-effect mutations involving adaptive phenotypes took place initially, followed by many small-effect mutations as the phenotypes moved closer to the optimum.

It is noted that the distribution of effect sizes of the QTLs identified in this study followed roughly an exponential model proposed by Orr's (1998a) and that nine out of 12 major-effect loci and as many as 74 small-effect loci underlay 12 putatively adaptive traits (Figures 4 and 5, Table 2). Such a polygenic and complex genetic basis of adaptive traits and an effect size distribution with significantly skewed toward the right appears to prevail during ecological speciation in

TABLE 3 Summary of the 23 significant epistatic interactions for 11 of the 16 quantitative traits.

Trait	QTLs	LOD	PVE (%)	*F* value
Reproduction-related trait (RR trait)				
ANL	ANL1.a×ANL9	2.72	0.7781	2.96*
	ANL3.c×**ANL5.b**	3.58	1.0271	3.91**
FH	**FH3**×**FH6**[a]	18.21	8.0955	21.63***
	FH3×FH7	2.66	1.1152	2.98*
	FH6×FH7[a]	3.00	1.2588	3.36**
Habitat-related trait (HR trait)				
AWL	AWL1.a×AWL2	2.81	0.7536	3.05*
	AWL1.b×AWL3.a	3.61	0.9711	3.92**
	AWL3.a×AWL6.a	2.46	0.6598	2.67*
	AWL4×AWL6.b	3.20	0.8610	3.48**
	AWL6.a×AWL7	2.96	0.7956	3.22*
CL	CL6.b×CL9	2.53	1.3191	2.84*
CD	CD1×CD6	2.30	1.3949	2.59*
	CD3×CD8	2.87	1.7447	3.24*
FLW	**FLW1**×FLW5	3.30	1.4945	3.72**
GL	GL2.a×GL2.b	3.98	1.7605	4.43**
	GL3.b×GL12	2.29	1.0068	2.53*
	GL6×GL12	2.57	1.1321	2.85*
GWE	GWE1.a×GWE10[a]	3.40	2.1314	3.85**
	GWE2×GWE10	3.20	2.0029	3.62**
GWI	GWI4.a×GWI8	2.36	1.2880	2.68*
PL	**PL1**×PL3	2.50	0.8406	2.77*
	PL1×PL4	2.26	0.7578	2.50*
PS	PS1×PS8	3.80	1.7088	4.28**

Note: Those with boldface indicate the QTLs with major effect.

[a]The QTL interactions that remain significant after applying false discovery rate correction at 5% level.

*$p<0.05$; **$p<0.01$; ***$p<0.001$.

plants and animals (e.g., Feng et al., 2019; Ferris et al., 2017; Fishman et al., 2002; Jacobs et al., 2017; Lexer et al., 2005; Lowry et al., 2015; Milano et al., 2016; Nakazato et al., 2013; and reviewed in Bomblies & Peichel, 2022; Dittmar et al., 2016; Hall et al., 2016). Overall, our findings supported accumulating empirical studies that indicated the polygenic basis of adaptive traits and multiple genetic regions underlying trait differentiations during adaptation and speciation (Bomblies & Peichel, 2022; Nosil et al., 2021; Saltz et al., 2017; Slate, 2005).

4.2 | QTL hotspots and the genetic basis underlying trait divergence

It is hypothesized that mechanisms facilitating coinheritance of adaptive phenotypes are favoured when organisms under divergent selection are adapting to multiple different aspects of new environments (Bomblies & Peichel, 2022; Nosil et al., 2021). Despite many studies on plants and animals, empirical investigations on ecological speciation remain rare with limited knowledge on the relationship between clustering of QTLs/genes and phenotypic divergence between species (Archambeault et al., 2020; Bomblies & Peichel, 2022). In this study, we identified 18 QTL hotspots on 11 chromosomes, with each involving multiple traits (3–13) and found that 11 of all 12 major-effect QTLs were in the QTL hotspots (Figure 4, Table S7). These results suggest that the formation of QTL hotspots or coinheritance of loci/genes, particularly the hotspots involving QTLs with large effect size, may play important roles in adaptation and speciation, as evidenced in many plants (Ferris et al., 2017; Grillo et al., 2009; Hall et al., 2006; Lowry et al., 2015; Onishi et al., 2007) and animals (Archambeault et al., 2020; Jacobs et al., 2017). Interestingly, the QTLs controlling three RR traits (ANL, FH, and PE) colocalized in the hotspots of chromosomes 1 and 6 simultaneously (Figure 4, Table S7), suggesting that a series of traits related reproductive isolation were selected together by either tight linkage of loci or pleiotropy. This notion was supported by correlation analyses in which significant correlations were detected between pairwise combinations of ANL, FH, and PE (Figure 3). QTLs

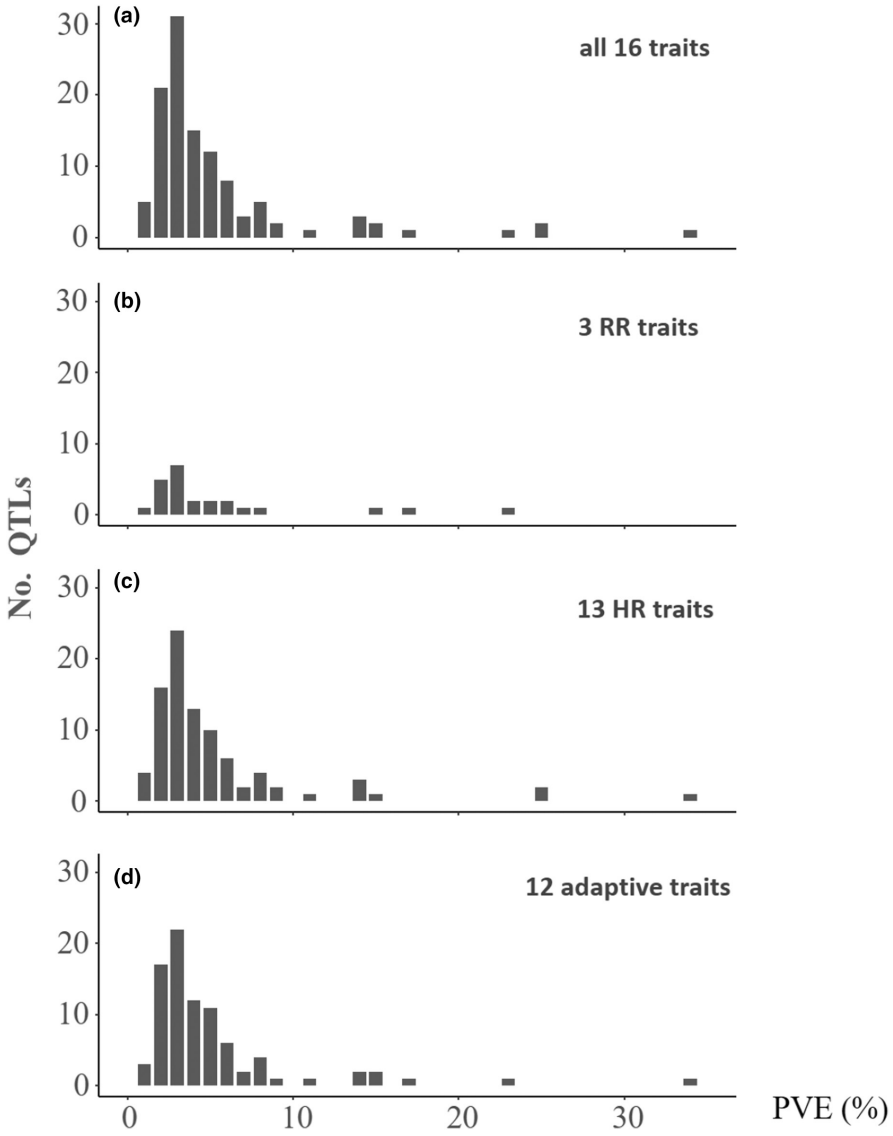

FIGURE 5 The distribution of effect sizes of QTLs identified in the F_2 mapping populations. (a) 113 QTLs for all 16 quantitative traits; (b) 24 QTLs for three reproductive-related (RR) traits; (c) 89 QTLs for 13 habitat-related (HR) traits; and (d) 86 QTLs for 12 putatively adaptive traits. The description of the trait categories is present in Table 1.

controlling two panicle traits (PL and PS) were also found in multiple hotspots (8 hotspots each) across different chromosomes (Figure 4, Table S7). Moreover, these traits correlated with almost all other traits (Figure 3), demonstrating that QTL clustering in which multiple loci evolved together was the preferential strategy of adaptation and speciation (Nosil et al., 2021; Peichel & Marques, 2017).

Early studies indicated that phenotypic differences, of particular the traits involving reproductive isolation, tended to go hand-in-hand (Orr, 2001). Later theoretical (Hoffmann & Rieseberg, 2008; Peichel & Marques, 2017) and empirical studies (Archambeault et al., 2020; Ferris et al., 2017; Grillo et al., 2009; Linnen et al., 2013; Lowry et al., 2015) suggested that the presence of QTL hotspots would facilitate adaptation to new environments and accelerate the process of speciation. Many studies have found that chromosomal inversions, usually colocalized with identified QTL hotspots, contributed to genetic differentiation between ecotypes/species in many plants and animals (Hoffmann & Rieseberg, 2008; Wellenreuther & Bernatchez, 2018). A large number of QTL hotspots, as evidenced in our study, suggest that formation of QTL hotspots plays an important role during ecological speciation to evolve new species that have optimum fitness, although we do not know yet whether these hotspots contain a single pleiotropic locus or many tightly linked loci or both. Elucidating genetic underpinning of QTL hotspots is critical to understanding the molecular mechanisms and the factors facilitating ecological adaptation and speciation.

4.3 | Divergent selection and its role in ecological speciation of *O. nivara*

It is widely acknowledged that natural selection is the primary force shaping the phenotypic differences that evolve during adaptation and speciation (Nosil, 2012; Nosil et al., 2021; Orr, 1998b; Rieseberg et al., 2002; Schluter & Rieseberg, 2022; Seehausen et al., 2014). Based on previous studies of interspecific phenotypic divergence using variance and Q_{ST}–F_{ST} analyses, Guo et al. (2016) found that 9 of 24 phenotypic traits measured showed significant divergence between *O. rufipogon* and *O. nivara*. Similarly, Cai et al. (2019) showed

that 11 of 18 phenotypic traits exhibited significantly higher values for Q_{ST} than for F_{ST} across six species pairs. These results demonstrated the roles of natural selection rather than random genetic drift in the phenotypic differentiation between *O. rufipogon* and *O. nivara*. However, Q_{ST}–F_{ST} analysis was performed under several assumptions (Fraser, 2020) and was unable to distinguish between direct selection acting on the traits and indirect selection due to correlations with other selected traits given that neutral traits correlating with traits under selection may show correlated responses in evolutionary change (Muir et al., 2014; Via & Hawthorne, 2005).

Analysis of the direction of allelic effects of QTLs can facilitate inference on whether trait divergence is consistent with directional natural selection (Muir et al., 2014; Orr, 1998b; Rieseberg et al., 2002). Specifically, if the effects of most QTLs for a trait move the phenotype in the same direction (positive), it is most likely that directional selection has caused the trait divergence. By contrast, weak selection or drift may be the main drivers for the divergence if a mixed of positive and negative (i.e., antagonistic) effects are present for the QTLs (Ferris et al., 2017; Muir et al., 2014; Nakazato et al., 2013; Orr, 1998b). In the present study, we found a high proportion of positive QTLs for nine of 12 putatively adaptive traits (Table 2), that is, the QTL effects were predominantly in the direction of expected species divergence. Of the nine traits, seven were also significant in Fraser's *v*-test (Table 1). These results suggest that interspecific differentiations of most traits were driven by divergent natural selection, supporting the importance of natural selection in phenotypic divergence between *O. rufipogon* and *O. nivara*. Nevertheless, it is premature to claim all these traits experience directional selection because indirect selection cannot be excluded given the fact that significant correlations have been found for pairwise combinations of these traits (Figure 3).

Accumulating evidence indicates that adaptation to new environments often involves shifts of many ecologically important traits and results in covariation of these traits across species (Erickson et al., 2004; Lowry et al., 2015; Muir et al., 2014). In our case, a suite of traits associated with the derived *O. nivara*, such as earlier flowering, shorter culm, annual life history, and prominently selfing mating system, are most likely to evolve to adapt to drier habitats (Banaticla-Hilario et al., 2013; Cai et al., 2019; Grillo et al., 2009; Guo et al., 2016). Trait shifts associated with xeric/mesic divergence have been evident in adaptive adaptation/speciation of many other plant species (e.g., Ferris et al., 2017; Hall et al., 2006; Lowry et al., 2015; Milano et al., 2016). Such trait covariation could be due to coselection, whereby each trait improves the adaptive capacity in dry environments and alternatively, arose from a shared genetic basis (pleiotropy/linkage of genes), in which traits evolve in concert (Erickson et al., 2004; Muir et al., 2014). Our previous studies suggested that the change of first heading (FH) might have been the first step in the evolution of *O. nivara*, because flowering time contributes to both local adaptation to avoid drought and reproductive isolation to block the gene exchange between species (Cai et al., 2019; Guo et al., 2016; Xu et al., 2020). Interestingly, we found FH in the two largest QTL hotspots with the number of QTLs over 10 (HS1 and HS10) (Figure 4, Table S7), implying the potential interactions between loci/genes controlling flowering time and those responsible for species divergence of other adaptive traits. Our correlation analyses also observed significantly positive correlation between FH and other adaptive traits (Figure 3). These findings suggested a common genetic basis underlying the co-adaptation of flowering time and other adaptive traits. Indeed, several studies have cloned/identified the QTLs that pleiotropically control flowering time, plant height, number of spikelets and drought escape in rice (e.g., Yan et al., 2011) and other grass species (Lowry et al., 2015). These co-adaptive traits are therefore excellent candidates for future research in terms of genetic and functional perspectives. Collectively, our QTL analysis added to a growing body of evidence that the annual *O. nivara* evolved from the perennial *O. rufipogon* as an adaptation to new environments due to divergent natural selection that favours co-adapted traits (Cai et al., 2019; Grillo et al., 2009; Guo et al., 2016; Huang et al., 2013). These results are consistent with the argument that adaptation and diversification would proceed rapidly under coordinated selection of multiple traits (Erickson et al., 2004; Saltz et al., 2017). To further investigate the prevalence and molecular basis of genomic coupling may be a key to understanding ecological adaptation and speciation.

AUTHOR CONTRIBUTIONS

S. G. conceived and designed project. Q-L. Meng, C-G. Qiang, M-F. Geng, N-N. Ren, M-X. Wang and Z-H. Jiao conducted the fieldwork and phenotyping; Q-L. Meng, C-G. Qiang, J-L Li, and X-J. Song performed the QTL mapping; Q-L. Meng, C-G. Qiang, Z. Cai, F-M. Zhang performed data analyses. S. Ge, Q-L. Meng, C-G. Qiang, and J-L, Li wrote the manuscript.

ACKNOWLEDGEMENTS

We thank Xun Xu, Xin Wang, Ning Li, Lian Zhou, Yu-Su Du, Xiu-Hua Wang, Jing-Dan Han, and other members of Ge's laboratory for help in phenotyping and lab assistances, and Hua-Zhao Liu of the CAS Field Station (Lingshui, China) for assistance with the field experiment. We also thank the International Rice Research Institute (IRRI) (Los Banos, Philippines) for providing seed samples and the CAS Field Station (Lingshui, China) for providing the experimental field. This work was financially supported by National Natural Science Foundation of China (NSFC) (32130008), Strategic Priority Research Program of Chinese Academy of Sciences (XDB31000000), NSFC (31900198 and 31800186), and Ministry of Science and Technology (2021YFD1200101-02).

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare for this article.

DATA AVAILABILITY STATEMENT

Raw sequence data from this study can be found on NCBI Sequence Read Archive (accession no. PRJNA988162). The genomic SNP data for F_2 and their parents are available on Dryad (doi: https://doi.org/10.5061/dryad.9zw3r22ng). Scripts for genotyping and linkage analysis are openly available on GitHub at https://github.com/qiang-cg/F2-QTL-mapping.

ORCID

Song Ge https://orcid.org/0000-0002-7683-5579

REFERENCES

Archambeault, S. L., Bartschi, L. R., Merminod, A. D., & Peichel, C. L. (2020). Adaptation via pleiotropy and linkage: Association mapping reveals a complex genetic architecture within the stickleback *Eda* locus. *Evolution Letters*, *4*, 282–301.

Arnegard, M. E., McGee, M. D., Matthews, B., Marchinko, K. B., Conte, G. L., Kabir, S., Bedford, N., Bergek, S., Chan, Y. F., Jones, F. C., Kingsley, D. M., Peichel, C. L., & Schluter, D. (2014). Genetics of ecological divergence during speciation. *Nature*, *511*, 307–311.

Banaticla-Hilario, M. C. N., Sosef, M. S. M., McNally, K. L., Hamilton, N. R. S., & van den Berg, R. G. (2013). Ecogeographic variation in the morphology of two Asian wild rice species, *Oryza nivara* and *Oryza rufipogon*. *International Journal of Plant Sciences*, *174*, 896–909.

Barbier, P. (1989). Genetic variation and ecotypic differentiation in the wild rice species *Oryza rufipogon*. I. Population differentiation in life-history traits and isozymic loci. *Japanese Journal of Genetics*, *64*, 259–271.

Barton, N. H., & Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nature Reviews Genetics*, *3*, 11–21.

Bioversity International, IRRI, and Africa Rice Center. (2007). *Descriptors for wild and cultivated Rice (Oryza spp.)*. Bioversity International.

Bomblies, K., & Peichel, C. L. (2022). Genetics of adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2122152119.

Bradshaw, H. D., Otto, K. G., Frewen, B. E., McKay, J. K., & Schemske, D. W. (1998). Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics*, *149*, 367–382.

Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/QTL: QTL mapping in experimental crosses. *Bioinformatics*, *19*, 889–890.

Burke, J. M., Tang, S., Knapp, S. J., & Rieseberg, L. H. (2002). Genetic analysis of sunflower domestication. *Genetics*, *161*, 1257–1267.

Cai, H. W., & Morishima, H. (2002). QTL clusters reflect character associations in wild and cultivated rice. *Theoretical and Applied Genetics*, *104*, 1217–1228.

Cai, H. W., Wang, X. K., & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity*, *92*, 409–417.

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., … Ge, S. (2019). Parallel speciation of wild rice associated with habitat shifts. *Molecular Biology and Evolution*, *36*, 875–889.

Connallon, T., & Hodgins, K. A. (2021). Allen Orr and the genetics of adaptation. *Evolution*, *75*, 2624–2640.

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the R package *sommer*. *PLoS ONE*, *11*, e0156744.

Dabravolski, S. A., & Isayenkov, S. V. (2023). The role of anthocyanins in plant tolerance to drought and salt stresses. *Plants-Basel*, *12*, 2558.

Davey, M. W., Kenis, K., & Keulemans, J. (2006). Genetic control of fruit vitamin C contents. *Plant Physiology*, *142*, 343–351.

Dittmar, E. L., Oakley, C. G., Conner, J. K., Gould, B. A., & Schemske, D. W. (2016). Factors influencing the effect size distribution of adaptive substitutions. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20153065.

Eizenga, G. C., Kim, H., Jung, J. K. H., Greenberg, A. J., Edwards, J. D., Naredo, M. E. B., Banaticla-Hilario, M. C. N., Harrington, S. E., Shi, Y. X., Kimball, J. A., Harper, L. A., McNally, K. L., & McCouch, S. R. (2022). Phenotypic variation and the impact of admixture in the *Oryza rufipogon* species complex (ORSC). *Frontiers in Plant Science*, *13*, 787703.

Erickson, D. L., Fenster, C. B., Stenoien, H. K., & Price, D. (2004). Quantitative trait locus analyses and the study of evolutionary process. *Molecular Ecology*, *13*, 2505–2522.

Faria, R., Renaut, S., Galindo, J., Pinho, C., Melo-Ferreira, J., Melo, M., Jones, F., Salzburger, W., Schluter, D., & Butlin, R. (2014). Advances in ecological speciation: An integrative approach. *Molecular Ecology*, *23*, 513–521.

Feng, C., Feng, C., Yang, L. H., Kang, M., & Rausher, M. D. (2019). Genetic architecture of quantitative flower and leaf traits in a pair of sympatric sister species of *Primulina*. *Heredity*, *122*, 864–876.

Ferris, K. G., Barnett, L. L., Blackman, B. K., & Willis, J. H. (2017). The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Molecular Ecology*, *26*, 208–224.

Fishman, L., Kelly, A. J., & Willis, J. H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution*, *56*, 2138–2155.

Frary, A., Frary, A., Daunay, M. C., Huvenaars, K., Mank, R., & Doganlar, S. (2014). QTL hotspots in eggplant (*Solanum melongena*) detected with a high resolution map and CIM analysis. *Euphytica*, *197*, 211–228.

Fraser, H. B. (2020). Detecting selection with a genetic cross. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 22323–22330.

Geng, M. F., Wang, X. H., Wang, M. X., Cai, Z., Meng, Q. L., Wang, X., Zhou, L., Han, J. D., Li, J., Zhang, F. M., Guo, Y.-L., & Ge, S. (2021). Genome-wide investigation on transcriptional responses to drought stress in wild and cultivated rice. *Environmental and Experimental Botany*, *189*, 104555.

Gonda, I., Ashrafi, H., Lyon, D. A., Strickler, S. R., Hulse-Kemp, A. M., Ma, Q. Y., Sun, H. H., Stoffel, K., Powell, A. F., Futrell, S., Thannhauser, T. W., Fei, Z., van Deynze, A. E., Mueller, L. A., Giovannoni, J. J., & Foolad, M. R. (2019). Sequencing-based bin map construction of a tomato mapping population, facilitating high-resolution quantitative trait loci detection. *Plant Genome*, *12*, 180010.

Grillo, M. A., Li, C. B., Fowlkes, A. M., Briggeman, T. M., Zhou, A. L., Schemske, D. W., & Sang, T. (2009). Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution*, *63*, 870–883.

Guo, J., Liu, R., Huang, L., Zheng, X. M., Liu, P. L., Du, Y. S., Cai, Z., Zhou, L., Wei, X. H., Zhang, F. M., & Ge, S. (2016). Widespread and adaptive alterations in genome-wide gene expression associated with ecological divergence of two *Oryza* species. *Molecular Biology and Evolution*, *33*, 62–78.

Hall, D., Hallingback, H. R., & Wu, H. X. (2016). Estimation of number and size of QTL effects in forest tree traits. *Tree Genetics & Genomes*, *12*, 110.

Hall, M. C., Basten, C. J., & Willis, J. H. (2006). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. *Genetics*, *172*, 1829–1844.

Hoffmann, A. A., & Rieseberg, L. H. (2008). Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics*, *39*, 21–42.

Huang, L., Du, Y. S., Zheng, X. M., Liu, R., Zhou, H. F., & Ge, S. (2013). Nucleotide diversity of 11S seed storage protein gene and its implications for ecological adaptation of *Oryza nivara*. *Journal of Systematics and Evolution*, *51*, 641–651.

Huang, X. H., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A. H., Guan, J. P., Fan, D. L., Weng, Q. J., Huang, T., Dong, G., Sang, T., & Han, B. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Research*, *19*, 1068–1076.

Jacobs, A., Womack, R., Chen, M., Gharbi, K., & Elmer, K. R. (2017). Significant synteny and colocalization of ecologically relevant quantitative trait loci within and across species of salmonid fishes. *Genetics*, *207*, 741–754.

Jakobson, C. M., & Jarosz, D. F. (2020). What has a century of quantitative genetics taught us about nature's genetic tool kit? *Annual Review of Genetics*, *54*, 439–464.

Jing, C.-Y., Zhang, F.-M., Wang, X.-H., Wang, M.-X., Zhou, L., Cai, Z., Han, J.-D., Geng, M.-F., Yu, W.-H., Jiao, Z.-H., Huang, L., Liu, R., Zheng, X.-M., Meng, Q.-L., Ren, N.-N., Zhang, H.-X., Du, Y.-S., Wang, X., Qiang, C.-G., … Ge, S. (2023). Multiple domestications of Asian rice. *Nature Plants*, *9*, 1221–1235.

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., Schwartz, D. C., Tanaka, T., Wu, J. Z., Zhou, S. G., Childs, K. L., Davidson, R. M., Lin, H., Quesada-Ocampo, L., Vaillancourt, B., Sakai, H., Lee, S. S., Kim, J., Numa, H., … Matsumoto, T. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*, *6*, 4.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology*, *35*, 25–34.

Kim, H., Jung, J., Singh, N., Greenberg, A., Doyle, J. J., Tyagi, W., Chung, J. W., Kimball, J., Hamilton, R. S., & McCouch, S. R. (2016). Population dynamics among six major groups of the *Oryza rufipogon* species complex, wild relative of cultivated Asian rice. *Rice*, *9*, 56.

Kitano, J., Ishikawa, A., Ravinet, M., & Courtier-Orgogozo, V. (2022). Genetic basis of speciation and adaptation: From loci to causative mutations. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *377*, 20200503.

Kumar, J., Sen Gupta, D., Gupta, S., Dubey, S., Gupta, P., & Kumar, S. (2017). Quantitative trait loci from identification to exploitation for crop improvement. *Plant Cell Reports*, *36*, 1187–1213.

Lexer, C., Rosenthal, D. M., Raymond, O., Donovan, L. A., & Rieseberg, L. H. (2005). Genetics of species differences in the wild annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics*, *169*, 2225–2239.

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics*, *26*, 589–595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Proc, G. P. D. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Linnen, C. R., Poh, Y. P., Peterson, B. K., Barrett, R. D. H., Larson, J. G., Jensen, J. D., & Hoekstra, H. E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, *339*, 1312–1316.

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*. *Molecular Ecology*, *24*, 5211–5228.

Lowry, D. B., Hernandez, K., Taylor, S. H., Meyer, E., Logan, T. L., Barry, K. W., Chapman, J. A., Rokhsar, D. S., Schmutz, J., & Juenger, T. E. (2015). The genetics of divergence and reproductive isolation between ecotypes of *Panicum hallii*. *New Phytologist*, *205*, 402–414.

Luo, X. D., Zhao, J., Dai, L. F., Zhang, F. T., Zhou, Y., Wan, Y., & Xie, J. K. (2016). Linkage map construction and QTL mapping for cold tolerance in *Oryza rufipogon* Griff. at early seedling stage. *Journal of Integrative Agriculture*, *15*, 2703–2711.

Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*, 68–78.

Meng, Q. L. (2021). QTL mapping of the traits related to species divergence between two wild rice species (*Oryza rufipogon* and *O. niavra*). PhD thesis, Institute of Botany, the Chinese Academy of Sciences, Beijing, China.

Milano, E. R., Lowry, D. B., & Juenger, T. E. (2016). The genetic basis of upland/lowland ecotype divergence in switchgrass (*Panicum virgatum*). *G3: Genes, Genomes, Genetics*, *6*, 3561–3570.

Morishima, H., Chang, W. T., & Oka, H. I. (1961). Directions of differentiation in populations of wild rice, *Oryza perennis* and *O. sativa* f. *spontanea*. *Evolution*, *15*, 326–339.

Morishima, H., Sano, Y., & Oka, H. I. (1984). Differentiation of perennial and annual types due to habitat conditions in the wild rice *Oryza perennis*. *Plant Systematics and Evolution*, *144*, 119–135.

Muir, C. D., Pease, J. B., & Moyle, L. C. (2014). Quantitative genetic analysis indicates natural selection on leaf phenotypes across wild tomato species (*Solanum* sect. *Lycopersicon*; Solanaceae). *Genetics*, *198*, 1629–1643.

Murray, M. G., & Thompson, W. F. (1980). Rapid isolation of high molecular-weight plant DNA. *Nucleic Acids Research*, *8*, 4321–4325.

Nakazato, T., Rieseberg, L. H., & Wood, T. E. (2013). The genetic basis of speciation in the *Giliopsis* lineage of *Ipomopsis* (Polemoniaceae). *Heredity*, *111*, 227–237.

Nosil, P. (2012). *Ecological speciation*. Oxford University Press.

Nosil, P., Feder, J. L., & Gompert, Z. (2021). How many genetic changes create new species? *Science*, *371*, 777–779.

Oakley, C. G., Savage, L., Lotz, S., Larson, G. R., Thomashow, M. F., Kramer, D. M., & Schemske, D. W. (2018). Genetic basis of photosynthetic responses to cold in two locally adapted populations of *Arabidopsis thaliana*. *Journal of Experimental Botany*, *69*, 699–709.

Onishi, K., Horiuchi, Y., Ishigoh-Oka, N., Takagi, K., Ichikawa, N., Maruoka, M., & Sano, Y. (2007). A QTL cluster for plant architecture and its ecological significance in Asian wild rice. *Breeding Science*, *57*, 7–16.

Oren, E., Tzuri, G., Vexler, L., Dafna, A., Meir, A., Faigenboim, A., Kenigswald, M., Portnoy, V., Schaffer, A. A., Levi, A., Buckler, E. S., Katzir, N., Burger, J., Tadmor, Y., & Gur, A. (2019). The multi-allelic *APRR2* gene is associated with fruit pigment accumulation in melon and watermelon. *Journal of Experimental Botany*, *70*, 3781–3794.

Orr, H. A. (1998a). The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution*, *52*, 935–949.

Orr, H. A. (1998b). Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics*, *149*, 2099–2104.

Orr, H. A. (2001). The genetics of species differences. *Trends in Ecology & Evolution*, *16*, 343–350.

Peichel, C. L., & Marques, D. A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*, 20150486.

Qin, L. M., Sun, L., Wei, L., Yuan, J. R., Kong, F. F., Zhang, Y., Miao, X., Xia, G. M., & Liu, S. W. (2021). Maize SRO1e represses anthocyanin synthesis through regulating the MBW complex in response to abiotic stress. *Plant Journal*, *105*, 1010–1025.

R Development Core Team. (2020). *R: A language and environment for statistical computing, V4.0.2*. R Foundation for Statistical Computing. https://www.R-project.org/

Ren, N. N. (2019). Studies on the phenotypic variation and its evolutionary significance if wild rice Oryza rufipogon and O. Nivara. PhD thesis, Institute of Botany, the Chinese Academy of Sciences, Beijing, China.

Rieseberg, L. H., Widmer, A., Arntz, A. M., & Burke, J. M. (2002). Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 12242–12245.

Roda, F., Walter, G. M., Nipper, R., & Ortiz-Barrientos, D. (2017). Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower. *Molecular Ecology*, *26*, 3687–3699.

Saltz, J. B., Hessel, F. C., & Kelly, M. W. (2017). Trait correlations in the genomics era. *Trends in Ecology & Evolution*, *32*, 279–290.

Sang, T., & Ge, S. (2007). Genetics and phylogenetics of rice domestication. *Current Opinion in Genetics & Development*, *17*, 533–538.

Sano, Y., Morishima, H., & Oka, H. I. (1980). Intermediate perennial-annual populations of *Oryza perennis* found in Thailand and their evolutionary significance. *Botanical Magazine-Tokyo*, *93*, 291–305.

Schluter, D., & Rieseberg, L. H. (2022). Three problems in the genetics of speciation by selection. *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2122153119.

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G. P., Bank, C., Brannstrom, A., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., … Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*, 176–192.

Slate, J. (2005). Quantitative trait locus mapping in natural populations: Progress, caveats and future directions. *Molecular Ecology*, *14*, 363–379.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genome-wide studies. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 9440–9445.

Sun, X. W., Liu, D. Y., Zhang, X. F., Li, W. B., Liu, H., Hong, W. G., Jiang, C. B., Guan, N., Ma, C. X., Zeng, H. P., Xu, C., Song, J., Huang, L., Wang, C., Shi, J., Wang, R., Zheng, X., Lu, C., Wang, X., & Zheng, H. (2013). SLAF-seq: An efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE*, *8*, e58700.

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, *27*, 205–233.

Thomson, M. J., Tai, T. H., McClung, A. M., Lai, X. H., Hinga, M. E., Lobos, K. B., Xu, Y., Martinez, C. P., & McCouch, S. R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical and Applied Genetics*, *107*, 479–493.

Uga, Y., Fukuta, Y., Cai, H. W., Iwata, H., Ohsawa, R., Morishima, H., & Fujimura, T. (2003). Mapping QTLs influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *Oryza rufipogon* Griff. *Theoretical and Applied Genetics*, *107*, 218–226.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). Genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Bioinformatics*, *28*, 254–255.

Vaughan, D. A. (1994). *The wild relatives of rice: A genetic resources handbook*. International Rice Research Institute.

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

Via, S., & Hawthorne, D. J. (2005). Back to the future: Genetic correlations, adaptation and speciation. *Genetica*, *123*, 147–156.

Visscher, P. M., Thompson, R., & Haley, C. S. (1996). Confidence intervals in QTL mapping by bootstrapping. *Genetics*, *143*, 1013–1020.

Wang, L., Wang, A. H., Huang, X. H., Zhao, Q. A., Dong, G. J., Qian, Q. A., Sang, T., & Han, B. (2011). Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombinant inbred lines. *Theoretical and Applied Genetics*, *122*, 327–340.

Wang, Q. J., Liao, Z. Q., Zhu, C. T., Gou, X. J., Liu, Y. X., Xie, W. B., Wu, F. K., Feng, X. J., Xu, J., Li, J. W., & Lu, Y. (2022). Teosinte confers specific alleles and yield potential to maize improvement. *Theoretical and Applied Genetics*, *135*, 3545–3562.

Wellenreuther, M., & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution*, *33*, 427–440.

Westerbergh, A., & Doebley, J. (2004). Quantitative trait loci controlling phenotypes related to the perennial versus annual habit in wild relatives of maize. *Theoretical and Applied Genetics*, *109*, 1544–1553.

Xiong, L. X., Liu, K. D., Dai, X. K., Xu, C. G., & Zhang, Q. F. (1999). Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theoretical and Applied Genetics*, *98*, 243–251.

Xu, X., Meng, Q. L., Geng, M. F., Ren, N. N., Zhou, L., Du, Y. S., Cai, Z., Wang, M. X., Wang, X., Wang, X. H., Han, J.-D., Jiang, S., Jing, C.-Y., Liu, R., Zheng, X.-M., Yang, Q.-W., Zhang, F.-M., & Ge, S. (2020). Divergence in flowering time is a major component contributing to reproductive isolation between two wild rice species (*Oryza rufipogon* and *O. nivara*). *Science China Life Sciences*, *63*, 1714–1724.

Yan, W. H., Wang, P., Chen, H. X., Zhou, H. J., Li, Q. P., Wang, C. R., Ding, Z. H., Zhang, Y. S., Yu, S. B., Xing, Y. Z., & Zhang, Q. F. (2011). A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Molecular Plant*, *4*, 319–330.

Yang, M. H., Wu, D. H., & Kao, C. H. (2019). A statistical procedure for genome-wide detection of QTL hotspots using public databases with application to rice. *G3: Genes, Genomes, Genetics*, *9*, 439–452.

Zhang, L. Y., Wang, S. Q., Li, H. H., Deng, Q. M., Zheng, A. P., Li, S. C., Li, P., Li, Z. L., & Wang, J. K. (2010). Effects of missing marker and segregation distortion on QTL mapping in F_2 populations. *Theoretical and Applied Genetics*, *121*, 1071–1082.

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*). *Molecular Ecology*, *19*, 2439–2454.

Zuo, J. F., Niu, Y., Cheng, P., Feng, J. Y., Han, S. F., Zhang, Y. H., Shu, G. P., Wang, Y. B., & Zhang, Y. M. (2019). Effect of marker segregation distortion on high density linkage map construction and QTL mapping in soybean (*Glycine max* L.). *Heredity*, *123*, 579–592.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Meng, Q.-L., Qiang, C.-G., Li, J.-L., Geng, M.-F., Ren, N.-N., Cai, Z., Wang, M.-X., Jiao, Z.-H., Zhang, F.-M., Song, X.-J., & Ge, S. (2024). Genetic architecture of ecological divergence between *Oryza rufipogon* and *Oryza nivara*. *Molecular Ecology*, *33*, e17268. https://doi.org/10.1111/mec.17268