JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12230

Research Article

Increased novel single nucleotide polymorphisms in weedy rice populations associated with the change of farming styles: Implications in adaptive mutation and evolution

Dong-Yuan Song[1†], Zhe Wang[1†], Zhuo-Jun Song[1], Cheng-Chuan Zhou[1], Peng-Hao Xu[1], Jie Yang[2], Ji Yang[1], and Bao-Rong Lu[1*]

[1]Ministry of Education, Key Laboratory for Biodiversity Science and Ecological Engineering, Department of Ecology and Evolutionary Biology, Fudan University, Shanghai 200438, China

[2]Institute of Food Crops, Jiangsu Academy of Agricultural Sciences/Nanjing Branch of Chinese National Center for Rice Improvement/Jiangsu High Quality Rice R&D Center, Nanjing 210014, China

[†]These authors contributed equally to this work.

*Author for correspondence. E-mail: brlu@fudan.edu.cn. Tel./Fax: 86-21-51630692.

Received 7 September 2016; Accepted 4 November 2016; Article first published online 4 January 2017

Abstract Substantial genetic variation is found in weedy rice (*Oryza sativa* f. *spontanea* Roshev.) populations from different rice-planting regions with the change of farming styles. To determine the association of such genetic variation with rice farming changes is critical for understanding the adaptive evolution of weedy rice. We studied weedy-rice specific novel single nucleotide polymorphisms (SNPs) by genome-wide comparison between DNA sequences of weedy and cultivated rice, in addition to polymerase chain reaction fingerprinting at 22 selected novel SNP loci in weedy rice populations. A great number of novel SNPs were identified across the weedy rice genome. High frequencies of the novel SNPs were determined at the 22 selected loci, although with considerable variation among weedy rice populations in different rice-planting regions. The highest frequency (~57%) of novel SNPs was identified in weedy rice populations from Jiangsu that experienced the most dramatic changes in rice farming styles, including the shift from transplanting to direct seeding, and from *indica* to *japonica* varieties. The lowest frequency (~29%) was detected in weedy rice populations from Northeast China, where rice farming has undergone relatively less change. The association between frequencies of novel SNPs in weedy rice populations and the extent of changes in rice farming styles suggests the critical role of adaptive mutation and accumulation of the mutation influenced by human activities in the rapid evolution of weedy rice.

Key words: adaptive evolution, change of farming practice, genome sequence, human influence, novel mutant, weed.

1 Introduction

Domestication of wild plants and animals is a critical milestone for human civilization (Ross-Ibarra et al., 2007). During this process many plant species with improved characteristics suitable for human uses have been selected and cultivated in the human managed environment – agro-ecosystems (Gepts, 2010). Domestication of crops reflects the rapid evolution of plant species adapting to natural and human-influenced environmental habitats (Gepts, 2010; Gross & Olsen, 2010). Simultaneously, particular types of weeds that are the same biological species of crops have also evolved in the agro-ecosystems. These weeds are referred to as conspecific weeds (Thurber et al., 2010; Xia et al., 2011a, 2011b; Lu et al., 2016). For example, weedy rice (*Oryza sativa* f. *spontanea* Roshev.), weedy barley (*Hordeum spontaneum* K. Koch), and weedy rye (*Secale cereale* L.) are conspecific weeds of cultivated rice (*O. sativa* L.), barley (*H. vulgare* L.), and rye (*S. cereale* L.), respectively (Molina-Cano et al., 1982; Cao et al., 2006; Burger & Ellstrand, 2014). Conspecific weeds share a close relationship with their cultivated and wild progenitors (see fig. 1 in Lu et al., 2016). Given such a unique role of genetic bridges to both their crop and wild progenitors, conspecific weeds can provide a useful system for studying the rapid adaptive evolution of plants occurring in human-influenced agro-ecosystems with different extents of environmental changes, including farming styles (Xia et al., 2011a, 2011b; Vigueira et al., 2013; Lu et al., 2016).

From the molecular viewpoint, evolution reflects the change of frequencies of alleles between populations of parents and their descendants over time (Wilson & Bossert, 1971; Li, 1997). Estimating the change of allelic frequencies can help to determine evolution of populations under a particular situation (Kirkpatrick, 1982; Li, 1997). Evolution is driven by four key factors: mutation, selection, genetic drift, and gene flow, affecting allele frequencies of given populations during evolution (Ellstrand, 2014). Usually, beneficial mutations, including both adaptive mutations (Rosenberg, 2001) and

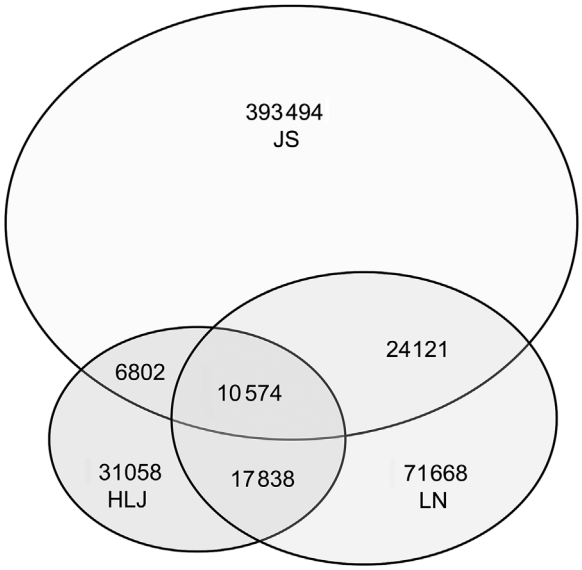

Fig. 1. Number of weedy rice-specific (novel) single nucleotide polymorphisms (SNPs) in three SNP databases of weedy rice populations from Jiangsu (JS), Liaoning (LN), and Heilongjiang (HLJ), China, across the entire genome. The analysis was based on a genome-wide comparison between the weedy rice SNP databases and the SNP-Seek Database (Alexandrov et al., 2015) developed from >3000 rice sequences.

non-adaptive mutations that pre-exist at the time organisms are exposed to a favorable environment (Tlsty et al., 1989), can increase the frequency of alleles associated with the mutations in a population under favorable selection. Single nucleotide polymorphisms (SNPs) can be used to detect the frequencies of these mutations. Therefore, it is possible to gain information about the driving forces of evolution through analyzing SNP frequencies of populations (Qiu et al., 2014). The availability of full genome sequencing data provides tremendous opportunities to identify SNPs through genome-wide comparison (Venter et al., 2001; Yu et al., 2002; Li et al., 2014). The technology has significantly increased the power for addressing evolutionary questions (Shastry, 2002; Konishi et al., 2006; Seeb et al., 2011).

Weedy rice is a typical conspecific weed infesting rice fields, causing substantial yield losses of cultivated rice (Delouche et al., 2007; Rao et al., 2007; Xia et al., 2011a). In some regions where no wild rice is distributed, weedy rice evolved from cultivated rice through dedomestication (Ishikawa et al., 2005; Cao et al., 2006; Reagon et al., 2010; Thurber et al., 2010; Xia et al., 2011b). Adaptive mutations and accumulation of beneficial mutants might be the major sources for the evolution of cultivar-originated weedy rice, apart from other sources such as gene flow (Xia et al., 2011a; He et al., 2014; Qiu et al., 2014; Lu et al., 2016). Thus, to detect novel SNPs and their variation patterns will help us to understand the role of mutations and the accumulation of mutants in the adaptive evolution of weedy rice populations (Qiu et al., 2014). Using a crop–weed–wild system (see fig. 1 in Lu et al., 2016) to study adaptive evolution of weedy rice involving SNPs has the following advantages. First, information on genome-wide DNA sequences is freely available in *Oryza*, providing great opportunities to fully elucidate variation of crop–weed–wild rice (Yu et al., 2002; Caicedo et al., 2007; McCouch et al., 2010; Li et al., 2014). Second, the SNP-Seek Database of Cultivated Rice (Alexandrov et al., 2015) based on genomic DNA sequences of >3000 rice cultivars (Li et al., 2014) enables the accurate identification of novel SNPs specific to weedy rice through comparison of the crop–weed SNP databases.

Previous studies have revealed significant genetic divergence of weedy rice populations within the same rice-planting regions (Cao et al., 2006; Jiang et al., 2012) and among different regions in China (Xia et al., 2011a; Liu et al., 2012b; Song et al., 2015). Identified genetic divergence of weedy rice is associated with the differentiation of *indica* and *japonica* cultivars grown in these regions (Liu et al., 2012b; Song et al., 2015), as well as the geographical distribution of the weedy rice populations (Xia et al., 2011a; Song et al., 2015). A relatively high level of genetic diversity was found in weedy rice populations from Jiangsu, compared with those in other regions such as in Guangdong and Northeast China (Song et al., 2015). The authors concluded, based on the genetic diversity pattern, that human influences have played an important role in shaping the pattern of weedy rice populations in these agro-ecosystems (Song et al., 2015). Here, we further hypothesize that the extent of changes in rice farming likely determines the allele frequencies associated by the mutations or/and selection in the particular environment. In fact, Jiangsu experienced the most dramatic change in rice farming by the shift from transplanting to direct seeding (Sun et al., 2014) and from *indica* to *japonica* varieties (Nai et al., 2012). Guangdong experienced relatively less change by the shift from transplanting to direct seeding of only *indica* varieties (Li et al., 2013). Northeast China represents the region with the least change by retaining transplanting of *japonica* varieties (Wang et al., 2009).

If the above hypothesis is true, we shall expect more weedy rice-specific SNPs (novel SNPs) in weedy rice populations in Jiangsu rice-planting regions, followed by those in Guangdong and Northeast China. Therefore, the objective of this study is to determine novel SNPs in weedy rice populations from these regions that experienced various changes in rice farming. In this study, we intend to address the following questions: (i) Are there any novel SNPs that are weedy rice-specific compared with the SNP-Seek Database of 3000 sequences from rice varieties (Alexandrov et al., 2015)? (ii) Do many novel SNPs present at particular loci in weedy rice populations collected from rice fields? and (iii) Does the frequency of novel SNPs show significant variation at all loci in weedy rice populations from different rice-planting regions? The answers to these questions will increase our understanding of human influences on the adaptive mutations and accumulation of pre-existing mutations, as well as the adaptive evolution of weedy rice populations in human managed habitats.

2 Material and Methods

2.1 Weedy rice populations used for identifying novel SNPs

For identifying novel SNPs that are specific to weedy rice, we collected a total of 90 weedy rice samples from rice fields in Jiangsu (JS), Liaoning (LN), and Heilongjiang (HLJ) provinces in China. For studying variation of novel SNPs in

different rice-planting regions in China, a total of 480 weedy rice plants included in 15 populations were collected from JS and Guangdong (GD) provinces and Northeast China (NE) (Table 1). More than 30 weedy rice plants (>10 m apart) were randomly sampled from each rice field. The spatial distances between the sampled weedy rice populations were >5 km. The mature panicles from each sampled weedy rice plant were collected in a paper bag as an independent sample. The weedy rice populations represented three rice-planting regions (JS, GD, and NE) with different extents of human influences/disturbances in terms of the styles of rice farming.

2.2 Identification of novel SNPs in weedy rice

To identify novel SNPs that were specific to weedy rice, we used total genomic DNA extracted from weedy rice samples collected in JS, LN, and HLJ, following the CTAB protocol of Murray & Thompson (1980). Resequencing of the nuclear genomes of weedy rice was based on three sets (each containing 30 individual plants) of pooled samples representing JS, LN, and HLJ regions. The obtained three weedy rice DNA sequences from resequencing were compared with the total genomic DNA sequence of a standard *japonica* rice variety Nipponbare (Os-Nipponbare-Reference-IRGSP-1.0, Kawahara et al., 2013) to generate three SNP datasets (our unpublished data) that were specific to weedy rice for further analyses. The SNP calling was completed by the software PoPoolation2, with the minimum coverage that was set as 10, the maximum coverage as 100, and minimum count of reads as 3 (Kofler et al., 2011). Using the previously published SNP-Seek Database of Cultivated Rice by Alexandrov et al. (2015; http://oryzasnp.org/), which contained approximately 29 million SNPs, we generated a new weedy rice SNP database by comparing the three obtained weedy rice SNP datasets (our unpublished data) with the SNP-Seek Database (Alexandrov et al., 2015). The new SNP database only retained weedy rice-specific SNPs that were not detected in the SNP-Seek Database. The new SNP database was considered to contain novel SNPs specific to weedy rice, relative to a total of >3000 cultivated rice varieties (including *indica* and *japonica*). This is because the cultivated rice SNP-Seek Database was developed from the genomic DNA sequences of 3024 rice cultivars (Li et al., 2014) by excluding the identical SNP loci of all included rice cultivars.

2.3 Primer design and validation

A total of 32 random-distributed novel SNP loci shared by the three obtained weedy rice DNA sequences were selected to design primer pairs that could be used to determine the occurrence of novel SNPs in weedy rice populations collected from rice fields. The primer pairs were designed following the protocols of Little (2001) and Liu et al. (2012a). The corresponding DNA sequences of *japonica* rice genome (Os-Nipponbare-Reference-IRGSP-1.0) were applied to verify the specificity of the 32 primer pairs using Primer-Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). All the forward primers were designed to be allele-specific to the cultivar nucleotide by the complementary base at the 3′-terminal. Additional deliberate mismatches were introduced at the penultimate or antepenultimate base of the forward primers to improve the amplification specificity (Little, 2001). The 32 primer pairs were validated by the polymerase chain reaction (PCR) method using the two standard rice varieties (*indica* 93-11 and *japonica* Nipponbare) as DNA templates. Eventually, only 22 primer pairs (Table 2) that amplified PCR products were retained for further analyses of novel SNPs in weedy rice populations. The retained 22 primer pairs were fluorescently labeled by FAM (blue), ROX (red), or JOE (green) on the reverse primers (Oetting et al., 1995; Schuelke, 2000). Such labeling could facilitate the detection of PCR products from multiple loci (mixed samples) when these loci were simultaneously included in electrophoreses (Oetting et al., 1995;

Table 1 Information on 15 weedy rice populations collected from Guangdong (GD1–GD3), Jiangsu (JS1–JS6) and Northeast China (NE1–NE6) used for polymerase chain reaction fingerprinting

Population code	No. of plants	Locality (city, province; all in China)	Latitude (N)	Longitude (E)
GD1	32	Leizhou, Guangdong	20°48′	110°11′
GD2	32	Leizhou, Guangdong	20°54′	110°08′
GD3	32	Leizhou, Guangdong	20°53′	110°04′
JS1	32	Taizhou, Jiangsu	32°33′	119°59′
JS2	32	Yangzhou, Jiangsu	32°28′	119°24′
JS3	32	Changzhou, Jiangsu	31°49′	120°02′
JS4	32	Lianyungang, Jiangsu	34°29′	119°08′
JS5	32	Yancheng, Jiangsu	33°48′	120°11′
JS6	32	Huaian, Jiangsu	33°32′	119°11′
NE1	32	Panjin, Liaoning	41°43′	122°06′
NE2	32	Kaiyuan, Liaoning	42°35′	123°57′
NE3	32	Dunhua, Jilin	43°18′	128°10′
NE4	32	Yanji, Jilin	42°50′	129°29′
NE5	32	Jiamusi, Heilongjiang	46°51′	130°29′
NE6	32	Haerbin, Heilongjiang	45°54′	126°49′

Table 2 DNA sequences of 22 primer pairs used to detect the presence or absence of novel single nucleotide polymorphisms in weedy rice

Primer pair (loci)	Forward primer (5′–3′)†	Reverse primer (5′-3′)	Position	Length of PCR fragments	Tm (°C)
L1	AAGGCAGACACGGTAGAAgA	AAGCAGAAGAACAAGACGCA	Chr-1	209	58
L2	CGATGGGTTTGATGGCTAtT	ATGGGATGGAGGGAGTAGTAGA	Chr-1	303	55
L3	TGCTCCTACAAAGGAGGGcA	CACATACTCCACCTGTCTCAAAA	Chr-2	338	54
L4	TTGGACCTTCAAGAGCACaA	CCTTCAGCCATCAAGATAAT	Chr-2	187	56
L5	GTCTAAGAAATAGTAGTAGTAGTCAcG	CCTGTATTAGACCTACGCAC	Chr-3	181	51
L6	CGTCCAAAATTACTGATGAAcT	GCATAAAGCGAAGACCTAAAC	Chr-4	205	59
L7	CCTATCTCACTCAACTCCTgTA	TTGAGGTGATGCAATACTTA	Chr-4	306	56
L8	CCTCATCCCCTGTGTTCTgG	TTGCCGAATCCAGGATAAGA	Chr-5	158	58
L9	TCAGTTCGCAAGTCTTTGCtC	GGAGTCTGCCCTGGCTGTTG	Chr-6	192	58
L10	CATTGTGATCATAGTTAAGAtC	ATCCCATTTATGTAGAGGCG	Chr-6	237	56
L11	GCGACACGGACAGTCGCaGT	TGCCTGTGCTTATTGTGCCTGAT	Chr-7	121	53
L12	GCTTTGTTAGTGCTACTAAAAaG	TCCCACAATCAACGATACCA	Chr-7	288	57
L13	TCAAAATGTTATGAATTGCcT	ATAGCTTGTGGCTCTAACCG	Chr-7	291	56
L14	TGGTTTGTGTGTGAGAGAGAcA	ATGCGTTAGAAACAAAGTGC	Chr-7	255	54
L15	AATGGCGACACCATTGTAgGC	ACGGTGGATTTGGGTGGTAT	Chr-9	209	56
L16	AGTGAAAAAGAAAAAGAAAAAgT	ACTTCCTACAATCGCTCCTAA	Chr-9	259	56
L17	TATGCCAAGCAACAGGAgAG	TAGAAACTTGAAAGCGAACCC	Chr-10	221	51
L18	CATGACCTCAGGTATGTCATgA	TTCAGCCATCTGTTGTAAACTC	Chr-10	208	57
L19	CCATATCTACAACCCAATgA	TTGATTTTGTGTAATTGGTTC	Chr-11	188	52
L20	AAAATGAAGACACTGCATTTGtT	GTAACCAGAAATGAACAAGGGAA	Chr-11	330	59
L21	CGCGTGTCACTGGCAGaT	AAGTAGCAAACCCCTCCAAAT	Chr-11	203	59
L22	CGAAAAAAAGAAAAGAAAACtG	CGCATCAACGAAGAATAACCTG	Chr-11	214	59

†Lowercase letters in DNA sequences of the forward primers indicate the additional mismatch bases.
Chr-, chromosome; Tm, annealing temperature.

Schuelke, 2000). All the primer pairs were designed using the software Primer Premier 5 (Lalitha, 2000).

2.4 Polymerase chain reaction amplification for confirmation of novel SNPs in weedy rice populations

To confirm novel SNPs that presented in weedy rice populations from different rice planting regions, we used total genomic DNA extracted from weedy rice samples of 15 populations in JS, GD, and NE, following the CTAB protocol of Murray & Thompson (1980). The PCR was carried out using the selected 22 primer pairs in a 2720 Thermal Cycler (Applied Biosystems, Foster, CA, USA), in a volume of 10 μL containing 5 μL Taq PCR Master Mix buffer (Sangon Biotech, Shanghai, China), 0.4 μmol/L forward primer, 0.4 μmol/L reverse primer, and 10 ng genomic DNA. The PCR procedures were set as follows: a denaturation period of 4 min at 94 °C, followed by 32 cycles of 30 s at 94 °C, 30 s at 51–59 °C, and 40 s at 72 °C, with a final extension at 72 °C for 7 min. All PCR products were separated and analyzed on a capillary electrophoresis genotyper (ABI 3130; Applied Biosystems).

2.5 Data score and analysis

The separated DNA fragments from electrophoreses were identified and scored using the software Genemapper version 4.1 (Applied Biosystems). The presence of the target DNA fragment at a specific locus (primer pair) indicated the same SNP as in cultivated rice, and was therefore scored as "0". In contrast, the absence of the target DNA fragment at a specific locus suggested a novel SNP that was specific to weedy rice, and was scored as "1". The generated "0, 1" data matrix was subjected to analyses for the allelic frequencies of novel SNPs in the weedy rice populations. Differences in the average frequencies of novel SNPs between rice-planting regions were compared based on one-way ANOVA, using the software SPSS 19.0 (IBM, New York, NY, USA).

3 Results

3.1 Novel SNPs in resequenced weedy rice

To estimate the distribution of novel SNPs in weedy rice samples representing Jiangsu, Liaoning, and Heilongjiang rice-planting regions, we generated a new SNP database by removing all SNPs that were identical to those in the >3000 rice cultivars after comparison with the Cultivated Rice SNP-Seek Database. Consequently, a great number of weedy rice-specific novel SNPs (>555 000) were identified from the DNA sequences of pooled weedy rice samples (Fig. 1). In addition, substantial variation in the number of novel SNPs was found through the comparison between DNA sequences representing weedy rice from different rice-planting regions (Fig. 1). Unexpectedly, a substantially greater number of novel SNPs were identified in the weedy rice DNA sequence representing Jiangsu (JS: 434 991) than those representing Liaoning (LN: 124 201) and Heilongjiang (HLJ: 66 272). Furthermore, the number of shared novel SNPs also showed considerable variation among the DNA sequences based on the three pooled weedy rice samples (Fig. 1). The considerable variation in numbers of novel SNPs observed

between weedy rice samples from Jiangsu and Liaoning/Heilongjiang (NE) might be associated with the differences in rice farming styles.

3.2 Variation of novel SNPs in weedy rice populations based on selected loci

To examine the variation of weedy rice-specific novel SNPs, we included the 22 selected novel SNP loci (primer pairs) to determine their frequencies and distribution in weedy populations from rice-planting regions (JS, GD, and NE) with different human influences on rice farming. The frequencies and distribution of weedy rice-specific novel SNPs were determined by identifying their presence or absence in particular loci through PCR.

In general, relatively high frequencies (up to 100%) of novel SNPs at the 22 selected loci were detected in the 15 weedy rice populations (JS1–6, GD1–3, and NE1–6) across the three rice-planting regions (see Table S1). However, great variation in the average frequencies of novel SNPs at the 22 loci was found among weedy rice populations representing the three rice-planting regions, based on one-way ANOVA ($P < 0.05$, Fig. 2). The highest average frequencies (~57%) of novel SNPs were detected in the six JS weedy rice populations, whereas the lowest average frequencies (~29%) were found in the six NE weedy rice populations (Fig. 2). An intermediated level of average frequencies (~41%) of novel SNPs were found in the GD weedy rice populations, although only three populations were included for analyses (Fig. 2). These results suggested a high level and great variation of novel SNPs in weedy rice populations from the three rice-planting regions. Such a variation pattern may also be associated with the differences in rice farming styles among regions.

In addition, some loci (e.g., L11 and L17) showed significantly different frequencies ($P < 0.05$) of novel SNPs in weedy rice populations among the three rice-planting regions (JS, GD, and NE) based on one-way ANOVA (see Figs. 3A, 3B; Table S1). However, other loci (e.g., L5 and L10) showed similar frequencies of novel SNPs in weedy rice populations from different regions (see Figs. 3C, 3D; Table S1). These results suggested that some weedy rice novel SNP loci were sensitively responding to the change of rice farming styles in different regions, but other novel SNP loci might not be sensitive to the change of rice farming styles in these regions.

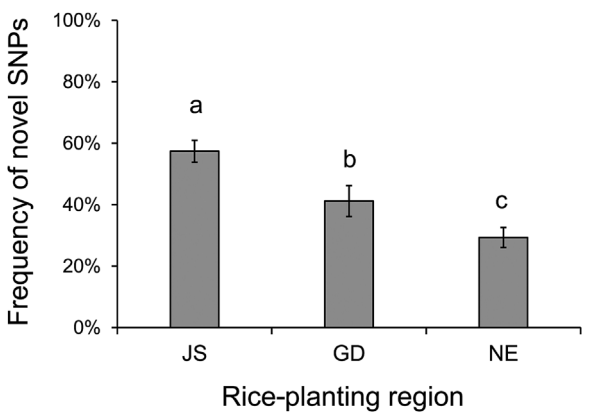

Fig. 2. Average frequencies of novel single nucleotide polymorphisms (SNPs) at the 22 loci in weedy rice populations from three rice-planting regions, Jiangsu (JS), Guangdong (GD), and Northeast China (NE), based on one-way ANOVA. Different letters above columns indicate significant difference ($P < 0.05$). Bars represent standard errors.

4 Discussion

Our results indicated the existence of a great number of weedy rice-specific novel SNPs at multiple loci in weedy rice samples/populations collected from different rice-planting regions, similar to Qiu et al. (2014) who also found many novel SNPs in weedy rice based on DNA sequence comparison using Nipponbare as the reference. This finding is based on two independent sets of experiments: the genome-wide comparison of DNA sequences and PCR fingerprinting at the selected loci. The genome-wide comparison was made between the SNP databases generated from the total genomic DNA sequences of weedy rice samples (our unpublished data) and >3000 rice cultivars (Alexandrov et al., 2015), where a tremendous number of novel SNPs were identified across the total genome of weedy rice samples. Polymerase chain reaction fingerprinting that included 22 primer pairs (loci) showed weedy rice-specific novel SNPs at all the selected loci in weedy rice populations from the different rice-planting regions. The presence of weedy rice-specific novel SNPs may suggest adaptive mutations being accumulated in weedy rice during the dedomestication and evolutionary processes. The observation is based on the fact that these novel SNPs have not been detected in the pool of cultivated rice in which >3000 rice cultivars were included for the Cultivated Rice SNP-Seek Database (Alexandrov et al., 2015). Alternatively, these novel SNPs may also be generated in the pre-existing mutations of cultivated rice. Positive selection on the pre-existing mutations during the dedomestication processes increased their frequencies to considerable extents under favorable environmental conditions. No matter which of the above pathways that increased the frequencies of novel SNPs, results from this study indicate the critical role of the adaptive mutations and accumulation of pre-existing mutations in the origin and evolution of weedy rice populations adapting to various rice-planting ecosystems. In addition, some of the novel SNPs potentially associated with beneficial mutations in weedy rice should be important for exploring useful genetic resources to improve cultivated rice in breeding (Lee et al., 2011; Qiu et al., 2014).

The explanations for novel SNPs present in weedy rice are based on the fact that only some (>3000) rice cultivars are included in the comparison for identifying weedy rice-specific SNPs. Nevertheless, our finding of a great number of weedy rice-specific novel SNPs indicates the important role of mutation and selection in the origin and adaptive evolution of weedy rice deviated from its cultivated progenitor. The explanation of our finding about the increased number of weedy rice-specific SNPs is supported by the previous study of Qiu et al. (2014) who also detected novel SNPs in weedy rice. The authors predicted, based on their findings, that the increased frequencies of novel SNPs in weedy rice are likely associated with functional genes that were responsible for the adaptation to the change in environment (e.g., genes

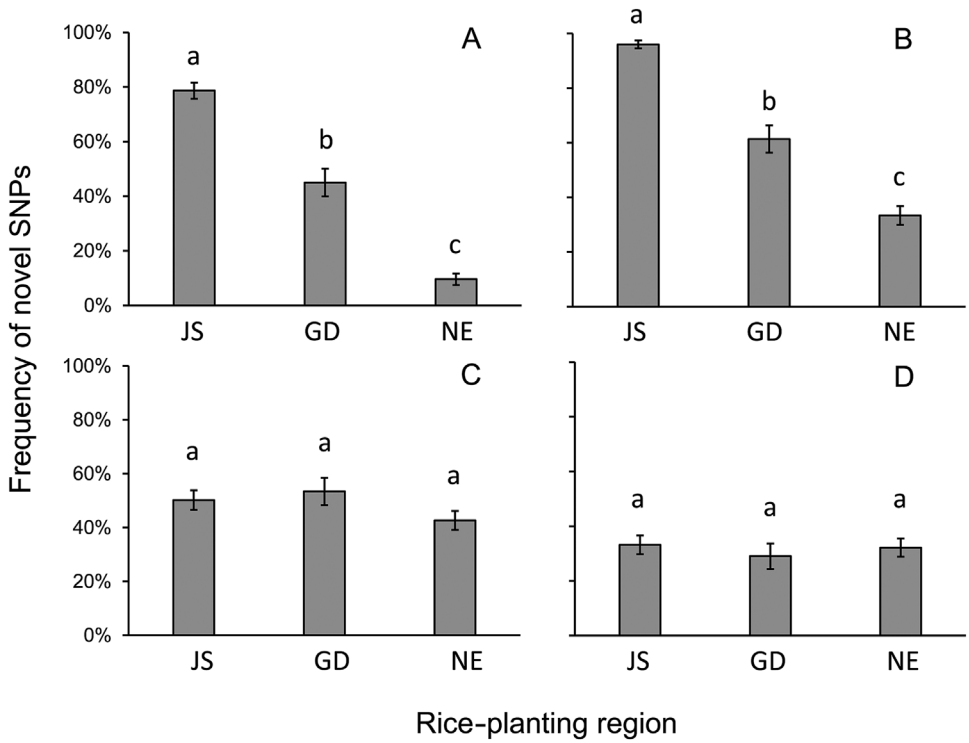

Fig. 3. Frequencies of novel single nucleotide polymorphisms (SNPs) in weedy rice populations from three rice-planting regions, Jiangsu (JS), Guangdong (GD), and Northeast China (NE), at four selected loci: **A,** L11; **B,** L17; **C,** L5; and **D,** L10. Different letters above columns indicate significant difference ($P < 0.05$) based on one-way ANOVA (analysis of variance). Bars represent standard errors.

tolerant to low-temperature and high-salt stresses, see Qiu et al., 2014 for details). Such beneficial genes from mutations have enabled weedy rice plants to adapt to the particular habitats of rice fields to compete with coexisting crop plants, cope with the changing environment (e.g., *japonica* cultivar brought diseases when introduced into Jiangsu), and escape from farmers' weeding (Delouche et al., 2007; Qiu et al., 2014). Therefore, results from the present study suggest that adaptive mutations, as indicated by the high frequencies of weedy-rice specific novel SNPs, and the accumulation of the pre-existing mutations that were inherited from rice cultivars through preferential selection have played an important role in the origin and evolution of weedy rice.

Noticeably, our analyses in this study further indicated that the average frequencies of weedy-rice specific novel SNPs varied significantly among weedy rice samples or populations collected from different rice-planting regions in China. The genome-wide comparison of DNA-sequence based SNP databases between weedy and cultivated rice indicated a substantially greater number of novel SNPs in weedy rice from Jiangsu than that from Northeast China. The PCR fingerprinting data also showed significantly greater average frequencies of novel SNPs in weedy rice populations from Jiangsu than those from Guangdong and Northeast China. Interestingly, the average frequency of novel SNPs in weedy rice populations is positively associated with the extent of change in rice farming styles in the three regions. In other words, the change of rice farming styles determined by human disturbance/activities will largely affected the amount of novel SNPs in weedy rice populations. For example, the highest average frequency of novel SNPs (~57%) was detected in weedy rice populations from Jiangsu province that experienced the most dramatic changes by transmission from cultivation of *indica* to *japonica* rice varieties, and from transplanting to direct seeding (Nai et al., 2012; Sun et al., 2014). In contrast, the lowest average frequency of novel SNPs (~29%) was detected in weedy rice populations from Northeast China where rice farming still maintains the traditional transplanting of *japonica* rice varieties (Wang et al., 2009). The intermediate average frequency of novel SNPs (~41%) was detected in weedy rice populations from Guangdong, where the shift from transplanting to direct seeding of *indica* rice varieties is the major features of rice farming (Li et al., 2013).

If the frequency of novel SNPs in weedy rice populations is determined by climatic factors such as temperature and photoperiod, a gradient variation pattern of novel SNPs associated with latitudes (from south to north or from north to south) should be observed. However, our observed results did not support such a gradient variation pattern of novel SNPs and, instead, supported the association of influences with the extent of farming style changes, although the role of transitional climatic factors in influencing the variation pattern of novel SNPs cannot be completely ruled out. Based on the above deduction, we further suggest that the adaptive mutations as indicated by the frequencies of novel SNPs of weedy rice are associated with the extent of change of farming styles—a type of environmental change, although selection can also influence the frequencies of novel SNPs. In

other words, the change of environment (e.g., farming styles) can induce adaptive mutations, promoting the adaptation of organisms to the changing environment. With increased knowledge of epigenetics, many studies have indicated the importance of environment-associated increases in mutations (Rosenberg, 2001). For example, an advantageous mutation of the functional gene lacking the ability to use lactose as energy was induced in *Escherichia coli* when *E. coli* populations were subjected to hunger stresses on cultural media containing lactose as energy sources. Consequently, the induced mutation enables the *E. coli* populations to utilize lactose as energy sources (Cairns et al., 1988). Similar results of environment change-induced mutations or adaptive mutations were found in many studies, including *E. coli* and other organisms (Tlsty et al., 1989; Bjedov et al., 2003; Roth et al., 2006; Jiang et al., 2014; Skinner et al., 2015). These findings suggest that widespread environment change-induced mutations are important for the origin and adaptive evolution of organisms.

Furthermore, we also found that some of the studied loci (78%) produced significantly different frequencies of novel SNPs in weedy rice populations among different regions. However, other loci (22%) showed similar frequencies of novel SNPs in weedy rice populations among regions. These results may indicate that the function of genes at the loci with significantly different frequencies of novel SNPs among regions is likely associated with adaptation to the change of environment. The accumulation of novel SNPs at these loci is most likely caused by selection under the change of environmental conditions, resulting in significant variation in the frequency of novel SNPs. This deduction is supported by the previous conclusion that alleles associated with the ability of ecological adaptation tend to have higher frequencies in weedy rice populations (Konishi et al., 2006; Thurber et al., 2010; Sun et al., 2013). However, the loci with similar frequencies of novel SNPs in the three rice-planting regions may indicate their important function for the growth and development of weedy rice plants, but may not be associated with the responses to change of environment.

In conclusion, we found a great number of weedy rice-specific novel SNPs at multiple loci in weedy rice samples/populations collected from different rice-planting regions in China, based on genome-wide comparison of DNA sequences and PCR fingerprinting at selected loci. Analyses further indicated significant variation in frequencies of the novel SNPs at the 22 selected loci among rice-planting regions. The highest average frequency (~57%) of the novel SNPs was identified in Jiangsu populations, followed by the populations in Guangdong (~41%) and Northeast China (~29%). Importantly, the frequencies of novel SNPs in weedy rice populations is closely associated with the extent of change in rice farming practices in different regions. The findings suggest the possible impact of human activity-associated environmental changes (e.g., farming styles) on mutations. Such an impact is illustrated by the identified variation in novel SNPs among the rice-planting regions that have experienced different changes in rice farming, although this conclusion needs confirmation by further studies. Alternatively, variation in frequencies of the novel SNPs among different regions may also be linked with natural or/and human selection on the genes containing the novel SNPs under various environmental conditions. No matter which of the explanations holds true, the human-influenced environmental change certainly plays a critical role in the origin and evolution of weedy rice, as well as in other weedy taxa occurring in agro-ecosystems with dramatic environmental changes, through adaptive mutations and selection.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant Nos. 31330014 and 31271683) and the National Program of Development of Transgenic New Species of China (Grant No. 2016ZX08011-006).

References

Alexandrov N, Tai S, Wang W, Mansueto L, Palis K, Fuentes RR, Ulat VJ, Chebotarov D, Zhang G, Li ZK, Mauleon R, Hamilton RS, McNally KL. 2015. SNP-Seek database of SNPs derived from 3000 rice genomes. *Nucleic Acids Research* 43: 1023–1027.

Bjedov I, Tenaillon O, Gerard B, Souza V, Denamur E, Radman M, Matic I. 2003. Stress-induced mutagenesis in bacteria. *Science* 300: 1404–1409.

Burger JC, Ellstrand NC. 2014. Rapid evolutionary divergence of an invasive weed from its crop ancestor and evidence for local diversification. *Journal of Systematics and Evolution* 52: 750–764.

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD. 2007. Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genetics* 3: 1745–1756.

Cairns J, Overbaugh J, Miller S. 1988. The origin of mutants. *Nature* 335: 142–145.

Cao QJ, Lu B-R, Xia H, Rong J, Sala F, Spada A, Grassi F. 2006. Genetic diversity and origin of weedy rice (*Oryza sativa* f. *spontanea*) populations found in north-eastern China revealed by simple sequence repeat markers. *Annals of Botany* 98: 1241–1252.

Delouche JC, Burgos NR, Gealy DR, Zorilla-San MG, Labrada R, Larinde M, Rosell C. 2007. Weedy rices: Origin, biology, ecology and control. Rome: FAO of the United Nations, Plant Production and Protection Paper, FAO Press.

Ellstrand NC. 2014. Is gene flow the most important evolutionary force in plants? *American Journal of Botany* 101: 737–753.

Gepts P. 2010. Crop domestication as a long-term selection experiment. *Plant Breeding Reviews* 24: 1–44.

Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends in Plant Science* 15: 529–537.

He ZX, Jiang XQ, Ratnasekera D, Grassi F, Perera U, Lu B-R. 2014. Seed-mediated gene flow promotes genetic diversity of weedy rice within populations: Implications for weed management. *PLoS ONE* 9: e112778.

Jiang C, Mithani A, Belfield EJ, Mott R, Hurst LD, Harberd NP. 2014. Environmentally responsive genome-wide accumulation of de novo *Arabidopsis thaliana* mutations and epimutations. *Genome Research* 24: 1821–1829.

Jiang Z, Xia H, Basso B, Lu B-R. 2012. Introgression from cultivated rice influences genetic differentiation of weedy rice populations at a local spatial scale. *Theoretical and Applied Genetics* 124: 309–322.

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu JZ, Zhou SG, Childs KL, Davidson RM, Lin HN, Vaillancourt B, Sakai H, Lee SS, Kim J, Numa H, Itoh T, Buell CR, Matsumoto T. 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6: 1–10.

Kirkpatrick M. 1982. Sexual selection and the evolution of female choice. *Evolution* 36: 1–12.

Kofler R, Pandey RV, Schlötterer C. 2011. PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27: 3435–3436.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M. 2006. An SNP caused loss of seed shattering during rice domestication. *Science* 312: 1392–1396.

Ishikawa R, Toki N, Imai K, Sato YI, Yamagishi H, Shimamoto Y, Sato T. 2005. Origin of weedy rice grown in Bhutan and the force of genetic diversity. *Genetic Resources and Crop Evolution* 52: 395–403.

Lalitha S. 2000. Primer premier 5. *Biotech Software & Internet Report* 1: 270–272.

Lee S, Jia Y, Jia M, Gealy DR, Olsen KM, Caicedo AL. 2011. Molecular evolution of the rice blast resistance gene *Pi-ta* in invasive weedy rice in the USA. *PLoS One* 6: e26260.

Li GJ, He ZX, Wang G, Ke BN, Luo L, Lu B-R. 2013. Occurrence, damage and control strategies of weedy rice in Leizou, Guangdong Province. *Weed Science* 31: 20–25. (in Chinese)

Li WH. 1997. *Molecular evolution*. Sunderland: Sinauer Associates.

Li ZK, Zhang G, McNally KL, Wang WS, Li J, Alexandrov NA, Liu X. 2014. The 3000 rice genome project. *GigaScience* 3: 7.

Little S. 2001. Amplification refractory mutation system (ARMS) analysis of point mutations. *Current Protocols in Human Genetics* 9: 1–12.

Liu J, Huang S, Sun M, Liu S, Liu Y, Wang W, Zhang XR, Hua W. 2012a. An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods* 8: 1–9.

Liu P, Cai XX, Lu B-R. 2012b. Single seeded InDel fingerprints in rice: An effective tool for *indica-japonica* rice classification and evolutionary studies. *Journal of Systematics and Evolution* 50: 1–11.

Lu B-R, Yang X, Ellstrand NC. 2016. Fitness correlates of crop transgene flow into weedy populations: A case study of weedy rice in China and other examples. *Evolutionary Applications* 9: 857–870.

McCouch SR, Zhao K, Wright M, Tung CW, Ebana K, Thomson M, McClung A. 2010. Development of genome-wide SNP assays for rice. *Breeding Science* 60: 524–535.

Molina-Cano JL, Gomez-Campo C, Conde J. 1982. *Hordeum spontaneum* C. Koch as a weed of barley fields in Morocco. *Journal of Plant Breeding* 88: 161–167.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321–4326.

Nai J, Zhang HC, Lu JF. 2012. Regional pattern changes of rice production in thirty years and its influencing factors in Jiangsu province. *Scientia Agricultura Sinica* 45: 3446–3452.

Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA, King RA. 1995. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* 30: 450–458.

Qiu J, Zhu J, Fu F, Ye CY, Wang W, Mao L, Qiang S, Liu TL, Fan LJ. 2014. Genome re-sequencing suggested a weedy rice origin from domesticated *indica-japonica* hybridization: A case study from southern China. *Planta* 240: 1353–1363.

Rao AN, Johnson DE, Sivaprasad B, Ladha JK, Mortimer AM. 2007. Weed management in direct-seeded rice. *Advances in Agronomy* 93: 153–255.

Reagon M, Thurber CS, Gross BL, Olsen KM, Jia Y, Caicedo AL. 2010. Genomic patterns of nucleotide diversity in divergent populations of US weedy rice. *BMC Evolutionary Biology* 10: 1–16.

Rosenberg SM. 2001. Evolving responsively: Adaptive mutation. *Nature Reviews Genetics* 2: 504–515.

Ross-Ibarra J, Morrell PL, Gaut BS. 2007. Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences USA* 104: 8641–8648.

Roth JR, Kugelberg E, Reams AB, Kofoid E, Andersson DI. 2006. Origin of mutations under selection: The adaptive mutation controversy. *Annual Review of Microbiology* 60: 477–501.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW. 2011. Single nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* 11: 1–8.

Shastry BS. 2002. SNP alleles in human disease and evolution. *Journal of Human Genetics* 47: 561–566.

Skinner MK, Guerrero-Bosagna C, Haque MM. 2015. Environmentally induced epigenetic transgenerational inheritance of sperm epimutations promote genetic mutations. *Epigenetics* 10: 762–771.

Song ZJ, Wang Z, Feng Y, Yao N, Yang J, Lu B-R. 2015. Genetic divergence of weedy rice populations associated with their geographic location and coexisting conspecific crop: Implications on adaptive evolution of agricultural weeds. *Journal of Systematics and Evolution* 53: 330–338.

Sun J, Qian Q, Ma DR, Xu ZJ, Liu D, Du HB, Chen WF. 2013. Introgression and selection shaping the genome and adaptive loci of weedy rice in northern China. *New Phytologist* 197: 290–299.

Sun TQ, Yang HJ, Li J, Deng JP. 2014. Retrospect and disadvantages analysis of Jiangsu direct-sowing rice and its countermeasures. *China Rice* 20: 5–9.

Thurber CS, Reagon M, Gross BL, Olsen KM, Jia Y, Caicedo AL. 2010. Molecular evolution of shattering loci in US weedy rice. *Molecular Ecology* 19: 3271–3284.

Tlsty TD, Margolin BH, Lum K. 1989. Differences in the rates of gene amplification in nontumorigenic and tumorigenic cell lines as measured by Luria-Delbrück fluctuation analysis. *Proceedings of the National Academy of Sciences USA* 86: 9441–9445.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Gocayne JD. 2001. The sequence of the human genome. *Science* 291: 1304–1351.

Vigueira CC, Olsen KM, Caicedo AL. 2013. The red queen in the corn: Agricultural weeds as models of rapid adaptive evolution. *Heredity* 110: 303–311.

Wang S, Wang W, Wang BL. 2009. Quality traits of main japonica rice cultivars in Northeastern China. *Journal of Huazhong Agricultural University* 28: 130–135. (in Chinese)

Wilson EO, Bossert WH. 1971. *A primer of population biology*. Sunderland: Sinauer Associates.

Xia HB, Wang W, Xia H, Zhao W, Lu B-R. 2011a. Conspecific crop-weed introgression influences evolution of weedy rice (*Oryza sativa* f. *spontanea*) across a geographical range. *PLoS One* 6: e16189.

Xia HB, Xia H, Ellstrand NC, Yang C, Lu B-R. 2011b. Rapid evolutionary divergence and ecotypic diversification of germination behavior in weedy rice populations. *New Phytologist* 191: 1119–1127.

Yu J, Hu S, Wang J, Wong G, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W,

Cong L, Geng J, Han Y, Li L, Hu G, Huang X, Li W, Liu Z, Liu J, Qi Q, Liu J, Li T, Lu X, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Ren X, Chen X, He J, Liu D, Tian C, Xia H, Bao Q, Li G, Gao H, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Long Z, Zhen C, Chen R, Hao B, Zheng W, Chen S, Guo W, Liu, Tao M, Wang J, Zhu L, Yuan L, Yang H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79–92.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.111/jse.12230/suppinfo:

Table S1. The average frequency (%) of novel single nucleotide polymorphisms at 22 loci (L1–L22) determined in 15 weedy rice populations from Jiangsu (JS1–JS6) and Guangdong (GD1–GD3) provinces, and Northeast China (NE1–NE6). Numbers in parentheses indicate standard errors.