Theor Appl Genet (2008) 116:835–843
DOI 10.1007/s00122-008-0717-3

ORIGINAL PAPER

A preliminary study on population genetic structure and phylogeography of the wild and cultivated *Zizania latifolia* (Poaceae) based on *Adh1a* sequences

Xin-wei Xu · Wei-dong Ke · Xiao-ping Yu · Jun Wen · Song Ge

Received: 29 October 2007 / Accepted: 12 January 2008 / Published online: 19 February 2008

Abstract Recent decades have witnessed growing interests in exploring the population genetics and phylogeography of crop plants and their wild relatives because of their important value as genetic resources. In this study, sequence variation of the nuclear *Adh1a* gene was used to investigate the genetic diversity and phylogeographic pattern of the wild and cultivated *Zizania latifolia* Turcz. Sequence data were obtained from 126 individuals representing 21 wild populations in China and 65 varieties of the cultivated *Zizania latifolia*. Low to medium level nucleotide diversity was found in the wild populations, with northeastern populations being the most variable. We detected significant population subdivision ($F_{ST} = 0.481$) but no significant phylogeogaphical structure, suggesting limited gene flow and dispersal among populations. The current pattern of genetic variation in the wild populations might be explained by a fragmentation of ancient populations due to habitat destruction and degradation during recent decades. The heterogeneous levels and spatial apportionment of genetic diversity among wild populations also suggested a history of gradual colonization of *Zizania latifolia* populations from the northeast to the south of China. Interestingly, all 65 varieties of the cultivated *Zizania latifolia* possessed a single identical genotype, implying a single domestication associated with very few initial individuals.

Communicated by J. Yu.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0717-3) contains supplementary material, which is available to authorized users.

X.-w. Xu · S. Ge (✉)
State Key Laboratory of Systematic and Evolutionary Botany,
Institute of Botany, Chinese Academy of Sciences,
Beijing 100093, China
e-mail: gesong@ibcas.ac.cn

X.-w. Xu
Freshwater Ecological Field Station of Liangzi Lake,
Wuhan University, Wuhan 430072, China

W.-d. Ke
National Garden of Aquatic Vegetable,
Wuhan Institute of Vegetable Sciences, Hubei 430065, China

X.-p. Yu
Plant Protection and Microbe Institute,
Zhejiang Academy of Agricultural Sciences,
Hangzhou 310021, China

J. Wen
Department of Botany, National Museum of Natural History,
Smithsonian Institute, MRC 166, Washington,
DC 20013-7012, USA

S. Ge
Graduate School, Chinese Academy of Sciences,
Beijing 100039, China

Introduction

Crop plants and their wild progenitors often form a wild-cultivated complex and constitute important genetic resources for plant breeding (Muller et al. 2006; Sang and Ge 2007). The genus *Zizania* belongs to the rice tribe (Poaceae: Oryzeae) and is an aquatic genus with a discontinuous distribution between eastern Asia and North America (Wen 1999; Wu et al. 2006). Of the four species in the genus, two are field crops with their seeds or young shoots harvested for food or vegetable, i.e., the annual *Zizania palustris* L. native to North America and the perennial *Zizania latifolia* native to Asia (Oelke 1993; Guo et al. 2007). In China, the grains of *Zizania latifolia*

have been used for offering tribute to the emperors and nobles during the Zhou Dynasty (from 771 to 221 B.C.) and later emerged as a vegetable, called Jiaobai, about 1,500 years ago (Sun 1905; Zhang 2006; Guo et al. 2007). It has also been considered a Chinese medicine since the Tang Dynasty (619–907) (Zhai et al. 2001). Because the young shoots of the plant become swollen, soft and edible after being infected by the fungus *Ustilago esculenta* and have high nutritional and economical value, *Zizania latifolia* has been cultivated widely in China, and at scattered locations in Japan and Korea (Guo et al. 2007). Particularly in China, over 100 local varieties of Jiaobai have been bred and this vegetable is now only second to the lotus (*Nelumbo nucifera* ssp. *nucifera*) among all 12 aquatic vegetables cultivated in China in terms of surface area of cultivation (Guo et al. 2007).

Wild populations of *Zizania latifolia* are widely distributed across China and are of importance as genetic resources for breeding and forage and because of its potential ecological functions such as purification of water (Zhai et al. 2000; Guo et al. 2007). To date, many studies have been conducted on *Zizania latifolia*, including its systematic position (Ge et al. 2002; Guo and Ge 2005), utilization as the tertiary gene pool of rice (Liu et al. 1999), nutritional value (Zhai et al. 2001), cultivar classification and breeding (Ke et al. 2000; Yu et al. 2005a, b; Guo et al. 2007) and ecology (Yang et al. 1999). No investigation has been conducted on the genetic diversity and population genetic structure of wild populations of this species. In addition, the origin of the cultivated *Zizania latifolia*, Jiaobai, has never been addressed and remains unknown.

Recent decades have witnessed the widespread use of DNA sequences in population genetic and phylogeographic studies in plant species (Schaal et al. 1998; Petit et al. 2005). In particular, in conjunction with advances in molecular techniques and analytical methods on haplotype data, nuclear gene regions provide enough intraspecific variation and have been successfully used in addressing various population and evolutionary questions on plant species including crops and their wild relatives (i.e., Olsen and Schaal 1999; Caicedo and Schaal 2004; Londo et al. 2006; Muller et al. 2006). In this study, we used sequence variation of the nuclear *Adh1a* gene to examine the genetic diversity and population structure of 21 *Z. latifolia* populations sampled throughout its distribution range in China. Our specific aims are to (1) assess and compare the level of genetic variation of the wild and cultivated *Zizania latifolia*; (2) examine the population structure and genealogical pattern of *Zizania latifolia* and the potential factors behind the patterns; and (3) explore the origin and domestication of the cultivated *Zizania latifolia*. Such information is indispensable for exploitation and utilization of wild resources and genetic improvement of the cultivars and is vitally important for developing conservation strategies for this economically important aquatic species.

Materials and methods

Plant materials

Twenty-one populations of *Zizania latifolia* were collected throughout the natural distribution of the species in China, from Heilongjiang province to Guangdong province (Fig. 1; Supplementary Table 1). Fifteen to thirty individuals per population were randomly sampled at an interval of at least 10 m to prevent collecting ramets from a single genet. Young and healthy leaves were harvested in the field and dried with silica gel for subsequent DNA extraction. Based on previous morphological and RAPD studies (Ke et al. 2000; Yu et al. 2005b), the cultivated *Zizania latifolia*, Jiaobai, can be classified into two main groups: the single-season crop that can be harvested once each year and the two-season crop that is harvested twice a year in the fall and summer (Guo et al. 2007). In this study, a total of 65 accessions of the cultivated Jiaobai were collected from 12 provinces covering the main cultivated areas in China, including 39 single-season varieties and 26 two-season varieties. A list of the cultivated accessions used in this study is shown in the supplementary materials (Supplementary Table 2).

DNA extraction, amplification and sequencing

Total genomic DNA was extracted from silica gel-dried leaves using the CTAB method (Doyle and Doyle 1987). To find suitable markers with sufficient information at the population level, we screened portions of 18 chloroplast, two mitochondrial and nine nuclear DNA regions by sequencing three to nine individuals from different populations (Supplementary Table 3). No polymorphism was found in the 20 chloroplast and mitochondrial DNA regions with the total length of 17,100 base pairs (bp). Of the nine nuclear loci, *Adh1a* showed the highest level of polymorphism (Supplementary Table 3) and evolved neutrally (see Sect. 'Results') and was thus chosen for further study. The *Adh1* gene encodes alcohol dehydrogenase I (alcohol NAD^+: oxidoreductase, EC 1.1.1.1), an important protein in the process of anaerobic metabolism. It is a single copy and located in the short arm of chromosome 11 in cultivated rice (Tarchini et al. 2000). Hass et al. (2003) found, however, that both *Adh1* and *Adh2* were duplicated in *Zizania palustris* and named them as *Adh1a*, *Adh1b*, *Adh2a*, and *Adh2b*, respectively. To ensure that the orthologous locus was used in this study, we designed a pair of *Adh1a*-specific primers (Adh1aF: 5′-TTACCATTTCGTTGGGACT-3′,

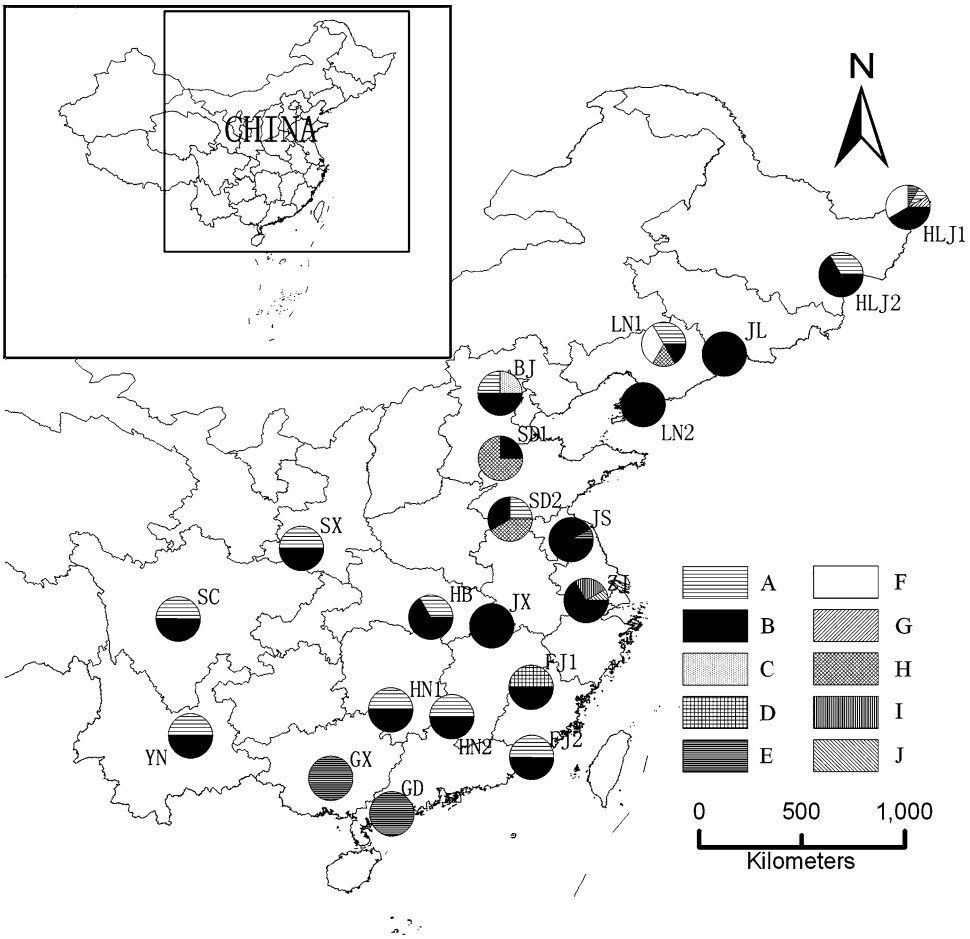

Fig. 1 Sampling localities of 21 *Zizania latifolia* populations and geographical distribution of ten haplotypes (A–J) of *Adh1a* sequenced

Adh1aR: 5′-GAACGCCGTGTTGATCTCT-3′) to amplify a region spanning from exon 4 to exon 10 of *Adh1a*. Polymerase chain reaction (PCR) was performed in a total volume of 25 µl, which contained 10–30 ng of template DNA, 0.2 µM of each primer, 200 µM of each dNTP, 10 mM Tris–Cl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.75 U ex*Taq* DNA polymerase (TaKaRa). Amplification was carried out in a T-personal thermocycler (Biometra, Germany) under the following cycling conditions: 95°C (1 min); 35 cycles of 95°C (45 s), 53°C (30 s), 72°C (1 min 30 s); and then 72°C (10 min).

Amplified products were purified from agarose gel with a Dingguo purification kit (Dingguo, Beijing, China). Purified PCR products were sequenced using the dideoxy chain termination method with the DYEnamic ET Terminator Kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA). DNA sequencing was performed by a MegaBACE 1000 automated sequencer (Amersham Pharmacia Biotech), following the manufacturer's instructions. In addition to the forward and reverse primers for amplification, an internal primer (Adh1R9: 5′-CGGGTGGGTCTTGAACTC-3′) was designed to sequence the PCR products. Individuals in *Zizania latifolia* can be either homozygous or heterozygous at a locus. For a heterozygote, 'double peaks' at a polymorphic site can be detected in the chromatogram when point mutation exists between two alleles. Alternatively, the chromatogram may be unable to read directly when length difference exists between two alleles due to insertions or deletions. In such cases, purified PCR products were cloned using pGEM T-easy vectors (Promega, Madison, WI, USA), and then two alleles were determined separately by sequencing multiple clones. Because 'singletons', i.e., the polymorphisms occurred in only one sequence and can represent either true sequence variation or *Taq* polymerase artifact, we confirmed all the singletons by repeating DNA extraction, PCR amplification, cloning, and sequencing (Zhang and Ge 2007). All sequences of different haplotypes were deposited into GenBank, with accession numbers EU252519–EU252528.

Data analysis

DNA sequences were aligned with CLUSTAL W (Thompson et al. 1994) and refined manually. The number of polymorphic sites (S), haplotype diversity (H_d), and genetic variation measured by average pairwise differences per base pair between sequences (π) (Nei and Li 1979) and Watterson's estimates (θ_W) from S (Watterson 1975), were

calculated using DNASP version 4.10 (Rozas et al. 2003). Tajima's *D* (Tajima 1989) and Fu and Li *D** (Fu and Li 1993) neutrality tests were used to determine whether *Adh1a* evolves in a neutral manner. The minimum number of recombination events was assessed using the algorithm of Hudson and Kaplan (1985) in the DNASP program.

A haplotype network was constructed by using TCS version 1.18 (Clement et al. 2000), which implements statistical parsimony to connect haplotypes constrained by 95% confidence intervals. A hierarchical analysis of population subdivision was performed using an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) as implemented in ARLEQUIN version 3.0 (Excoffier et al. 2005). Two measures of population differentiation G_{ST} and N_{ST} were compared by the program HAPLONST (Pons and Petit 1996). G_{ST} makes use of haplotype frequencies, while N_{ST} takes into account the differences between haplotypes. A higher N_{ST} than G_{ST} usually indicates the presence of phylogeographical structure with closely related haplotypes being found more often in the same area than the less closely related haplotypes (Pons and Petit 1996). Mantel test performed in ARLEQUIN was used to examine the correlation between geographical distance and Slatkin's measure M [$M = (1/F_{ST} - 1)/2$], a measure of the extent of gene flow in an island model at equilibrium (Slatkin 1993).

We also calculated a pairwise mismatch distribution to test for population expansion using DNASP. Populations at demographic equilibrium should present a multimodal mismatch distribution, whereas expanding populations are expected to be unimodal (Harpending 1994).

Results

Sequence variation and genetic diversity

A total of 252 sequences from wild *Zizania latifolia* and 130 sequences from the cultivated *Zizania latifolia*, Jiaobai, were obtained, with two sequences per individual/accession. The total length of the aligned sequences was 1,045 bp, including 515 bp in exon and 530 bp in intron. In total, 38 polymorphic sites were observed, including two indels (1 and 2 bp in length, respectively) and 36-point mutations. Of them, 25 polymorphic sites occurred in introns and 13 in exons including two that encode amino acid substitutions (Lys/Arg and Ser/Ala). Ten haplotypes were found in the wild populations and cultivated accessions. Two most common haplotypes were B (observed 133 times; 52.8%) and A (observed 55 times; 21.8%) and present in 19 and 12 populations, respectively. Haplotypes E (observed 26 times), H (16 times), and F (8 times) were detected in four, three and two populations, respectively. The remaining five haplotypes were each detected in a single population with haplotypes G and J found only once (Table 1). Haplotype diversity varied widely among populations, ranging from 0.0 (five populations: JL, LN2, JX, GD, and GX) to over 0.7 (three populations: HLJ1, LN1, and SD2). Nucleotide diversity presented a similar trend (Table 1). Remarkably, only a single genotype, consisting of haplotypes A and B, was found in all 65 varieties of Jiaobai, with the nucleotide diversity in Jiaobai being merely 9.3% of that found in wild populations (Table 1).

Tests of neutrality and haplotype network analysis

Neutrality tests were performed to determine whether *Adh1a* locus is subject to selection or has evolved neutrally. Neither Tajima's *D* ($D = 0.299$, $P > 0.1$) nor Fu and Li *D** ($D^* = 1.469$, $P > 0.05$) rejected the null hypothesis of neutral evolution. Estimated with four-gamete test, the minimum number of historical recombination events at the *Adh1a* locus was 0. Because *Adh1a* provided sufficient variation without recombination and evolved neutrally, this locus may reliably trace the population history of *Zizania latifolia* and thus is a suitable marker for phylogeographic study.

A haplotype network was constructed, which can be split into three distinct clades (Fig. 2). The first clade consists of three haplotypes (F, G, and J) and is mainly restricted in two populations (HLJ1 and LN1) in northeastern China. Haplotypes in this clade differ from the haplotypes in the other two clades by 23 mutations. The second clade includes haplotypes C, E, and H, which occur in the north, east, and south of China. The third clade includes four haplotypes (A, B, D, and I) and involves all the populations except for two (GD and GX) in southern China. These two clades differ from each other by eight mutations. The most common haplotypes (A and B) occur in this clade and associate with the cultivated Jiaobai (Fig. 2). The haplotype network, together with the haplotype distribution (Table 1), indicates that haplotypes B, C, and F are the ancestral haplotypes because they locate at internal nodes in the network. Although haplotypes A, E, and H have relatively high frequency, they are on the tips of the network and are thus considered as derived. Note that seven out of 21 populations (HLJ1, LN1, BJ, SD1, SD2, JS, and ZJ) consist of haplotypes from distinct lineages. Interestingly, exactly like the cultivated Jiaobai, six populations (SX, SC, HN1, HN2, FJ2, and YN) consist exclusively of a single heterozygote with haplotype (allele) composition of A and B. Although these populations are distributed widely in northwestern, southwestern, and southern China (Fig. 1), they most likely have originated from the escaped Jiaobai (see Sect. 'Discussion'). Therefore, they were not included in the following population structure analysis.

Table 1 Haplotype distribution and measures of diversity in wild and cultivated *Zizania latifolia* populations

Population	*n*	Haplotype										*H*	*S*	H_d	π	π_s	θ_w	θ_{ws}
		A	B	C	D	E	F	G	H	I	J							
1 HLJ1	12	1	5			1	4	1				5	32	0.758	0.0147	0.0234	0.0102	0.0164
2 HLJ2	12	4	8									2	1	0.485	0.0005	0	0.0003	0
3 JL	12		12									1	0	0	0	0	0	0
4 LN1	12	4	2				4		2			4	32	0.788	0.0148	0.0230	0.0102	0.0164
5 LN2	12		12									1	0	0	0	0	0	0
6 BJ	12	3	6	3								3	12	0.682	0.0047	0.0069	0.0038	0.0056
7 SD1	12		3						9			2	12	0.409	0.0047	0.0076	0.0038	0.0061
8 SD2	12	3	4						5			3	13	0.712	0.0065	0.0098	0.0041	0.0061
9 JS	12		11			1						2	11	0.167	0.0018	0.0028	0.0035	0.0056
10 ZJ	12		8							3	1	3	29	0.530	0.0049	0.0072	0.0092	0.0143
11 HB	12	4	8									2	1	0.485	0.0005	0	0.0003	0
12 JX	12		12									1	0	0	0	0	0	0
13 GD	12					12						1	0	0	0	0	0	0
14 GX	12					12						1	0	0	0	0	0	0
15 FJ1	12		6		6							2	1	0.545	0.0005	0.0008	0.0003	0.0005
16 FJ2	12	6	6									2	1	0.545	0.0005	0	0.0003	0
17 SX	12	6	6									2	1	0.545	0.0005	0	0.0003	0
18 SC	12	6	6									2	1	0.545	0.0005	0	0.0003	0
19 HN1	12	6	6									2	1	0.545	0.0005	0	0.0003	0
20 HN2	12	6	6									2	1	0.545	0.0005	0	0.0003	0
21 YN	12	6	6									2	1	0.545	0.0005	0	0.0003	0
Total	252	55	133	3	6	26	8	1	16	3	1	10	36	0.660	0.0054	0.0082	0.0057	0.0089
22 CUL	130	65	65									2	1	0.504	0.0005	0	0.0002	0

n number of sequences sampled, *H* number of haplotypes, *S* number of segregating sites, H_d haplotype diversity, π and π_s nucleotide diversity for the total and silent sites, respectively, θ_w and θ_{ws} Watterson's parameter for the total and silent sites, respectively, *CUL* cultivars

Population genetic structure

An AMOVA revealed significantly high amounts of variation both between and within populations at *Adh1a* locus. Although more than half nucleotide diversity (51.93%, $P < 0.001$) was attributable to variation within populations, a global F_{ST} value (0.481, $P < 0.001$) indicated a high level of genetic differentiation among the sampled populations. Pairwise comparisons of F_{ST} between populations suggested that three populations (GD, GX, and SD1) significantly differentiated from all other populations, and contributed most to population differentiation (Supplementary Table 4).

A test for phylogeographical structure of haplotype variation showed that the difference between N_{ST} (0.490) and G_{ST} (0.462) was not significant ($P > 0.05$), indicating that closely related haplotypes were not more likely to co-occur in the same area than less closely related haplotypes. A Mantel's test did not support the isolation by distance model across populations ($r = 0.154$, $P = 0.069$), suggest-

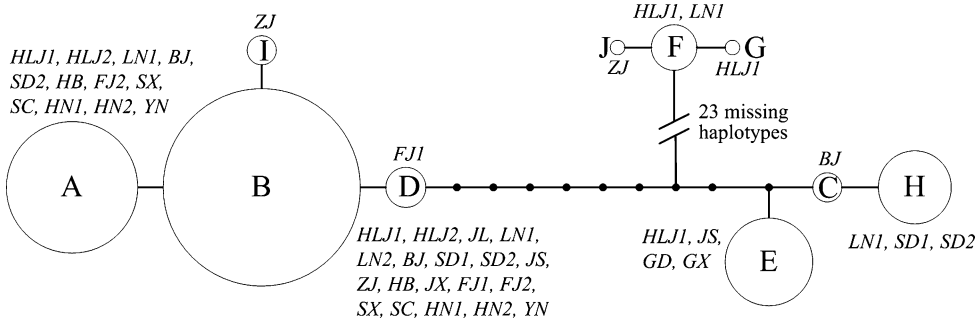

Fig. 2 Haplotype network of the nuclear *Adh1a* fragment. *Circles* represent different haplotypes (A–J) with size proportional to their relative frequency. *Black dots* represent inferred interior nodes that were absent in the samples. Population names are *italic* and indicated besides the haplotypes

ing that gene flow between populations cannot be predicted by geographical distance. When the six populations with identical genotype to Jiaobai were included in the analyses, almost the same results were obtained. To test for the hypothesis of population expansion in *Zizania latifolia*, we calculated the frequency distribution of pairwise nucleotide differences among individual haplotypes. The mismatch distribution for the species was clearly multimodal and inconsistent with the bell-shape curve expected for an expanding population (Fig. 3), suggesting that recent population expansion was unlikely in this species.

Discussion

Genetic diversity in the wild populations of *Zizania latifolia*

The *Adh* gene family has two or three loci in most plant groups and their levels and pattern of nucleotide variation have been extensively investigated across a variety of plant groups including grasses (e.g., Gaut and Clegg 1993; Lin et al. 2001; Chiang et al. 2003; Yoshida et al. 2004). To date, most studies have found a low level of nucleotide variation in *Adh* loci due to purifying selection although a wide range of variation was observed across different species (Chiang et al. 2003). Using sequences of multiple genes, Zhu et al. (2007) and Zhang and Ge (2007) have investigated the species-wide nucleotide diversity of wild species in *Oryza*, a close relative of *Zizania*. At *Adh1* locus, they observed that the silent nucleotide diversity (θ_{sil}) of four *Oryza* species ranged from 0.0042 in *O. officinalis* to 0.0118 in *O. rufipogon*. In this study, we detected a comparable level of nucleotide diversity in *Zizania latifolia* ($\theta_{sil} = 0.0089$). Given that *Adh1* is the most variable locus among the nine nuclear genes screened in this study (Supplementary Table 3), *Zizania latifolia* populations maintain a low to medium level of variation. The relatively low level of genetic diversity in this species is also supported by our preliminary survey of cytoplasmic DNA variation. Of the 18 chloroplast and two mitochondrial fragments screened, no single mutation was found across at least three distinct populations, even though the total length of the sequenced regions was up to 17,100 bp (Supplementary Table 3). Because different factors such as effective population size, mating system, and population dynamics may influence the level of genetic diversity, further investigations using additional molecular markers are needed to gain additional insights into the genetic diversity in this species.

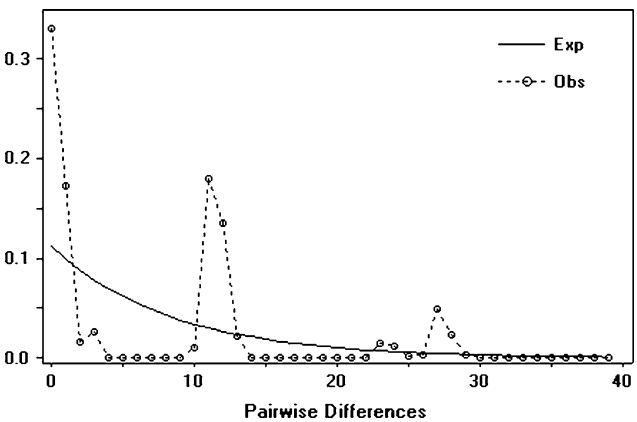

Fig. 3 Mismatch distribution for the wild *Zizania latifolia* populations showing the observed pairwise nucleotide site differences (*dotted line*) and the expected (*solid line*) obtained with DNASP

In previous studies on common wild rice (*Oryza rufipogon*), Gao et al. (2000) and Zhou et al. (2003) found that Guangxi and Guangdong provinces in southern China maintained higher genetic diversity than other regions. In this study, we observed that the most variable populations were HLJ1 and LN1 from northeastern China while populations from Guangxi (GX) and Guangdong (GD) possessed the lowest levels of nucleotide diversity (Table 1). The higher level of diversity in the northern populations than those from southern China may result from the origin and dispersal history of *Zizania latifolia* (see discussion below). It is noteworthy that each of the 11 populations that are distributed in almost entire distribution of the species in China comprises only a single genotype, either as a homozygote with one allele (JL, LN2, JX, GD, and GX) or as a heterozygote with two alleles (SX, SC, HN1, HN2, FJ2, and YN) (Table 1). This suggests that these populations are founded by single or a few individuals and might be attributed to the high proportion of asexual reproduction in *Zizania latifolia*, inconsistent with the observation that this species expands rapidly through clonal reproduction (Yang et al. 1999).

Population structure and evolutionary history

An important feature for *Zizania latifolia* is its high level of genetic differentiation among populations ($F_{ST} = 0.481$), indicating that almost half of total nucleotide diversity resides among populations. High genetic differentiation may ascribe to a low rate of gene flow among populations. Although *Zizania* species is wind-pollinated and may carry pollen to long distances, it has been shown that most effective pollination occurs at a local scale (Oelke 1993; Lu et al. 2005). Long distance dispersal of seeds is also not effective for *Zizania* species because of the easily shattering of the seeds after maturation (Oelke 1993; Guo et al. 2007). In *Zizania latifolia*, the movement of vegetative propagules such as rhizome fragments might be the important means for population expansion because the spreading rhizomes may float in the water and disperse with water current and wind. However, such a clonal strategy mainly occurs in the

same water body and is unlikely between populations spatially isolated on a large geographic scale. In a study on the American wild rice (*Zizania palustris* var. *palustris*) using 13 isozyme markers, Lu et al. (2005) revealed a low level of genetic diversity but high genetic differentiation ($F_{ST} = 0.30$) among 17 populations in northern Wisconsin in spite of a limited coverage of the samples. Limited gene flow among populations was considered to be the explanation for the high population differentiation and was consistent with the observed weak correlation between genetic and geographic distances in *Zizania palustris* var. *palustris* (Lu et al. 2005). Although this result is not strictly comparable to our case in *Zizania latifolia* because of the different molecular markers and sampling scales, both studies indicate limited potential for long-distance dispersal in *Zizania* species. The lack of a 'star-like' phylogeny of haplotypes, together with the multimodal mismatch distribution (Figs. 2 , 3) also suggested that population expansion did not occur in *Zizania latifolia*, which is consistent with the limited dispersal hypothesis.

An interesting finding is that no significant association was detected between genetic and geographical distances in *Zizania latifolia* populations, as determined by both the Mantel test and the result of geographical distribution of *Adh1a* haplotypes where the populations with geographical proximity did not share closely related haplotypes. For example, populations LN1 and LN2 have diverged significantly ($F_{ST} = 0.317$, $P < 0.01$; Supplementary Table 3) although they were closely related geographically with only 260 km distance between them. This pattern of variation is most likely to have resulted from incomplete lineage sorting due to a fragmentation of ancient populations. Although the Pleistocene glaciations have enormous effects on the species dynamics of plants in China (Axelrod et al. 1996) and cannot be ruled out as a factor, habitat destruction and degradation from decades of agriculture and urbanization throughout China may have had more profound impacts on the *Zizania latifolia* populations. Like many aquatic species, *Zizania* populations are subject to different ecological constraints and evolutionary processes than terrestrial plants and relied on specific habitats of water bodies (Barrett et al. 1993; Oelke 1993; Lu et al. 2005). It has been shown that the habitats suitable for aquatic species in China have been severely destroyed and fragmented because of water pollution and the degeneration and disruptions of hydrological regimes (Gao et al. 2000; Meng et al. 2004). For example, in the Dongping Lake in Shangdong province, *Zizania latifolia* changed from a dominant species in 1979 to local extinction in 1994 (Yu et al. 2005a). The drastic decrease in the number and sizes of the extant *Zizania latifolia* populations have led to alternation of population genetic structure and random loss of genetic diversity through drift. This explanation is further supported by the fact that seven populations (HLJ1, LN1, BJ, SD1, SD2, JS, and ZJ) comprise distinct haplotypes and that some populations (e.g., GD and GX) were fixed to a single allele or haplotype (Table 1).

This study provides additional insights into the evolutionary history of this species. Based on a multiple-gene phylogenetic study of the rice tribe, Guo and Ge (2005) suggested that *Zizania* species originated in New World and dispersed from North America to Asia because all closely related genera to *Zizania* were found in the Americas. It is most likely that *Zizania latifolia* arrived in eastern Asia from North America through the Bering Land Bridge that was suitable for exchanges for many temperate plants during the Miocene (Wen 1999), an argument consistent with the timeframe of diversification of rice tribe (Guo and Ge 2005). Subsequently, the species colonized southward and achieved its distribution across the entire eastern Asia. The southward colonization was supported by the fact that all three main clades of haplotypes occurred in northeastern China while two clades and one clade occurred in eastern and central/southern China, respectively (Fig. 2). This finding is in accordance with the spatial distribution of genetic diversity where most variable populations are in northeastern China (Table 1). It should be noted that although 18 cpDNA, 2 mtDNA, and 9 nuclear fragments have been screened, only *Adh1a* has been used in this study because of lack of information in all other fragments. Such a strategy may lead to bias in parameter estimation, particularly for genetic diversity. Further investigations with more extensive sampling and additional molecular markers would be necessary to elucidate the population dynamics and evolutionary history of this species.

Implications for the origin of cultivated *Zizania latifolia*

During the process of domestication, various degrees of reduction of genetic diversity were found in crops relative to their wild ancestors because of genetic bottleneck and artificial selection (Buckler et al. 2001; Wright and Gaut 2005). For example, maize (*Zea mays* ssp. *mays*) and cultivated sunflower (*Helianthus annuus*) contained ~80 and 50% of the genetic diversity retained in their wild progenitors, respectively (Wright and Gaut 2005; Liu and Burke 2006); whereas rice (*O. sativa*) possessed only 20–30% of the diversity in its wild progenitors (Zhu et al. 2007). In this study, it is remarkable that all varieties of the cultivated *Zizania latifolia* have an identical genotype consisting of two alleles or haplotypes (A and B) although the 65 varieties include almost all main cultivars from 12 provinces in China. This suggests a single domestication of the cultivated *Zizania latifolia*, probably arising from a single or a few closely related individuals. This conclusion is consistent with the observations that the cultivated *Zizania latifolia* is incapable of flowering due to the infection of the fungus *U.*

esculenta and vegetative propagation by perennial rhizomes is the only way for its survival (Guo et al. 2007).

According to historical records, the earliest cultivation of Jiaobai as a vegetable occurred in the Taihu Lake basin about 1,500 years ago (Zhang 2006). In this study, two populations from the Taihu Lake basin (JS and ZJ) contained only haplotype B while six populations to the north of the region possessed the A and B that existed in Jiaobai (A and B) (Table 1). Therefore, it is possible that the cultivated *Zizania latifolia* originated in the region north to the Taihu Lake. Alternatively, the Taihu Lake basin was the region associated with the origin of Jiaobai but the wild populations with haplotype A were not sampled in our study. Taken together, it is still premature to conclude on the origin of the cultivated *Zizania latifolia* before a phylogeographic study with additional sensitive molecular markers are used and a more extensive sampling is made.

A notable characteristic for the cultivated *Zizania latifolia* is that this vegetable can escape from cultivation into the wild (personal observations in the field). If mycelia of *U. esculenta* in the main stem fails to invade new tillers during the tillering period, which is sporadic, the plants without infection (called 'male' Jiaobai) can occur during cultivation (Ding et al. 1991). Consequently they can flower and show no obvious morphological difference from wild plants. Therefore, it is difficult to discriminate the escaped Jiaobai from true wild populations in many cases. This probably happened in the six populations (SX, SC, HN1, HN2, FJ2, and YN) because they have the identical genotype with Jiaobai and they grow near the area where this vegetable is cultivated. In addition, these populations are distributed in different river valleys and are spatially isolated by thousands of kilometers, and therefore are unlikely to have originated through natural dispersal. Consequently, we have reasons to argue that these populations were originally introduced by humans as crops and subsequently escaped into the wild. It is thus important to pay close attention to germplasm collections in population studies.

Acknowledgments We thank Dan Yu and Sheng-lin Xia for help in field survey, and Feng Li and Xu-song Zheng for help in collection of cultivated varieties. We are also grateful to Liang Tang, Xin-hui Zou and Li-li Zang for providing primers, to Jin-xiu Wang for discussion, and to two anonymous reviewers for helpful comments and suggestions. This study was supported by the National Basic Research Program of China (2007CB815704) and National Natural Science Foundation of China (30121003) to Song Ge and a grant of the Excellent Overseas Young Scientists Program to Jun Wen.

References

Axelrod DI, Al-Shehbaz I, Raven PH (1996) History of the modern flora of China. In: Zhang A, Wu S (eds) Floristic characteristics and diversity of east Asian plants. Springer, New York, pp 43–55

Barrett SCH, Eckert CG, Husband BC (1993) Evolutionary processes in aquatic plant populations. Aquat Bot 44:105–145

Buckler ES, Thornsberry JM, Kresovich S (2001) Molecular diversity, structure and domestication of grasses. Genet Res 77:213–218

Caicedo AL, Schaal BA (2004) Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. Mol Ecol 13:1871–1882

Chiang YC, Schaal BA, Chou CH, Huang S, Chiang TY (2003) Contrasting selection modes at the *Adh1* locus in outcrossing *Miscanthus sinensis* vs. inbreeding *Miscanthus condensatus* (Poaceae). Am J Bot 90:561–570

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659

Ding XY, Xu XS, Chen WP (1991) A preliminary study on the development of 'male' and smutted tillers of *Zizania latifolia*. J Wuhan Bot Res 9:115–120

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf material. Phytochem Bull 19:11–15

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0 An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133:693–709

Gao L, Ge S, Hong DY (2000) Allozyme variation and population genetic structure of common wild rice, *Oryza rufipogon* Griff. in China. Theor Appl Genet 101:494–502

Gaut BS, Clegg MT (1993) Molecular evolution of the *Adh* locus in the genus *Zea*. Proc Natl Acad Sci USA 90:5095–5099

Ge S, Li A, Lu BR, Zhang SZ, Hong DY (2002) A phylogeny of the rice tribe Oryzeae (Poaceae) based on *mat*K sequence data. Am J Bot 89:1967–1972

Guo YL, Ge S (2005) Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. Am J Bot 92:1548–1558

Guo HB, Li SM, Peng J, Ke WD (2007) *Zizania latifolia* Turcz. cultivated in China. Genet Resour Crop Evol 54:1211–1217

Harpending HC (1994) Signature of ancient population-growth in a low-resolution mitochondrial-DNA mismatch distribution. Hum Biol 66:591–600

Hass BL, Pires JC, Porter R, Phillips RL, Jackson SA (2003) Comparative genetics at the gene and chromosome levels between rice (*Oryza sativa*) and wildrice (*Zizania palustris*). Theor Appl Genet 107:773–782

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164

Ke WD, Zhou GL, Peng J, Huang XF, Liu YP, Li SM (2000) Hierarchical cluster analysis and comprehensive evaluation of *Zizania latifolia* germplasm. China Veg 6:18–21

Lin J, Brown AHD, Clegg MT (2001) Heterogeneous geographic patterns of nucleotide sequence diversity between two alcohol dehydrogenase genes in wild barley (*Hordeum vulgare* subspecies *spontaneum*). Proc Natl Acad Sci USA 98:531–536

Liu AZ, Burke JM (2006) Patterns of nucleotide diversity in wild and cultivated sunflower. Genetics 173:321–330

Liu B, Liu ZL, Li XW (1999) Production of a highly asymmetric somatic hybrid between rice and *Zizania latifolia* (Griseb): evidence for inter-genomic exchange. Theor Appl Genet 98:1099–1103

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci USA 103:9578–9583

Lu YQ, Waller DM, David P (2005) Genetic variability is correlated with population size and reproduction in American wild-rice (*Zizania palustris* var. *palustris*, Poaceae) populations. Am J Bot 92:990–997

Meng W, Su YB, Zheng BH (2004) Analysis of current situation of water pollution and its control strategy for Chinese river basins. J China Inst Water Resour Hydropower Res 2:242–246

Muller MH, Poncet C, Prosperi JM, Santoni S, Ronfort J (2006) Domestication history in the *Medicago sativa* species complex: inferences from nuclear sequence polymorphism. Mol Ecol 15:1589–1602

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Oelke EA (1993) Wild rice: domestication of a native North American genus. In: Janick J, Simon JE (eds) New crops. Wiley, New York, pp 235–243

Olsen KM, Schaal BA (1999) Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proc Natl Acad Sci USA 96:5586–5591

Petit RJ, Duminil J, Fineschi JS, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14:689–701

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144:1237–1245

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

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

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47:264–279

Sun YR (1905) Correct meaning of the book of rites of the Zhou Dynasty. China Press, Beijing

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

Tarchini R, Biddle P, Wineland R, Tingey S, Rafalski A (2000) The complete sequence of 340 kb of DNA around the rice *Adh1-Adh2* region reveals interrupted colinearity with maize chromosome 4. Plant Cell 12:381–391

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276

Wen J (1999) Evolution of Eastern Asian and Eastern North American disjunct distributions in flowering plants. Annu Rev Ecol Syst 30:421–455

Wright SI, Gaut BS (2005) Molecular population genetics and the search for adaptive evolution in plants. Mol Biol Evol 22:506–519

Wu ZY, Raven PH, Hong DY (2006) Flora of China. Poaceae. vol 22. Science Press/Missouri Botanical Garden, Beijing/St Louis, pp 186–187

Yang YF, Wang SZ, Lang HQ (1999) Study on the structures of clone population *Zizania latifolia* in wetland of the Songnen Plain of China. Acta Prat Sin 8:66–71

Yoshida K, Miyashita NT, Ishii T (2004) Nucleotide polymorphism in the *Adh1* locus region of the wild rice *Oryza rufipogon*. Theor Appl Genet 109:1406–1416

Yu SP, Sun GY, Dou SZ, Ma YM, Ge HM (2005a) The declining of hydrophyte of Dongping Lake and the influence of the south-to-north water transfer project. China Environ Sci 25:200–204

Yu XP, Zhang JF, Lu ZX, Chen JM, Zheng XS, Xu HX, Chen LZ (2005b) DNA polymorphism of 12 selected water oat varieties (*Zizania latifolia*) collected from Zhejiang. Bull Sci Technol 21:137–141

Zhai CK, Sun GJ, Lu ZM, Jiang ZK, Zhang XQ (2000) Chinese *Zizania* L. resource and its utilization value. Resour Sci 22:22–26

Zhai CK, Lu CM, Zhang XQ, Sun GJ, Lorenz KJ (2001) Comparative study on nutritional value of Chinese and North American wild rice. J Food Compost Anal 14:371–382

Zhang XF (2006) A general introduction of agricultural economy during the period of Six Dynasties. In: Hu AX (ed) Socioeconomic studies in the regions to the south of Yangtze River, volume of Six Dynasties and Sui and Tang Dynasty. China Agriculture Press, Beijing

Zhang LB, Ge S (2007) Multilocus analysis of nucleotide variation and speciation in *Oryza officinalis* and its close relatives. Mol Biol Evol 24:769–783

Zhou HF, Xie ZW, Ge S (2003) Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*Oryza rufipogon* Giff.) in China. Theor Appl Genet 107:332–339

Zhu QH, Zheng XM, Luo JC, Gaut BS, Ge S (2007) Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: Severe bottleneck during domestication of rice. Mol Biol Evol 24:875–888