



Inferring the Origin of Cultivated *Zizania latifolia*, an Aquatic Vegetable of a Plant-Fungus Complex in the Yangtze River Basin

Yao Zhao^{1,2}, Zhiping Song², Lan Zhong³, Qin Li^{1,2}, Jiakuan Chen² and Jun Rong^{1*}

¹ Jiangxi Province Key Laboratory of Watershed Ecosystem Change and Biodiversity, Center for Watershed Ecology, Institute of Life Science and School of Life Sciences, Nanchang University, Nanchang, China, ² The Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University, Shanghai, China, ³ Institute of Vegetable, Wuhan Academy of Agriculture Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:

Eric Von Wettberg,
University of Vermont,
United States

Reviewed by:

Michael Benjamin Kantar,
University of Hawaii,
United States
Zhe Cai,
Institute of Botany, China

*Correspondence:

Jun Rong
rong_jun@hotmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 24 June 2019

Accepted: 10 October 2019

Published: 08 November 2019

Citation:

Zhao Y, Song Z, Zhong L, Li Q,
Chen J and Rong J (2019) Inferring
the Origin of Cultivated *Zizania*
latifolia, an Aquatic Vegetable of
a Plant-Fungus Complex in the
Yangtze River Basin.
Front. Plant Sci. 10:1406.
doi: 10.3389/fpls.2019.01406

Crop domestication is one of the essential topics in evolutionary biology. Cultivated *Zizania latifolia*, domesticated as the special form of a plant-fungus (the host *Zizania latifolia* and the endophyte *Ustilago esculenta*) complex, is a popular aquatic vegetable endemic in East Asia. The rapid domestication of cultivated *Z. latifolia* can be traced in the historical literature but still need more evidence. This study focused on deciphering the genetic relationship between wild and cultivated *Z. latifolia*, as well as the corresponding parasitic *U. esculenta*. Twelve microsatellites markers were used to study the genetic variations of 32 wild populations and 135 landraces of *Z. latifolia*. Model simulations based on approximate Bayesian computation (ABC) were then performed to hierarchically infer the population history. We also analyzed the ITS sequences of the smut fungus *U. esculenta* to reveal its genetic structure. Our results indicated a significant genetic divergence between cultivated *Z. latifolia* and its wild ancestors. The wild *Z. latifolia* populations showed significant hierarchical genetic subdivisions, which may be attributed to the joint effect of isolation by distance and hydrological unconnectedness between watersheds. Cultivated *Z. latifolia* was supposedly domesticated once in the low reaches of the Yangtze River. The genetic structure of *U. esculenta* also indicated a single domestication event, and the genetic variations in this fungus might be associated with the diversification of cultivars. These findings provided molecular evidence in accordance with the historical literature that addressed the domestication of cultivated *Z. latifolia* involved adaptive evolution driven by artificial selection in both the plant and fungus.

Keywords: domestication, *Zizania latifolia*, *Ustilago esculenta*, genetic structure, approximate Bayesian computation

INTRODUCTION

Cultivated plants are fundamental materials for ensuring the survival and development of human beings (Hajjar and Hodgkin, 2007; Ronald, 2014; FAO, 2019). Through domestication, humans are able to change the morphological and physiological performances of wild plants. With ongoing selective pressure, wild plants would finally adapt to human-derived cultivation environments

and provide products that humans demand (Zeder et al., 2006; Gross and Olsen, 2010; Giovannoni, 2018). Plant domestication is regarded as a complex plant-animal co-evolution process, and it is an ideal model to study evolution (Purugganan and Fuller, 2009; Tufto, 2017; Whitehead et al., 2017; Zeder, 2017). As a typical scenario of recent rapid evolution, domestication has demonstrated how a species diverges from its ancestor through artificial selection (Zeder et al., 2006; Purugganan and Fuller, 2009; Diez et al., 2015; Wang et al., 2018).

For a cultivated plant, it is of great importance to decipher the puzzle of its domestication to understand the genetic basis for further improvements. Regardless of the domesticated species we study, there are several questions that need to be answered: what the ancestor is, when and where the domestication occurred, and what happened during the wild-to-cultivated transition (Doebley et al., 2006; Zeder, 2015). Benefitting from cumulative archaeological evidence and the advances of molecular biology, these questions appear to be well resolved for some major crops (Doebley et al., 2006). For example, archaeological and genetic evidence indicated that maize (*Zea mays*) had been domesticated once from its ancestor species teosinte (*Zea mexicana*) in Mexico (reviewed in van Heerwaarden et al., 2011), and common bean (*Phaseolus vulgaris*) had more than two independent domestication events (reviewed in Bitocchi et al., 2012). However, the origins of well-studied crops still may be under debate due to the uncertainty and complexity of historical human activities, such as Asian cultivated rice (*Oryza sativa*) (Ge and Sang, 2011; Huang et al., 2012; Choi and Purugganan, 2018) and a few grain crops generally thought to have originated from the Fertile Crescent (*Hordeum vulgare*, *Triticum aestivum*, and *Avena sativa*) (Fuller et al., 2012; Heun et al., 2012). Moreover, the answers to the domestication puzzles for a given species could also be blurred by post domestication migrations or genetic improvements (eg. cross-regional introduction, diversification of varieties, and hybridizations with wild progenitors) (Burger et al., 2008; Diez et al., 2015). Therefore, cultivated plants may be domesticated in their own way, which should be dissected independently.

The genus *Zizania* belongs to the rice tribe (Poaceae, Oryzaceae), and it 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 harvested for food (the annual *Zizania palustris* native to North America) or enlarged young shoots as vegetable (the perennial *Zizania latifolia* native to Asia) (Oelke, 1993; Xu et al., 2010). Because the young shoots of *Z. latifolia* become swollen, soft, and edible after being infected by the obligatory smut fungus *Ustilago esculenta* and have high nutritional and economic value, it had been domesticated as a vegetable called Jiaobai which is now widely cultivated in the Yangtze River Basin (Guo et al., 2007; Ke, 2008). The cultivated *Z. latifolia* is highly domesticated. The phenotypic differences between wild and cultivated *Z. latifolia* are distinct: the domesticated plant usually has erected plant architecture, broader leaves and shortened internodes; and it could generate much larger swollen shoots containing few black spots (Guo et al., 2007). Once *Z. latifolia* is infected by *U. esculenta*, it loses the ability to undergo sexual reproduction *via* seeds because the fungus hinders the development of the flower primordium.

Thus, the breeding of new accessions can only rely on the natural somatic mutants in tillers or rhizomes. In each generation, individuals with good-quality products (swollen shoots) should be screened out for clonal propagation to ensure the stability of agronomic traits of offspring (Guo et al., 2007). The plants that generate flowers or those with too many black spots in the swollen shoots are removed during harvest. Although precautions have been made, unfavorable mutants always occur in the fields, which may be attributed to the high variability of the smut fungus (Guo et al., 2007). To date, there are two main ecotypes including more than 100 local varieties with distinct phenotypic differences. One ecotype is a single-season plant that only can be harvested once each year in the fall (from August to November). This ecotype is a strictly short-day plant that can produce swollen shoots only when the days become short in fall. The other is a double-season crop that is planted in the spring and then can be harvested twice in the fall and the summer thereafter, and it is insensitive to light and endemic to the Yangtze River Basin. The single-season ecotype is taller and has fewer underground rhizomes, lower yield, and a wider distribution relative to the double-season ecotype.

Historical records can help to trace the domestication process of *Z. latifolia* in China (Ke, 2008). The oldest record showed that grains of *Z. latifolia* had been harvested to offer as tribute to the nobles in the Zhou Dynasty (from 771 to 221 BC). Until the Tang Dynasty (618–907 AD), *Z. latifolia* was famous for the tender taste of its grains. During the period between the Zhou and Tang Dynasties, the swollen shoots were rarely reported to be used for food. Only in the low reaches of the Yangtze River was it occasionally eaten by local residents. After the Tang Dynasty, people no longer harvested the grains of *Z. latifolia*. Alternatively, *Z. latifolia* was domesticated to be an aquatic vegetable (Guo et al., 2007; Chen et al., 2012; Chen et al., 2017). Although the utilization history of *Z. latifolia* has provided important clues, several blanks still remain. It was possible that *Z. latifolia* underwent pre-domestication during a long-time use as a food. In the meantime, *U. esculenta* acted as a plant disease during grain production. Then, the plant-fungus complex was domesticated. In another scenario, this vegetable was directly domesticated once or twice after the Tang Dynasty and then generated two main ecotypes with many cultivated varieties. However, these scenarios are just assumptions, and they still lack solid molecular evidence.

Cultivated *Z. latifolia* is the one and only crop that was domesticated in the form of a plant-fungus complex. Undoubtedly, its domestication was associated with the responses of the plant and the fungus to artificial selection. Thus, the cultivated *Z. latifolia* could not only provide a unique case of domestication of a plant-fungus complex but also an ideal model to investigate changes in the genetic interactions of two closely linked species under artificial selective pressures. The plant and the fungus have been independently studied. Wild populations of *Z. latifolia* are widely distributed across eastern China. This plant is of importance as a genetic resource for breeding and forage (Chen et al., 2006). Many studies have been conducted on *Z. latifolia*, including investigation of its phylogeny (Ge et al., 2002; Guo and Ge, 2005; Xu et al., 2010), ecology (Yang et al., 1999), population structure (Xu et al., 2008; Chen et al., 2012), utilization as a tertiary gene pool of rice (Liu et al., 1999), nutritional value (Zhai et al., 2001), cultivar

classification, and breeding (Guo et al., 2007). The phylogeny and cytology of *U. esculenta* have been investigated as well (Stoll et al., 2005; Zhang et al., 2012). The genetic interactions between plant and fungus were recently revealed by whole-genome sequencings and comparative transcriptome analyses (Guo et al., 2015; Wang et al., 2017; Ye et al., 2017; Zhang et al., 2017). However, the origin of cultivated *Z. latifolia* has seldom been genetically addressed and remains unknown.

This study focused on dissecting the genetic structure of *Z. latifolia* and the parasitic *U. esculenta*. Twelve microsatellites markers were used to investigate the genetic variations in *Z. latifolia* populations, and model simulations with approximate Bayesian computation (ABC) were then performed to infer the domestication history. The pattern of genetic variations for *U. esculenta* was identified based on ITS sequence analyses. By combining the genetic structures of host and endophyte, the domestication scenario of cultivated *Z. latifolia* could be unveiled.

MATERIALS AND METHODS

Plant Collection

Thirty-two natural populations of *Z. latifolia* were collected across the species distribution range in China (Figure 1, Table 1). In each population, the individuals were randomly sampled with an interval

of 10 M between plants to reduce the chance of collecting samples from the same genotype. Fresh young leaves were collected from each plant and placed into plastic bags containing silica gel for dehydration. A total of 824 samples of wild *Z. latifolia* were finally obtained. The leaf samples of 135 landraces of cultivated *Z. Latifolia* (the detail information of landraces were attached in **Supplementary Table 2**) were provided by the National Aquatic Vegetable Germplasm Nursery of Chinese Academy of Agriculture Science, Wuhan.

Infected Plant Collection and Fungi Isolation

Among the 32 wild populations, we found five individuals infected by the smut fungus *U. esculenta*. A total of 37 strains of *U. esculenta* were isolated from 5 wild accessions and 32 cultivars following the method described by Zhang et al. (2012). These strains were grown on potato dextrose agar (PDA) slant at 26°C and stored in a 10°C incubator.

DNA Extraction and PCR Assays

Total genomic DNA was extracted following the protocol described by Song et al. (2006). Twelve microsatellites were selected from 100 Asian cultivated rice microsatellites (www.gramene.org) and 16 *Z. latifolia* specific microsatellites (Quan et al., 2009) (Table S1). The PCR products were labeled with fluorescent dyes using 5'-tagged forward primers (FAM, JOE,

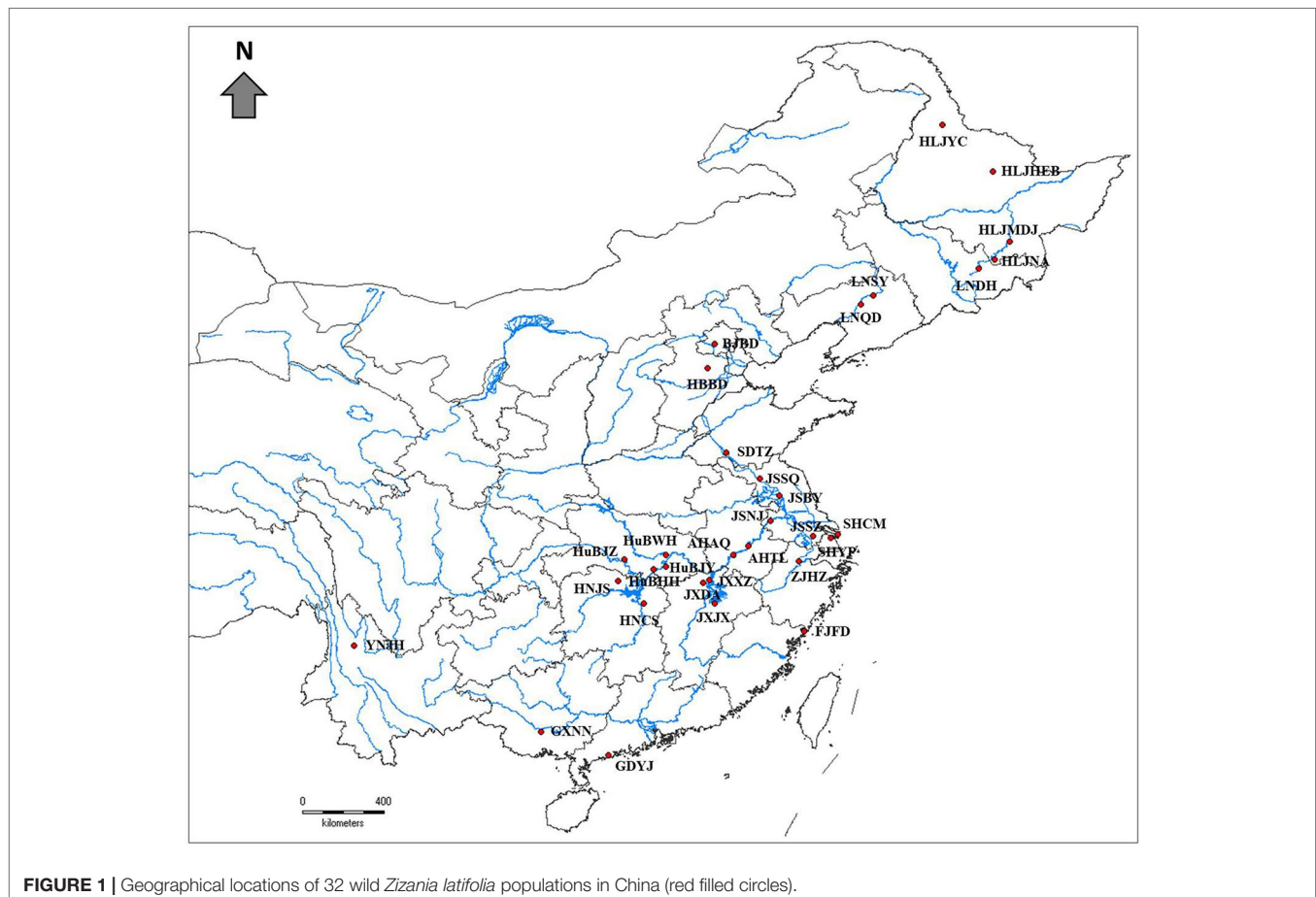


FIGURE 1 | Geographical locations of 32 wild *Zizania latifolia* populations in China (red filled circles).

TABLE 1 | Parameters of genetic diversity of *Zizania latifolia* populations and cultivars based on 12 microsatellites.

Code	Population	<i>N</i>	<i>A_r</i>	<i>A_e</i>	<i>H_o</i>	<i>H_e</i>	<i>F</i>	<i>N_e</i>	<i>N_i</i>	<i>Q</i> (> 0.05)	Bottleneck	Habitat
1	HLJYC	30	2.909	1.979	0.383	0.383	0.022	226				Pond
2	HLJHEB	6	3.182	2.402	0.273	0.563	0.450	152		0.054	S	River
3	HLJMDJ	26	2.727	1.647	0.270	0.310	0.188	210				Pond
4	HLJNA	30	3.273	2.078	0.415	0.457	0.069	191			Y	Lake
5	JLDH	42	3.273	1.965	0.346	0.464	0.275	305				River
6	LNSY	30	4.545	2.895	0.483	0.585	0.111	233				River
7	LNQD	28	4.818	2.653	0.478	0.539	0.051	276				Stream
8	BJHD	40	2.000	1.407	0.243	0.229	-0.049	256				Stream
9	HBBD	12	3.091	2.177	0.269	0.494	0.407	171				Marsh
10	SDTZ	34	3.727	2.140	0.395	0.472	0.108	254				Marsh
11	JSBY	39	5.727	2.893	0.567	0.588	-0.004	312				Lake
12	JSSQ	25	4.000	2.059	0.449	0.425	-0.068	255				River
13	JSNJ	14	2.818	2.116	0.483	0.487	0.013	169			S	Pond
14	JSSZ	30	4.273	1.972	0.524	0.441	-0.102	251				Lake
15	SHYP	18	1.636	1.544	0.533	0.284	-0.816	172			Y, S	Stream
16	SHCM	12	2.818	2.538	0.618	0.570	-0.114	150	4	0.330	Y, S	Stream
17	ZJHZ	27	3.455	1.973	0.461	0.438	0.001	192	3	0.114		Marsh
18	AHTL	14	2.091	1.705	0.571	0.371	-0.468	208			Y, S	Pond
19	AHAQ	33	4.545	2.493	0.498	0.546	0.058	280				Marsh
20	JXXZ	15	4.364	2.660	0.467	0.511	-0.004	187				Lake
21	JXDA	29	1.818	1.302	0.250	0.167	-0.170	210				Stream
22	JXJX	39	3.364	2.264	0.461	0.465	-0.039	304			Y	Stream
23	HuBWH	31	3.545	2.240	0.538	0.505	-0.098	278			Y	Lake
24	HuBJY	27	3.364	2.397	0.477	0.439	-0.133	231				Stream
25	HuBHH	32	4.727	2.545	0.542	0.546	-0.038	297				Marsh
26	HuBJZ	33	4.455	2.632	0.494	0.526	0.030	278				Stream
27	HuNJS	38	3.000	1.994	0.589	0.459	-0.222	273				Pond
28	HuNCS	35	1.455	1.450	0.434	0.229	-0.918	120	35	0.994	Y, S	Pond
29	FJFD	14	1.818	1.676	0.539	0.304	-0.868	112	14	0.987	Y, S	Stream
30	GDYJ	15	1.636	1.636	0.636	0.329	-1.000	150			Y, S	Stream
31	GXNN	15	1.636	1.636	0.636	0.329	-1.000	158			Y, S	Stream
32	YNJH	11	2.182	2.013	0.355	0.374	0.023	137	5	0.471	Y, S	Lake
Overall			3.196	2.096	0.459	0.432	-0.135					
SD			1.118	0.434	0.113	0.112	0.383					
	Cultivars	135	2.182	1.519	0.455	0.242	-0.751	100	—	0.997	Y	

N, sample size; *A_r*, allelic richness; *A_e*, number of effective alleles; *H_o*, observed heterozygosity; *H_e*, expected heterozygosity; *F*, fixation index; *N_e*, the effective population size; *N_i*, the number of individuals that genetically identical to cultivars; *Q*, the introgression coefficient; the significant bottleneck effect is labeled by Y (TPM model) and S (Mode shift). The *F* values that significantly deviated from 0 are shown in bold font.

and ROX) and GS350 as an internal size standard labeled with TAMRA. The PCR products were sequenced on an ABI 3730 (Applied Biosystems) automated sequencer. The resulting chromatograms were visualized and analyzed using GENEMAPPER v4.0 software (Applied Biosystems).

The ITS region of smut fungus was amplified using ITS1 and ITS4. PCR products were purified and sequenced on an ABI 3130XL capillary sequencer (Applied Biosystems). All ITS sequences were deposited into GeneBank, with accession numbers MK811211-MK811247.

Genetic Data Analyses

The frequencies of putative null alleles were estimated using the software FREENA (Chapuis and Estoup, 2007). Null alleles were detected in 34 of 384 population-locus combinations, and the presence of null alleles was not associated with particular populations or loci. We did not find a significant bias in the estimations of *F_{st}* based on the data of all loci compared to the estimate based on corrected genotype data for null alleles (*t*-test, *P* = 0.419). The parameters of population genetic

variations were estimated by expected heterozygosity (*H_e*), observed heterozygosity (*H_o*) and Wright's fixation index *F* using GENALEX 6.5 (Peakall and Smouse, 2012). We also calculated the mean effective allele number (*A_e*) and allelic richness (*A_r*). Hardy-Weinberg equilibrium departures were tested using exact tests implemented in GENEPOP v4.0 for each population-locus combination. A global test across loci for departures from Hardy-Weinberg equilibrium was conducted using Fisher's method. Significant deviations were further evaluated using the sequential Bonferroni test.

BOTTLENECK (Piry et al., 1999) was used to screen each population for bottleneck signatures over the last two to four *N_e* generations or in a recent period. A significant heterozygote excess or deficit due to a bottleneck was assessed by a two-tailed Wilcoxon signed rank test with 1,000 iterations under the assumption of a stepwise mutation model. At the same time, the distribution of allelic frequencies was drawn to identify whether there had been a mode shift away from an L-shaped distribution, indicating recent population bottlenecks.

MIGRATE (Beerli, 2008) was used to estimate the effective population size *N_e*. MIGRATE uses coalescent theory to jointly

estimate the mutation scaled population size θ ($4N_e\mu$) over a long period of time ($\sim 4N_e$ generations). Three runs were conducted using MIGRATE. First, two shorter runs (10 short chains of 10,000 sampled, 500 record and three final chains of 100,000, 5,000 recorded) were performed and used to verify that the MCMC estimated the parameters correctly. Then, a final run (10 short chains of 10,000 sampled, 500 recorded and three final chains of 500,000 sampled and 25,000 recorded) was performed, and θ values from this final run are reported. The initial run used an estimate of F_{st} as a starting parameter to calculate θ . Each subsequent run used the ML estimates from the previous run as new starting parameters.

Principal coordinate analysis (PCoA), as implemented in GENALEX 6.5, was conducted to investigate the genetic divergence between cultivated and wild *Z. latifolia* plants. The genetic relationship between the cultivated and wild plants was also estimated using STRUCTURE 2.3.4 (Pritchard et al., 2000) with $K = 2$ and following the method described by Harter et al. (2004). If the proportion of an individual's genome with ancestry (the introgression coefficient Q) from the cultivated group was $>50\%$, we defined that finding as a sign of cultivar fertility.

The degree of population divergence was quantified using F_{st} with GENEPOP 4.0. A PCoA was performed with the wild populations only to investigate the pattern of genetic divergence within this species. The genetic structure was further explored using the Bayesian clustering algorithm implemented in STRUCTURE 2.3.4. The program was given no prior information on ancestral populations and run 10 times for each value of K ancestral populations, under the admixture model with correlated allele frequencies, using 200,000 Markov chain Monte Carlo iterations and a burn-in of 100,000 iterations. We inferred K using the *ad hoc* statistic ΔK (Evanno et al., 2005). The resulting matrices of estimated cluster membership coefficients were permuted with CLUMPP (Jakobsson and Rosenberg, 2007). The final matrix for each K value was visualized with DISTRUCT (Rosenberg, 2004).

The genetic distance (GD) matrix for all pairs of individuals and populations of wild *Z. latifolia* was estimated by using MSA 3.0 (Dieringer and Schlotterer, 2003). Unrooted neighbor-joining trees were constructed using the PHYLIP package 3.6 with 1,000

bootstraps. The relatedness between individuals and populations was visualized with the software Dendroscope 3.

Isolation by distance (IBD) among the wild populations was tested using the Mantel test implemented in GENALEX 6.5. The significance of IBD values was assessed using 9,999 permutations.

Inference of the Domestication History of Cultivated *Z. latifolia*

We tested possible models of the demographic history of wild and cultivated accessions to infer the domestication history of cultivated *Z. latifolia* using Approximate Bayesian Computing (ABC) (Beaumont et al., 2002), which allows model choice and parameter estimation without the need to calculate the likelihood function of the models. Data sets are generated under the considered models and are reduced to a set of summary statistics. Simulated data sets with the summary statistics closest to those of the observed data are used to determine the posterior distribution of the parameters and the relative posterior probability of each model. DIYABC v2.0 (Cornuet et al., 2014) was used to compare different models and infer historical parameters. According to the geographical distribution and Bayesian clustering approach implemented in STRUCTURE, the *Z. latifolia* populations could be assigned into two genetic groups, and the southern group showed watersheds-related genetic subdivisions (see the Results section for more details). To reduce the complexity of scenario settings, hierarchical methods were applied following Lombaert et al. (2011). Competing scenarios first were evaluated between the two groups to identify the one with the highest posterior probability and then repeated within the supporting group among subgroups.

We then used an ABC approach to infer demographical parameters for the best scenario. The prior parameters and designs of the scenarios are given in Table 2 and Figures 3 and 4. To construct the reference table, we calculated 10,000,000 simulated data sets for each competing scenario in DIYABC. The posterior probabilities of the competing scenarios were estimated by polychotomous logistic regression based on the 1% of data sets of the simulated reference table, according to the method described in Cornuet et al. (2014).

TABLE 2 | Prior distribution of parameters for hierarchical ABC analyses and posterior parameters estimation.

Prior parameters		Value	Posterior parameters			Quantile 2.5%	Quantile 97.5%
Parameter	Distribution		Mean	Median	Mode		
Analysis 1 (2 scenarios)							
N_s	Uniform	10–20,000	9,180	8,650	7,070	3,330	11,300
N_1	Uniform	10–1,000	247	228	199	49.3	319
t_1	Uniform	10–2,000	754	697	537	166	1,010
t_2	Uniform	10–2,000	971	903	668	250	1,320
Analysis 2 (9 scenarios)							
N_e	Uniform	10–20,000	8,040	7,840	7,630	2,660	9,750
N_d	Uniform	10–1,000	373	340	269	54.2	505
t_1	Uniform	10–2,000	1,600	1,400	1,030	335	2,040
t_2	Uniform	10–2,000	1,980	1,710	1,260	508	2,470
t_3	Uniform	10–2,000	2,550	1,930	1,330	426	3,290

The Phylogenetic Analysis of *U. esculenta* ITS Sequences

The ITS sequences were aligned with CLUSTAL W (Thompson et al., 1994) and adjusted manually. The number of polymorphic sites (S), haplotype diversity (H_d), and genetic variation measured by average pairwise differences per base pair between sequences (θ) (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). Maximum parsimony (MP) searches were performed with 1,000 random taxon addition replicates followed by tree bisection-reconnection branch swapping in PAUP* 4.0b10 (Swofford, 2002). Gaps were treated as missing data. Parsimony bootstrap (PB) for the clades was examined with 1,000 bootstrap replicates using the same options as above.

RESULTS

Genetic Variations in *Z. latifolia*

We analyzed plant genotypes for 824 wild and 135 cultivated accessions. Based on the 12 SSR loci, the 959 *Z. latifolia* accessions yielded a total of 138 alleles with 11.46 alleles per locus on average. The mean effective number of alleles (A), allelic richness (A_r), observed and expected heterozygosity (H_o and H_e), and Wright's fixation index (F) are listed in **Table 1**. H_e ranged from 0.167 (JXDA) to 0.588 (JSBY) among wild populations. The genetic diversity of cultivated accessions ($H_e = 0.242$) was much lower than that of wild accessions ($H_e = 0.432$). Deviations from Hardy-Weinberg equilibrium were detected in 90 of 384 comparisons after applying the sequential Bonferroni test, which might be due to inbreeding. The Wright's fixation indices varied greatly among populations (**Table 1**). Several populations (HLJHEB, HLJMDJ, JLDH, HBBB, SHYP, AHTL, JXDA, HuNCS, FJFD, GDYJ, GXNN, and YNJH) displayed signatures of heterozygote excess or deficit.

The PCoA indicated a distinct genetic divergence between wild and cultivated accessions. However, several putative wild accessions were identified as cultivars. This identification was further supported by the result of STRUCTURE analysis ($K = 2$, **Figures 2A, B**). The ΔK analysis indicated that the wild populations of *Z. latifolia* would demonstrate fine genetic structures with an increasing K value, which were shown in **Figures 2C, D**. The HuNCS and FJFD populations were thoroughly composed of plants with cultivar-like genotype. In addition, another 12 accessions from 3 wild populations (SHCM, ZJHZ, and YNJH) were identified as cultivars (**Table 1**). The neighbor-joining tree showed that the cultivars were genetically close to the populations in the lower reaches of the Yangtze River (**Figure S1**).

Genetic Structure

The effective population size N_e estimated by MIGRATE ranged from 100 to 312 ($\mu = 10^{-4}$, **Table 1**). Thirteen of 32 wild populations showed the evidence for historical or recent genetic bottlenecks, suggesting that the effects of genetic drift on these populations were prominent. The cultivars also showed a strong genetic bottleneck due to domestication (**Tables 1 and 2**).

After excluding the individuals with cultivar-like genotype (SHCM, ZJHZ, HuNCS, YNJH, and FJFD), F -statistics revealed that the wild populations of *Z. latifolia* were genetically divergent ($F_{st} = 0.320$, **Table S1**). The Mantel test suggested a weak but significant sign of isolation by distance among *Z. latifolia* populations (**Figure S2**, $r = 0.21$, $P < 0.05$). The outcomes of a Bayesian assignment with STRUCTURE indicated that $K = 2$ is the best for describing the actual genetic structure of wild populations (**Figure 3A**). One group included the northern populations in the north and Northeast Plain of China, and the remaining southern populations formed another group mainly located in the Yangtze River Basin. However, the PCoA showed no distinct genetic divergence among wild populations, and the first three principal components described 28.6% of the total variances. In addition, the Bayesian assignment of the southern populations revealed watershed-related genetic subdivisions in the Yangtze River Basin: the Hongzhe-Tai Lake region (Cluster I), the Poyang Lake region (Cluster II) and the Dongting Lake region (Cluster III) (**Figure 4A**). Two southern populations located in the Pearl River Basin (GDYJ and GXNN) showed a close ancestry with the populations in the Poyang Lake region cluster (**Figure 4A**).

In particular, we tested the genetic divergence between two ecotypes of cultivars. By using the software STRUCTURE, we found no significant genetic divergence with $K = 2$. Instead, the PCoA revealed the genetic differences between ecotypes (**Figure S3**). The microsatellite marker ZM24 could be an effective molecular marker to distinguish two ecotypes of cultivars.

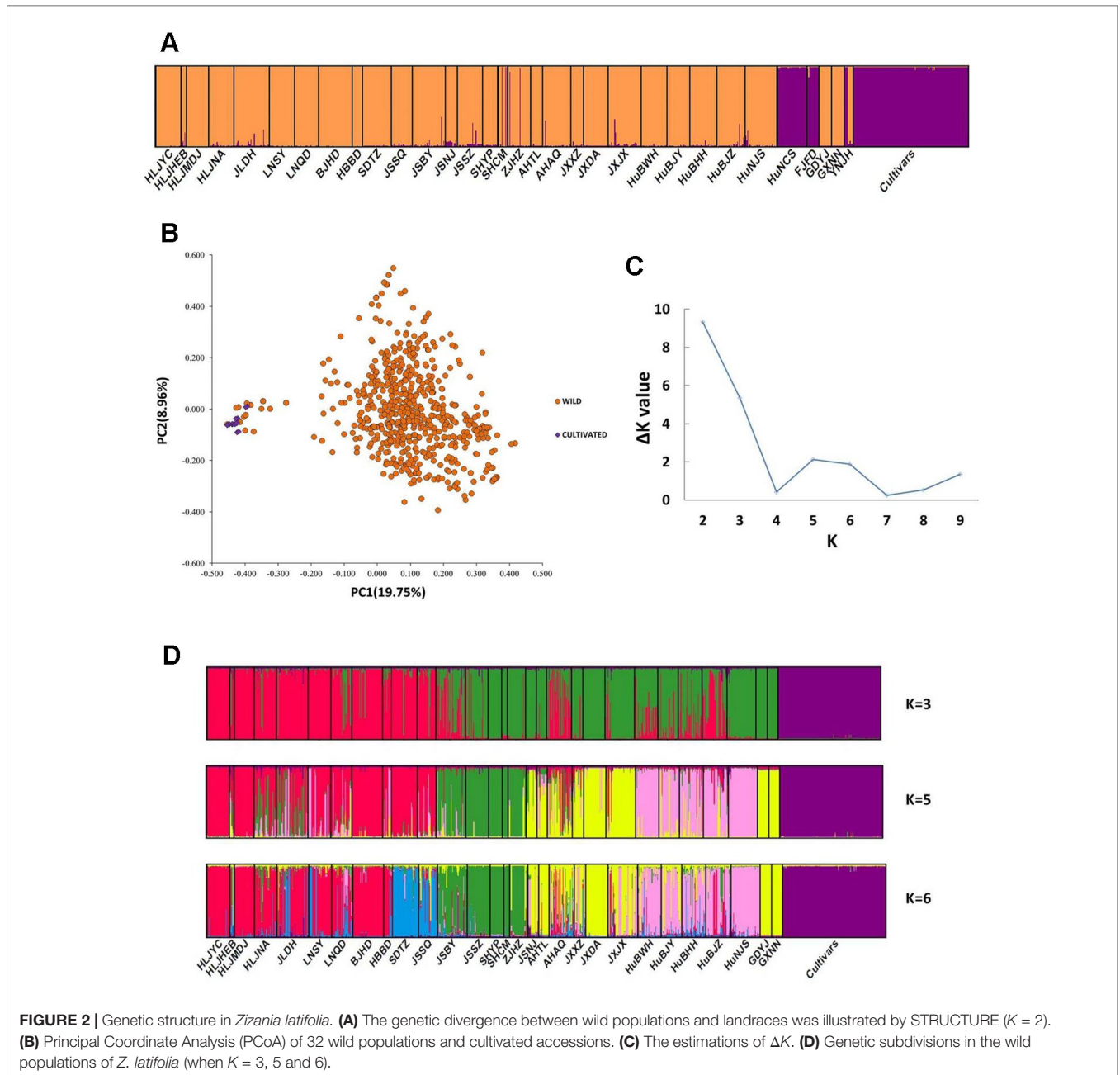
Modeling the Demography of *Z. latifolia* Domestication

Following our hierarchical ABC strategy, we first compared the two predesigned scenarios in Analysis 1. Then the posterior parameters of the best-supported scenario were estimated, and the model checking was accessed between the posterior data set and observed data (**Table 2** and **Figure S4A**). Scenario A had a higher posterior probability (0.700) (**Figure 3B**), and it indicated that the cultivars diverged from the group of southern populations. In Analysis 2, 9 competing scenarios were further investigated (**Table 2** and **Figure S4B**). Scenario 7 showed the highest posterior probability (0.295), which indicated that the Hongze-Tai Lake region cluster firstly diverged from the other clusters in the southern group 1930 generations (t_3) ago. The Poyang Lake region cluster and Dongting Lake cluster then diverged 1,710 generations (t_2) ago. The cultivars were single-domesticated in the Hongze-Tai Lake region 1,400 generations (t_1) ago (**Figure 4B** and **Table 2**).

Sequence Variation and Phylogenetic Analysis

A total of 37 strains of *U. esculenta*, including 5 wild and 32 cultivated strains, were sequenced. The total length of the aligned sequences was 712 bp. Twenty-seven polymorphic sites were observed. Thirty-one haplotypes were found in the wild and cultivated accessions, and the haplotype diversity (H_d) was 0.991. The nucleotide diversity (θ) was 0.01, and θ_w was 0.01.

An ITS sequence of *U. alcornii* was downloaded from GeneBank and used as the outgroup taxon. Another ITS sequence of *U. esculenta* (Genebank accession number: JX219372.1) was used as



the control. The MP tree is shown in **Figure 5**. The structure of the MP tree showed significant divergence between the wild fungi and those from cultivars, indicating a single domestication event.

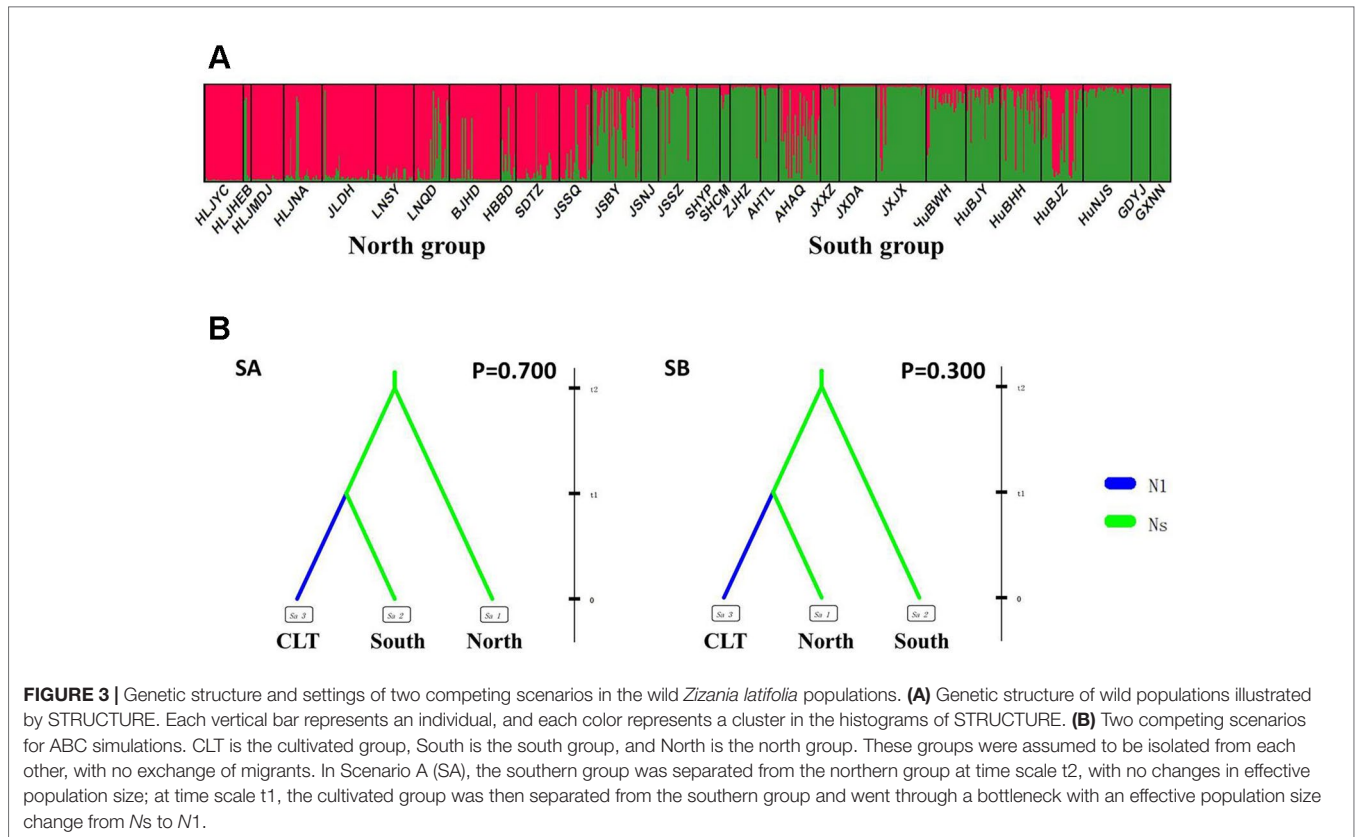
DISCUSSION

The domestication of cultivated *Z. latifolia* is a typical example that can demonstrate how a plant-fungus complex rapidly evolved under artificial selections. Our results showed that there was prominent genetic divergence between wild populations and cultivated accessions after long-time cultivation. A single domestication event occurred in the lower reaches of the Yangtze River at the

Tang Dynasty (approximately 1,400 years ago), then this vegetable was constantly cultivated and bred. The genetic divergence of the plant may be responsible for the differences between the ecotypes of cultivars, and the genetic variations in *U. esculenta* are likely associated with the diversification of cultivated varieties.

The Possible Domestication Process of *Z. latifolia*

There are distinct phenotypic differences between wild and cultivated *Z. latifolia*. Compared with its wild ancestor, the cultivated plant usually has an erected plant architecture, broader leaves, and shortened internodes (Guo et al., 2007). The cultivated



plants were reported to be vigorous in photosynthetic efficiency (Yan et al., 2013). These morphological and physiological changes together constituted the Domestication Syndrome of cultivated *Z. latifolia*, which would be referenced in distinguishing wild and cultivated plants.

We found significant genetic divergence between wild and cultivated *Z. latifolia* (Figures 2A, B). The extent of genetic differentiation between wild and cultivated plants ($F_{st} = 0.288$) was even greater than the differentiation among wild populations ($F_{st} = 0.175$). This phenomenon is usually observed in highly domesticated species such as wheat, rice or maize (Heun et al., 1997; Harter et al., 2004; Zeder et al., 2006; Zhao et al., 2013). Unlike the annual grain crops, the breeding of *Z. latifolia* relies on asexual reproduction. By utilizing natural mutants that occur in tillers or rhizomes, cultivated *Z. latifolia* was able to generate more than 100 varieties. Theoretically, when compared with annual crops bred *via* sexual reproduction, the evolutionary rate of cultivated *Z. latifolia* is relatively low. The low evolutionary rate and short domestication time (<2,000 years) should result in low genetic divergence between wild and cultivated plants, which is contrary to our finding. The unexpected high genetic divergence between wild and cultivated plants could also be attributed to the breeding system of cultivated *Z. latifolia*. Because cultivated *Z. latifolia* would occasionally degenerate to the wild phenotype, farmers had to screen out the elite plants as genets every year. Although the clonal reproduction of cultivated *Z. latifolia* greatly restricted new mutants stemmed from sexual propagation and blocked gene flows from wild accessions to cultivars, the

annually strong artificial selection would somehow accumulate mutants, inducing significant genetic divergence between wild and cultivated plants.

Previous studies had inferred that the cultivated *Z. latifolia* may be domesticated in the Yangtze River Basin, but they did not provide any convincing evidence. (Xu et al., 2008; Xu et al., 2015). Using microsatellite markers and model inference, our study located the origin site of cultivated *Z. latifolia* in the areas around Hongze Lake and Tai Lake (Figure 4). By ABC modeling, the cultivated *Z. latifolia* diverged from ancestor populations 1,400 generations ago. Given 1 year per generation, the domestication event occurred 1,400 years ago at the Tang Dynasty. Furthermore, the genetic relationship of fungi hinted that the single-season ecotype might have been domesticated first, and then the double-season ecotype was bred from the single-season ecotype. These results were in accordance with the historical records (Ke, 2008).

Based on current molecular data, we could not be sure whether *Z. latifolia* underwent historical pre-domestication as a food crop. There is no historical record or remained landraces that can support the existence of pre-domestication. We never found the plants with any domestication traits similar to cereal crops in the wild, too. However, the vegetable cultivars of *Z. latifolia* could escape to natural habitats. We found typical escaped individuals in five provinces that were the main production areas of cultivated *Z. latifolia* (Table 1). Cultivated plants that cannot survive on their own in natural habitats often have maladaptive domestication traits (Campbell and Snow, 2009). Cultivated *Z. latifolia* has lost the ability to undergo sexual reproduction, which would

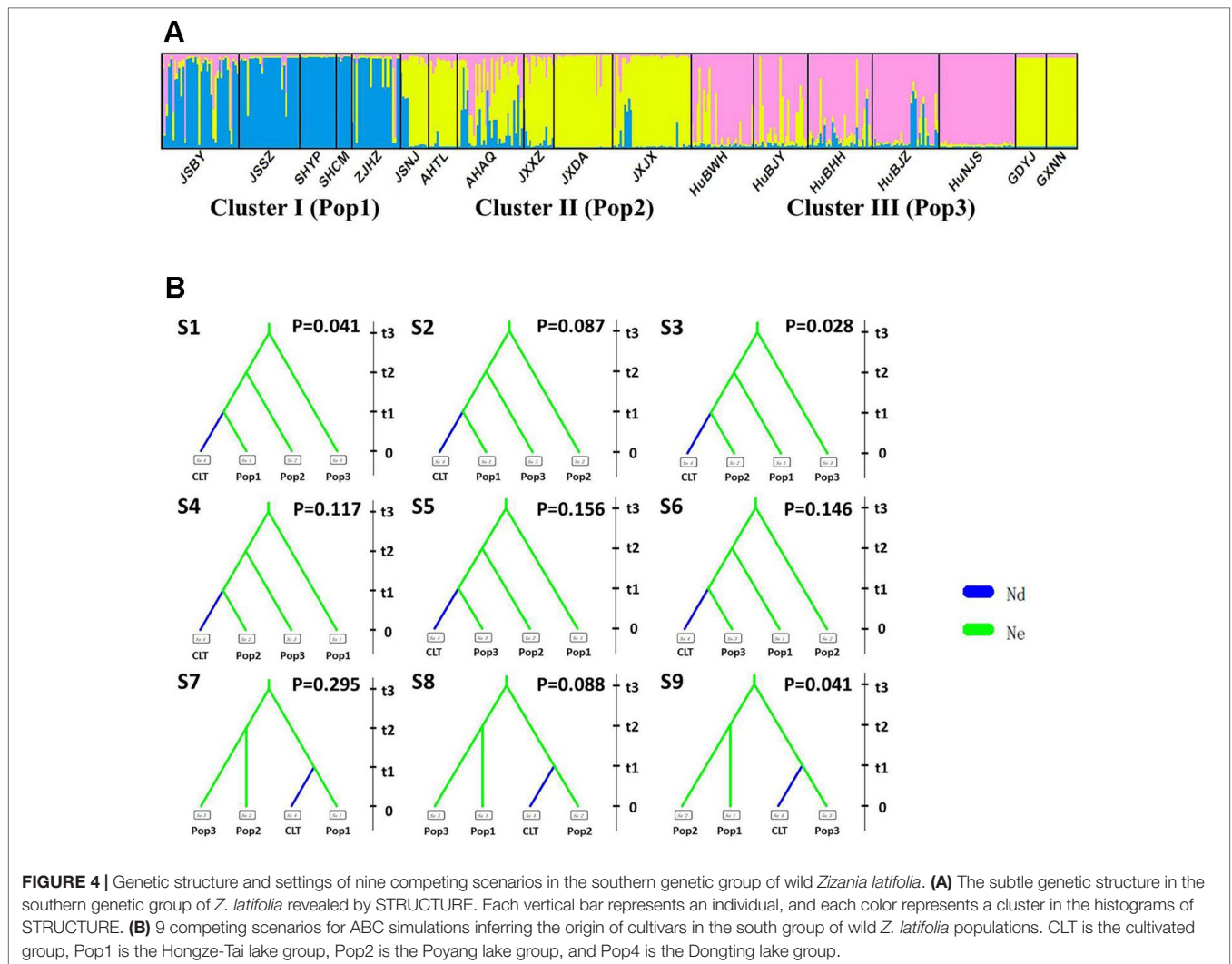


FIGURE 4 | Genetic structure and settings of nine competing scenarios in the southern genetic group of wild *Zizania latifolia*. **(A)** The subtle genetic structure in the southern genetic group of *Z. latifolia* revealed by STRUCTURE. Each vertical bar represents an individual, and each color represents a cluster in the histograms of STRUCTURE. **(B)** 9 competing scenarios for ABC simulations inferring the origin of cultivars in the south group of wild *Z. latifolia* populations. CLT is the cultivated group, Pop1 is the Hongze-Tai lake group, Pop2 is the Poyang lake group, and Pop4 is the Dongting lake group.

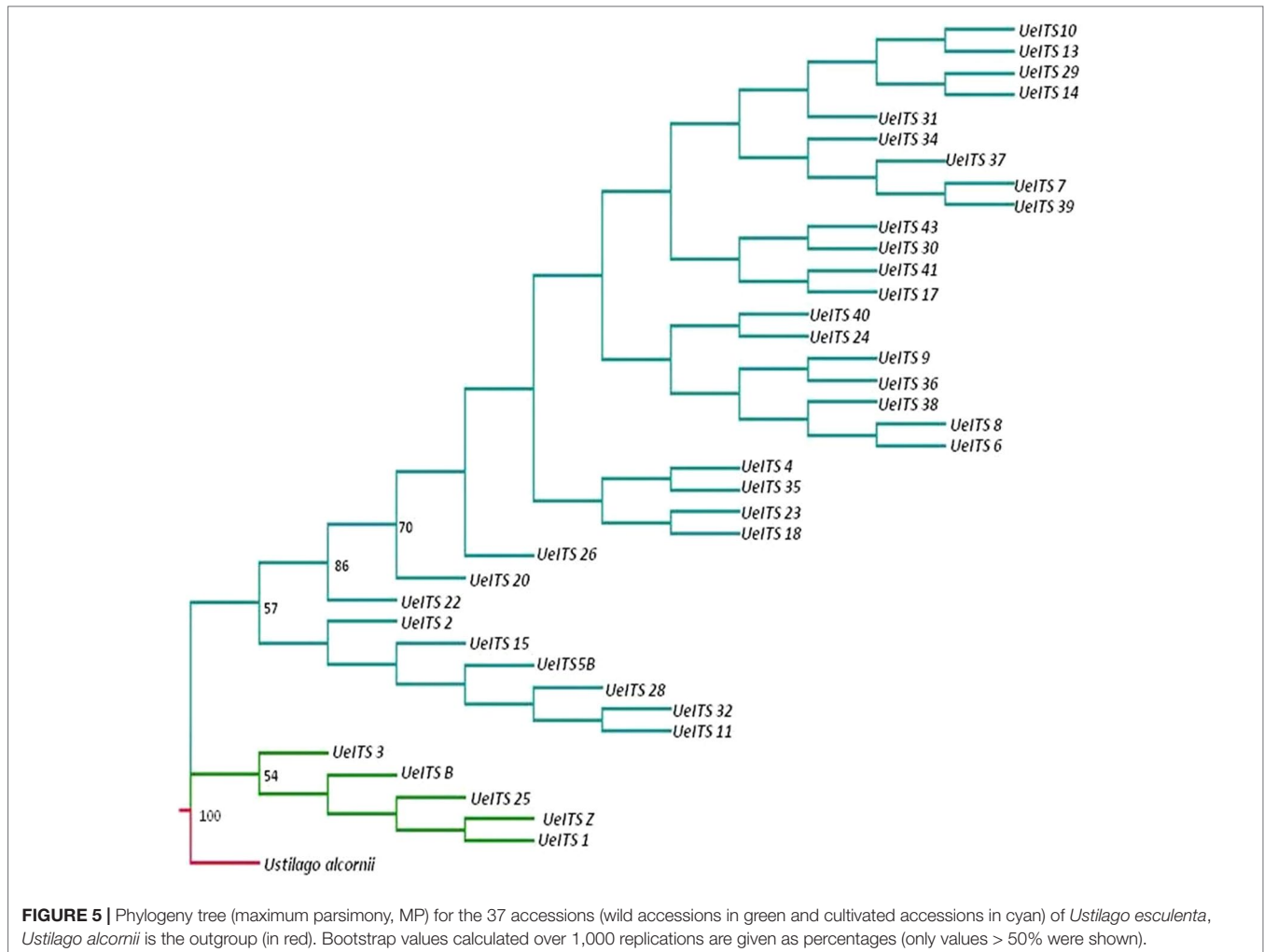
be unfavorable for persistence in the wild. Alternatively, feral *Z. latifolia* relies on asexual growth to reproduce in natural habitats. In fact, it often occupies the shores of ponds or streams near fields by clonal propagation *via* underground rhizomes and detached shoots, just as we had noticed in FJFD and HuNCS. Moreover, without pollen flow, it seems unlikely that wild populations would be introgressed by feral cultivars.

Genetic Diversity and Population Structure of Wild *Z. latifolia*

Xu et al. (2015) had characterized the genetic diversity of the four species in the genus *Zizania* using 3 cross-specific amplified microsatellite markers. They found that *Z. latifolia* showed a relatively low genetic diversity relative to its annual relative *Z. palustris* in North America ($H_e = 0.374$ vs $H_e = 0.630$). However, they might have underestimated the genetic diversity of both species due to insufficient genetic markers. In contrast, another study investigated the genetic variations of *Z. latifolia* restricted to the middle reaches of the Yangtze River by using

10 microsatellites and found a high level of genetic diversity ($H_e = 0.532$) (Chen et al., 2012). In our study, a relatively high genetic diversity ($H_e = 0.497$) was detected.

Most of the fixation indices (F) for the high-latitude populations were significantly greater than 0 (Table 1), indicating a heterozygous deficit that could be attributed to inbreeding or outbreeding among ramets. However, several populations in low latitudes were composed by a few dominant heterozygous genotypes, which resulted in negative F values. This phenomenon might be attributed to local adaptation or a rapid population expansion based on a few heterozygous ancestors *via* clonal growth. For example, 10 years ago, a few ancestors floating down from upstream colonized and formed the population JXDA (Zhao et al., field surveys). Previous phylogeographic analysis suggested that the genus *Zizania* originated in North America and then dispersed into eastern Asia *via* the Bering land bridge during the Tertiary (Xu et al., 2010). Hewitt (2004) had proposed that the range-wide patterns of population genetic diversity are usually shaped by past climate-driven range dynamic. In support of this view, the genealogical pattern based on the nuclear *Adh1a*



gene of *Z. latifolia* indicated that high-latitude populations harbored more abundant haplotypes (Xu et al., 2008). However, in our study, the populations in high latitudes (the northern group, $H_e = 0.447$) did not show a higher genetic diversity than those in low latitudes (the southern group, $H_e = 0.444$).

Our results and previous studies all indicated strong genetic differences among populations (Figure 2). The Mantel test further revealed the effect of isolation by distance (IBD, Figure S2). IBD is expected to be the primary factor leading to a high level of genetic divergence, due to limited gene flow and the effects of genetic drift (Slatkin, 1987). This process was supported by our study on wild *Z. latifolia* populations. The aquatic habitat of *Z. latifolia* is discrete and patchy. These isolated wetlands restrict migration between populations, and as a consequence, aquatic-living plants usually exhibit a high level of genetic divergence among populations (Barrett et al., 1993). The dispersal of seeds or detached shoots by water flow is unpredictable, and these are unlikely to disperse via water currents between spatially highly isolated populations. In wind-pollinated species, the pollen can be carried for long distances, but most effective pollination occurs at a local scale

(Willson, 1983). Nevertheless, the isolation between wetlands might be counteracted by seasonal flooding, resulting in a temporal connection within a water system. The genetic structure of *Z. latifolia* provided evidence that supported the effects of hydrological connectivity in shaping the genetic structure of aquatic plant species. Isolation by geographical barriers and genetic similarity along rivers was found in regional populations of *Z. latifolia* (Chen et al., 2012; Chen et al., 2017). The effects of hydrological unconnectedness on genetic relationship of populations within a water system were also reported in another emergent macrophyte, *Oryza rufipogon* (Wang et al., 2008).

The genetic assignments and morphological performances of two populations (GDYJ and GXNN) at the leading-edge seemed to be rather special (Table 1 and Figure 2D). On one hand, these populations are cultivated-like in phenotype, but they showed no sign of *U. esculenta* infection; on the other hand, they are genetically different from cultivated *Z. latifolia*, and have heterozygous genotypes. Their small population size and low diversity might be attributed to the “Founder effect.” Bayesian assignment showed they could be grouped to the Poyang Lake

region cluster, implying a migration through the geographical barrier (Nanling Mountains).

The Association of Plant and Fungus Under Artificial Selection

As an obligate parasitic fungus, *U. esculenta* is harmful to its host under natural conditions. The fungus not only inhibits the emergence of inflorescence but also requires a large amount of carbohydrates to form the swollen shoot attractive to herbivorous insects. *U. esculenta* infection is not very common in wild populations, as we have only collected 5 samples across 32 range-wide populations. In our common garden trials, the fitness of infected wild plants is significantly lower than that of healthy ones (Zhao et al., data not shown), which implies a weak competitive ability of the infected plants.

The Red Queen hypothesis proposed that the parasite and host should keep on adapting and evolving to gain reproductive advantages (Van Valen, 1973). The antagonistic co-evolution between species would finally be terminated either by the extinction of one species or the emergence of symbiosis. The domestication of infected *Z. latifolia* seemed to accelerate the process of achieving symbiosis. On the one hand, the persistent selection force of mankind delayed the development of the fungus and extended the mycolic stage (Zhang et al., 2017). On the other hand, humans gradually screened out plants with the genotype that adapted to cultivated conditions (Yan et al., 2013). Moreover, with the efforts exerted by mankind, the distribution range of this plant-fungus complex was largely expanded.

We detected minor but significant genetic differentiation between the hosts of two ecotypes of cultivated *Z. latifolia* (Figure S3), which was consistent with the pattern revealed by the sequences of *Adh1* (Xu et al., 2008). However, the ITS sequences of *U. esculenta* strains from cultivated *Z. latifolia* showed unexpectedly high levels of diversity compared with a previous study (You et al., 2011). There might be inconsistent evolutionary rates in the parasite and host. The evolution of the host determines the responses to photoperiod and temperature which resulted in the differences between two ecotypes in cultivated *Z. latifolia* (Guo et al., 2015), and the faster evolutionary rate of the parasite may lead to the diversification of varieties (Figure 5, also see Ye et al., 2017).

REFERENCES

- Barrett, S. C. H., Eckert, C. G., and Husband, B. C. (1993). Evolutionary processes in aquatic plant populations. *Aquat. Bot.* 44 (2-3), 105–145. doi: 10.1016/0304-3770(93)90068-8
- Beaumont, M.A., Zhang, W., and Balding, D.J. (2002). Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–2035. doi: 10.0000/PMID12524368
- Beerli, P. (2008). Migrate version 3.0: a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available from <http://popgen.scs.edu/migrate.html>.
- Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., et al. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *P. Natl. Acad. Sci. U. S. A.* 109 (14), E788–E796. doi: 10.1073/pnas.1108973109

Currently, it is still difficult to dissect the relative contributions from the host and parasite in shaping Domestication Syndrome. We could only infer the changes in morphological traits during the domestication process based on the phylogeny of cultivated *U. esculenta* (Figure 5). First, the single-season ecotype was domesticated from a wild plant-fungus complex. This ecotype retained a few primary traits, such as light sensitivity, long internodes, and small swollen shoot. Then, artificial selections on one of the single-season varieties led to the emergence of the double-season ecotype, which was insensitive to light and could be harvested twice. The double-season ecotype was more advanced than the single-season ecotype, but it could only be cultivated in a restricted region. There is an urgent need to improve the fitness of the double-season ecotype.

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

YZ, ZS, JC and JR conceived the idea and designed the research project. YZ, QL and LZ collected the data. YZ performed the data analysis. YZ drafted the initial manuscript with contribution from JR and ZS. All the authors contributed critically to the discussion and edited the manuscript before submission.

FUNDING

This study was supported by the National Natural Science Foundation of China (31600293) and the China Postdoctoral Science Foundation Grant (2015M571483).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.01406/full#supplementary-material>

- Burger, J. C., Chapman, M. A., and Burke, J. M. (2008). Molecular insights into the evolution of crop plants. *Am. J. Bot.* 95 (2), 113–122. doi: 10.3732/Ajb.95.2.113
- Campbell, L. G., and Snow, A. A. (2009). Can feral weeds evolve from cultivated radish (*Raphanus sativus*, Brassicaceae)? *Am. J. Bot.* 96 (2), 498–506. doi: 10.3732/ajb.0800054
- Chapuis, M. P., and Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* 24 (3), 621–631. doi: 10.1093/molbev/msl191
- Chen, Y., Long, L. K., Lin, X. Y., Guo, W. L., and Liu, B. (2006). Isolation and characterization of a set of disease resistance-gene analogs (RGAs) from wild rice, *Zizania latifolia* Griseb. I. Introgression, copy number lability, sequence change, and DNA methylation alteration in several rice-*Zizania* introgression lines. *Genome* 49 (2), 150–158. doi: 10.1139/G05-097

- Chen, Y. Y., Chu, H. J., Liu, H., and Liu, Y. L. (2012). Abundant genetic diversity of the wild rice *Zizania latifolia* in central China revealed by microsatellites. *Ann. Appl. Biol.* 161 (2), 192–201. doi: 10.1111/j.1744-7348.2012.00564.x
- Chen, Y. Y., Liu, Y., Fan, X. R., Li, W., and Liu, Y. L. (2017). Landscape-scale genetic structure of wild rice *Zizania latifolia*: the roles of rivers, mountains and fragmentation. *Front. Ecol. Evol.* 5, 17. doi: 10.3389/fevo.2017.00017
- Choi, J. Y., and Purugganan, M. D. (2018). Multiple origin but single domestication led to *Oryza sativa*. *G3-Genes. Genom. Genet.* 8 (3), 797–803. doi: 10.1534/g3.117.300334
- Cornuet, J. M., Pudlo, P., Veysier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., et al. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30 (8), 1187–1189. doi: 10.1093/bioinformatics/btt763
- Dieringer, D., and Schlotterer, C. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* 3 (1), 167–169. doi: 10.1046/j.1471-8286.2003.00351.x
- Diez, C. M., Trujillo, I., Martínez-Urdiroz, N., Barranco, D., Rallo, L., Marfil, P., et al. (2015). Olive domestication and diversification in the Mediterranean Basin. *New Phytol.* 206 (1), 436–447. doi: 10.1111/nph.13181
- Doebly, J. F., Gaut, B. S., and Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell* 127 (7), 1309–1321. doi: 10.1016/j.cell.2006.12.006
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14 (8), 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- FAO (2019). *The State of the World's Biodiversity for Food and Agriculture*. Bélanger, J., and Pilling, D., editors. Rome: FAO Commission on Genetic Resources for Food and Agriculture Assessments.
- Fuller, D. Q., Asouti, E., and Purugganan, M. D. (2012). Cultivation as slow evolutionary entanglement: comparative data on rate and sequence of domestication. *Veg. Hist. Archaeobot.* 21 (2), 131–145. doi: 10.1007/s00334-011-0329-8
- Ge, S., Li, A., Lu, B. R., Zhang, S. Z., and Hong, D. Y. (2002). A phylogeny of the rice tribe Oryzaceae (Poaceae) based on matK sequence data. *Am. J. Bot.* 89 (12), 1967–1972. doi: 10.3732/ajb.89.12.1967
- Ge, S., and Sang, T. (2011). Inappropriate model rejects independent domestications of indica and japonica rice. *P. Natl. Acad. Sci. U. S. A.* 108 (39), E55–E55. doi: 10.1073/pnas.1111601108
- Giovannoni, J. (2018). Tomato multiomics reveals consequences of crop domestication and improvement. *Cell* 172 (1–2), 6–8. doi: 10.1016/j.cell.2017.12.036
- Gross, B. L., and Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends Plants Sci.* 15 (9), 529–537. doi: 10.1016/j.tplants.2010.05.008
- Guo, H. B., Li, S. M., Peng, J., and Ke, W. D. (2007). *Zizania latifolia* Turcz. cultivated in China. *Genet. Resour. Crop. Ev.* 54 (6), 1211–1217. doi: 10.1007/s10722-006-9102-8
- Guo, L. B., Qiu, J., Han, Z. J., Ye, Z. H., Chen, C., Liu, C. J., et al. (2015). A host plant genome (*Zizania latifolia*) after a century-long endophyte infection. *Plant J.* 83 (4), 600–609. doi: 10.1111/pj.12912
- Guo, Y. L., and Ge, S. (2005). Molecular phylogeny of Oryzaceae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. *Am. J. Bot.* 92 (9), 1548–1558. doi: 10.3732/ajb.92.9.1548
- Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica.* 156 (1–2), 1–13. doi: 10.1007/s10681-007-9363-0
- Harter, A. V., Gardner, K. A., Falush, D., Lentz, D. L., Bye, R. A., and Rieseberg, L. H. (2004). Origin of extant domesticated sunflowers in eastern North America. *Nature* 430 (6996), 201–205. doi: 10.1038/nature02710
- Heun, M., Abbo, S., Lev-Yadun, S., and Gopher, A. (2012). A critical review of the protracted domestication model for Near-Eastern founder crops: linear regression, long-distance gene flow, archaeological, and archaeobotanical evidence. *J. Exp. Bot.* 63 (12), 4333–4341. doi: 10.1093/jxb/ers162
- Heun, M., Schafer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., et al. (1997). Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278 (5341), 1312–1314. doi: 10.1126/science.278.5341.1312
- Hewitt, G. M. (2004). Biodiversity: a climate for colonization. *Heredity (Edinb)* 92 (1), 1–2. doi: 10.1038/sj.hdy.6800365
- Huang, X. H., Kurata, N., Wei, X. H., Wang, Z. X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490 (7421), 497–501. doi: 10.1038/nature11532
- Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23 (14), 1801–1806. doi: 10.1093/bioinformatics/btm233
- Ke, W. D. (2008). “Jiaobai,” in *Crops and Their Wild Relatives in China*, vol. Vegetable Crops. Eds. Dong, Y. C., and Liu, X. (Beijing: China Agriculture Press), 1084–1095.
- Liu, B., Liu, Z. L., and Li, X. W. (1999). Production of a highly asymmetric somatic hybrid between rice and *Zizania latifolia* (Griseb): evidence for inter-genomic exchange. *Theor. Appl. Genet.* 98 (6–7), 1099–1103. doi: 10.1007/s001220051173
- Lombaert, E., Guillemaud, T., Thomas, C. E., Handley, L. J. L., Li, J., Wang, S., et al. (2011). Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: case study of the invasive ladybird *Harmonia axyridis*. *Mol. Ecol.* 20 (22), 4654–4670. doi: 10.1111/j.1365-294X.2011.05322.x
- Nei, M., and Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* 76 (10), 5269–5273. doi: 10.1073/pnas.76.10.5269
- Oelke, E. A. (1993). “Wild rice: domestication of a native North American genus,” in *New crops*. Eds. Janick, J., and Simon, J. E. (New York: Wiley), 235–243.
- Peakall, R., and Smouse, P. E. (2012). GenAler 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28 (19), 2537–2539. doi: 10.1093/bioinformatics/bts460
- Piry, S., Luikart, G., and Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* 90 (4), 502–503. doi: 10.1093/jhered/90.4.502
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155 (2), 945–959. doi: 10.1111/j.1471-8286.2004.00684.x
- Purugganan, M. D., and Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature* 457 (7231), 843–848. doi: 10.1038/nature07895
- Quan, Z. W., Pan, L., Ke, W. D., Liu, Y. M., and Ding, Y. (2009). Sixteen polymorphic microsatellite markers from *Zizania latifolia* Turcz. (Poaceae). *Mol. Ecol. Resour.* 9 (3), 887–889. doi: 10.1111/j.1755-0998.2008.02357.x
- Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249. doi: 10.1046/j.1420-9101.1995.8030385.x
- Ronald, P. C. (2014). Lab to farm: applying research on plant genetics and genomics to crop improvement. *PLoS Biol.* 12 (6), e1001878. doi: 10.1371/journal.pbio.1001878
- Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4 (1), 137–138. doi: 10.1046/j.1471-8286.2003.00566.x
- Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19 (18), 2496–2497. doi: 10.1093/bioinformatics/btg359
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science* 236 (4803), 787–792. doi: 10.1126/science.3576198
- Song, Z. P., Zhu, W. Y., Rong, J., Xu, X., Chen, J. K., and Lu, B. R. (2006). Evidence of introgression from cultivated rice to *Oryza rufipogon* (Poaceae) populations based on SSR fingerprinting: implications for wild rice differentiation and conservation. *Evol. Ecol.* 20 (6), 501–522. doi: 10.1007/s10682-006-9113-0
- Stoll, M., Begerow, D., and Oberwinkler, F. (2005). Molecular phylogeny of *Ustilago*, *Sporisorium*, and related taxa based on combined analyses of rDNA sequences. *Mycol. Res.* 109, 342–356. doi: 10.1017/S0953756204002229
- Swofford, D. L. (2002). *PAUP*: Phylogenetic analysis using parsimony and other methods Version 4b10*. Sunderland, USA: Sinauer Associates.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (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 (22), 4673–4680. doi: 10.1093/nar/22.22.4673
- Tufto, J. (2017). Domestication and fitness in the wild: a multivariate view. *Evolution* 71 (9), 2262–2270. doi: 10.1111/evo.13307
- van Heerwaarden, J., Doebly, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., Gonzalez, J. D. S., et al. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *P. Natl. Acad. Sci. U. S. A.* 108 (3), 1088–1092. doi: 10.1073/pnas.1013011108
- Van Valen, L. (1973). Body size and numbers of plants and animals. *Evolution* 27 (1), 27–35. doi: 10.1111/j.1558-5646.1973.tb05914.x

- Wang, M., Li, W. Z., Fang, C., Xu, F., Liu, Y. C., Wang, Z., et al. (2018). Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nat. Genet.* 50 (10), 1435–1441. doi: 10.1038/s41588-018-0229-2
- Wang, M. X., Zhang, H. L., Zhang, D. L., Qi, Y. W., Fan, Z. L., Li, D. Y., et al. (2008). Genetic structure of *Oryza rufipogon* Griff. in China. *Heredity* 101 (6), 527–535. doi: 10.1038/hdy.2008.61
- Wang, Z. D., Yan, N., Wang, Z. H., Zhang, X. H., Zhang, J. Z., Xue, H. M., et al. (2017). RNA-seq analysis provides insight into reprogramming of culm development in *Zizania latifolia* induced by *Ustilago esculenta*. *Plant Mol. Biol.* 95 (6), 533–547. doi: 10.1007/s11103-017-0658-9
- Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7, 256–276. doi: 10.1016/0040-5809(75)90020-9
- Wen, J. (1999). Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annu. Rev. Ecol. Syst.* 30, 421–455. doi: 10.1146/annurev.ecolsys.30.1.421
- Whitehead, S. R., Turcotte, M. M., and Poveda, K. (2017). Domestication impacts on plant - herbivore interactions: a meta-analysis. *Philos. T. R. Soc. B.* 372 (1712), 20160034. doi: 10.1098/rstb.2016.0034
- Willson, M. F. (1983). *Plant reproductive ecology*. New York, USA: Wiley.
- Wu, Z. Y., Raven, P. H., and Hong, D. Y., (2006). “Flora of China,” in *Poaceae*, vol. 22. (Beijing/St Louis: Science Press/Missouri Botanical Garden).
- Xiang, Q. Y., Crawford, D. J., Wolfe, A. D., Tang, Y. C., and DePamphilis, C. W. (1998). Origin and biogeography of *Aesculus* L. (Hippocastanaceae): a molecular phylogenetic perspective. *Evolution* 52 (4), 988–997. doi: 10.1111/j.1558-5646.1998.tb01828.x
- Xu, X. W., Ke, W. D., Yu, X. P., Wen, J., and Ge, S. (2008). A preliminary study on population genetic structure and phylogeography of the wild and cultivated *Zizania latifolia* (Poaceae) based on Adh1a sequences. *Theor. Appl. Genet.* 116 (6), 835–843. doi: 10.1007/s00122-008-0717-3
- Xu, X. W., Walters, C., Antolin, M. F., Alexander, M. L., Lutz, S., Ge, S., et al. (2010). Phylogeny and biogeography of the eastern Asian–North American disjunct wild-rice genus (*Zizania* L., Poaceae). *Mol. Phyl. Evol.* 55, 1008–1017. doi: 10.1016/j.ympev.2009.11.018
- Xu, X. W., Wu, J. W., Qi, M. X., Lu, Q. X., Lee, P. F., Lutz, S., et al. (2015). Comparative phylogeography of the wild-rice genus *Zizania* (Poaceae) in eastern Asia and North America. *Am. J. Bot.* 102 (2), 239–247. doi: 10.3732/ajb.1400323
- Yan, N., Wang, X. Q., Xu, X. F., Guo, D. P., Wang, Z. D., Zhang, J. Z., et al. (2013). Plant growth and photosynthetic performance of *Zizania latifolia* are altered by endophytic *Ustilago esculenta* infection. *Physiol. Mol. Plant.* P. 83, 75–83. doi: 10.1016/j.pmp.2013.05.005
- Yang, Y. F., Wang, S. Z., and Lang, H. Q. (1999). Study on the structures of clone population *Zizania latifolia* in wetland of the Songnen Plain of China. *Acta Prataculturalence* 8, 66–71.
- Ye, Z., Pan, Y., Zhang, Y., Cui, H., Jin, G., McHardy, A. C., et al. (2017). Comparative whole-genome analysis reveals artificial selection effects on *Ustilago esculenta* genome. *DNA Res.* 24 (6), 635–648. doi: 10.1093/dnares/dsx031
- You, W. Y., Liu, Q., Zou, K. Q., Yu, X. P., Cui, W. Y., and Ye, Z. H. (2011). Morphological and molecular differences in two strains of *Ustilago esculenta*. *Curr. Microbiol.* 62, 44–54. doi: 10.1007/s00284-010-9673-7
- Zeder, M. A. (2015). Core questions in domestication research. *P. Natl. Acad. Sci. U. S. A.* 112 (11), 3191–3198. doi: 10.1073/pnas.1501711112
- Zeder, M. A. (2017). Domestication as a model system for the extended evolutionary synthesis. *Interface Focus* 7 (5), 20160133. doi: 10.1098/rsfs.2016.0133
- Zeder, M. A., Emshwiller, E., Smith, B. D., and Bradley, D. G. (2006). Documenting domestication: the intersection of genetics and archaeology. *Trends Genet.* 22 (3), 139–155. doi: 10.1016/j.tig.2006.01.007
- Zhai, C. K., Lu, C. M., Zhang, X. Q., Sun, G. J., and Lorenz, K. J. (2001). Comparative study on nutritional value of Chinese and North American wild rice. *J. Food Compos. Anal.* 14 (4), 371–382. doi: 10.1006/jfca.2000.0979
- Zhang, J. Z., Chu, F. Q., Guo, D. P., Hyde, K. D., and Xie, G. L. (2012). Cytology and ultrastructure of interactions between *Ustilago esculenta* and *Zizania latifolia*. *Mycol. Prog.* 11 (2), 499–508. doi: 10.1007/s11557-011-0765-y
- Zhang, Y. F., Cao, Q. C., Hu, P., Cui, H. F., Yu, X. P., and Ye, Z. H. (2017). Investigation on the differentiation of two *Ustilago esculenta* strains - implications of a relationship with the host phenotypes appearing in the fields. *BMC Microbiol.* 17, 228. doi: 10.1186/S12866-017-1138-8
- Zhao, Y., Vrieling, K., Liao, H., Xiao, M., Zhu, Y., Rong, J., et al. (2013). Are habitat fragmentation, local adaptation and isolation-by-distance driving population divergence in wild rice *Oryza rufipogon*? *Mol. Ecol.* 22 (22), 5531–5547. doi: 10.1111/mec.12517

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Zhao, Song, Zhong, Li, Chen and Rong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.