Journal of Integrative Agriculture 2016, 15(9): 1991–2001

Available online at www.sciencedirect.com

ScienceDirect

RESEARCH ARTICLE

Genetic diversity of pepper (*Capsicum* spp.) germplasm resources in China reflects selection for cultivar types and spatial distribution

ZHANG Xiao-min[1*], ZHANG Zheng-hai[1*], GU Xiao-zhen[1], MAO Sheng-li[1], LI Xi-xiang[1,2*], Joël Chadœuf[3], Alain Palloix[3#], WANG Li-hao[1], ZHANG Bao-xi[1]

[1] *Key Laboratory of Vegetable Genetics and Physiology, Ministry of Agriculture/Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

[2] *National Gene Bank of Vegetables, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

[3] *Génétique et Amélioration des Fruits et Légumes, BP94, UR1052, PACA, INRA, Montfavet F-84143, France*

Abstract

Pepper (*Capsicum* spp.) is an important vegetable crop in the world. Now the pepper in China contributes one-third of the world's peppers production. Genetic diversity of the pepper germplasm of China is expected interesting to know. To explore the structure of genetic diversity in Chinese pepper germplasm resources and possible relationship with cultivar types or geographic origin, we sampled and compared 372 GenBank pepper accessions (local cultivars and landraces) from 31 provinces, autonomous regions and municipalities of China and 31 additional accessions from other countries. These accessions were genotyped using 28 simple sequence repeat (SSR) markers spanning the entire pepper genome. We then investigated the genetic structure of the sampled collection using model-based analysis in STRUCTURE v2.3.4 and examined genetic relationships by the unweighted pair-group method of mathematical averages (UPGMA) in MEGA. In addition to geographic origin, we evaluated eight plant and fruit traits. In total, 363 alleles were amplified using the 28 SSR primers. Gene diversity, polymorphism information content and heterozygosity of the 28 SSR loci were estimated as 0.09–0.92, 0.08–0.92 and 0.01–0.34, respectively. The UPGMA cluster analysis clearly distinguished *Capsicum annuum* L. from other cultivated pepper species. Population structure analysis of the 368 *C. annuum* accessions uncovered three genetic groups which also corresponded to distinct cultivar types with respect to the plant and fruit descriptors. The genetic structure was also related to the geographic origin of the landraces. Overall results indicate that genetic diversity of Chinese pepper landraces were structured by migration of genotypes followed by human selection for cultivar types in agreement with consumption modes and adaptation to the highly diversified agro-climatic conditions.

Received 4 August, 2015 Accepted 11 March, 2016
ZHANG Xiao-min, Tel: +86-10-82109551,
E-mail: zhangxiaomin1987@yeah.net;
Correspondence WANG Li-hao, Tel: +86-10-82109551,
Fax: +86-10-62174123, E-mail: wanglihao@caas.cn
[*] These authors contributed equally to this study.
[#] Deceased.

doi: 10.1016/S2095-3119(16)61364-3

Keywords: *Capsicum* spp. germplasm collection, genetic diversity, population structure, microsatellite

1. Introduction

Pepper (*Capsicum* spp.) is one of the most globally important vegetable crops because of its pungency and high nutritional value. It is becoming increasingly popular among consumers, with industrial applications also rising worldwide. *Capsicum* L. pepper is a genus in the Solanaceae family, comprising 27 recognized species (Baral and Bosland 2002). Studies have provided evidence for the existence of five distinct cultivated species: *Capsicum annuum* L., *Capsicum frutescens* L., *Capsicum chinens*e Jacq., *Capsicum baccatum* L., and *Capsicum pubescens* Ruiz et Pav (Pickersgill *et al.* 1979, 1991).

Peppers were introduced from the West Indies into Europe in 1493 following the first voyage of Christopher Columbus, after which they were rapidly distributed to Africa and Asia (Purseglove 1968; Andrews 1984; Somos 1984; Nicolai *et al*. 2013). The first historic record of peppers in China appeared in 1591 in a book entitled *Tsung-sheng-pa-chien* (Kao 1591), where they were described aspaint-brush-shaped, highly ornamental fruits with a hot taste and red color. In 1688, Chen (1688) described chili peppers in his book *Hua Jing*, calling them Kadsura Pepper Stemsor, the common name of spicy tomatoes.

Two routes of initial introduction into China have historically been hypothesized for peppers (Zheng 2006). The first is *via* the Silk Road extending from Europe through the Middle East to northwestern and then central China, an ancient trade network exploited by Zhang Qian during the Han Dynasty for nearly 2000 years ago. The other possible route was the Maritime Silk Road, which stretched across the Indian Ocean from the Persian Gulf to Southeast Asia (Andrews 1984). In the 17th–18th centuries, introductions may also have occurred from the west coast of South America (Peru) across the Pacific Ocean to Southeast Asia. This would have brought pepper into China at the end of the Ming Dynasty and the beginning of the Qing Dynasty, especially after 1683 with the opening of a sea trade route (Zou 2009). Further introductions occurred during the 20th century, when a large number of cultivar accessions were introduced and integrated into modern breeding programs in China. Because of the preference by farmers for a long cultivation history, many local landraces were distributed across the country. Since the 1960s, pepper accessions in China, including foreign lines, have been collected in the National Genebank of China (Zhang *et al.* 2010).

The genetic diversity of pepper accessions in China has not yet been extensively analyzed, even though such studies could provide information about sources of genetic differentiation following both natural and human selection. Some studies using random amplified polymorphic DNA, and inter-simple sequence repeat or simple sequence repeat (SSR) markers have been reported (Chen 2006; Luo 2006; Zhou *et al.* 2009), but these surveys only considered a few (up to 89) accessions, so did not evaluate the entire Chinese germplasm. In the early 1990s, genetic analyses using isozyme, nuclear, and chloroplast DNA markers confirmed the botanic structure of the genus *Capsicum* (Rodriguez *et al.* 1999; Walsh and Hoot 2001; Toquica *et al.* 2003; Ince *et al.* 2010; Jeong *et al.* 2010; Ibiza *et al.* 2012). Genetic diversity has also been explored in restricted geographic areas (mainly Mexico and the Andean area), which has yielded evidence of a genetic shift from wild to cultivated environments (Hernández-Verdugo *et al.* 2001; Votava *et al.* 2002; Aguilar-Meléndez *et al.* 2009; Albrecht *et al.* 2012; González-Jara *et al.* 2012; Pacheco-Olvera *et al.* 2012). A larger collection, including 1352 non-redundant accessions of 11 *Capsicum* species from 89 countries, was recently analyzed by Nicolai *et al.* (2013) using 28 SSR loci. Their study confirmed the strong interspecific structure of the *Capsicum* genus and offered new information about the wild origin of the cultivated *C. annuum* species. They also revealed that the genetic structure of the domesticated species *C. annuum* correlates with cultivar types rather than geographic origins, and thus mostly results from human selection. Another recent study by Hill *et al.* (2013) analyzed 46 accessions worldwide for genetic diversity and population structure using a microarray design of 30000 unigenes in a GeneChip.

In this paper, we report the first comprehensive exploration of the genetic diversity of Chinese pepper germplasm. We used 28 SSR markers to survey 372 accessions sampled from 31 provinces, autonomous regions and municipalities of China. The uncovered genetic diversity was tentatively related to both agricultural traits and geographic origins. The genetic structure of the Chinese pepper germplasm resources revealed the influence of geographic origin and long-term artificial selection.

2. Materials and methods

2.1. Materials

A total of 403 accessions, mostly landraces of the culti-

vated species *C. annuum* with a few lines of *C. chinense*, *C. baccatum*, and *C. frutescens* as out groups (Table 1 and Appendix A), were analyzed in this study. All accessions were obtained from the China National Vegetable Germplasm Bank and the pepper research group located at the Institute of Vegetables and Flowers of the Chinese Academy of Agricultural Sciences (IVF-CAAS). A total of 372 accessions (369 of *C. annuum*, two of *C. chinense*, and one of *C. frutescens*) were originally collected from 31 provinces, autonomous regions and municipalities of China. The remaining 31 accessions originated from the United States (*n*=7), Bulgaria (*n*=5), Austria (*n*=3), Japan (*n*=3), France (*n*=2), Peru (*n*=2), Brazil (*n*=1), Canada (*n*=1), Ethiopia (*n*=1), India (*n*=1), Italy (*n*=1), Mexico (*n*=1), Myanmar (*n*=1), the Netherlands (*n*=1), and Romania (*n*=1).

2.2. Phenotypic trait measurements

Fifteen plants of each accession were grown in a greenhouse in 2012 and 2013. Eight fruit and plant traits related to International Plant Genetic Resources Institute descriptors for *Capsicum* (PGRI *et al.* 1995) were measured from three plants per year. The four evaluated fruit traits were fruit weight (g), fruit length (cm), fruit diameter at the widest point (cm), and pericarp thickness (mm). The three primary plant traits were days to flowering date (after transplantation), plant height (cm), and estimated leaf area (leaf length×width, cm^2). The pungency of each accession was independently assessed by three individuals. Photographs were taken of the fruit to record their shape.

2.3. SSR analyses

DNA was extracted from one young plantlet randomly selected from six individuals of each accession using the modified cetyltrimethylammonium bromide-based method of Fulton *et al.* (1995). After determining the quality and concentration of genomic DNA quality with a Biospec-nano microspectrophotometer (Shimadzu, Tokyo, Japan), DNA samples were diluted to 25 ng L^{-1}. For SSR analysis, we used 28 microsatellite markers publicly available from Nicolai *et al.* (2013) that spanned 11 of the 12 pepper chromosomes (Appendix B). Each SSR forward primer was labeled with one of four fluorescent dyes: 6-carboxy-fluorescein, 6-carboxy-tetramethylrhodamine, hexachloro-6-carboxy-fluorescein, or carboxy-X-rhodamine. All primers were synthesized by Sangon Company (Shanghai, China).

PCR amplifications were performed in 10 μL reaction volumes containing 2 μL of 25 ng μL^{-1} genomic DNA as a template, 5 μL of 2× GoTaq R colorless master mix, 0.25 μL each of forward and reverse primers (10 μmol L^{-1}), and 2.5 μL of sterilized ddH_2O. Following the reaction, 8 μL formamide and an internal size standard (10% Liz-500; LIZ, USA) were added to each mixture, which was then heated at 95°C for 5 min and then cooled in ice water. Electrophoresis was carried out on an ABI 3130xl genetic analyzer (Applied Biosystems, now Life Technology, USA).

2.4. Data analysis

Genetic diversity analysis GeneMapper 4.0 software was used to evaluate the sizes of amplified fragments (SSR alleles) from different samples. The number of alleles, allele frequencies, number of genotypes, Nei's unbiased gene diversity index (He), observed heterozygosity (Ho), and polymorphism information content (PIC) were calculated using Power Marker v3.25 software (http://www.powermarker.net). Genetic distances (chord distances) were calculated by the

Table 1 Geographic distribution of Chinese pepper germplasm samples analyzed in this study

Distribution (provinces, autonomous regions and municipalities)	Number of accessions	Distribution (provinces, autonomous regions and municipalities)	Number of accessions
Anhui	18	Jiangxi	9
Beijing	5	Shaanxi	4
Chongqing	8	Shandong	14
Fujian	21	Shanghai	3
Gansu	5	Shanxi	16
Guangdong	2	Sichuan	19
Guangxi	6	Tianjin	6
Guizhou	30	Xinjiang	9
Hainan	1	Tibet	11
Hebei	10	Yunnan	23
Heilongjiang	12	Zhejiang	5
Henan	15	Jilin	24
Hubei	7	Liaoning	12
Hunan	36	Ningxia	5
Inner Mongolia	11	Qinghai	5
Jiangsu	20		

method of Cavalli-Sforza and Edwards (1967) and used to construct a phylogenetic tree based on the unweighted pair-group method of mathematical averages (UPGMA) as implemented in MEGA (Tamura*et al.* 2007).

Genetic structure To infer the population structure of the 369 *C. annuum* accessions from China, we used the model-based program STRUCTURE v2.3.4 (Pritchard *et al.* 2000). The optimal number of genetic groups (K) was computed by performing eight Markov chain Monte Carlo runs for each value of K from 1 to 10. Each run consisted of 100 000 iterations, with a burn-in period of 10 000 iterations, and used an admixture model allowing for Hardy-Weinberg equilibrium, correlated allele frequencies, and independent loci. A Pr(X|K) index with respect to each K was used to calculate ΔK (Evanno *et al.* 2005), where X denotes the genotypes of the sampled individuals. According to this approach, the optimal K is estimated as the location of the first peak of A $\Delta K=|L^2(K)|/s[\Pr(x|k)]$, where $|L^2(K)|$ denotes the absolute value of the second order rate of change of Pr(X|K), and s[Pr(x|k)] is the standard deviation of the Pr(X-|K), s means standard deviation of L(K).

2.5. Spatial data analysis

Testing for spatial dependence between genotypic data was performed through a Mantel test (Manly 1991) by computing a matrix of geographic distances between sample pairs and a genotypic similarity matrix whose elements are the sum of markers with the same alleles for each sample pair, and then testing the correlation between matrix elements. The P-value of the correlation under the assumption of independence between genotypic data and geographic positions was computed using 1 000 random permutations of the genotypic plant observations.

To test whether the groups identified from the STRUCTURE analysis were spatially structured, we computed empirical variograms of the probability of membership in each STRUCTURE-defined group and their confidence bands under the assumption of the absence of geographic dependence by randomizing these probabilities (Cressie 1993). Kriging was then used to construct maps of the spatial probability of membership in a given group (Wackernagel 2003).

3. Results

3.1. Analysis of SSR marker polymorphism in the Chinese pepper germplasm

The analysis of 28 SSR loci across the 372 Chinese pepper accessions uncovered 363 alleles. The number of distinct alleles per locus ranged from 2 to 28, with an average of 13.79 (Table 2, Fig. 1). Gene diversity and PIC indices, which represent the number of alleles and their distribution, were highly variable across the SSR loci, ranging from 0.09 to 0.92 (average, 0.62) and 0.08 to 0.92 (average, 0.59), respectively. Observed heterozygosity was also highly variable between loci (0.02–0.33) and between accessions (0.01–0.34; data not shown), indicating that the accessions (local landraces) were not fixed to homozygosity.

3.2. UPGMA analysis of the pepper accessions

At a genetic distance of 0.34, the 403 pepper accessions originating from China (n=372) and foreign countries (n=31) were divided into five groups (Appendix C). Group 1 contained only the single accession of *C. baccatum* (408), while group 2 comprised five *C. chinense* accessions (405, 402, 404, 406, and 68). Group 3 included only one accession: Bigrice (316) from Yunnan, China. Group 4 consisted of two accessions: 407 (PI634826) (*C. frutescens)* and 56 (Hot-Wuping; *C. annuum*). Group 5 comprised the remaining 394 accessions, which were all *C. annuum* except one accession which was classified as *C. frutescens* (370; Millet).

3.3. Genetic structure analysis of samples from the Chinese germplasm collection

STRUCTURE analysis using Evanno's (2005) method indicated that the best fit of the genetic structure of the 368 *C. annuum* accessions from China was a division into three groups (K=3; Appendix D and Fig. 2). Several accessions showed evidence of admixture: They partially belonged to two or three groups, as inferred by the proportion of their genome attributed to each. Group 1 included 89 accessions, of which 69 (78%) had more than 80% of their genomes represented by this group. Group 2 included 101 accessions, with 69 (68%) having more than 80% of their genomes assigned to it. Finally, 178 accessions were members of group 3, with 137 (77%) of these having more than 80% of their genomes represented by this group. Group 2 thus appeared to be somewhat more highly admixed. Allele numbers, gene diversity, and observed heterozygosity were homogeneous between groups (Table 3). Despite its higher number of alleles, group 3 had a slightly lower diversity index, which most likely represents its larger size.

3.4. Phenotypic comparison of the three STRUCTURE groups

We compared phenotypic traits among the three STRUCTURE groups using an analysis of variance, taking into account only those accessions with low admixture levels (i.e., with more than 80% of their genome belonging to a

Table 2 Characteristics of the 28 analyzed microsatellite loci

Marker	Chromosome	Number of genotypes	Number of alleles	Gene diversity (He)	Heterozygosity (Ho)	Polymorphism information content (PIC)
Epms397	P1	45	20	0.85	0.17	0.84
Gpms178	P1	74	29	0.92	0.23	0.92
Gpms6	P2	7	6	0.12	0.08	0.11
Epms755	P2	9	7	0.34	0.16	0.30
Gpms169	P2	15	9	0.41	0.17	0.38
Gpms100	P2	9	6	0.45	0.15	0.38
HpmsE1-111	P3	23	11	0.77	0.33	0.74
Epms386	P3	37	18	0.77	0.19	0.73
HpmsE008	P3	12	10	0.79	0.02	0.76
Gpms93	P3	13	10	0.79	0.08	0.76
Gpms165	P5	40	27	0.79	0.07	0.78
HpmsE088	P6	9	6	0.09	0.04	0.09
Hpms1-5	P6	53	29	0.80	0.12	0.78
Epms426	P7	22	13	0.67	0.20	0.63
Gpms161	P7	44	21	0.80	0.17	0.78
Epms310	P8	29	15	0.72	0.28	0.67
Epms342	P8	37	18	0.75	0.14	0.73
Hpms2-24	P9	14	11	0.55	0.22	0.47
Epms419	P9	25	16	0.77	0.09	0.74
HpmsE051	P9	10	7	0.48	0.17	0.42
HpmsE013	P10	17	10	0.73	0.11	0.69
Epms101	P11	11	7	0.65	0.07	0.61
Epms391	P11	6	5	0.18	0.03	0.18
Gpms29	P11	34	19	0.85	0.21	0.83
Epms331	P11	26	13	0.75	0.15	0.72
HpmsE128	P12	22	14	0.64	0.08	0.61
HpmsE064	P12	10	7	0.44	0.02	0.40
Gpms197	P12	37	22	0.70	0.08	0.69
Mean		24.64	13.79	0.63	0.14	0.60

single group) (Figs. 3 and 4). Members of group 1 were characterized by a larger plant leaf area, later flowering date, and shorter plant height. With respect to fruit traits, this group displayed a heavier fruit weight, shorter fruit length with higher pericarp thickness values, and larger fruit diameter. A large proportion of fruit from these accessions was triangular, half-long quadrangular, or square (blocky) types of sweet pepper cultivars. Group 3 mainly comprised significantly taller plants with long fruits characterized by a very high fruit length:width ratio. Fruit weights were also lower than in the other groups. This group mainly included cultivar types with small, elongated, pungent fruits. With the exception of slightly smaller leaves, group 2 members were intermediate between the other two groups with respect to all plant traits. Although their fruits also appeared to be intermediate, their thin pericarps and small, elongated fruits more closely resembled those of group 3.

3.5. Spatial analysis

The relationship between pairwise genetic similarity and geographic distance of the Chinese *C. annuum* accessions is shown in Fig. 5. According to a Mantel test, this relationship was significant (P<0.001), indicating that similarity decreased with distance. At any given distance, however, the genetic variability among pairs was large. As shown in Fig. 6, maps of the probabilities of the observed plants belonging to a given STRUCTURE-computed group clearly revealed a non-random spatial distribution of these probabilities. This finding was confirmed by the variograms computed from the observed data (Fig. 6, middle panel), which did not lie inside their confidence bands under the assumption of independence. Accessions from STRUCTURE group 1 mainly corresponded to regions of northeastern and northern China, where continental climatic conditions are predominant; these samples were mostly from Hebei, Jilin, Liaoning, Heilongjiang, and Beijing, with some from Jiangsu, Shandong, Zhejiang, Yunnan, and Tibet. Accessions from group 2 were primarily collected in central southern regions (Sichuan lowlands, Guizhou) and coastal southeastern regions that are generally humid. These regions surrounded the area of the highest probability for group 3. Accessions from group 3 were strongly concentrated in a restricted area between Guangzhou and Zhengzhou, with an additional

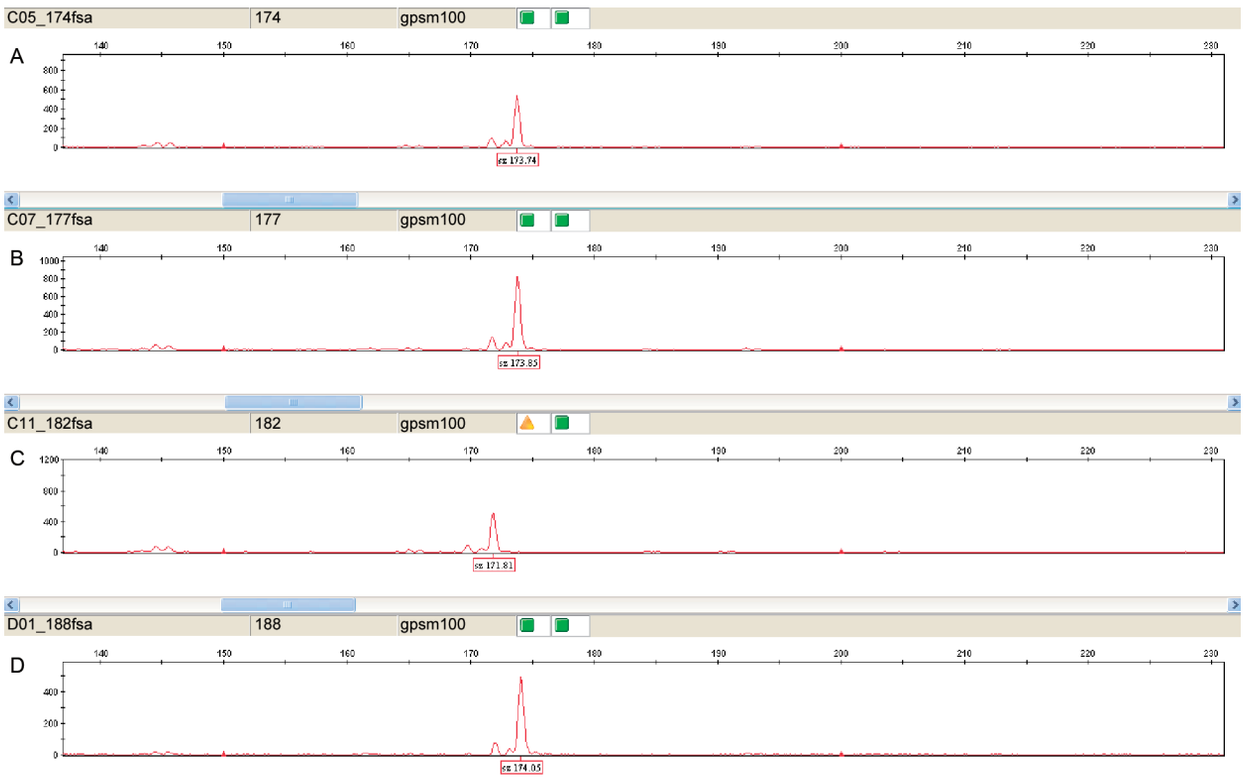

Fig. 1 Polymorphysis of simple sequence repeat (SSR) marker Gpms100 in accession. A, Chenjiasmall. B, Sifangtou-jilin. C, Ox horn Big. D, Sandaojin-Yanbian.

Table 3 Genetic diversity of the three *Capsicum annuum* groups identified from the model-based STRUCTURE analysis

Groups from structure	Number of genotypes	Number of alleles	Gene diversity (He)	Heterozygosity (Ho)	Polymorphism information content (PIC)
Group 1	11.36	7.93	0.60	0.13	0.56
Group 2	13.46	8.89	0.61	0.14	0.57
Group 3	15.46	9.50	0.55	0.15	0.52

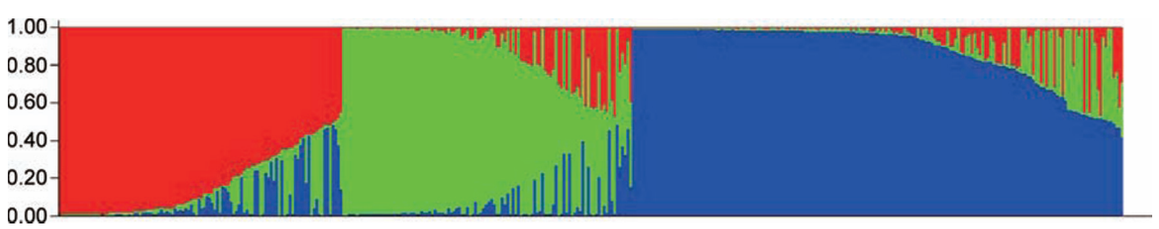

Fig. 2 Three *Capsicum annuum* groups are inferred by STRUCTURE analysis. The distribution of the 363 *C. annuum* accessions from China into distinct groups by the model-based method (STRUCTURE 2.3.4) is indicated by the color code (red=group 1; green=group 2; blue=group 3).

spot of concentration north of Shanghai.

3.6. Comparative polymorphism analysis of *Capsicum* spp. germplasm resources inside and outside of China

The screening of 28 SSR loci in 29 non-Chinese accessions yielded 192 alleles, with 4–12 distinct alleles observed per locus, and 6.86 on average. Heterozygosities ranged from 0 to 0.28, with an average of 0.07; because of the small sample size, these values cannot be directly compared to the Chinese accessions. Gene diversity index and PIC values were generally high, ranging from 0.13 to 0.84 (average, 0.64) and 0.13 to 0.82 (average, 0.61), respectively.

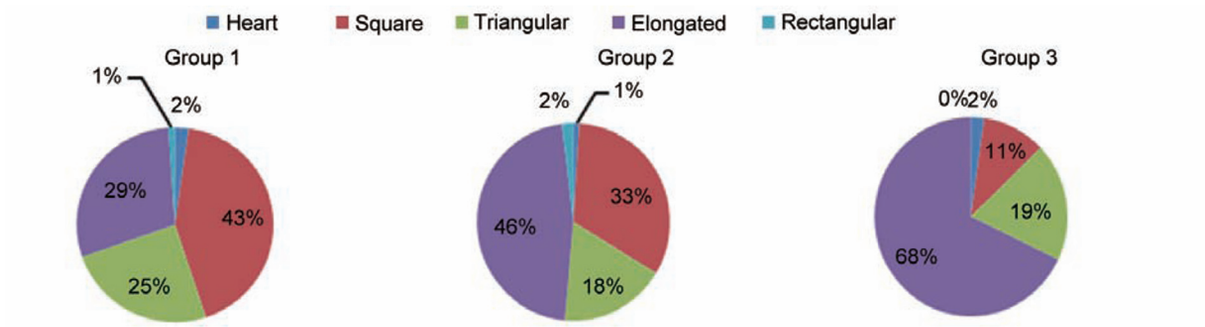

Fig. 3 Distribution of the fruit type in the three *C. annuum* groups inferred from STRUCTURE analysis.

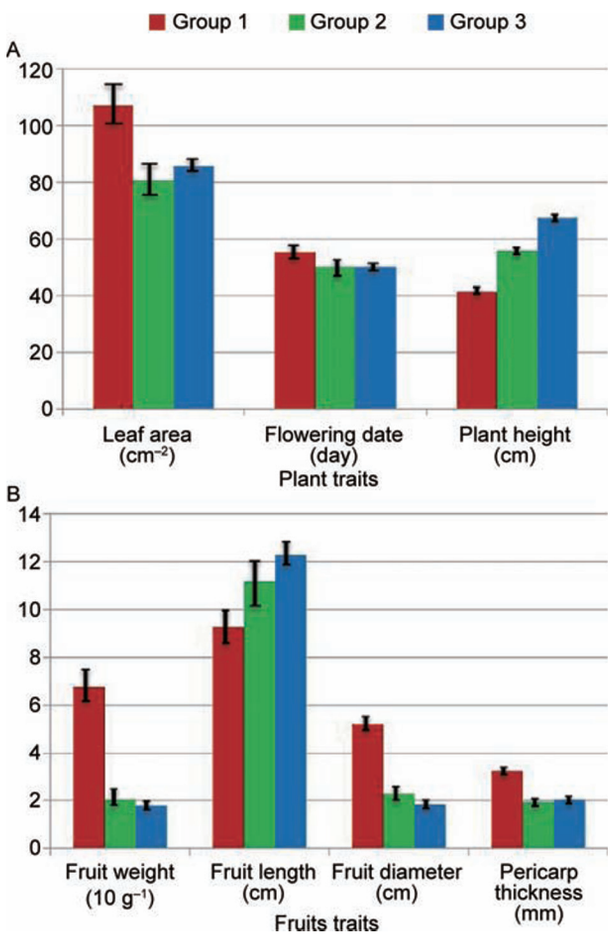

Fig. 4 Average values of seven phenotypic traits in the three *C. annuum* groups (at $P>0.9$) defined by STRUCTURE analysis. A, plant traits. B, fruit traits. Vertical bars represent confidence intervals of the mean values at $P=0.95$.

These averages were slightly higher than those of the 372 Chinese *C. annuum* accessions, suggesting that the genetic diversity among Chinese landraces is close to that of this panel of foreign accessions.

4. Discussion

Capsicum accessions were carefully chosen from more than 2 200 germplasm resources housed at the National Genebank of China (Zhang *et al.* 2010) to ensure they were representative of the 31 provinces, autonomous regions and municipalities of China from where they were originally collected. The 368 accessions of *C. annuum*, currently the most widespread species in production, were selected to avoid redundancy and maximize diversity between accessions on the basis of phenotypic traits and geographic origins. Accessions of *C. chinense* and *C. frutescens*, namely Yellow Lantern from Hainan, Xiaohuanglantern from Guizhou, and Millet from Sichuan, were also included. The collection was fingerprinted using 28 highly polymorphic SSR loci distributed over the entire pepper genome.

SSR fingerprinting has been highly successful in exploring the genetic diversity of known *Capsicum* taxonomic groups. Assigned to *C. annuum* on the basis of phenotype characteristics, the markers used in this study effectively distinguished the species *C. baccatum* from *C. chinense* and *C. annuum*, but *C. frutescens* and *C. annuum* were less distinct, possibly reflecting misclassifications or interspecific origins because these two species are cross-fertile. Because a very large majority of Chinese pepper landraces belongs to *C. annuum*, the analysis of factors impacting on diversity has to focus on this species.

Within Chinese *C. annuum*, gene flow is frequent because of breeding and natural crossing. In our study, accessions from the same province generally grouped together, with those having the same origin showing closer relationships at the molecular level. This was shown to be significant from the Mantel test (Fig. 5), in which genetic similarity between pairs of accessions decreased with geographic distances. As an example, accessions 143, 144, 149, 150, 152, 162, 165, and 169, all from Hunan Province, clustered together. Accessions between different provinces exhibited larger

genetic distances, with this genetic variation possibly shaped by different ecological conditions in the different regions or a consequence of evolution under farmer selection for horticultural and consumption quality traits. Some cultivars with distinct names collected from different provinces, such as Dandongshizi and Neimeng (persimmon), displayed very close genotypes. Their similarities may either be caused by close kinship or indicate that they are derivatives from the same cultivar. Other accessions had different names and geographic origins, but similar genotypes; examples include Ganzhou (point), Kailu (small point), Ledu (line), and Ankang (ten sisters) from Jiangxi, Inner Mongolia, Qinghai, and Shaanxi, respectively. These samples probably reflect seed exchanges between regions or multiple introductions of some very close genotypes, which differ only by a few genes impacting the phenotype (fruit color, pungency, and shape) but not the SSR haplotypes. The admixture observed between the three STRUCTURE groups demonstrates that genetic material has been exchanged between cultivars through breeding and natural recombination. Further evidence for this exchange is suggested by the UPGMA tree, where the *C*. *annuum* accessions showed a continuous distribution.

The genetic diversity of the *C*. *annuum* accessions was shown to be related to cultivar type, with two extreme phenotypes. Group 1 mainly included large, sweet-fruited cultivars with shorter plants, whereas group 3 mostly included cultivars with smaller, elongated fruits, often pungent, and taller plants. Group 2 was more difficult to characterize with intermediate types and a frequently admixed genome. This genetic structure within Chinese landraces is close to that observed in the worldwide pepper collection analyzed by Nicolai *et al.* (2013), as well as to that reported by Hill *et al.* (2013). These authors structured the 46 accessions into three groups with *K*=3; one group includes most bell peppers, another includes most small hot peppers, while the third is mixed with different types of pepper. Further, our research reveals the impact of human (farmer) selection in primary and secondary centers of diversification, driven by local adaptation but also consumption. Considering the allelic richness and diversity index (He), the values within each of the three groups of Chinese accessions (0.55 to 0.61) are close to those estimated from the worldwide accession panel of Nicolai *et al.* (2013) (0.40 to 0.64). This represents the large genetic diversity of Chinese peppers that may result from multiple and diverse introductions into the country, as well as from highly diverse agroclimatic constraints in such a large cultivation area. One surprising result was the observed higher diversity of large-fruited cultivars in China (group 1, He=0.60) compared with the corresponding cluster 3 of the worldwide panel (He=0.40),

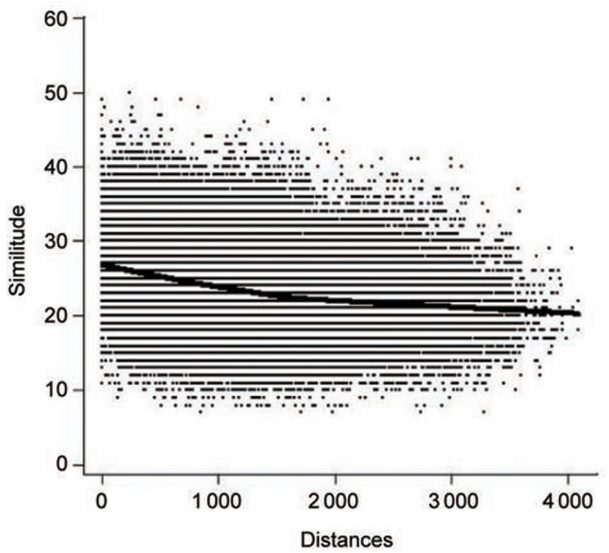

Fig. 5 Changes in genotypic similarity with respect to geographic distance. Each point represents a pair of accessions positioned according to their geographic distance and genetic similarity. The line represents the local regression between observed similarity and geographic distance.

and the reverse effect in the small, hot-fruited group 3 from China (He=0.55) compared with the corresponding cluster 1 of the worldwide panel (He=0.64). Indeed, small and pungent peppers were shown to be much more diverse at the worldwide scale than large fruited peppers, which displayed a restricted genetic variability. In China, it seems that this difference is less obvious, with a lower diversity of small-fruited landraces but higher variability in large-fruited local cultivars that may reflect importation as well as local selection events.

The genetic diversity of *C*. *annuum* accessions was also related to geographic distribution. Because the groups defined by STRUCTURE analysis were not distributed randomly with respect to geographic locations, group 1, mostly cultivated in northern and north-eastern regions, probably corresponds to the large-fruited cultivars with short and early flowering plants that are better adapted to the temperate and continental conditions of these regions. Groups 2 and 3 were more heavily represented in tropical and subtropical southeastern regions which are hotter and more humid, and where hot peppers are more frequently consumed. Group 2 was also more prevalent in coastal and central highlands, whereas group 3 was more strongly associated with the inner lowlands of Henan, Hunan, and Hubei provinces, where two crops per year are common (Li *et al.* 2014).

The Maritime Silk Road was historically one of the first trade routes into China. As expected from this, our results

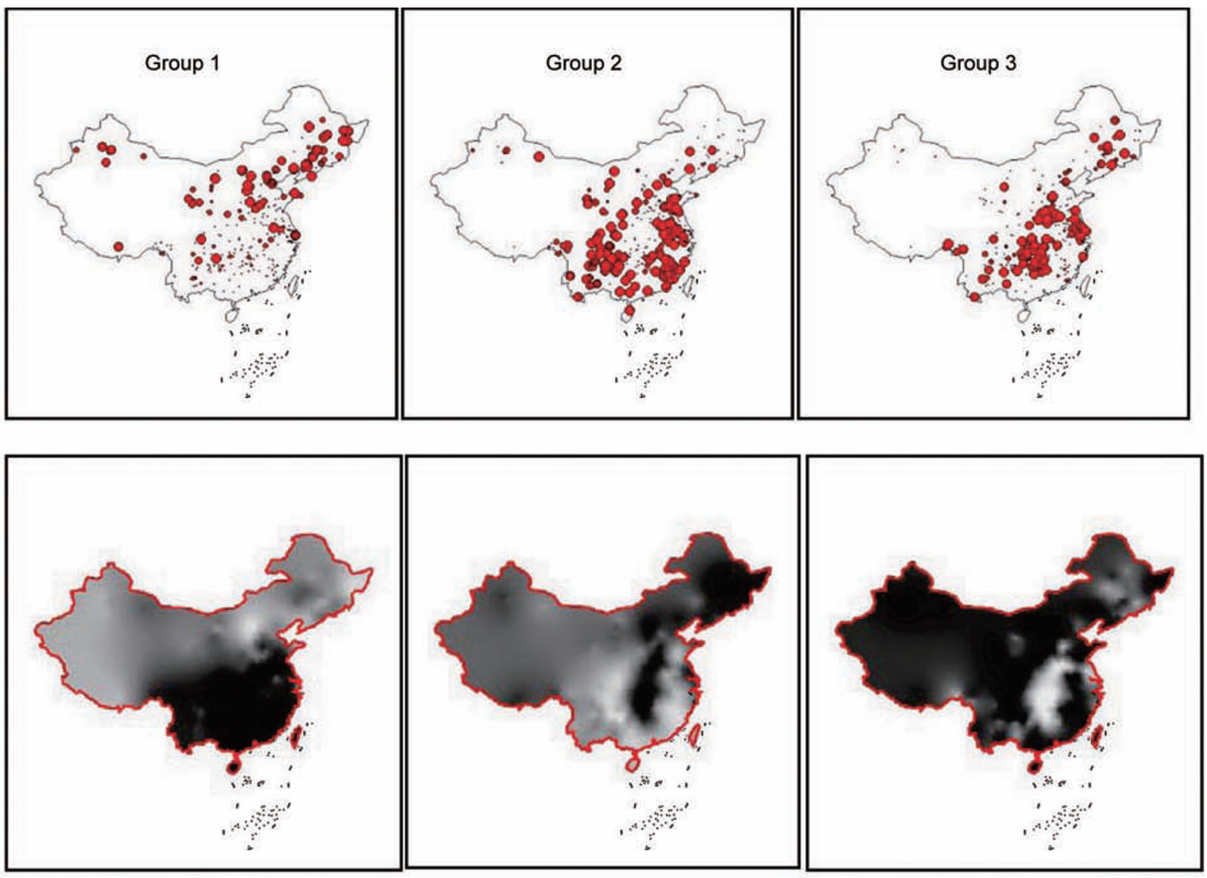

Fig. 6 Membership probabilities of accessions in each group in the mainland of China. Upper panel: spatial partitioning of estimated probabilities of membership in a given group defined by STRUCTURE; the higher the probability, the larger the circle radius. Lower panel, spatial distribution of the predicted probability of membership in a given group; the lower the probability, the darker the area.

reveal that the highest Chinese pepper genetic diversity is found in southern and southeastern China with a higher number of alleles in groups 2 and 3. For example, Millet, Xiaohuanglantern, Yellow Lantern and many other peppers are from Yunnan, Guizhou, Fujian, and Hainan provinces. To provide further insights into the introduction and evolution of peppers in China, a larger-scale evaluation of the Chinese pepper germplasm is needed. Moreover, a comparison of more germplasm resources and the incorporation of additional non-Chinese accessions should contribute to a more complete understanding of Chinese pepper germplasm diversity.

This work has potential implications for pepper breeding. With knowledge of genetic distances between accessions, we can choose those with broader genetic diversity to produce a higher degree of heterosis from crosses. Our analysis, and future related studies, will enable a core collection of the Chinese pepper collection to be established. This study therefore provides a sound basis for the further characterization of the biodiversity of this important vegetable.

5. Conclusion

The highest Chinese pepper genetic diversity is found in southern and southeastern China reveals the Maritime Silk Road was historically one of the first pepper trade routes into China. Geographic distribution corresponds to different fruits size characters. Gene flow is frequent because of breeding and natural crossing.

Acknowledgements

This work was supported by the National Key Technology Research and Development Program of China (2011BAZ01732), the Project of the Ministry of Agriculture of China (CARS-25), the Core Research Budget of the Non-profit Governmental Research Institute, China (ICS, CAAS; 1610032011011), and the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS).

Appendix associated with this paper can be available on

http://www.ChinaAgriSci.com/V2/En/appendix.htm

References

Aguilar-Meléndez A, Morrell P L, Roose M L, Kim S C. 2009. Genetic diversity and structure in semiwild and domesticated chili peppers (*Capsicum annuum*; Solanaceae) from Mexico. *American Journal of Botany*, **96**, 1190–1202.

Albrecht E, Zhang D, Saftner R A, Stommel J R. 2012. Genetic diversity and population structure of *Capsicum baccatum* genetic resources. *Genetic Resources and Crop Evolution*, **59**, 517–538.

Andrews J. 1984. *Pepper — The Domesticated Capsicum*. University of Texas Press, Austin, Texas.

Baral J B, Bosland P. 2002. An updated synthesis of the *Capsicum* genus. *Capsicum and Eggplant Newsletter*, **21**, 11–21.

Cavalli-Sforza L L, Edwards A W. 1967. Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, **19**, 233.

Chen H Z. 1688. *Hua Jing*. China. (in Chinese)

Chen X J, Chen J F, Di H, Lou Q F. 2006. RAPD diversity analysis of 5 domesticated species. *Acta Horticulturae Sinica*, **33**, 751–756.

Cressie N. 1993. *Statistics for Spatial Data*. John Wiley and Sons, Chichester.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, **14**, 2611–2620.

Fulton T M, Chunwongse J, Tanksley S D. 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter*, **13**, 207–209.

Kao L. 1591. *Tsung Sheng Pa Chien*. People's Medical Publishing House, Beijing. p. 852. (in Chinese)

González-Jara P, Moreno-Letelier A, Fraile A, Piñero D, García-Arenal F. 2012. Impact of human management on the genetic variation of wild pepper *Capsicum annuum* var. *glabriusculum*. *PLoS ONE*, **6**, 28715.

Hernández-Verdugo S, Luna-Reyes R, Oyama K. 2001. Genetic structure and differentiation of wild and domesticated populations of *Capsicum annuum* (Solanaceae) from Mexico. *Plant System Evolution*, **226**, 129–142.

Hill T A, Ashrafi H, Wo S R C, Yao J, Stoffel K, Truco M J, Kozik A, Michelmore R W, Deynze A V. 2013. Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30k unigene pepper GeneChip. *PLOS ONE*, **8**, e56200.

Ibiza V P, Blanca J, Cañizares J, Nuez F. 2012. Taxonomy and genetic diversity of domesticated *Capsicum* species in the Andean region. *Genetic Resources and Crop Evolution*, **I59**, 1077–1088.

Ince A G, Karaca M, Onus AN. 2010. Genetic relationships within and between *Capsicum* species. *Biochemistry Genetics*, **48**, 83–95.

Jeong H J, Jo Y D, Park S W, Kang B C. 2010. Identification of *Capsicum* species using SNP markers based on high resolution melting analysis. *Genome*, **53**, 1029–1040.

Li L, Mark A, Friedl M A, Xin Q, Gray J, Pan Y, Frolking S. 2014. Mapping crop cycles in China using MODIS-EVI time series. *Remote Sensing*, **6**, 2473–2493.

Li Y. 1984. *History of Crops in China*. China Science and Technology Publisher, Beijing. (in Chinese)

Luo Y D, Li J G, Li M F. 2006. Analysis of genetic diversity of *Capsicum* germplasm resources by using SSR markers. *Biotechnology Bulletin*, **S1**, 337–341.

Manly F J B. 2003. *Randomization*, *Bootstrap and Monte-Carlo Methods in Biology*. Chapman and Hall/CRC, Melbourne. p. 1991.

Nicolai M, Cantet M, Lefebvre V, Sage-Palloix A M, Palloix A. 2013. Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. *Genetic Resources and Crop Evolution*, **60**, 2375–2390.

Pacheco-Olvera A, Hernandez-Verdugo S, Rocha-Ramírez V, Gonzalez-Rodríguez A, Oyama K. 2012. Genetic diversity and structure of pepper (*Capsicum annuum* L.) from Northwestern Mexico analyzed by microsatellite markers. *Crop Science*, **52**, 231–241.

PGRI (Plant Genetic Resources Institute), AVRDC (Asian Vegetable Research and Development Center), CATIE (Center Agricultural Tropical de Investigacion Ensenanza). 1995. *Descriptors for Capsicum* (*Capsicum* spp.). International Plant Genetic Resources Institute, Rome, Italy.

Pickersgill B. 1991. Cytogenetics and evolution of *Capsicum* L. In: Tsuchiya T, Gupta P K, eds., *Chromosome Engineering in Plants*: *Genetics*, *Breeding*, *Evolution*. Elsevier, Amsterdam. pp. 139–160.

Pickersgill B, Heiser C B, McNeill J. 1979. Numerical taxonomic studies of variation and domestication in some species of *Capsicum*. In: Hawkes J G, ed., *The Biology and Taxonomy of the Solanaceae*. Academic Press, New York. pp. 679–690.

Pritchard J K, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Purseglove J M. 1968. *Tropical Crops*. vol. 2. John Wiley and Sons, New York.

Rodriguez J M, Berke T, Engle L, Nienhuis J. 1999. Variation among and within *Capsicum* species revealed by RAPD markers. *Theoretical and Applied Genetics*, **99**, 147–156.

Somos A. 1984. *The Paprika*. Akademiai Kiado, Budapest. p. 302.

Tamura K, Dudley J, Nei M, Kumar S. 2004. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology Evolution*, **24**, 1596–1599.

Toquica S P, Rodriguez F, Martinez E, Duque M C, Tohme J. 2003. Molecular characterization by AFLPs of *Capsicum* germplasm from the Amazon department in Colombia. *Genetic Resources and Crop Evolution*, **50**, 639–647.

Votava E J, Nabhan G P, Bosland P W. 2002. Genetic diversity

and similarity revealed *via* molecular analysis among and within an *in situ* population and *ex situ* accessions of chiltepin (*Capsicum annuum* var. *glabriusculum*). *Conservation Genetics*, **3**, 123–129.

Wackernagel H. 2003. *Multivariate Geostatistics*. Springer, Berlin. p. 899.

Walsh B M, Hoot S B. 2001. The phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two non-coding regions: The chloroplast atpB-rbcL spacer region and nuclear waxy introns. *International Journal of Plant Sciences*, **162**, 1409–1418.

Yi G, Lee J M, Lee S, Choi D, Kim B D. 2006. Exploitation of pepper EST-SSRs and an SSR-based linkage map. *Theoretical and Applied Genetics*, **114**, 113–130.

Zhang B X, Wang L H, Mao S L, Zhang Z H. 2010. Research progress on pepper breeding and genetic during China's Eleventh Five-year Plan. *China Vegetables*, **24**, 1–9. (in Chinese)

Zheng N. 2006. Thought and discussion on the introduction *Capsicum*. *Agricultural Archaeology*, **4**, 177–184.

Zhou J, Shen H L, Yang W C, Tan F, Wang Y L, Guo S. 2009. Analysis of genetic diversity of *Capsicum* germplasm by using SSR markers. *Acta Agriculturae Boreali-Sinica*, **24**, 62–67. (in Chinese)

Zou X X. 2009. *Genetics and Breeding of Pepper*. Scientific Publisher, Beijing. (in Chinese)

(Managing editor WENG Ling-yun)