Genes & Genomics (2010) 32: 207-215
DOI 10.1007/s13258-009-0888-6

SHORT COMMUNICATION

Analysis of genetic diversity among indigenous landraces from sesame (*Sesamum indicum* L.) core collection in China as revealed by SRAP and SSR markers

Yan-xin Zhang · Xiu-rong Zhang · Wei Hua · Lin-han Wang · Zhuo Che

Received: 22 December 2009 / Accepted: 12 February 2010 / Published online: 30 June 2010

Abstract
The molecular genetic diversity of 404 indigenous landraces from sesame core collection in China were evaluated by 11 SRAP and 3 SSR markers, 175 fragments were generated, of which 126 were polymorphic with an average polymorphism rate of 72%. Jaccard's genetic similarity coefficients (GS=0.7130), Nei's gene diversity (h=0.2418) and Shannon's Information index (I=0.3847) were calculated, a dendrogram of the 404 landraces was made, landraces from various zones were distributed throughout the dendrogram, accessions from different agro-ecological zones were indistinguishable by cluster analysis, geographical separation did not generally result in greater genetic distance, a similar pattern was obtained using principal coordinates (PCO) analysis. As to seven agro-ecological zones, the maximum Nei's gene diversity (h = 0.2613) and Shannon index (I = 0.3980) values in zone VII indicated that they were genetically more diverse than those in other zones, while the least genetically diverse region was zone III (h = 0.1772, I = 0.2858). Nei's genetic identity and genetic distance among landraces from seven agro-ecological zones were also analyzed, the genetic relationship of seven zones was inferred using the UPGMA method. This study demonstrated that SRAP and SSR markers were appropriate for evaluation of sesame genetic diversities. There existed extensive genetic diverse among indigenous landraces and the abundance of genetic diversity of landraces in different agro-ecological zones was various. Understanding of these characteristics of indigenous landraces in China can provide theoretical foundation for further collection, effective protection and reasonable utilization of these sesame landraces in breeding.

Keywords Core collection; SSR; Genetic diversity; Landraces; Sesame; SRAP

Y.-X. Zhang · X.-R. Zhang(✉) · W. Hua · L.-H. Wang · Z. Che
Oil Crops Research Institute Chinese Academy of Agricultural Sciences, Key Laboratory of Oil Crop Biology of the Ministry of Agriculture, No.2 Xudong 2nd Rd, Wuhan 430062, China
e-mail: zhangxr@oilcrops.cn

Introduction

Sesame (*Sesamum indicum* L.) is one of the four major oil crops (rapeseed, soybean, peanut and sesame) with nutritional value in China. It has been cultivated for about 2200 years, while China has been identified as one of the five sesame diversity centres in classical studies (Zeven and Zhukovsky, 1975; Hawkes, 1983).

A sesame core collection containing 453 accessions (404 indigenous landraces, 18 indigenous released cultivars, 31 exotic accessions) was established from 4251 accessions (Zhang et al., 2000). The 404 indigenous landraces were from seven agro-ecological zones (29 provinces) in China, which were divided according to climatic and geographic characteristics, as well as the planting system. The 404 indigenous landraces in this sesame core collection represent well the genetic variability of the whole indigenous sesame landraces in China.

Though the quantity of indigenous sesame landraces is considerable in China, the diversities in morphological traits, phenotypic traits, agronomic traits, quality traits and resistance to stress or diseases among them have been evaluated, three sesame germplasm catalogues were edited and published (Chen, 1981; Feng, 1992, 1997). Dong (1990) identified resistance to *Fusarium oxysporium* of several Chinese sesame germplasm, Li et al. (1991) identified some Chinese sesame germplasm of their resistance to *Macrophomina phaseolina*, Feng et al. (1991) identified and evaluated tolerance to water-logging stress of several Chinese sesame germplasm, which were

Table 1. Description of 404 accessions used in this study

Pop. code	Acc. No.	Varietal type	Origin	Annotation
A	42	Landrace	Northeast and northwest of China	agro-ecological zone I
B	78	Landrace	Northern China	agro-ecological zone II
C	55	Landrace	The Yellow River and Huai River valley	agro-ecological zone III
D	69	Landrace	The Yangtze and Han River valley	agro-ecological zone IV
E	56	Landrace	Middle and lower of Yangtze valley	agro-ecological zone V
F	57	Landrace	South-central of China and southern China	agro-ecological zone VI
G	47	Landrace	Southwest of China	agro-ecological zone VII

the prior studies of sesame accessions resistance to disease or stress. Xiao et al. (1992) analyzed the main characteristics of 372 black seed coat sesame accessions in China, including morphological traits, agronomic traits, quality traits and resistances, Feng et al. (1996) carried out further characterization and evaluation of 225 pre-selected sesame accessions in China, with 3 repeats at 14 experimenta1 stations in different ecological regions from 1992 to 1994. They screened for a set of accessions with certain excellent agronomic traits or resistance characters which could be used as parents material in breeding. Zhang et al. (2001) evaluated resistance to *Macrophomina phaseolina* and *Fusarium oxysporium* of 85 sesame accessions primarily selected, elucidated resistance and application potential of disease resistant accessions. They also found that there were isoenzyme bands similarity and differences between disease resistant and susceptible accessions in some zone. Although such many traits among Chinese sesame germplasm have been studied, morpho-agronomic markers may not reflect true genetic identities and diversities (Ferdinandez et al., 2001). Information on Chinese sesame germplasm genetic diversity at molecular level is limited and evaluation of molecular genetic diversity among the whole indigenous sesame landraces is especially absent. However, molecular markers are simple, informative, and portable compared to morphological and phenotypic traits. The utilization of molecular markers have been effective in evaluating genetic variation within species, because they are not affected by the environmental conditions in which the plants are grown (Powell et al., 1996).

This paper represents the latest study of the molecular genetic diversity among indigenous landraces from sesame core collection in China. The objectives of this investigation were to explore the characteristics of genetic diversity and their relationships among landraces from different agro-ecological zones in China, then provide theoretical foundation for utilizing these landraces effectively and carrying out the molecular biological studies purposefully.

Materials and Methods

Plant Materials

All of the 404 indigenous landraces from the sesame core collection in China were used in this study (Table 1). They were divided into 7 populations (with population code A to G) according to the 7 agro-ecological zones (I to VII) in China (Zhang et al., 2000). Young healthy leaves from plants of each accession were collected and conserved under -80℃ until DNA extraction.

DNA extraction, polymerase chain reactions (PCR) amplification and electrophoresis

Total genomic DNAs of 404 accessions were prepared from young leaves according to CTAB method (Doyle and Doyle, 1987) with some modification.

A total of 36 SRAP primer combinations and 10 SSR primer pairs were used (Table 2). The SRAP and SSR primer sequences were referred to Li et al. (2001) and Dixit et al. (2005), respectively. PCR of SRAP markers were conducted in 20 μL solution containing 80 ng of DNA, 50 ng of forward primers, 50 ng of reverse primers, 1 × buffer (MBI), 4 mmol of $Mg2^{+}$, 0.40 mmol of dNTPs, and 1 U Taq polymerase (MBI). The PCR profile was an initial denaturation of 94℃ for 2 min, followed by 4 cycles of 94℃ for 1 min, 35℃ for 1 min, 72℃ for 1 min, then 34 cycles of 94℃ for 1 min, 50℃ for 1 min, 72℃ for 1 min, and a final incubation at 72℃ for 5 min, then 4℃ thereafter. According to Zhang et al. (2008), PCR of SSR markers were conducted in 20 μL solution containing 25 ng of DNA, 4 μmol of forward primers, 4 μmol of reverse primers, 1 × buffer (MBI), 2 mmol of Mg^{2+}, 0.25 mmol of dNTPs, and 0.80 U Taq polymerase (MBI). The PCR profile was an initial denaturation of 94℃ for 3 min, followed by 34 cycles of denaturing at 94℃ for 50 sec, annealing at appropriate temperature for 45 sec, extension at 72℃

Table 2. Sequences of primers used in this study.

Primer code	Forward primer		Reverse primer
	Primer sequence (5'—3')	Primer code	Primer sequence (5'—3')
Me01	TGAGTCCAAACCGGATA	Em01	GACTGCGTACGAATTAAT
Me02	TGAGTCCAAACCGGAGC	Em02	GACTGCGTACGAATTTGC
Me03	TGAGTCCAAACCGGAAT	Em03	GACTGCGTACGAATTGAC
Me04	TGAGTCCAAACCGGACC	Em04	GACTGCGTACGAATTTGA
Me05	TGAGTCCAAACCGGAAG	Em05	GACTGCGTACGAATTAAC
Me06	TGAGTCCAAACCGGTAA	Em06	GACTGCGTACGAATTGCA
Me07	TGAGTCCAAACCGGTCC	Em07	GACTGCGTACGAATTCAA
Me08	TGAGTCCAAACCGGTGC	Em08	GACTGCGTACGAATTCTG
Me09	TGAGTCCAAACCGGACG	Em09	GACTGCGTACGAATTCGA
Me10	TGAGTCCAAACCGGACT	Em10	GACTGCGTACGAATTCAG
Me11	TGAGTCCAAACCGGAGG	Em11	GACTGCGTACGAATTCCA
GBssr-sa-05-F	TCATATATAAAAGGAGCCCAAC	GBssr-sa-05-R	GTCATCGCTTCTCTCTTCTTC
GBssr-sa-08-F	GGAGAAATTTTCAGAGAGAAAAA	GBssr-sa-08-R	ATTGCTCTGCCTACAAATAAAA
Sesame-09-F	CCCAACTCTTCGTCTATCTC	Sesame-09-R	TAGAGGTAATTGTGGGGGA
GBssr-sa-33-F	TTTTCCTGAATGGCATAGTT	GBssr-sa-33-R	GCCCAATTTGTCTATCTCCT
GBssr-sa-72-F	GCAGCAGTTCCGTTCTTG	GBssr-sa-72-R	AGTGCTGAATTTAGTCTGCATAG
GBssr-sa-108-F	CCACTCAAAATTTTCACTAAGAA	GBssr-sa-108-R	TCGTCTTCCTCTCTCCCC
GBssr-sa-123-F	GCAAACACATGCATCCCT	GBssr-sa-123-R	GCCCTGATGATAAAGCCA
GBssr-sa-173-F	TTTCTTCCTCGTTGCTCG	GBssr-sa-173-R	CCTAACCAACCACCCTCC
GBssr-sa-182-F	CCATTGAAAACTGCACACAA	GBssr-sa-182-R	TCCACACACAGAGAGCCC
GBssr-sa-184-F	TCTTGCAATGGGGATCAG	GBssr-sa-184-R	CGAACTATAGATAATCACTTGGAA

for 1 min, and a final incubation at 72℃ for 10 min, then 4℃ thereafter. PCR were carried out in a 96 well plate in a PTC-100 thermocycler (MJ Research, Watertown, MA).

PCR products were size-separated on 6% denaturing polyacrylamide gels. The electrophoresis parameters and silver staining of gels were based on the protocols of Lin et al. (2005).

Data analysis

After silver staining, all major DNA band sizes were recorded using 100 bp DNA ladder as the reference. The amplified DNA fragments were recorded as either 1 or 0, respectively, representing presence or absence of the band.

Pairwise genetic similarity coefficient (GS) was calculated using Jaccard's coefficient by the SIMQUAL program of the NTSYS-pc software package version 2.1 (Rohlf, 2000). Afterwards, a dendrogram (Sokal and Michener, 1958) was constructed based on the genetic similarity matrix by the UPGMA algorithm and generated by the SHAN clustering program (Sneath and Sokal, 1973) of the NTSYS-pc. Subsequently, principal coordinate (PCO) analysis was carried out using principal component analysis programs such as DCENTER and EIGEN of the NTSYS-pc and based on genetic similarity matrices. Finally a scatter plot was obtained. Nei's gene diversity (*h*), Shannon's information index (I), Nei's genetic identity and genetic distance were estimated using POPGENE version 1.32 (Yeh and Boyle, 1997). Duncan's multiple range test was performed by SPSS software. Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

Results

Level of polymorphism

Total of 36 SRAP primer combinations and 10 SSR primer pairs were employed randomly to screen polymorphism between 12 accessions (typical accessions from population A to G). Among of them, 11 SRAP primer combinations and 3 SSR primer pairs (Table 3) amplified abundant, clear and repeatable fragments, then they were employed to evaluate the genetic diversity of 404 landraces.

A total of 175 amplified fragments were detected and 126 of them were polymorphic, the polymorphism rate was 72%.

Table 3 Amplification of primers used in this study

Primer	Loci size (bp)	Total loci	Monomorphic loci No.	Polymorphic loci No.	Polymorphism rate (%)
Me01Em01	440-750	10	1	9	90.00
Me02Em02	450-540	11	5	6	54.55
Me02Em03	350-640	8	2	6	75.00
Me03Em02	220-1300	14	3	11	78.57
Me06Em05	390-700	10	6	4	40.00
Me07Em06	480-900	26	14	12	46.15
Me08Em04	150-610	14	4	10	71.43
Me08Em05	210-590	23	6	17	73.91
Me09Em06	180-650	14	4	10	71.43
Me09Em08	150-650	16	2	14	87.50
Me10Em10	180-920	14	2	12	85.71
GBssr-sa-123	290-360	5	0	5	100.00
GBssr-sa-173	250-270	2	0	2	100.00
GBssr-sa-182	250-550	8	0	8	100.00
Total	150-1300	175	49	126	72.00

The number of fragments detected by each primer ranged from 2 (GBssr-sa-173) to 26 (Me07Em06), with an average of 12.5. The number of polymorphic fragments detected by each primer ranged from 2 (GBssr-sa-173) to 17 (Me08Em05), with an average of 9, Product sizes ranged from 150 to 1300 bp (Table 3).

Genetic diversity analysis

Genetic similarity coefficient (GS), Nei's gene diversity (h) and Shannon's information index (I) for variation were estimated for all agro-ecological zones (Table 4). The average pairwise GS (0.7865) was the highest in zone III (also with the highest minimum pairwise GS 0.5405), indicated that genetic relationship between landraces in this zone was the nearest. The followings were average pairwise GS in zone I (0.7447), zone II (0.7339), zone IV (0.7214), zone VI (0.7208), zone V (0.7030) and zone VII (0.6857), the pairwise GS of the 404 landraces in all of the seven zones was ranged from 0.3937 to 0.9892, with an average of 0.7130. For all of the 404 landraces, the average Nei's gene diversity and Shannon index for the 175 loci was 0.2418 and 0.3847, respectively. The maximum Nei's gene diversity (h = 0.2613) and Shannon index (I = 0.3980) values among the landraces in zone VII indicated that they were genetically more diverse than those in other zones, while the least genetically diverse region was zone III (h = 0.1772, I = 0.2858). The h and I values of the other five zones decreased ordinally from zone V to zone VI, IV, II and I (data shown in Table 4). The trend of this variation consistent with that of pairwise GS above, showing that genetic diversity of the seven zones were various, and the abundant of genetic diversity were decreased from zone VII to zone V, VI, IV, II, I and III.

Table 4. Description of genetic diversity of each agro-ecological zone

Zone	Max. GS*	Min. GS	Aver. GS	na*	ne*	*h**	I*
I	0.9892	0.4426	0.7447	1.8333	1.3319	0.2119	0.3397bcd
II	0.9765	0.4370	0.7339	1.9444	1.3294	0.2138	0.3408bcd
III	0.9697	0.5405	0.7865	1.7857	1.2731	0.1772	0.2858d
IV	0.9789	0.4344	0.7214	1.8968	1.3593	0.2293	0.3600ab
V	0.9892	0.4000	0.7030	1.8333	1.4193	0.2511	0.3829ab
VI	0.9789	0.4359	0.7208	1.8570	1.3770	0.2320	0.3626ab
VII	0.9592	0.4711	0.6857	1.8333	1.4375	0.2613	0.3980a
I-VII	0.9892	0.3937	0.7130	2.0000	1.3792	0.2418	0.3847

Letters marked with * were abbreviations: GS indicates Genetic similarity coefficient; Na indicates Observed number of alleles; Ne indicates Effective number of alleles; h indicates Nei's gene diversity; I indicates Shannon's Information index.
Different letters show significant difference by Duncan's multiple range test ($P<0.05$) for the values within the eighth column.

Duncan's multiple range test of the I values from seven zones was performed (Table 4). Although differences between zone I, II and III, and difference between zone IV, V, VI and VII were not significant, zone III differed significantly from zone IV, V, VI and VII, and zone VII differed significantly from zone I, II and III.

Clustering analysis based on UPGMA method

The UPGMA-based dendrogram obtained from the SRAP and SSR data was shown in Figure 1, the distribution of landraces (origin from population A to G) on the dendrogram was shown

Table 5. Description of the distribution of accessions in the dendrogram.

Cluster	Group	Sub-group	Frequency							Total No. of Acc
			Pop. A	Pop. B	Pop. C	Pop. D	Pop. E	Pop. F	Pop. G	
	I-1	I-1-1	23	30	38	37	24	30	22	204
		I-1-2	7	29	13	15	12	11	5	92
		I-1-3	.	.	.	1	.	.	.	1
		I-1-4	.	2	.	.	.	.	1	3
		I-1-5	.	.	1	.	.	.	.	1
		I-1-6	4	.	.	.	1	.	.	5
		I-1-7	.	1	1	.	1	.	.	3
I		I-1-8	.	.	.	.	3	2	4	9
		Subtotal	34	62	53	53	41	43	32	318
		I-2-1	1	1	.	2	.	2	.	6
		I-2-2	.	2	1	2	1	.	.	6
	I-2	I-2-3	1	.	.	.	.	.	.	1
		I-2-4	1	.	.	.	.	.	.	1
		Subtotal	3	3	1	4	1	2	.	14
	Subtotal		37	65	54	57	42	45	32	332
		II-1-1	5	2	.	2	9	5	2	25
		II-1-2	.	.	.	.	.	.	5	5
	II-1	II-1-3	.	1	1	1	.	1	2	6
		II-1-4	.	3	.	5	4	4	1	17
		Subtotal	5	6	1	8	13	10	10	53
		II-2-1	.	2	.	.	.	.	.	2
II		II-2-2	.	3	.	1	.	.	.	4
	II-2	II-2-3	.	1	.	2	.	1	.	4
		II-2-4	.	.	.	.	.	.	3	3
		Subtotal	.	6	.	3	.	1	3	13
	II-3		.	.	.	1	1	1	2	5
	II-4		.	1	.	.	.	.	.	1
	Subtotal		5	13	1	12	14	12	15	72

in Table 5. The dendrogram divided the 404 landraces into two clusters (cluster I and cluster II) at genetic similarity of 0.59. Cluster I was a robust cluster (with 332 landraces), which was more than three times larger than cluster II (with 72 landraces). It could be grouped into two groups (I-1, I-2) at genetic similarity of 0.66, and consisted of twelve sub-groups (I-1-1 to I-1-8, I-2-1 to I-2-4),. Cluster II could be grouped into four groups (II-1, II-2, II-3, II-4) at genetic similarity of 0.66, and group II-1 and II-2 consisted of eight sub-groups (II-1-1 to II-1-4, II-2-1 to II-2-4).

The obvious character of the dendrogram was that group I-1 in cluster I was represented by the largest number (78.71%) of accessions, but this sub-group exhibited the least amount of genetic diversity. The sub-group I-1-1 comprised 204 landraces (more than 50%) out of 404 landraces analyzed, which were landraces from zone II, III and IV. There were 92 landraces in sub-group I-1-2 and it was also mainly consisted of landraces from zone II, III and IV, as well as including some landraces from zone II. The other sub-groups in cluster I comprised of only a small quantity of landrace each. By summed up all the distribution of 332 landraces in cluster I, it is definite that most of landraces from zone II, III and IV were included in cluster I, accounting for 83.33%, 98.18% and 82.61% of landraces respectively. There were 25 landraces in sub-group II-1-1, nearly 36% of them were from zone V. Sub-group II-1-2 was consisted of only 5 landraces from zone VII. Except for sub-group II-1-4 comprising 17 landraces, the remaining sub-groups and group II-3, II-4 of cluster II comprised of only a few of accessions each. Summarized the distribution of 72 landraces in cluster II, it was mainly consisted of landraces

Table 6. Unbiased measures of identity and genetic distance among seven agro-ecological zones of sesame landraces.

	Zone I	Zone II	Zone III	Zone IV	Zone V	Zone VI	Zone VII
Zone I		0.9786	0.9750	0.9770	0.9721	0.9696	0.9621
Zone II	0.0216		0.9755	0.9835	0.9727	0.9672	0.9571
Zone III	0.0253	0.0248		0.9780	0.9652	0.9720	0.9524
Zone IV	0.0233	0.0166	0.0222		0.9863	0.9799	0.9694
Zone V	0.0283	0.0276	0.0354	0.0138		0.9809	0.9794
Zone VI	0.0309	0.0333	0.0284	0.0204	0.0193		0.9785
Zone VII	0.0387	0.0439	0.0488	0.0311	0.0208	0.0217	

from zone V and VII. Accessions from various zones were distributed throughout the dendrogram.

Principal coordinates (PCO) analysis

The PCO analysis indicated that, the first and the second principal component accounted for 22.06% and 10.09% of the total variation (Fig. 2), respectively, while the third principal component accounted for an additional 4.59%. The contributions of the remaining principal components were less than 3.18% individually. The scatter plot of the first two principal components is presented in Figure 2. The scattering pattern in Figure 2 agrees closely with the clustering displayed in the dendrogram (Fig. 1). Accessions from the major cluster in the dendrogram seemed to form a very close cluster in the PCA scatter plot (indicated by ellipses numbered with A to F). Accessions clustered in ellipses A, B, C, D, E and F were basically from subgroup I-1-1, I-1-2, I-1-3 to I-1-8 and group I-2, II-1, II-2 to II-4 of the dendrogram correspondingly.

Analysis of Nei's genetic identity and genetic distance among seven agro-ecological zones

Nei's genetic identity and genetic distance among seven agro-ecological zones of sesame landraces were analyzed and summarized in Table 6. Genetic distances among agro-ecological zones are very low, zone IV and zone V pursued the lowest genetic distance (0.0138) as well as the highest genetic identity (0.9863), while zone III and zone VII pursued the highest genetic distance (0.0488) as well as the lowest genetic identity (0.9524).

According to the Nei's genetic distance shown in Table 6, the genetic relationship of landraces in seven agro-ecological zones was inferred using the UPGMA method, while optimal dendrogram was shown in Figure 3. It could be grouped into two clusters, the first cluster comprised landraces from Northern China (zone I and zone II), as well as the Yellow River and Huai River valley (zone III), while the second cluster was consisted of landraces from the Yangtze River valley (zone IV and zone V), Southern China (zone VI), and Southwest of China (zone VII). Two distinct zones, zone III and zone VII, formed two separate branches. However, zones with similar eco-geographical factors were grouped into the same or adjacent groups, such as zone I and II, as well as zone IV, V, and VI.

Discussion

Advantages of markers used in this study

DNA markers are powerful and reliable tools for evaluating genetic variation within or between populations (Powell et al., 1996; Qamaruz et al., 1998). Genetic variability in sesame has been previously studied by isozymes (Isshiki and Umezake, 1997) and many DNA markers, including RAPD (Bhat et al., 1999; Ercan et al., 2004; Pham et al., 2009), ISSR (Kim et al., 2002) and AFLP (Hernan and Petr, 2006). However, this is the latest study of utilization of the SRAP and SSR markers on sesame genetic diversity. SRAP is a promising technique for the characterization of genetic diversity because it possesses a high degree of reproducibility and discriminatory

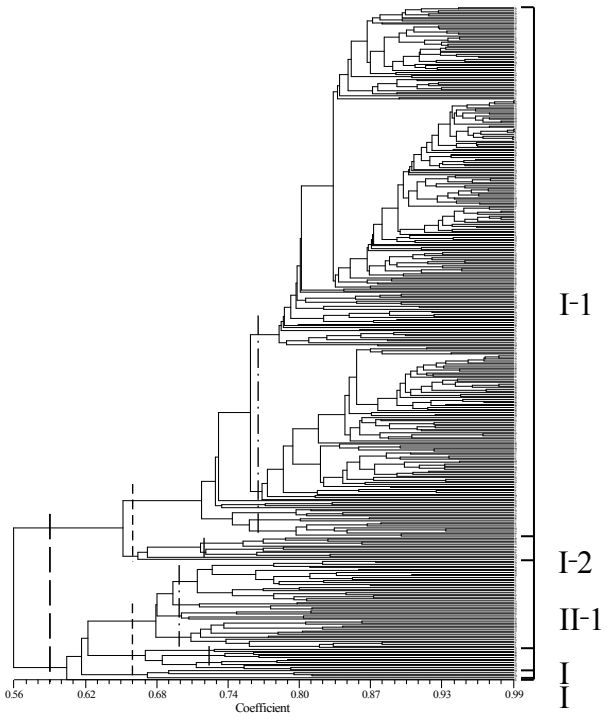

Figure 1. Dendrogram for indigenous landraces in sesame core collection. Dendrogram was constructed by UPGMA method based on SRAP and EST-SSR markers. Accessions were showed only by their population code according to Table 1

power, as well as high polymorphism rate. SSR can be developed easily at a lower cost, often have putative functions, and are more valuable for genetic diversity analysis, comparative mapping and marker assisted selection (MAS) breeding since they are derived from transcripts. SRAP and SSR markers have been successfully applied to many cultivated and wild plants, however, the application of them on sesame was scarce. This study was the latest report of estimating genetic diversity in sesame by SRAP and SSR markers, the polymorphic fragments amplified by each primer was with an average of 9, the polymorphism rate of amplification reached 72%, so they represented as excellently as in other plants.

Analysis of various genetic diversity characteristics in zone III and zone VII

The diverse abundant of genetic diversity of landraces in different agro-ecological zones of China resulted from many reasons, including ecogeographic factors and sesame cultivation history and so on. China is famous for its vast territory, therefore various planting systems were formed in different agro-ecological zones with different climatic, ecologic and geographic characters. Thus it is probable that main crops and planting systems in different provinces from the same agro-ecological zone were diverse. Sesame has been cultivated in China for about 2200 years. The agro-ecological zone III, which comprised of Henan province and Anhui province, was the main area of sesame production in China from old. The eco-geographic factors were really suitable for sesame growth, its sesame sowing area in zone III to produce nearly half of the total sesame production in China. The crop planting system in zone III was two-maturity per year and sesame was summer planting, which was highly popular in this zone. Farmers reciprocally introduced landraces frequently, this could deeply decrease the genetic diversity of landraces there. Since agro-ecological zone III was the main area of sesame production in China, cultivar improvement was regarded and started from the 1950s, many excellent released cultivars such as "786" and "332" which suited for the production of that time were bred, and consequently some landraces were substituted gradually by released cultivars. In China the indigenous sesame germplasm resources were collected since the 1950s and 3200 accessions were assembled up to the year of 1963. It is a pity that only less than half of them were conserved by the end of the 1970s, subsequently large scale of sesame germplasm further collection was carried out with support of the Chinese government, but many old landraces (especially those from zone III) had disappeared then, so the genetic diversity of landraces from zone III was limited (h=0.1772, I=0.2858). In contrast with it in zone III, landraces in zone VII displayed prevalence of higher diversity (h=0.2613, I=0.3980). The vast area frome sichuan, Guizhou, Yunnan to Tibet, the particularity and complexity of geographical and climatic factors in zone VII could partly account for the higher genetic diversity there, diverse planting system were formed. Summer planting, spring planting and autumn planting of sesame existed simultaneously in different regions. Furthermore, since sesame was not the main crop in zone VII, cultivar improvement was not regarded as much, many old landraces had been being cultivated, while the higher genetic diversity could be retained.

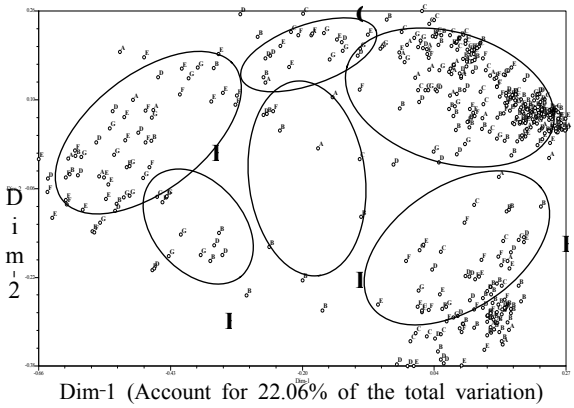

Figure 2. The scatter plot of principal coordinates analysis for indigenous landraces in sesame core collection. The scatter plot was constructed based on SRAP and EST-SSR markers. The major clusters were indicated by ellipses numbered with A to F.

Impact of geographical factors on sesame landraces in China

In this study, accessions from various zones were distributed throughout the dendrogram (Fig. 1), indicating accessions from different agro-ecological zones were indistinguishable by cluster analysis, which agreed with others in showing that geographical separation did not generally result in greater genetic distance (Ercan et al., 2004; Hernan and Petr, 2006; Venkataramana et al., 1999). This could be ascribed to the long cultivation history of sesame in China, landraces intoduce, exchanged, or traded from different agro-ecological zones. This study also analyzed characteristics about the genetic diversity, genetic identity and genetic distance among seven agro-ecological zones in China (Table 4, Table 6 and Figure 3), which displayed a decrease trend of genetic diversity from southwest region to southern region and northern region. These results provided some information in DNA level, which may be helpful reference in future study on the exploration the origin, the genetic diversity centre and the spreading of sesame landraces in China.

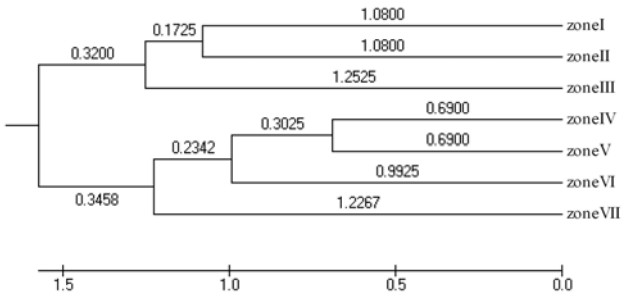

Figure 3. Dendrogram showing genetic relationships among sesame landraces from the seven agro-ecological zones in China.

Theoretical significance for sesame germplasm and breeding

As the depth of genetic diversity of landraces in different agro-ecological zones of China was various, and geographical separation did not generally result in greater genetic distance, it is necessary to reinforce the collection and protection of sesame germplasm resources from agro-ecological zones with higher diversity, such as southwest region of China. Furthermore, parent materials of sesame breeding could be selected from the regions mentioned above, genetic distance of parents should be considered firstly, while geographic distance is not so important.

Acknowledgements This research was financially supported by the Item of National Science and Technology Infrastructure Construction (2005DKA21001-20), the Item of National Plant Germplasm Resources Protection (NB08-2130135-34,35), and the Bioversity International was also appreciated.

References

Bhat V, Babrekar P and Lakhanpaul S (1999) Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Euphytica 110: 21-33.

Chen CY (1981) *The Catalogue of Sesame Germplasm Research in China*. Agricultural Scientech Press, Beijing.

Dixit A, Jin MH, Chung JW, Yu JW, Chung HK, Ma KH, Park YJ and Cho EG (2005) Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). Mol. Ecol. Notes 5: 736-738.

Dong ZX (1990) Identification of Chinese sesame germplasm resistance to *Fusarium oxysporium*. Oil Crops China 3: 74-76.

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

Ercan AG, Taskin M and Turgut K (2004) Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers. Genet. Resour. Crop Evol. 51: 599-607.

Feng XY (1992) *The Catalogue of Sesame Germplasm Research in China (continuation 1)*. Agricultural Scientech Press, Beijing.

Feng XY (1997) *The Catalogue of Sesame Germplasm Research in China (continuation 1)*. Agricultural Scientech Press, Beijing .

Feng XY, Zhang XR and Liu YY (1996) Further characterization and evaluation of pre-selected sesame germplasm. Oil Crops China 18: 63-66.

Feng XY, Zhang XR and Xiao TH (1991) Identification and evaluation of Chinese sesame germplasm tolerance to water logging stress. Oil Crops China 3: 12-15.

Ferdinandez YSN, Somers DJ and Coulman BE (2001) Estimation the genetic relationship of hybrid bromegrass to smooth bromegrass and meadow bromegrass using RAPD markers. Plant Breed. 120: 149-153.

Hawkes J (1983) *The diversity of crop plants*. Harvard University Press, Cambridge.

Hernan EL and Petr K (2006) Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). BMC Genet. 7: 1-10.

Isshiki S and Umezake T (1997) Genetic variations of isozymes in cultivated sesame. Euphytica 93: 375-377.

Kim DH, Zur G, Danin Y, Lee SW, Shim KB, Kang CW and Kashi Y (2002) Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. Plant Breed. 121: 259-262.

Li G, Quiros CF (2001) Sequence-related amplified polymorphism (SRAP) a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. Theor. Appl. Genet. 103: 455-461.

Li LL, Wang SY and Fang XP (1991) Identification of Chinese sesame germplasm resistance to *Macrophomina phaseolina*. Oil Crops China 1: 3-6.

Lin ZX, He DH, Zhang XL, Nie YC, Guo XP, Feng CD and Stewart JM (2005) Linkage map construction and mapping QTL for cotton fiber quality using SRAP, SSR and RAPD. Plant Breed. 124: 180-187.

Pham TD, Bui TM, Werlemark G, Bui TC, Merker A and Carlsson AS (2009) A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. Genet. Resour. Crop Evol. 56: 679-690.

Powell W, Machray GC and Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1: 215-222.

Qamaruz ZF, Michael FF and Parker JS (1998) Molecular techniques employed in the assessment of genetic diversity: A review focusing on orchid conservation. Lindleyana 13: 259-283.

Rohlf FJ (2000) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1, User Guide. Exeter Software, New York.

Sneath PH and Sokal RR (1973) Numerical Taxonomy: The Principal and Practice of Numerical Classification. W. H. Freeman and Company, San Francisco.

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 28: 1409-1438.

Tamura K, Dudley J and Nei M (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Bio. Evo. 24: 1596-1599.

Venkataramana BK, Prashant BP and Suman L (1999) Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Euphytica 110: 21-33.

Xiao TH, Feng XY and Zhang XR (1992) Analysis of the distribution and main characteristics of black seed coat sesame germplasm in China. Oil Crops China 14: 34-37.

Yeh FC and Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgium J. of Botany 129: 157.

Zeven A and Zhukovsky P (1975) Dictionary of cultivated plants and their centres of diversity. PUDOC, Wageningen.

Zhang XR, Cheng Y, Liu SY, Feng XY, Jin LM, Jin QL, Huang CY, Wang ZA, Xu XF and Sun MY (2001) Evaluation of sesame germplasm resistance to *Macrophomina phaseolina* and *Fusarium oxysporium*. Chin. J. Oil Crop Sci. 23: 23-27.

Zhang XR, Zhao YZ, Cheng Y Feng XY, Guo QY, Zhou MD and Hodgkin T (2000) Establishment of sesame germplasm core collection in China. Genet. Resour. Crop Evol. 47: 273-279.

Zhang YX, Lin ZX, Xia QZ, Zhang MJ and Zhang XL (2008) Characteristics and analysis of simple sequence repeats in the cotton genome based on a linkage map constructed from a BC1 population between *Gossypium hirsutum* and *G. barbadense*. Genome 51: 534-546.