Analysis of intraspecific variation of Chinese *Carthamus tinctorius* L. using AFLP markers

ZHANG Lei[1], HUANG Bei-bei[1], KAI Guo-yin[2], GUO Mei-li[1*]

(1. *School of Pharmacy, Second Military Medical University, Shanghai 200433, China*; 2. *School of Life Science and Technology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiaotong University, Shanghai 200030, China*)

Abstract: Aim To investigate the intraspecific variation of *Carthamus tinctorius* L. (safflower) and establish foundation for further breeding of safflower germplasm resource and screening the quality correlation genes. **Methods** Amplified fragment length polymorphism (AFLP) was carried out to analyze genetic variation of 28 safflower populations collected in China. Unweighed pair-group method of with arithmetical averages (UPGMA) cluster analysis was used to construct a dendrogram and to estimate the genetic distances among the populations. **Results** All populations could be uniquely distinguished using 12 selected primer combinations. Similarity coefficients ranged from 0.48 to 0.96 among the populations. Dendrogram revealed distinct segregation of all the cultivars into three main groups and one midst group. **Conclusion** Limited genetic diversity exists within the tested 28 collections at intra specific level and AFLP-based phylogeny was not absolutely consistent with that based on morphological characters may be due to the interaction effect between genotype and environment.

Key words: AFLP; Chinese *Carthamus tinctorius* L.; dendrogram; genetic diversity

CLC number: Q81 **Document code:** A **Article ID:** 0513 - 4870(2006)01 - 0091 - 06

中国红花遗传多样性的 AFLP分子标记

张 磊[1], 黄蓓蓓[1], 开国银[2], 郭美丽[1*]

(1. 第二军医大学 药学院, 上海 200433; 2. 上海交通大学 植物生物技术研究中心, 农业与生物学院, 复旦-交大-诺丁汉植物生物技术研发中心, 生命科学技术学院, 上海 200030)

摘要: 目的 考察红花的种内变异, 为进一步进行红花种质资源选育和筛选品质相关基因奠定基础。方法 采用扩增片段长度多态性(AFLP)技术, 研究中国红花 28个不同栽培居群在 DNA水平的多态性。采用 UPGMA构建系统树图进行聚类分析和计算遗传距离。结果 筛选得到的 12对引物的扩增片段具有丰富的多态性。各居群间的遗传相似系数在 0.48 ~ 0.96。聚类分析显示它们分为 3大主要类群和一个中间类群。结论 红花种内存在一定的遗传多样性。基于 AFLP条带统计的聚类结果和表型特征并不完全一致, 可能是基因型和环境共同作用于表型的结果。

关键词: 扩增片段长度多态性; 中国红花; 系统树图; 遗传多样性

Flos Carthami, the dried flower petals of Safflower (*Carthamus tinctorius* L.), is an important crude drug in traditional Chinese medicine for promoting blood circulation and removing blood stasis[1]. It is used in the prevention and treatment of cardio-vascular and thrombotic diseases in China. Many molecular genetic analysis methods have been used for the quality control of safflower[2]. Knowledge of genetic diversity among the available safflower germplasm is an important

Received date: 2005-04-26.

Foundation item: National Natural Science Foundation of China (30271588); State Administration of Traditional Chinese Medicine of China (02-03ZP54); Key task project in scientific and technological research of Shanghai (043919313).

* Corresponding author Tel / Fax: 86 - 21 - 25074576, E-mail: mlguo@smmu.edu.cn

prerequisite for breeding and genetic improvement programs in the future. Molecular markers allow the analysis of variation and detection of genetic variation at the molecular level, as is the case for somaclonal variation isolation. Marker assisted breeding for increasing the gene pool offers great advantages for genetic improvement[3]. Rapid unequivocal identification of intra-specific variation is as important for safflower as it is for other vegetatively propagated plants. AFLP technique is a recently developed system that combines the specificity of restriction enzyme analysis with the relative technical simplicity of the polymerase chain reaction (PCR)[4]. AFLP has been used to investigate genetic variation in a wide variety of micro-organism, plant and animal germplasm collections because this technique is highly efficient and reproducible, and can be used to survey overall genetic differences in the nuclear genome in a single assay without any prior sequence knowledge[5].

In the present study, AFLP analysis was performed with some modification by using non-radioactive labeling technique instead of classical radioactive labeling[6] to identify and characterize safflower varieties cultivated in China and to determine the genetic relationship among the populations.

Materials and methods

Plant material and DNA extraction Twenty-eight safflower populations (10 samples per accession) sampled in China were obtained from National Germplasm Resource Bank (NGRB). The samples were collected in the field and subsequently frozen in liquid nitrogen before they were lyophilized and ground to fine powder with mortar and pestle. Plant total DNA was isolated from 50 mg of lyophilized leaf material using CTAB protocol[7]. DNA concentrations were estimated and standardized against the known concentrations of λ-DNA on 1.5% (*w*/*v*) agarose gels. Genomic DNA was stored undiluted in TE (10 mmol·L^{-1} Tris, 1 mmol·L^{-1} EDTA) buffer (pH 8.0) at -20 ℃.

AFLP analysis Using AFLP™ (plant mapping kit) purchased from Applied Biosystems, the method reported by Vos et al[4] was used with modifications of labeling of the *Mse*I-primers. Near infrared (NIR) fluorescence technology was used for imaging labeled DNA fragments. Initially, genomic DNA (200 ng) was double digested using two restriction enzymes (*Eco*RI and *Mse*I, 1.25 U·μL^{-1} each) in a total volume of 20 μL. Simultaneously, double stranded adaptors were ligated to the restricted fragments overnight at 4 ℃. Pre-selective amplification was performed using the above ligation products with dilution (20-fold) by pre-selective primers (*Eco*RI: 5′-GACTGCGTACCAATTCA-3′ and *Mse*I: 5′-GATGAGTCCTGAGTAAC-3′). Pre-selective amplification was performed by 30 cycles amplification (94 ℃ for 30 s, 56 ℃ for 30 s and 72 ℃ for 80 s). The PCR product was diluted 20-fold and then used for further selective amplification with the following protocol: 1 cycle of 94 ℃ for 30 s, 65 ℃ for 30 s, 72 ℃ for 80 s; followed by 12 cycles of amplification with annealing temperature decreasing by 0.7 ℃ in each cycle (starting at 94 ℃ for 30 s, 64 ℃ for 30 s, 72 ℃ for 80 s) and ending at 23 rd cycles of amplification (94 ℃ for 30 s, 55 ℃ for 30 s, 72 ℃ for 80 s). The selective amplification primer combinations had an additional three nucleotide extension at the 3′ end. In all reactions, only the *Mse*I primers were labeled with a fluorescent dye at 5′ end, FAM (6-carboxyfluorescein) for blue (Applied Biosystems). Initially 49 primer combinations were checked for selective amplification and from those primer combinations, twelve were selected for further analysis since they amplified more and with distinct polymorphic loci. The PCR products were separated on polyacrylamide gel using an ABI prism 377 DNA sequencer. Electrophoresis was carried out under 3 V, 60 mA and 200 W electric field at 51 ℃. Molecular weights were estimated using 100 bp DNA ladder (Shenggong Inc., Shanghai, China). AFLP analysis was repeated at least twice for each DNA sample.

Data analysis AFLP amplification products were scored for their presence (1) or absence (0) among 28 populations for all the primer combinations employed to generate a binary matrix. Only those fragments that were distinct and visualized with medium or high intensity were considered for data analysis. Fragments common to all populations were considered as non-informative. The NTSYS-pc 2.0 software was used for statistical treatment of data. Genetic similarity (GS) was calculated by making a pairwise comparison among the populations using Nei and Li[8] coefficient. The formula is given as follows: $GS_{ij} = 2a/2a + b + c$, where GS_{ij} is the measure of GS between individuals *i* and *j*, *a* is the number of polymorphic fragments that are shared by *i* and *j*, *b* is the number of fragments present in *i* and absent in *j*, and *c* is the number of fragments present in *j* and absent in *i*. The similarity

matrix thus generated was used to construct a phenetic dendrogram by means of unweighed pair-group method with arithmetic averages (UPGMA) with PHYLIP 3.57.

Results and discussion

1 AFLP analysis

The 28 safflower populations evaluated included 10 from Xinjiang, 4 from Henan, 3 from Shandong, 2 from Gansu, Anhui and Ningxia, 1 from Shanxi, Hebei, Zhejiang, Yunnan, Sichuan respectively. In China, the populations differ in phylliform, leaf margin, leaf thorn, color of corolla and thorn of bract (Table 1). Preliminary phenotypic evaluation of the safflower populations analyzed indicated that the populations were different in many traits and tended to support the molecular analysis.

Using 8 *Eco*RI and 8 *Mse*I primers with a total of 64 combinations provided in Applied Biosystem AFLP kit, 12 of 49 primer combinations were selected with good reproducibility of the AFLP amplification patterns (Table 2). In total, 845 fragments with sizes ranging between 36 bp and 488 bp for all the 28 individuals were produced by the selected 12 primer combinations, among which 621 were polymorphic (73.5% polymorphism). This level of polymorphism is higher than that in many intraspecific AFLP studies reported[3,9,10]. The percentage of polymorphism ranged from a high of 93 in primer combination E-AAG + M-CTT to a low of 52 in primer combination E-AAG + M-CTA, the average number of fragments per plant per primer combination was 70.4. Table 2 can be used to restrict the choice of primer combinations for further characterizations of the varieties analyzed. A typical representative of AFLP profile generated by primer combination E-AAG + M-CAT was shown in Figure 1. A total of 65 fragments were obtained with this primer combination and the overall polymorphism percentage calculated was 70.8. The fingerprint patterns obtained provided unequivocal identification of each cultivar and the polymorphism

Table 1 Twenty eight safflower (*Carthamus tinctorius* L.) populations with related characters

No.	Local name	Locality	Phylliform	Leaf margin	Leaf thorn	Color of corolla	Thorn of bract	
							Number	Length
1	Hami Youci	Xinjiang	a	c	Many	Yellow	Many	Long
2	Hami Wuci	Xinjiang	a	d	No	Red	No	Middle
3	Jimusa'er Wuci	Xinjiang	a	d	No	Red	No	No
4	Tacheng Wu ci	Xinjiang	a	d	No	Orange	Little	Middle
5	Ruoqiang Wuci	Xinjiang	a	d	No	Red	No	Middle
6	Ruoqiang Youci	Xinjiang	a	c	Many	White	Many	Long
7	Zhangye Wuci	Gansu	a	d	No	Red	No	No
8	Zhangye Youci	Gansu	a	e	Less	Yellow	Little	Middle
9	Yanjin Dahongpao	Henan	a	f	Less	Red	Little	Middle
10	Wucihong	Ningxia	a	c	Middle	Red	Little	Long
11	Xinxiang Honghua	Henan	a	c	Middle	Yellow	Little	Long
12	Hezhe Honghua	Shandong	b	f	No	Yellow	Little	Short
13	Rucheng Honghua	Shanxi	a	d	No	Red	Little	Short
14	Shulu Honghua	Hebei	a	d	No	Yellow	No	Short
15	Hezhe Wuci	Shandong	a	c	Many	Red	Many	Long
16	Qixian Honghua	Henan	a	f	Middle	Orange	Little	Short
17	Boxian Honghua	Anhui	a	c	Middle	Orange	Little	Short
18	Yutai Honghua	Shandong	a	f	Middle	Yellow	Little	Short
19	Hefei Honghua	Anhui	a	g	Middle	Orange	Many	Short
20	Wuzhong Honghua	Ningxia	a	c	No	Yellow	Many	Middle
21	Yuyao Honghua	Zhejiang	a	e	Middle	Orange	Little	Middle
22	Wuci Dahongpao	Henan	b	f	No	Red	Little	Middle
23	Optional	Xinjiang	a	f	Many	Yellowy	Many	Long
24	Optional	Xinjiang	a	f	Many	Yellowy	Many	Long
25	Optional	Xinjiang	a	f	Many	White	Many	Long
26	Optional	Xinjiang	a	f	Many	White	Many	Long
27	Weishan Honghua	Yunnan	a	f	Less	Orange	Little	Middle
28	Jianyang Honghua	Sichuan	a	f	Less	Orange	Little	Middle

* a: Inversion Lanceolate; b: Elliptical; c: Serrature; d: Integer; e: Parted; f: Lobate; g: Deep Serrate

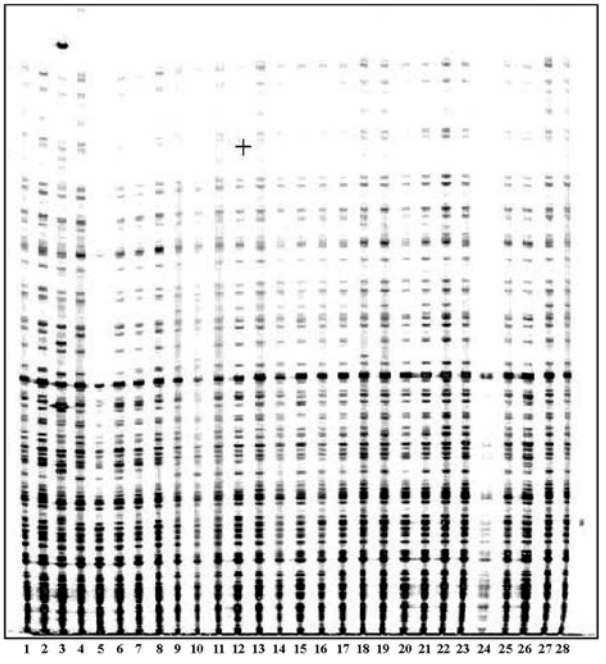

Figure 1 AFLP profiles generated by the E-AAG + M-CAT primer combination across 28 safflower populations (lanes 1 - 28)

Table 2 Selected primers and level of polymorphism among *Carthamus tinctorius* L. populations by AFLP analysis

Selective amplification primer combinations	Number of polymorphic fragments	Total number of fragments	Percentage of polymorphic markers/%
*Eco*R I-AAG + *Mse* I-CTC	32	62	52
*Eco*R I-AAG + *Mse* I-CAG	38	57	67
*Eco*R I-AAG + *Mse* I-CAT	46	65	71
*Eco*R I-AAC + *Mse* I-CAC	48	65	74
*Eco*R I-AAC + *Mse* I-CAG	54	77	70
*Eco*R I-AAC + *Mse* I-CAT	49	62	79
*Eco*R I-AAG + *Mse* I-CAC	77	87	89
*Eco*R I-AAG + *Mse* I-CTA	39	68	57
*Eco*R I-AAG + *Mse* I-CTT	57	61	93
*Eco*R I-ACA + *Mse* I-CAC	59	77	77
*Eco*R I-ACA + *Mse* I-CTC	75	104	72
*Eco*R I-ACA + *Mse* I-CTG	47	60	78
Total	621	845	73.5

observed among populations was appropriate to differentiate cultivars. The AFLP technique was confirmed to be a powerful tool for charactering genetic variation among cultivated populations of safflower in China.

2 Genetic similarity matrix and cluster analysis

The GS matrix based on Nei and Li's coefficient[8] was calculated for AFLP. Pairwise distances calculated with the two indexes revealed that there are exactly the same relationships among populations (Table 3). Although accession 27 (sampled from Yunnan) and 28 (sampled from Sichuan) were cultivated in different places with largely geographic distance, they shared the same phenotype (Table 1) and the maximum GS of genotype (0.91), while several cultivars were related distantly. Safflower genotypes were grouped based on genetic similarities in a phenetic dendrogram using UPGMA algorithm (Figure 2). The dendrogram separated all the tested cultivars, with the exception of 27 and 28, which were thus considered genetically closely related. Populations 1, 2, 4, 6, 7 and 8, showed high GS with 24 and 25. Populations 3, 5, 9, 10, 16 and 20 were clearly separated. In the case of populations 11 and 12, 14 and 21, 22 and 23, 15 and 19, no genetic variability were detected from each other, although comparable morphological characteristics were showed in Table 1. The results suggested that the availability of molecular tools to characterize and distinguish different cultivars could provide the possibility to distinguish the wide array of synonyms and homonyms within safflower cultivars and define genetic relationships among different populations. The AFLP methodology offers the possibility to screen a large number of anonymous *loci*: a single primer combination is frequently sufficient to ensure cultivar distinction[11,12].

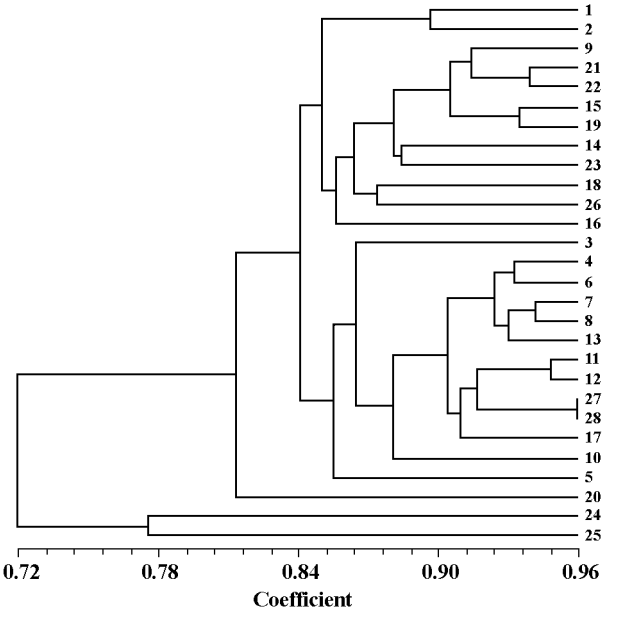

Figure 2 Neighbour-joining tree of 28 safflower populations based on genetic similarity from AFLP data generated by UPGMA method

In Table 1, these varieties were characterized by inverted lanceolate phylliform, many leaf thorns and many long thorns of bract. GS did not necessarily

Table 3 Genetic similarity matrix between populations of the *Carthamus tinctorius* L. (Numbers correspond to those of Table 1)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	1.00																											
2	0.81	1.00																										
3	0.62	0.70	1.00																									
4	0.67	0.78	0.79	1.00																								
5	0.60	0.68	0.72	0.80	1.00																							
6	0.66	0.74	0.77	0.86	0.78	1.00																						
7	0.68	0.75	0.78	0.86	0.78	0.86	1.00																					
8	0.69	0.76	0.75	0.86	0.75	0.85	0.88	1.00																				
9	0.70	0.70	0.70	0.71	0.64	0.71	0.75	0.76	1.00																			
10	0.62	0.67	0.71	0.76	0.68	0.77	0.77	0.81	0.69	1.00																		
11	0.69	0.75	0.75	0.81	0.72	0.80	0.85	0.83	0.78	0.80	1.00																	
12	0.66	0.74	0.77	0.82	0.73	0.79	0.84	0.84	0.75	0.79	0.89	1.00																
13	0.70	0.78	0.77	0.86	0.79	0.82	0.87	0.85	0.74	0.77	0.84	0.84	1.00															
14	0.75	0.76	0.65	0.72	0.65	0.71	0.72	0.72	0.79	0.66	0.74	0.73	0.75	1.00														
15	0.72	0.75	0.72	0.78	0.68	0.74	0.79	0.81	0.81	0.75	0.84	0.82	0.81	0.82	1.00													
16	0.77	0.71	0.60	0.65	0.53	0.62	0.66	0.67	0.74	0.65	0.73	0.71	0.67	0.73	0.76	1.00												
17	0.66	0.73	0.72	0.82	0.72	0.79	0.80	0.83	0.72	0.76	0.81	0.82	0.81	0.69	0.79	0.69	1.00											
18	0.75	0.78	0.66	0.73	0.63	0.73	0.71	0.76	0.71	0.69	0.75	0.74	0.73	0.71	0.75	0.73	0.77	1.00										
19	0.71	0.77	0.71	0.77	0.68	0.75	0.77	0.79	0.77	0.72	0.84	0.82	0.79	0.77	0.87	0.74	0.83	0.79	1.00									
20	0.63	0.68	0.65	0.71	0.68	0.74	0.71	0.71	0.64	0.70	0.70	0.70	0.74	0.67	0.71	0.59	0.71	0.68	0.71	1.00								
21	0.68	0.71	0.67	0.71	0.62	0.70	0.72	0.74	0.85	0.69	0.78	0.77	0.73	0.79	0.81	0.75	0.75	0.73	0.82	0.65	1.00							
22	0.70	0.72	0.70	0.74	0.65	0.72	0.74	0.76	0.82	0.72	0.82	0.78	0.75	0.77	0.87	0.75	0.78	0.74	0.84	0.66	0.88	1.00						
23	0.75	0.79	0.67	0.74	0.63	0.71	0.74	0.75	0.75	0.69	0.79	0.77	0.75	0.79	0.79	0.75	0.72	0.78	0.79	0.65	0.77	0.79	1.00					
24	0.49	0.55	0.55	0.54	0.57	0.58	0.54	0.54	0.51	0.54	0.56	0.54	0.56	0.52	0.52	0.45	0.55	0.56	0.54	0.61	0.50	0.52	0.54	1.00				
25	0.59	0.61	0.56	0.57	0.56	0.59	0.58	0.56	0.57	0.56	0.62	0.57	0.59	0.59	0.61	0.54	0.56	0.58	0.51	0.65	0.59	0.61	0.60	0.63	1.00			
26	0.70	0.76	0.72	0.72	0.65	0.71	0.72	0.74	0.75	0.70	0.76	0.77	0.74	0.72	0.80	0.71	0.75	0.77	0.78	0.69	0.75	0.78	0.77	0.60	0.61	1.00		
27	0.65	0.73	0.74	0.85	0.71	0.80	0.79	0.84	0.73	0.78	0.83	0.85	0.85	0.70	0.78	0.71	0.85	0.77	0.81	0.68	0.75	0.77	0.75	0.55	0.55	0.75	1.00	
28	0.64	0.72	0.77	0.84	0.73	0.78	0.79	0.82	0.71	0.79	0.82	0.85	0.84	0.68	0.78	0.67	0.83	0.74	0.80	0.69	0.74	0.75	0.72	0.55	0.55	0.75	0.91	1.00

reflect similarity or difference in outer traits, such as chemical components, or agronomic traits[13]. Usually, molecular markers reflecting differences in the whole plant genome are not necessarily related to a specific plant secondary compound[14]. Thus, the AFLP dendrogram may be more appropriate for universally taxonomic studies than that based on concrete traits such as chemical components (end-use related) with practical interest, but do not necessarily correlate with taxonomy. DNA genotyping offers the unique capacity to classify populations regardless of environmental conditions or plant growth stage[15].

In conclusion, classification of safflower germplasm was attempted earlier by using morphological characters, but analysis with this system failed to reveal the taxonomic affinities between populations, probably due to the environmental plasticity of morphological traits. It can be overcome by molecular analysis. The availability of a large number of fragments defining independent genetic loci with highly reproducible polymorphism detection provides the efficient evaluation of genetic diversity. AFLP marker technique was confirmed to be a powerful tool for characterizing intra-specific variation among cultivated populations of safflower. We have employed AFLP for the assessment of genetic diversity of safflower cultivars in China for the first time. The polymorphism observed among populations was useful and effective to differentiate cultivars. The fingerprint patterns obtained allowed unequivocal identification of each cultivar. Diversity assessment revealed that genetic diversity was rather limited among the collected cultivars, suggesting that a more extensive collection effort for safflower germplasm is required to broaden the genetic resource for the breeding program.

References

[1] State Pharmacopeia Committee of the PRC. Pharmacopeia of the People's Repbublic of China (subdivision) [S]. Beijing: Chemical Industry, 2005: 103.

[2] Ma XQ, Zhu DY, Li SP, et al. Authentic identification of stigma Croci (stigma of *Crocus sativus*) from its adulterants by molecular genetic analysis [J]. Planta Med, 2001, 67: 183 - 186.

[3] Balasaravanan T, Pius PK, Kumar RR, et al. Genetic diversity among south Indian tea germplasm (*Camellia sinensis*, *C. assamica* and *C. assamica* spp. lasiocalyx) using AFLP markers [J]. Plant Sci, 2003, 165: 365 - 372.

[4] Vos P, Hogers R, Bleeker M, et al. AFLP: a new

technique for DNA fingerprinting [J]. Nucleic Acids Res, 1995, 23: 4407 - 4414.

[5] Karp A, Seberg O, Buiatti M. Molecular techniques in the assessment of botanical diversity [J]. Ann Bot, 1996, 78: 143 - 149.

[6] Vignani R, Scali M, Masi E, et al. Genomic variability among *Vitis vinifem* L. "Sangiovese" assessed by microsatellite and non-radioactive AFLP test [J]. EJB, 2002, 5: 15.

[7] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue [J]. Phytochem Bull, 1987, 19: 11 - 15.

[8] Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonuclease [J]. Proc Natl Acad Sci USA, 1979, 76: 5269 - 5273.

[9] Sensi E, Vignani R, Scali M, et al. DNA fingerprinting and genetic relatedness among cultivated varieties of *Olea europaea* L. estimated by AFLP analysis [J]. Sci Hortic, 2003, 97: 379 - 388.

[10] Schnell RJ, Olano CT, Campbell RJ, et al. AFLP analysis of genetic diversity within a jackfruit germplasm collection [J]. Sci Hortic, 2001, 91: 261 - 274.

[11] Singh A, Chaudhury A, Srivastava PS, et al. Comparison of AFLP and SAMPL markers for assessment of intra-population genetic variation in *Azadirachta indica* A. Juss [J]. Plant Sci, 2002, 162: 17 - 25.

[12] Cervera MT, Cabazas JA, Sánchez-Escribano E, et al. Characterization of genetic variation within table grape varieties (*Vitis vinifem* L.) based on AFLP markers [J]. Vitis, 2000, 39: 109 - 112.

[13] Zhao MB, Deng XL, Wang YL, et al. Establishment of chromatographic fingerprint and quality assessment of *Carthamus tinctorius* L. by high performance liquid chromatography [J]. Acta Pharm Sin (药学学报), 2004, 39: 212 - 216.

[14] Nan P, Peng SL, Shi SH, et al. Interpopulation congruence in Chinese *Primula ovalifolia* revealed by chemical and molecular markers using essential oils and ISSRs [J]. Z Naturforsch, 2003, 58: 57 - 61.

[15] Massimo L, Mariangela M, Bernardetta L, et al. Morphological characterization, essential oil composition and DNA genotyping of *Ocimum basilicum* L. cultivars [J]. Plant Sci, 2004, 167: 725 - 731.