

## SSR 标记对野生大豆种群遗传结构的研究

徐立恒, 李向华

(银川科技职业学院 生物与农业工程系, 宁夏 银川 750105)

**摘 要:** 使用 40 对 SSR 引物分析了 8 个省区天然野生大豆种群的遗传结构和遗传多样性。结果表明: 15 个种群共检测到 633 个等位基因, 平均每对引物等位变异基因数平均值 15.83 个; 种群平均 Shannon 指数 ( $I$ ) 0.7835, 种群平均期望杂合度 ( $He$ ) 0.4070, 种群平均观察杂合度 ( $Ho$ ) 0.0012, 平均种群内遗传多样性 ( $Hs$ ) 0.4277, 平均种群间遗传多样性 ( $D_{ST}$ ) 0.4543, 种群分化系数 ( $G_{ST}$ ) 0.5151; 供试的 15 个天然野生大豆种群可分为 4 组, 基本存在于 2 个地理分布区, 天津以北地理区遗传多样性较高。

**关键词:** 野生大豆; SSR 标记; 遗传多样性

中图分类号: S565.1

文献标识码: A

文章编号: 1000-9841(2011)01-0041-05

## Analysis on Genetic Structure of Wild Soybeans Populations by SSR Markers

XU Li-heng, LI Xiang-hua

(School of Biological and Agricultural Engineering, Yinchuan Science and Technology Vocational College, Yinchuan 750105, Ningxia, China)

**Abstract:** Wild soybean (*Glycine soja*) is commonly accepted as the progenitor species of the cultivated soybean (*Glycine max*). They contain many characters potentially valuable for supplementing the soybean germplasm pool. Our objective was to evaluate the genetic diversities of wild soybean collected from 8 provinces in different latitudes of China. We sampled 15 representative natural populations in 2005. Forty public SSR primer pairs over the 20 linkage groups were applied to evaluate genetic diversity. A total of 633 bands were detected with an average number of 15.83 per locus. Mean expected heterozygosity per locus ( $He$ ) was 0.4070 for the populations, and the mean Shan-non index ( $I$ ) for the population was 0.7835. Mean observed heterozygosity per locus ( $Ho$ ) for the populations was 0.0012. Between-population genetic diversity ( $Hs$ ) averaged 0.4277, and within-population genetic diversity ( $D_{ST}$ ) averaged 0.4543. Mean coefficient of gene differentiation for loci ( $G_{ST}$ ) in the populations was estimated to be 0.5151. Cluster analysis showed that 15 populations fell into 4 groups, and existed mainly in geography distribution area, and the area with the latitude higher than of Tianjin had better genetic diversity.

**Key words:** Wild soybean (*Glycine soja*); SSR marker; Genetic diversity

野生大豆 (*Glycine soja* Sieb. & Zucc.) 是栽培大豆的祖先, 广泛分布于东亚地区, 以中国分布最为广泛。Maughan 和 Li 等<sup>[1-2]</sup> 研究表明野生大豆具有较高的遗传变异性。一些学者在种群水平上进行过不同空间尺度的研究。Kiang 等和 Fujita 等<sup>[3-4]</sup> 在河流生态系统的种群中发现地理距离与遗传距离和遗传多样性有相关性, 而府宇雷等<sup>[5]</sup> 在非河流生态系统种群间没有发现类似的相关性。种群是物种存在和进化的基本单位, 野生大豆遗传资源搜集或原位保护的方法、策略和措施都是以种群的遗传多样性为依据。尽管一些研究者使用同工酶分析了我国少数几处野生大豆天然种群的遗传结构, 但是, 对于中国广泛的种群分布及复杂的生态条件而言目前进行的研究还远远不够, 特别是对遗传资

源保护具有重要意义的利用 DNA 标记进行特定地区的种群遗传结构分析和评价研究更少。该文通过使用 SSR 标记对全国 8 个省区天然野生大豆种群的遗传多样性和遗传结构进行分析, 评价各地区的遗传变异性以及在相当的地域范围 SSR 标记对野生大豆分子生态遗传学研究的效果, 为更好地进行野生大豆生物学保护及合理有效利用提供依据。

### 1 材料与方法

#### 1.1 供试材料

供试野生大豆材料来自北纬 23° ~ 52° 的中国天然野生大豆分布区 (包括黑龙江、北京、天津、安徽、湖北、重庆、河南、广东)。采用随机取样方法, 株距 10 ~ 20 cm (形态差异明显的可小于 10 cm)。

收稿日期: 2010-10-14

第一作者简介: 徐立恒 (1980-), 女, 硕士, 主要从事分子生物学研究。E-mail: 42598919@qq.com。

15 个种群共 454 个单株 (表 1)。

表 1 15 个天然野生大豆种群收集点和生境

Table 1 Collection sites and habitats of 15 natural populations of wild soybean

序号 No.	采集地点 Collection	取样株数 Size of sample	纬度(N) Latitude	经度(E) Longitude	海拔 Altitude	原生境 Habitat
1	河南内乡 Neixiang, Henan	31	33.05°	116.87°	160	山脚 The foot of a hill
2	河南商城 Shangcheng, Henan	31	36.30°	116.65°	58	公路边 Highway side
3	重庆开县 Kaixian, Chongqing	30	31.60°	108.75°	165	山坡 Mountain slope
4	安徽五河 Wuhe, Anhui	31	32.95°	117.36°	22	河岸 Riverside
5	湖北宜昌 Yichang, Hubei	30	30.70°	113.70°	133.1	山坡 Mountain slope
6	广东英德 Yingde Guangdong	32	23.17°	113.33°	41	山脚 The foot of a hill
7	安徽巢湖 Chaohu, Anhui	31	31.65°	117.80°	115	河岸 Riverside
8	天津 Tianjin	31	39.42°	117.28°	67	湿地 Wetland
9	黑龙江塔河 Tahe, Heilongjiang	31	52.92°	124.75°	350	农田边 Farmland edge
10	黑龙江桦南 Huanan, Heilongjiang	31	46.47°	130.76°	520	野地 Wild lands
11	黑龙江望奎 Wangkui, Heilongjiang	31	46.97°	126.30°	500	野地 Wild lands
12	黑龙江延寿 Yanshou, heilongjiang	24	45.63°	128.25°	550	野地 Wild lands
13	黑龙江巴彦 Bayan, Heilongjiang	31	46.51°	127.43°	500	野地 Wild lands
14	安徽宿松 Susun, Anhui	30	30.53°	117.05°	19.8	河岸 Riverside
15	北京平谷 Pinggu, Beijing	30	40.51°	117.25°	89	极干旱河堤 Extremely drought embankment

## 1.2 试验方法

1.2.1 样本 DNA 的提取 选取幼苗的三出复叶叶片,采用 CTAB 法提取总 DNA。

1.2.2 SSR 标记 根据 Cregan 等<sup>[6]</sup>构建的大豆遗传连锁图谱,在 20 个连锁群上共计选择 40 对 SSR 引物(表 3)。

PCR 反应体系为每 20  $\mu\text{L}$  反应体积中含有 50 ng DNA 模板,引物( $50 \text{ ng} \cdot \mu\text{L}^{-1}$ )0.5  $\mu\text{L}$ ,10  $\times$  PCR Buffer 5.0  $\mu\text{L}$ ,Taq 酶( $5 \text{ U} \cdot \mu\text{L}^{-1}$ )0.5  $\mu\text{L}$ ,dNTP( $2 \text{ mmol} \cdot \text{L}^{-1}$ )0.5  $\mu\text{L}$ , $\text{H}_2\text{O}$  8.5  $\mu\text{L}$ 。PCR 在 PE9600 型 PCR 仪上进行,PCR 产物在 6% 聚丙烯酰胺凝胶上电泳分离,硝酸银染色。

1.2.3 数据分析方法 选择各种群的多态位点百分数( $P$ )、平均等位基因数( $A$ )、平均有效等位基因数( $A_e$ )、平均期望杂合度( $H_e$ )和实际观察杂合度( $H_o$ )、Shannon 多样性指数( $I$ )等指数对各种群的遗传变异进行检测,进而分析各种群的遗传多样性。以上数据均利用基于 Windows 平台的计算机软件 POPGENE 计算得出。种群间遗传分化评价采用 Nei 距离( $D$ )和遗传一致度( $I$ )来评价。SSR 位点基因分化系数采用  $G_{ST}$  评价( $G_{ST} = D_{ST}/H_T = (H_T - H_S)/H_T$ ),其中  $H_T$  为总群体的基因多样性, $H_S$  表示种群内基因多样性, $D_{ST}$  为种群间基因多样性。上述计算使用 POPGENE 3.0 软件。种群聚类分析根据种群间遗传一致度聚类,使用 NTSYSpc21-2 软件。

## 2 结果与分析

### 2.1 种群的遗传结构

多态位点百分率( $P$ )是反应遗传多态的重要指标之一。如表 2 所示,15 个供试野生大豆种群多态位点百分率( $P$ )平均值为 86%。种群 2、4、7、9 和 11 在 40 个位点上都表现为多态性,每个位点都存在遗传变异,其中,种群 4 多态位点百分率( $P$ )最高达 100%,境内湿地生态区,具备野生大豆生存、繁衍的良好生态环境。种群 5( $P = 17.50\%$ )仅有 7 对 SSR 位点表现多态性,其余 33 对 SSR 位点上的常见等位基因频率都大于 0.99,表现单态性。

15 个种群平均等位基因数( $A$ )为 3.4750,最高为 5.8500(13 号种群),最低为 1.2500(5 号种群)。平均有效等位基因数( $A_e$ )为 2.1978,最高为 3.4450(13 号种群),最低为 1.1080(5 号种群)。 $A_e$  在种群上的分布趋势与( $A$ )值有 13 个种群相同,有 2 个种群不一致(6 号和 9 号种群)。1 号种群与 8 号种群的平均等位基因数相同( $A = 2.6750$ ),但 8 号种群的  $A_e$ (1.8720)高于 1 号种群(1.7530),说明 1 号种群的极低频率的等位基因数要比 8 号种群的多,8 号种群的等位基因频率大小要相对均匀一些。而 5 号种群的遗传参数都很低,表现遗传上的单一化,说明该种群经受过自然选择。

表 2 SSR 检测的 15 个野生大豆种群遗传多样性

Table 2 Genetic diversity among 15 populations based on microsatellite loci

种群 Population	多态位点比例 <i>P</i> /%	种群等位变异数 <i>A'</i>	占总种群等位 变异比例 <i>Ap</i> /%	平均等位变异 <i>A</i>	有效等位变异 <i>Ae</i>	香农指数 <i>I</i>	平均期望 杂合度 <i>He</i>	实际观察 杂合度 <i>Ho</i>
1	95.00	107	17	2.6750	1.7530	0.6020	0.3610	0.0008
2	100.00	161	25	4.0250	2.4900	0.8198	0.4732	0.0000
3	80.00	94	15	2.3500	1.4680	0.4400	0.2558	0.0000
4	100.00	199	31	4.9750	3.0960	1.2280	0.6370	0.0000
5	17.50	50	8	1.2500	1.1080	0.0910	0.0539	0.0000
6	97.50	131	21	3.2750	1.7400	0.6540	0.3688	0.0000
7	100.00	176	28	4.4000	2.7490	1.0910	0.5847	0.0032
8	85.00	107	17	2.6750	1.8720	0.6020	0.5721	0.0000
9	100.00	165	26	4.1250	2.4870	1.0210	0.5511	0.0008
10	97.50	158	25	3.9500	2.6550	1.0370	0.5706	0.0056
11	100.00	178	28	4.4500	2.7500	1.0710	0.0048	0.0000
12	95.00	129	20	3.2250	2.0190	0.7650	0.4373	0.0021
13	97.50	234	37	5.8500	3.4450	1.3360	0.6633	0.0056
14	95.00	138	22	3.4500	2.1130	0.8110	0.4589	0.0000
15	30.00	58	9	1.4500	1.2240	0.1830	0.1131	0.0000
平均 Mean	86.00	139	22	3.4750	2.1978	0.7835	0.407	0.0012

Shannon 多样性指数 (*I*) 是常用来衡量某一种群或地区的遗传多样性高低的尺度之一。15 个种群的 *I* 值高低有相当大的差异 (表 2)。最高为 1.3360 (13 号种群), 最低为 0.0910 (5 号种群), 平均值为 0.783, 除了 2 种群的 *I* 值较低外, 其余 14 个种群值都在 0.440 以上。为了比较遗传多样性高低, 将 15 个种群按纬度从低到高排列的 Shannon 指数 *I* 作图比较 (图 1)。从图 1 可以看到, 随着不同地理种群纬度的升高, 种群的遗传多样性增幅没有一定的规律, 二者之间没有明显的相关性。种群 4、7、9、10、11 和 13 的遗传多样性比较高, 即安徽和黑龙江的野生大豆遗传多样性较高。

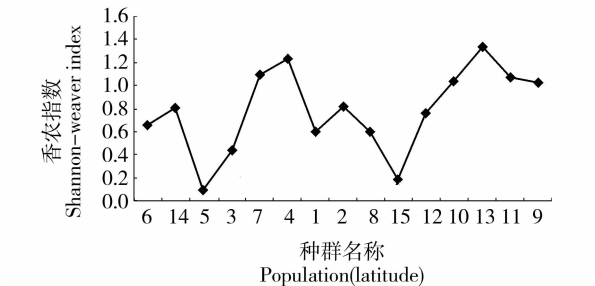


图 1 香农指数 (*I*) 与纬度关系趋势图

Fig.1 The relationship between Shannon-Weaver index and latitude

种群的平均期望杂合度 *He* 与 Shannon-Weaver 指数一样是反映遗传多样性的高低尺度, 趋势相同。其平均为 0.4070。可分为高中低 3 组: 种群 4、7、8、9、10 和 13 有较高的 *He* 值 (*He* > 0.500); 种群 2、6、12 和 14 次之 (0.362 ~ 0.500); 较低的为种群 1、3、5、11 和 15 (*He* < 0.362)。实际观测杂合度 (*Ho*) 平均为 0.0012, 是一个相当低的值, 这说明野

生大豆在自然界是严格自交的。图 2 和图 3 分别是平均期望杂合度 (*He*) 和实际观察杂合度 (*Ho*) 的趋势图, 在这 2 个图中也显示了与 Shannon 指数 (*I*) 同样的趋势, 也证明上述地区的遗传变异水平较高。

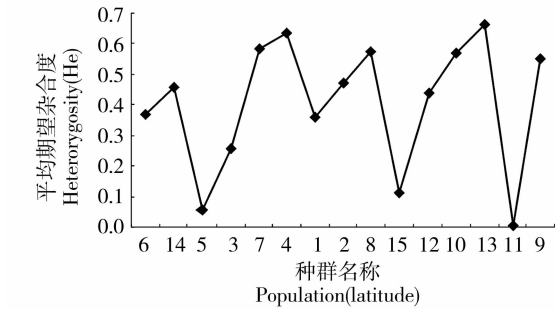


图 2 平均期望杂合度 (*He*) 与纬度关系趋势图

Fig.2 The trend of the expected heterozygosity (*He*) among different latitudes

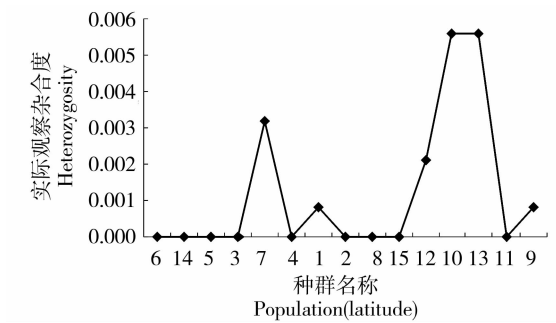


图 3 实际观察杂合度 (*Ho*) 与纬度关系趋势图

Fig.3 The trend of the observed heterozygosity (*Ho*) among different latitudes

2.2 种群基因分化

种群不同位点的遗传分化程度见表 3。种群内遗传多样性 *Hs* = 0.3225 (satt146) ~ 0.5614 (satt431),

平均为 0.427 7。种群间遗传多样性( $D_{ST}$ ) 在 40 个位点上的平均值为 0.4543。总体上,平均种群间  $D_{ST}$  值(0.4543)略大于种群内  $H_S$  值(0.4277),在 40 对引物中有 27 对引物表现  $D_{ST}$  值高于  $H_S$  值;有 12 对引物  $H_S$  值高于  $D_{ST}$  值。

表 3 不同 SSR 位点的遗传参数  
Table 3 Genetic parameters for SSR loci

位点 loci	连锁群 Linkage group	位点等位 变异数 A1	种群内遗 传多样性 $H_S$	种群间遗 传多样性 $D_{ST}$	总种群遗 传多样性 $H_T$	遗传分 化系数 $G_{ST}$
satt185	E	20	0.4707	0.4252	0.8959	0.4746
satt352	G	17	0.4636	0.4653	0.9289	0.5009
satt386	D2	15	0.4113	0.4154	0.8267	0.5025
satt173	O	16	0.3971	0.5015	0.8986	0.5581
satt596	J	11	0.3795	0.4627	0.8422	0.5494
satt002	D2	21	0.5102	0.4121	0.9223	0.4468
satt309	G	16	0.4169	0.4732	0.8901	0.5316
satt022	N	20	0.4080	0.4772	0.8852	0.5391
satt373	L	16	0.5088	0.3788	0.8876	0.4268
satt216	D1b + W	21	0.4010	0.5129	0.9139	0.5612
satt429	A2	14	0.4264	0.4392	0.8656	0.5074
satt530	N	18	0.4684	0.4348	0.9032	0.4814
satt453	B1	12	0.4131	0.4327	0.8458	0.5116
satt146	F	11	0.3225	0.5382	0.8607	0.6253
satt300	A1	15	0.3668	0.5034	0.8702	0.5785
satt242	K	17	0.3761	0.5226	0.8987	0.5815
satt180	C1	9	0.3226	0.4843	0.8069	0.6002
satt177	A2	23	0.4316	0.4746	0.9062	0.5237
satt588	K	16	0.3987	0.4624	0.8611	0.5370
satt434	H	11	0.3924	0.4833	0.8757	0.5519
satt308	M	21	0.3672	0.5570	0.9242	0.6027
satt590	M	18	0.3994	0.4876	0.8870	0.5497
satt267	D1a + Q	12	0.4124	0.4602	0.8726	0.5274
satt168	B2	13	0.4596	0.4173	0.8769	0.4759
satt281	C2	22	0.5585	0.3742	0.9327	0.4012
satt286	C2	15	0.3650	0.4900	0.8550	0.5731
satt239	I	18	0.3832	0.5075	0.8907	0.5698
satt556	B2	19	0.4296	0.4453	0.8749	0.5090
satt243	O	14	0.3958	0.4861	0.8819	0.5512
satt236	A1	12	0.4905	0.3628	0.8533	0.4252
satt431	J	20	0.5614	0.3603	0.9217	0.3909
satt586	F	13	0.4471	0.4355	0.8826	0.4934
satt307	C2	18	0.5073	0.4066	0.9139	0.4449
satt194	C1	15	0.4081	0.4576	0.8657	0.5286
satt268	E	11	0.4218	0.4213	0.8431	0.4997
satt226	D2	13	0.5010	0.3711	0.8721	0.4255
satt390	A2	13	0.4409	0.4362	0.8771	0.4973
sat_099	L	14	0.4129	0.4545	0.8674	0.5240
satt197	B1	16	0.4196	0.4734	0.8930	0.5301
satt571	I	17	0.4410	0.4717	0.9127	0.5168
平均 Mean		15.825	0.4277	0.4543	0.8820	0.5151

种群间遗传分化系数在各位点也不同,所有 40 对 SSR 位点的种群平均  $G_{ST}$  值为 0.5151。7 个位点上种群间的等位基因频率的分布差异较大。有 13 对引物种群间遗传分化程度低( $G_{ST} < 0.500$ )。其余引物的遗传分化程度为中等(0.500 ~ 0.560)。

不仅在不同连锁群上的位点的基因分化程度不同,在同一连锁群上的不同位点也可能存在明显差异。例如引物 satt431 和 satt596 同位于 J 连锁群上,satt431 的等位基因总数为 20 个, $G_{ST} = 0.3909$ ;而 satt596 的等位基因总数为 11, $G_{ST}$  为 0.5494,二者之间遗传变异水平差异很大。这可能是由于在染色体不同位置上的突变频率不同。

### 2.3 种群间遗传聚类分析

由图 4 可见,在遗传一致度  $I = 0.21$  处,15 个天然野生大豆种群所绘制出的树状图大致可分为 4 类:(I) 种群 1、2、3、4、6 和 7;(II) 种群 8、9、10、11、12 和 13;(III) 种群 5 单独为一类;(IV) 种群 14 和 15。第 I 类种群组全部为黄河以南至长江流域,包括 1 个广东种群。第 II 类种群组全部为北方群体,包含全部 5 个黑龙江种群和 1 个天津种群。第 III 类种群组仅包含 1 个种群 5,来自湖北宜昌,该种群的形成曾受到非自然因素的影响,形态类型非常单一。第 IV 类种群包含 2 个种群。其中种群 15 是来自北京的耐干旱群体,生长在干旱的沙石河坝上。种群 14 与种群 15 距离相对较远,与北京耐干旱种群聚在一起是个例外。聚类结果显示,野生大豆天然种群在大的空间尺度上明显存在地理相关性:包括天津在内的以北区域和黄河以南区域的两大地域的地理分化。

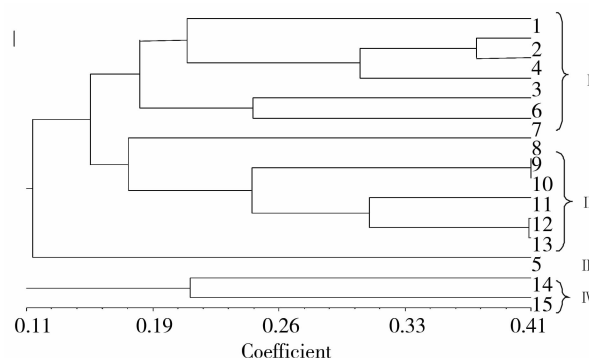


图 4 15 个天然野生大豆种群的聚类分析

Fig. 4 Cluster analysis of 15 natural populations of wild soybean based on genetic identity

## 3 讨论

Loveless 和 Hamrick<sup>[7]</sup> 认为自交植物遗传多样性特点是居群间遗传多样性( $D_{ST}$ ) 高于居群内遗传

多样性( $H_s$ )。该研究结果也表现出相同的趋势,但是,40 对引物之间存在差异,其中有 27 对引物  $D_{ST}$  值高于  $H_s$  值,有 1 对引物  $H_s$  和  $D_{ST}$  相等,12 对  $H_s$  高于  $D_{ST}$ 。居群平均基因分化度  $G_{ST}$  为 0.5151。总体上供试野生大豆遗传多样性分布趋势为黑龙江、安徽居群最高,北京和湖北的 2 个居群最低。

聚类分析显示地理纬度近的居群虽能聚在一起,但组内并不完全符合地理纬度靠近的居群首先聚在一起的规律。研究选用的材料只是该纬度带上的一个居群,而不是该纬度带上的一个基因池。例如,地理纬度较近的黑龙江 5 个居群(9、10、11、12、13)聚在一起,说明在一定地理范围内遗传距离与地理距离存在一定的相关性,这可能是因为在某种生态区域或生态条件下,光温等生态因子没有太大的变化,自然选择的压力趋于一致,种群组成进化必然有一定趋同性,遗传结构也趋向类似。居群 5 单独聚为一组,是因为在局部小生境下有可能发生漂变,自然选择等因素使某种优势类型(或个体)固定、繁殖。地理距离相隔较远的天津居群 8 和黑龙江的 5 个居群(9、10、11、12、13)聚在了一起,说明研究材料的北方居群在大的气候程度上,其遗传结构有更多的相似性。

该研究中的 7 个非保护点天然野生大豆居群的各项遗传指标均接近于 8 个原位保护点居群,其多态位点比例还高于原位保护点的居群。从总体上看,目前原位保护点的随机设置基本有代表性,但是个别保护点遗传多样性很单一,如湖北宜昌保护点(居群 5)。因此在设立原位保护点时在预先进行遗传多样性评价的基础上进行设立会提高保护效率和减少保护成本。

由于经济发展和对资源的过度开发利用,天然野生大豆资源不断遭到破坏,因此,开展天然野生

大豆资源的保护工作已刻不容缓。对 8 省区 15 个野生大豆居群的研究表明,尽管同一省份内大环境是一致的,但小环境对野生大豆遗传多样性的影响也不容忽视,如该试验的居群 4、7、14 生长在一个大环境中,但遗传差异较大,表明生长的小环境对野生大豆的遗传结构影响较大。因此,在取样进行异位保护时,除进行随机取样外,还应根据小生境变化进行人工选择。另外,这 3 个居群的遗传分化较为明显,可以推测每个居群具有自己独特的遗传结构,因此在进行原生境保护时,3 个居群应分别采取适宜的保护措施。

## 参考文献

- [1] Maughan P J, Saghai M A, Buss G R. Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean [J]. *Genome*, 1995, 38:715-723.
- [2] Li Z L, Nelson R L. RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces [J]. *Crop Science*, 2001, 41:1337-1347.
- [3] Kiang Y T, Chiang Y C, Kaizuma N. Genetic diversity in natural populations of wild soybean in iwate prefecture, Japan [J]. *Journal of Heredity*, 1992, 83:325-329.
- [4] Fujita R, Ohara M, Okazaki K, et al. The extent of natural cross-pollination in wild soybean (*Glycine soja*) [J]. *Journal of Heredity*, 1997, 88:124-128.
- [5] Fu Y L, Qian J, Ma Y H, Li J, et al. Genetic differentiation research on populations of wild soybeans in different scales [J]. *Acta Ecologica Sinica*, 2002, 22:176-184.
- [6] Cregan P B, Bush A L, Shoemaker R C, et al. An integrated genetic linkage map of the soybean genome [J]. *Crop Science*, 1999, 39:1464-1490.
- [7] Loveless M D, Hamrick J L. Ecological determinants of genetic structure in plant populations [J]. *Annual Review Ecology Systematics*, 1984, 15:93-99.