

中国一年生野生大豆(*Glycine soja*)核心资源构建

赵丽梅^{①②} 董英山^{①②③*} 刘宝^① 郝水^① 王克晶^④ 李向华^④

(① 东北师范大学遗传研究所, 长春 130024; ② 吉林省农业科学院生物技术研究中心, 公主岭 136100; ③ 吉林大学生命科学院, 长春 130021; ④ 中国农业科学院作物科学研究所, 北京 100081. * 联系人, <mailto:mysdong@cjaas.com>)

摘要 对我国国家种质资源库中保存的 6172 份一年生野生大豆(*Glycine soja*)资源进行了系统的遗传多样性分析, 比较了核心种质构建中常用的随机取样、恒量取样、比例取样、对数取样和遗传多样性取样 5 种取样策略, 确认对于一年生野生大豆, 遗传多样性取样策略是最简单有效的取样策略. 根据起源和生态类型对一年生野生大豆资源进行分组, 应用遗传多样性策略和层次聚类, 构建一个野生大豆核心资源. 该核心资源包括 652 份材料, 取样比例为 10.65%. 代表性验证表明, 质量性状指标代表性为 100%, 平均指标代表性为 98.4%; 13 个指标群体结构相似系数为 0.96; 遗传多样性代表性为 81.38%; 20 对 SSR 引物扩增 299 份材料, 分析结果代表性为 83.64%.

关键词 野生大豆 核心资源 遗传多样性 取样策略

作物的遗传多样性是作物改良的基础, 世界各国对作物资源收集、保存、鉴定评价、利用等研究工作非常重视. 作物种质资源收集量的迅速增加, 给资源的保存与管理带来了更大的困难. 与此同时, 由于资源收集过程中的异地、异人及异系统性, 使得整个资源收集品中存在着极大的重复性. 明确作物的遗传多样性及多样性分布, 用最少的资源收集品, 最大限度保存作物遗传多样性成为资源工作者关注热点. 为此, Frankel^[1]于 1984 年首次提出了“核心资源”的概念, 引起了种质资源工作者的广泛关注. 在过去的 20 余年中, 有关学者分别从各种作物的遗传多样性, 核心资源构建方法、构建程序及核心资源利用方面进行了广泛的研究^[2~11], 使各种作物遗传多样性研究得到长足发展, 核心资源构建方法逐步完善与成熟. 目前小麦、大麦、花生、芝麻、苜蓿、水稻及多年生野生大豆等作物均已构建了核心资源^[12~20].

一年生野生大豆(*Glycine soja*)是栽培大豆的野生近缘种, 是我国作物资源中的瑰宝, 它分布于东亚, 包括朝鲜、日本、前苏联的远东和我国的中部. 我国自 20 世纪 70 年代后期开始进行一年生野生大豆资源研究, 目前已收集、分类、评价和保存 6000 余份一年生野生大豆资源, 并建立了数据库. 所收集的资源包括了全球一年生野生大豆资源的约 90%^[21]. 本文报道了利用所收集的全部一年生野生大豆资源, 以形态学、生态地理学、生物化学资料为基础构建的核心资源.

1 材料与方法

(i) 材料. 本研究所用材料和数据由中国农业科学院提供, 包括 6172 份一年生野生大豆资源的鉴定数据和部分种子. 这些材料来自我国 25 个省市, 7 个生态区, 评价了野生大豆的 22 个指标, 其中 12 个指标数据完整, 包括地理分布、叶形、茸毛色、花色、种皮色、脐色、子叶色、泥膜、蛋白含量、油份含量、百粒重和熟期.

(ii) 取样策略和统计分析. 根据收集来源及生态区, 将我国一年生野生大豆资源分为 8 组(表 1). 前 7 组来自 7 个自然生态区^[22,23], 后 1 组为人工杂交创新资源. 在取样量相同的条件下, 比较目前国际上通用的 5 种取样策略, 即随机取样策略、比例取样策略、恒量取样策略、对数取样策略及遗传多样性取样策略. 随机取样策略: 从各组中随机取样; 比例取样策略: 根据各组群体的大小按一定比例取样; 恒量取样策略

表 1 研究材料分组

分组	纬度/°N	份数	组号
I 极早熟类型组	48 以上	487	1
II 早熟类型组	44~48	2410	2
III 中早熟类型组	40~44	229	3
IV 中熟类型组	35~40	1628	4
V 中晚熟类型组	30~35	634	5
VI 晚熟类型组	27~30	137	6
VII 极晚熟类型组	24~27	525	7
VIII 杂交创新种质		122	8
合计		6172	

略: 在所有各组中, 不管群体的样本数量多少, 取样数量都相等; 对数取样策略: 根据各组群体大小按对数比例取样; 遗传多样性取样策略: 按照各组群体的遗传多样性取样.

依据下列公式计算遗传多样性指数、相似系数和变异系数:

(1) Shannon指数(H')^[24]:

$$H' = -\sum_{i=1}^S P_i \ln(P_i);$$

S : 小区中类型数;

i : 小区中第 i 个类型;

P_i : 第 i 个类型在小区中所占比例.

(2) Simpson指数(D)^[24]:

$$D = 1 - \sum_{i=1}^S P_i^2;$$

S : 类型数;

P_i : 第 i 个类型占总类型数的比例.

(3) McIntosh指数(D_{Mc})^[25]:

$$D_{Mc} = \frac{N - \sqrt{\sum_{i=1}^S n_i}}{N - \sqrt{N}};$$

n_i : 第 i 个类型数;

N : 类型总数.

(4) 变异系数 CV ^[24]:

$$CV(\%) = \frac{1}{m} \sum_{j=1}^m \frac{S_j}{\bar{X}} \times 100;$$

m : 指标数;

S_j : 第 j 个指标的误差标准;

$\bar{X}$: 第 j 个指标的平均数.

(iii) 核心资源材料的筛选. 根据选取的资源收集品对原始资源收集品的代表性, 确定取样策略和取样比例. 按照确定的取样策略, 根据资源收集品的来源、地理分布、对材料进行聚类, 从中随机选择资源材料, 按照确定的取样比例构建核心资源.

(iv) 对核心资源的评价. 通过比较核心资源群体和原始资源群体的各项指标、群体结构、遗传多样性、地理分布和分子标记多态性等对核心资源进行评价. 选择第 2 组的核心资源材料和该组核心资源以外的材料共计 299 份, 其中核心资源材料 128 份, 核心资源以外材料随机选取 171 份, 按常规方法用 20 个 SSR 分子标记鉴定、评价.

2 结果与分析

2.1 取样策略和取样比例

采取什么样的取样策略从如此庞大的资源中选出最具代表性的材料是构建一年生野生大豆核心资源的关键. 本研究采用Brown^[2]提出的 10%的取样比例, 比较国际上通用的 5 种取样策略(表 2). 从各种取样策略在各组的取样比例看, 随机取样策略在各组取样数取样比例差异很大, 且各组的取样比例与各组的遗传多样性不成比例; 比例取样策略不仅保证了各组的群体与比例, 而且降低了材料少、多样性较高组的取样比例. 恒量取样与比例取样恰好相反. 对数取样在综合了比例取样和恒量取样特点后却过分

表 2 5 种取样策略各组取样比较

组号		取样策略					原始资源
		随机取样	比例取样	恒量取样	对数取样	多样性取样	
1	取样数	106	48	77	74	40	487
	比例	21.8	10	15.8	16	8.2	
2	取样数	333	241	77	99	181	2410
	比例	13.8	10	3.2	4.1	7.5	
3	取样数	10	23	77	69	29	229
	比例	3.3	10	33.6	30.1	12.7	
4	取样数	111	163	77	91	178	1628
	比例	6.8	10	4.7	5.6	10.9	
5	取样数	21	63	77	81	92	634
	比例	3.3	9.9	12.1	12.8	14.5	
6	取样数	2	14	77	60	18	137
	比例	1.5	10.2	56.2	44.4	13.1	
7	取样数	31	53	77	79	52	525
	比例	5.9	10.1	14.7	15	9.9	
8	取样数	4	12	77	60	16	122
	比例	3.3	10	63.1	49.2	13.1	

表 3 5 种取样策略的评价

各项指标	策略										原始群体
	随机		恒量		比例		对数		多样性		
	获得 指标数	符合度	获得 指标数	符合度	获得 指标数	符合度	获得 指标数	符合度	获得 指标数	符合度	
类型 ^{a)}	136	35.8	133	35.1	128	33.8	124	32.7	348	91.8	379
茸毛色	2	100	2	100	2	100	2	100	2	100	2
花色	2	100	2	100	2	100	2	100	2	100	2
叶形	6	85.7	7	100	5	71.4	5	71.4	7	100	7
粒色	5	100	5	100	5	100	5	100	5	100	5
脐色	2	66.7	2	66.7	3	100	3	100	3	100	3
子叶色	2	100	1	50	1	100	2	100	2	100	2
泥膜	3	100	3	100	3	100	3	100	3	100	3
蛋白	19	79.2	22	91.7	21	95.8	23	95.8	23	95.8	24
粗脂肪	17	94.4	17	94.4	17	100	18	100	18	100	18
百粒重	22	91.7	24	100	18	79.2	19	79.2	23	95.8	24
生育期	110	73.3	118	78.7	102	70.7	106	70.7	125	83.3	150
平均		85.6		84.7		82.9		88		97.2	

a) 根据野生大豆的农艺性状叶形、茸毛色、花色、种皮色、脐色、子叶色、泥膜, 将野生大豆划分为 379 个类型; 茸毛色为棕色、灰色; 花色为白色、紫色; 叶形分为线叶、卵圆、披针、上披针、椭圆、圆以及其他; 粒色分为褐色、双色、黑色、黄色以及其他; 脐色为褐、黑、兰、绿、双色、黄、无色; 子叶色为黄色或其他色; 泥膜为有、无、光滑; 蛋白质含量、粗脂肪含量以 1% 为一个级; 百粒重以 1 g 为一个级; 生育期以 10 d 为一个级。下同

放大了材料少的组的取样比例; 而遗传多样性策略则按各组材料多少及遗传多样性大小形成了各组比较合理的取样比例。进一步比较各种取样策略所获得资源收集品对原始收集品的符合度(表 3)可以看出, 遗传多样性取样策略的符合度明显高于其他 4 种取样策略, 其各项指标符合度达到 97.2%, 而其他 4 种取样策略即随机取样、恒量取样、比例取样及对数取样的符合度分别为 85.6%, 84.7%, 82.9% 及 88%。因此我们确定遗传多样性取样策略是构建一年生野生大豆资源核心群体的最佳取样策略。

适合的取样比例是构建核心资源的中心。在确定了构建一年生野生大豆核心资源最佳取样策略后, 以遗传多样性策略按不同的比例对一年生野生大豆进行取样, 计算不同取样比例所获资源收集品对原始收集品的符合度。在 9.8% 以下时随取样比例增加, 其符合度迅速增加(表 4)。高于 9.8%, 随取样比例增

加, 符合度增加缓慢。因而确定 9.8% 左右是构建一年后生野生大豆核心资源的适宜取样量。

2.2 核心资源收集品的筛选

根据收集来源及生态区, 可将我国一年生野生大豆分为 8 组(表 1)。在野外收集的 7 个组中, 早熟类型区(第 2 组)和中熟类型区(第 4 组)收集材料最多, 分别为 2410 份和 1628 份。晚熟类型区最少为 137 份。8 个组的遗传多样性指数(Shannon 指数)分别为 4.48, 4.56, 4.66, 5.11, 4.99, 4.73, 4.56 和 5.17。依据 8 个质量性状指标将所有收集品划分成 379 个类型。根据同一类型在同一组中的数量占资源收集品总量比例多于或少于 5%, 将各组分 2 部分(表 5)。少于 5% 的部分共有 2363 份材料, 占资源总量的 38.3%。组内同一类型随机选取 1 份作为预选核心收集品, 共选取 623 份材料建成预选核心库 C1。取样比例为 26.4%,

表 5 各组取样方式

组	取样方式	
	随机	聚类
1	203	284
2	852	1558
3	153	76
4	578	1050
5	271	363
6	31	106
7	153	372
8	122	0
合计	2363	3809

表 4 不同取样比例的符合度

取样比例(%)	符合度(%)
0.5	35.4
1	51
2	61.8
4	71.6
6	74.6
7.9	81.9
9.8	91.6
12.2	96.4
14.3	98.1
16.2	98.9

表 6 各组取样比例

组号	原始资源	预选资源		核心资源		
		份数	对原始资源比例(%)	份数	对预选资源比例(%)	对原始资源比例(%)
1	487	61	12.5	42	68.85	8.62
2	2410	248	10.3	197	79.43	8.17
3	229	46	20.1	30	65.22	13.10
4	1628	247	15.2	197	79.75	12.10
5	634	119	18.8	97	81.51	15.30
6	137	29	21.2	17	58.62	12.41
7	525	61	11.6	55	90.16	10.48
8	122	14	11.5	17	121.43	13.93
合计/平均	6172	825	13.4	652	79.03	10.56

库 C1 的变异系数为 34.1%，多样性指数为 4.2。多于 5%的材料共有 3809 份，占资源总量的 61.7%。对其进行聚类，同一聚类组内随机取一份；然后把这些收集品合起来进一步聚类，同一聚类组内随机取一份；如此往复，聚类后计算其多样性指数。当剩余 202 份材料时(C2)，其变异系数为 32.9%，多样性指数为 3.6，与 C1 接近，取样比例为 5.3%。合并 C1、C2 形成含 825 份材料的预选核心库 C。

对预选核心库 C 进行繁殖鉴定，用鉴定数据对其进行校正，然后对其进行聚类，从中筛选出 606 份材料。进一步补充极端及特殊类型，最后构建成拥有 652 份材料的我国野生大豆核心资源。总的取样比例为 10.52%，各组取样比例分别为 8.62%，8.17%，13.10%，12.10%，15.30%，12.41%，10.48%和 13.93% (表 6)。

2.3 核心资源的评价

(i) 各项指标覆盖率。用 21 个指标比较核心资源和原始资源(表 7)。质量性状指标比较有或无，数量性状指标分级比较同一级的有或无。分级标准为：蛋白质含量和脂肪含量 2%为一级，百粒重 1 g 为一级，生育期 10 d 为一级。其结果是 18 个指标的覆盖率为 100%，蛋白质、百粒重、生育期 3 个性状指标的覆盖率为 91.6%，95.8%和 78.7%。平均覆盖率为 98.4%。

(ii) 群体结构的比较。用 13 个指标计算核心资源群体与原始资源群体各项指标的变异系数，比较 2 个群体结构差异(图 1)。除主茎类型与百粒重外，其群体内其他各项指标变异极为相似，核心资源与基础群体的相似系数为 0.96。

(iii) 多样性的代表性。从 2 个群体的遗传多样性指数来看(表 8)，原始资源群体的多样性指数的 3 个指标 Simpson 指数、Shannon 指数、McIntosh 指数

表 7 原始资源与核心资源各指标的符合度

各项指标 a)	原始群体原有 指标数	核心群体获得 指标数	符合度(%)
生态区	8	8	100
茸毛色	2	2	100
花色	2	2	100
叶形	7	7	100
粒色	5	5	100
脐色	7	7	100
子叶色	2	2	100
泥膜	3	3	100
蛋白质	24	22	91.6
粗脂肪	17	17	100
类型	379	379	100
百粒重	24	23	95.8
生育期	150	118	78.7
病毒	3	3	100
孢囊 1	2	2	100
孢囊 2	3	3	100
孢囊 3	3	3	100
孢囊 4	2	2	100
孢囊 5	4	4	100
纬度	29	29	100
经度	35	35	100
平均			98.4

a) 生态区为野生大豆收集的 7 个生态区，加一个人工杂交材料(表 1)；病毒为花叶病毒的 3 个级别；孢囊 1~5 为抗孢囊线虫的 1~5 的生理小种；纬度、经度以 1 度为一个级别。下同

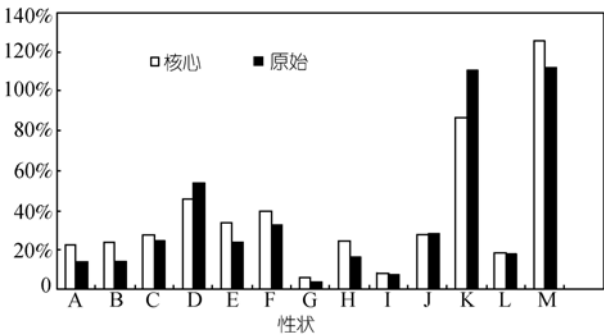


图 1 核心资源与原始资源各项指标变异系数比较

A, 茸毛色; B, 花色; C, 叶形; D, 主茎; E, 粒色; F, 脐色; G, 子叶色; H, 泥膜; I, 蛋白质; J, 脂肪; K, 百粒重; L, 生育期; M, 海拔高度

表 8 原始资源与核心资源多样性比较^{a)}

各项指标	Simpson(J)指数		Shannon(H) 指数		McIntosh 指数	
	原始群体	核心群体	原始群体	核心群体	原始群体	核心群体
茸毛色	0.15611	0.34124	0.42231	0.75611	0.0824	0.1957
花色	0.15028	0.35974	0.40856	0.78605	0.07919	0.20763
叶形	0.65494	0.69519	1.67446	1.88692	0.41781	0.46516
主茎	0.51129	0.57359	1.23525	1.34236	0.30474	0.36044
泥膜	0.27271	0.55695	0.73848	1.33273	0.14906	0.34734
子叶色	0.00259	0.00613	0.0143	0.03003	0.00131	0.00319
脐色	0.49626	0.51153	1.01018	1.09284	0.29394	0.31279
粒色	0.52443	0.71679	1.50025	2.08858	0.31432	0.48582
区号	0.7518	0.78174	2.37582	2.4869	0.50815	0.5532
蛋白质	0.9142	0.91541	3.77304	3.81276	0.71594	0.73555
脂肪	0.88471	0.91987	3.4692	3.76349	0.66875	0.74357
百粒重	0.57377	0.84539	2.14645	3.21354	0.35154	0.62981
生育期	0.85644	0.8696	3.15886	3.24228	0.62893	0.66302
海拔	0.82807	0.82097	2.97264	2.99818	0.59274	0.59885
合计	7.5776	8.91414	24.8998	28.83277	5.10882	6.30207
平均	0.54	0.64	1.78	2.06	0.37	0.45
代表性(%)	81.48		84.27		78.38	

a) 代表性=[1-(核心多样性指数-原始群体多样性指数)/原始群体多样性指数]×100

表 9 地理分布比较

水平分布(依据 2 个纬度, 3 个经度划分一个小区)									垂直分布		
小区号	取样量		小区号	取样量		小区号	取样量		海拔	取样量	
	原始资源	核心资源		原始资源	核心资源		原始资源	核心资源		原始资源	核心资源
1	3	3	28	11	1	55	4	1	<100	1803	214
2	7	3	29	25	1	56	36	8	100~200	1199	125
3	3	1	30	32	3	57	93	11	200~300	914	65
4	197	13	31	182	18	58	98	9	300~400	453	38
5	62	5	32	169	9	59	78	12	401~500	270	22
6	116	9	33	86	6	60	59	7	501~600	144	15
7	29	1	34	76	13	61	42	3	601~700	76	9
8	18	1	35	5	2	62	5	1	701~800	174	35
9	63	11	36	10	1	63	185	17	801~900	52	7
10	8	1	37	20	1	64	548	56	901~1000	49	8
11	13	1	38	18	2	65	60	2	1001~1100	68	11
12	11	1	39	55	8	66	51	4	1101~1200	38	6
13	22	5	40	193	41	67	11	3	1201~1300	23	1
14	60	7	41	76	6	68	1	1	1301~1400	39	6
15	5	1	42	185	35	69	5	1	1401~1500	15	0
16	16	1	43	16	1	70	4	1	1501~1600	9	1
17	9	1	44	7	1	71	54	8	1601~1700	10	1
18	62	1	45	288	17	72	431	34	1701~1800	5	1
19	32	3	46	76	5	73	511	31	1801~1900	2	1
20	34	13	47	28	2	74	446	40	1901~2000	1	0
21	108	17	48	96	16	75	236	22	2601~2700	1	1
22	167	39	49	30	2	76	80	7	3601~3700	1	1
23	22	1	50	37	1	77	34	3	无	826	84
24	23	6	51	23	6	78	34	3			
25	2	1	52	27	4	79	53	5			
26	51	4	53	2	1	80	1	1			
27	63	3	54	33	5						

分别为 0.54, 1.78, 0.37; 相对的核心资源的 3 个指标分别为 0.64, 2.06 和 0.45. 核心资源在 3 个指标分别丢掉 0.1, 0.28 和 0.08 的遗传多样性, 代表性分别为 81.48%, 84.27%, 78.38%, 平均遗传多样性符合度为 81.38%.

(iv) 地理分布的比较. 原始资源群体来自全国 7 个生态区 79 个地理小区, 而这 79 个生态小区的材料在构建的核心资源中均被包括(表 9), 水平地理分布的覆盖率为 100%.

垂直分布上, 原始资源群体的材料来自海拔 1.8~2650 m 的地区, 核心资源的垂直分布上与原始资源群体一致.

(v) 分子水平验证. 以核心资源第 2 组为分子标记检测对象, 在该组核心资源中随机选取 128 份材料, 占该组核心资源的 65%, 剔除核心资源后在该组随机取 171 份材料, 占该部分的 7.4%, 温室种植. 当植株长到 6 片叶时取材提 DNA.

用其中的 30 份材料对 600 对 SSR 引物进行筛选, 筛选出 60 对多态性好的 SSR 引物. 用此 60 对引物对

299 份材料进行扩增, 从中选择效果好的 20 对 SSR 引物进行分析(表 10). 20 对 SSR 引物共扩增出 549 条带, 平均每对引物扩增 27.5 条带. 核心资源相对于原始资源在这 20 对 SSR 引物所代表的等位基因中丢失 93 个, 平均每对引物丢失 4.7 条带. 在第 2 组中, 核心资源与原始资源的相似系数为 0.83, 核心资源对原始资源的代表性为 83.64%.

通过以上研究, 构建了一个完善的、有代表性的我国一年生野生大豆核心资源. 核心资源中的样品来自 7 个生态区及 1 个人工创造的材料组; 地理分布包括北起黑龙江省的塔河(53°N), 南到广西省的象州(24°N), 东起黑龙江的绕河(134°E), 西至西藏的察禹(97°E); 海拔高度 1.8~2650 m, 涵盖了原始资源中所有指标、变异范围、遗传多样性及 83.64%的等位基因, 为评价、管理和利用我国野生大豆资源提供了一个有价值的核心资源.

3 讨论

3.1 核心资源的构建

Brown^[2]认为, 构建核心资源的关键是材料的分组, 分组方法常见的有按植物分类地位、地理分布、农业生态区等分组. 一年生野生大豆为一个种, 虽然有关学者根据种粒大小、叶形等对其进行了分类, 但对其结果却有很大的争议. 已收集的一年生野生大豆资源主要来源于野外收集及人工杂交后代. 而野外收集材料按光温生态区分布可分成 7 个区, 据此将一年生野生大豆分为 8 个组. 保证了各来源及各生态区一年生野生大豆遗传多样性的完整性.

对各组材料的进一步筛选是核心资源构建的极为重要部分, 也是最大程度保存遗传多样性的关键. 目前常采用的策略有 5 种, 即随机取样策略(R 策略)、等比取样策略(C策略)、比例取样策略(P策略)、对数取样策略(L 策略)及多样性取样策略(G策略). Yonezawa^[11]研究了 5 种取样策略后指出, 当对所收集的资源掌握一定数据时, 遗传多样性策略是最佳策略. 其他策略是在对所收集资源掌握数据很少时应用. 我国从“六五”就开始进行一年生野生大豆资源研究, 掌握了大量的数据资料. 用 10%取样量研究比较 5 种策略, 结果证明, G策略的符合度明显高于其他 4 种策略.

核心资源的取样量是核心资源构建的中心. Brown^[2]认为取样量 10%作为核心资源, 可以保证 70%

表 10 原始资源与核心资源 SSR 分子标记相似系数比较

引物	相似系数	丢带数	总带数	代表性(%)
satt005	0.74	5	33	84.84
satt168	0.95	2	24	91.67
satt180	0.81	7	25	72
satt184	0.85	3	29	89.66
satt216	0.65	9	34	73.53
satt236	0.94	1	17	94.11
satt239	0.8	8	29	72.41
satt243	0.83	8	30	73.33
satt267	0.83	6	22	72.73
satt268	0.88	2	22	90.9
satt300	0.87	3	26	88.46
satt307	0.88	5	32	84.38
satt334	0.86	1	25	96
satt345	0.88	5	31	83.87
satt352	0.92	2	29	93.1
satt386	0.88	3	25	88
satt390	0.71	6	33	81.82
satt434	0.87	4	24	83.33
satt442	0.7	10	34	70.59
satt453	0.78	3	25	88
合计	16.63	93	549	1672.73
平均	0.83	4.7	27.5	83.64

的遗传背景的概率在 95%水平. 这个数据说明了一定的规律, 但并不是绝对的, 这要看所收集的基础收集品的重复程度. 我们采用遗传多样性策略进行不同的取样量比较, 发现 10%左右的取样量是一个最佳取样点, 在其以下, 随取样量增加, 代表性迅速增加; 在其以上, 随取样量增加, 代表性增加缓慢. 因而, 确定 10%为一年生野生大豆核心资源的取样量. 为了达到这一标准, 对预选核心收集品进行了进一步聚类, 当数量接近 10%左右时, 停止聚类, 进一步补充形成核心资源.

以上方法是在分析研究了国际上通用的核心资源构建方法, 结合我国一年生野生大豆资源的具体情况确定的, 实践证明是切实可行的. 所构建的核心资源良好的代表了原始资源, 也为其他作物核心资源的构建提供了参考.

3.2 核心资源的评价

核心资源的评价是检验核心资源代表性的标准. 对核心资源的验证有许多种, 最常见的有 2 类: 一类为比较预选核心资源与原始资源, 然后比较核心资源与预选核心资源; 另一类为直接比较核心资源与原始资源. 后一类更能反映出核心资源对原始资源的代表性. 比较的指标主要有农艺性状与数量性状的特征值、标准差、变异系数等, 也有比较生理生化指标. DNA 指标虽然是最佳指标, 但由于成本高, 很少有人应用. 本研究比较了核心资源与原始资源的 21 个指标的特征值、变异系数、3 个多样性指数、地理分布及 DNA 水平的符合度, 从性状有无、群体结构、遗传多样性及等位基因等几个方面验证了核心资源对原始资源的代表性, 更为完全.

3.3 核心资源的完善与发展

我国是一年生野生大豆的最主要起源地, 遗传多样性极为丰富^[25,26], 使我国在一年生野生大豆研究方面有得天独厚的优势. 核心种质的构建是保持我国这种优势的关键, 也是一年生野生大豆资源研究及利用的基础. 由于群体的极大缩小, 因而可对整个核心种质群体进行系统研究, 研究其遗传多样性、起源等, 筛选优异基因, 最大限度地发挥资源的作用. 然而, 核心资源并不是永恒的, 还必须在研究应用中不断发展、完善. 将研究筛选的优异材料应用于育种及生物技术, 再将此过程中创造的新的资源不断增加到核心资源中, 同时也把核心资源中已被更好资

源代替的材料淘汰. 总之, 核心资源研究是一个动态过程, 要在不断研究应用中完善.

致谢 中国农业科学院作物科学研究所的邱丽娟、贾继增研究员及“973”专家组的成员对本研究的完成提出了宝贵的意见, 特此致谢. 国家自然科学基金(批准号: 30370890, 30490250)、国家转基因专项(批准号: JY 03-B-16)和国家杰出青年科学基金(批准号: 30225003)资助项目.

参 考 文 献

- 1 Frankel O H. Genetic perspectives of germplasm conservation. In: Arber W, Llimensee K, Peacock W J, et al. eds. Genetic Manipulation: Impact on Man and Society. Cambridge: Cambridge University Press, 1984. 179~183
- 2 Brown A H D. Core Collections: A practical approach to genetic resources management. *Genome*, 1989, 31: 818~824
- 3 Chandra S, Huaman Z, Krishna H S, et al. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data: A simulation study. *Theor Appl Genet*, 2002, 104: 1325~1334[DOI]
- 4 Freyssinet G M. Structure and evolution of the small and large RuBisCo subunits from various photosynthetic organisms. *Plant Physiol Biochem*, 1991, 29: 517~529
- 5 Galwey N W. Verifying and validating the representativeness of a core collection. In: Hodgkin T, Brown A H D, van Hintum, et al. eds. Core Collection of Plant Genetic Resources. UK, London: IPGR Press, 1995. 187~198
- 6 Hamon S, Dussert S, Noirot M, et al. Core Collections: Accomplishments and challenges. *Plant Breeding Abstracts*, 1995, 65: 1125~1133
- 7 Hu J, Zhu J, Xu H M. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theor Appl Genet*, 2000, 101: 264~268[DOI]
- 8 Li C T, Shi C H, Wu J G, et al. Methods of developing core collections based on the predicted genotypic value of rice (*Oryza sativa* L.). *Theor Appl Genet*, 2004, 108: 1172~1176[DOI]
- 9 Ortiz R, Ruiz-Tapia E N, Mujica-Sanchez A. Sampling strategy for a core collection of Peruvian quinoa germplasm. *Theor Appl Genet*, 1998, 90: 475~483[DOI]
- 10 Skinner D Z, Bauchan G R, Auricht G, et al. A method for the efficient management and utilization of large germplasm collections. *Crop Sci*, 1999, 39: 1237~1242
- 11 Yonezawa K, Nomura T, Morishima H. Sampling strategies in core collections. In: Hodgkin T, Brown A H D, van Hintum Th J L, et al. eds. Core Collection of Plant Genetic Resources. UK, London: IPGR Press, 1995. 35~54
- 12 Basigalup D H, Barnes D K, Stucker R E. Development of a core collection for perennial Medicago plant introductions. *Crop Sci*,

- 1995, 35:1163~1168
- 13 Brown A H D, Grace J P, Speer S S. Designation of a "core" collection of perennial Glycin. Soybean Genetic Newsletter, 1987, 14: 59~70
- 14 Casler M D. Patterns of variation in collection of perennial ryegrass accessions. Crop Sci, 1995, 35: 1169~1177
- 15 Holbrook C C, Anderson W F, Pittman R N. Selection of a core collection from the U.S. Germplasm Collection of Peanut. Crop Sci, 1993, 33: 859~861
- 16 Hamon S, Noirot M, Anthony F D. Developing a coffee core collection using the principal components score strategy with quantitative data. In: Hodgkin T, Brown A H D, van Hintum Th J L, et al. eds. Core Collection of Plant Genetic Resources. UK, London: IPGR Press, 1995. 117~126
- 17 Martynov S P, Dobrotvorskaia T V, Dotlacil L, et al. Genealogical approach to the formation of the winter wheat core collection. Genetika, 2003, 39: 1091~1098
- 18 Diwan N, Bauchan G R, McIntosh M S. A core collection for the United States annual Medicago germplasm collection. Crop Sci, 1994, 34: 279~285
- 19 van Hintum T J L, Haalman D. Pedigree analysis for composing a core collection of modern cultivars with examples from barley (*Hordeum vulgares* Lat). Theor Appl Genet, 1998, 88: 70~74
- 20 Skroch P W, Nienhuis J, Tohme B S, et al. Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collection. Crop Sci, 1998, 38: 448~496
- 21 徐豹. 中国野生大豆(*Glycine soja*)蛋白质含量的多样性和地理分布. FAO/IBPGR 植物遗传资源通讯, 1993, 94/95: 1~3
- 22 李福山. 中国野生大豆资源的地理分布及生态分划研究. 中国农业科学, 1993, 26(2): 47~55
- 23 徐豹. 中国野生大豆生态类型的研究. 中国农业科学, 1987, 20(5): 29~35
- 24 唐启义, 冯明光. 实用统计分析及其 DPS 数据处理系统. 北京: 中国农业出版社, 1997. 119~120
- 25 Dong Y S, Zhuang B C, Zhao L M, et al. The genetic diversity of annual wild soybeans grown in China. Theor Appl Genet, 2001, 103: 98~103[DOI]
- 26 徐豹, 庄炳昌, 徐航, 等. 中国野生大豆(*Glycine soja*)脂肪含量多样性及地理分布. 大豆科学, 1993, 12(4): 269~274
- (2004-11-20 收稿, 2005-04-06 收修改稿)