

普通野生稻和亚洲栽培稻核心种质遗传多样性的检测研究*

孙传清 李自超 王象坤

(中国农业大学植物遗传育种系, 北京 100094)

提 要 以122份野生稻和75份栽培稻在44个RFLP位点的多态性为资料, 采用逐步聚类法和分组随机法, 按原始样本的50%、20%和10%, 分别构建初级核心样本、次级核心样本和核心样本, 用多态位点数(N_p)、等位基因数(N_a)、基因型数(N_g)及平均基因多样性(H_s)等参数, 检测其遗传多样性。结果表明, 初级核心样本的 N_p 、 N_a 和 N_g 分别达到原始样本的90%、90%和80%以上。次级核心样本的 N_p 、 N_a 、 N_g 仍分别可达原始样本的90%、80%和75%。核心样本的 N_p 可达原始样本的90%, 而 N_a 和 N_g 下降幅度大, 分别相当于原始样本的65%和55%。无论哪一级核心样本, 栽培稻遗传多样性的减少幅度比野生稻小。比较逐步聚类法和分组随机法表明, 逐步聚类法构建的核心样本比分组随机法更能保持较大的遗传多样性。作者认为检验核心种质遗传多样性的首选参数是等位基因数。

关键词 水稻; RFLP; 核心种质; 遗传多样性

Studies on Evaluation of the Genetic Diversity of Core Collection of Common Wild Rice (*Oryza rufipogon* Griff.) and Asian Cultivated Rice (*Oryza sativa* L.)

SUN Chuan-Qing LI Zi-Chao WANG Xiang-Kun

(Department of Plant Genetic and Breeding, China Agricultural University, Beijing 100094, China)

Abstract Primary core samples (PCS), secondary core samples (SCS) and core samples (CS) as 50%, 20% and 10% of the number of entries of original samples (OS) were constructed using the methods of sampling progressively by clustering (SPC) and sampling randomly after grouping (SA) based on RFLP data at 44 loci of 122 accessions of *Oryza rufipogon* and 75 varieties of *Oryza sativa*. The genetic diversity of PCS, SCS and CS were compared by parameters of number of polymorphic loci (N_p), number of alleles (N_a), number of genotypes (N_g) and average gene diversity (H_s) at the 44 loci detected. It was indicated that the N_p and N_a of PCS were about 90% of those of OS, and the N_p of PCS was about 80% of that of OS. The N_p , N_a and N_g of SCS could be at the level of 90%, 80% and 75% of OS respectively. It was found that the N_p of CS reached the level of 90% of OS while the N_a and N_g of CS were only about 65% and 55% of these of OS respectively. It was also found that the genetic diversity of cultivated rice decreased more slowly than that common wild rice in PCS, SCS and CS. Through comparison of the genetic diversity of SCS and CS constructed by SPC and SRG, it is indicated that core collection constructed by SPC could maintain more genetic diversity of the original resources. It was suggested that the number of alleles was the best parameter for detecting genetic diversity of core

* 本研究系国家重点基础研究规划(973计划)项目之一(G1998010201)

收稿日期: 1999-04-20, 接受日期: 1999-12-15

Received on: 1999-04-20, Accepted on: 2000-12-15

collection of rice

Key words Rice; RFLP; Core collection; Genetic diversity

植物遗传资源是作物新品种选育的重要物质基础, 世界各国都非常重视遗传资源的搜集与保护。据不完全统计, 全世界已有植物种质资源450多万份。然而, 种质库中资源数量的急剧膨胀, 给资源的保存、研究与利用带来了困难。Frankel于1984年提出了核心种质(Core collection)的概念, Brown^[1](1989a)将其进一步发展, 即通过一定的方法从整个种质资源中选择一部分样本, 以最小的资源数量, 尽可能最大限度的代表整个资源的多样性。至今, 已在30多种作物上构建了核心种质。

在核心种质的构建中取样方法是关键。取样方法一般来说分为两种, 即完全随机取样与系统取样。完全随机取样是在整个资源中随机取样, 而系统取样则考虑了材料的遗传结构, 给不同材料以不同权重, 采用分层取样方法。实际上完全随机的方法用得较少, 一般是将系统取样与随机取样相结合, 即先将所有材料分组或分类, 在组内或群内再用随机取样的方法。本研究是以122份野生稻和75份栽培稻在44个RFLP位点的多态性为资料, 试探不同取样方法和不同取样比例, 构建出的核心种质的遗传多样性的变化, 为今后栽培稻和野生稻核心种质的构建及遗传多样性的检测提供依据。

1 材料与方法

1.1 材料

供试材料由122份普野和75份栽培稻组成。122份普野中, 中国39份, 南亚40份, 东南亚43份; 75份栽培稻来自亚洲11个国家, 根据我们分类研究的结果^[2], 其中37份为籼稻, 38份为粳稻。每份材料取1个个体用于RFLP分析。

1.2 RFLP分析

用CTAB法^[3]提取总DNA, 每个样本取3 μg DNA用*Dra*I或*Hind*III消化后, 电泳、印迹及RFLP的检出按照文献^[2]方法进行。本研究选用44个单拷贝探针进行杂交。

1.3 核心样本的构建

1.3.1 逐步聚类法 先将197份材料分为野生稻和栽培稻两大组, 两大组分别按照UPGMA^[4]法进行聚类分析, 根据聚类图, 在同一小组内采用随机的方法, 按照50%的取样比例, 构建样本数为总数50%的初级核心样本。然后对初级核心样本进行聚类分析, 逐步压缩为原始样本20%的次级核心样本。再在次级核心种质的基础上进行聚类分析和逐步压缩为原始样本10%的核心种质。由50%到20%及10%的方法为, 两个遗传距离最小的样本在去掉1个时, 不是采用随机的方法, 而是将其中与相邻距离较小的1个去掉。

1.3.2 分组随机法 先将197份材料分为野生稻和栽培稻两大组, 野生稻按照来源分为南亚、东南亚和中国3个小组, 栽培稻则根据以前的分类结果分为籼粳两组, 每一组内采用随机的方法, 按照20%、10%的比例, 分别构建次级核心样本和核心种质, 并分析其遗传多样性。

1.4 遗传多样性的评价

对单拷贝的探针而言, 每个探针检测了一个位点, 每一条多态性带可视为一个等位基因, 每一种带型可视为一种基因型。用4个参数即多态位点数(N_p)、等位基因数(N_a)、基因型数(N_g)、平均基因多样性(H_s)评价原始样品和核心样品的遗传多样性^[5]。

2 结果与分析

2.1 逐步聚类取样法不同级别核心样本遗传多样性的比较

2.1.1 初级核心样本 从表1可以看出, 由原始样本压缩成初级核心样本, 无论是野生稻还是栽培稻, 多态位点数(N_p)能保持原始样本的98%以上, 变化很小。等位基因数(N_a)和基因型数(N_g)分别能保持在原始样本的90%和84%以上。平均基因多样性(H_s)除整个初级核心样本外, 栽培稻和野生稻内变化很小。说明将原始样本压缩50%后, 其遗传多样性基本与原始样本一致。

表1 原始样本、初级核心样本、次级核心样本及核心样本的遗传多样性比较
Table 1 Comparison of genetic diversity of original samples, primary core samples, secondary core samples and core samples

所有样品 All samples		野生稻 (<i>O. rufipogon</i>)				栽培稻 (<i>O. sativa</i>)			
		合计 Total	南亚 South A sia	东南亚 Southeast A sia	中国 China	合计 Total	籼稻 <i>Indica</i>	粳稻 <i>Japonica</i>	
N_p	OS	44	43	37	38	40	33	28	29
	PCS	43	42	34	37	40	33	28	29
	PCS/OS	0.98	0.98	0.92	0.97	1.00	1.00	1.00	1.00
	SCS	42	42	34	35	38	33	26	28
	SCS/OS	0.95	0.98	0.92	0.92	0.95	1.00	0.93	0.97
	CS	41	40	27	27	32	32	23	26
	CS/OS	0.93	0.93	0.73	0.71	0.80	0.97	0.82	0.90
N_a	OS	181	177	117	125	136	106	91	87
	PCS	163	160	106	112	124	104	87	86
	PCS/OS	0.90	0.90	0.91	0.90	0.91	0.98	0.96	0.99
	SCS	147	141	96	94	111	97	79	82
	SCS/OS	0.81	0.80	0.82	0.75	0.82	0.92	0.87	0.94
	CS	120	113	77	79	91	87	69	76
	CS/OS	0.66	0.64	0.66	0.63	0.67	0.82	0.76	0.87
N_g	OS	241	232	127	138	175	124	99	92
	PCS	210	194	114	121	153	113	88	89
	PCS/OS	0.87	0.84	0.90	0.88	0.87	0.91	0.89	0.97
	SCS	185	175	101	98	133	107	82	83
	SCS/OS	0.77	0.75	0.80	0.71	0.76	0.86	0.83	0.90
	CS	137	128	78	80	96	91	69	76
	CS/OS	0.57	0.55	0.61	0.58	0.55	0.73	0.70	0.83
H_s	OS	0.3567	0.3672	0.3131	0.3036	0.3543	0.2939	0.1849	0.1502
	PCS	0.3068	0.38	0.289	0.2901	0.3584	0.2992	0.1844	0.174
	SCS	0.3738	0.3856	0.3258	0.3212	0.3385	0.3129	0.1911	0.2301
	CS	0.3832	0.3963	0.2854	0.294	0.3459	0.3134	0.1747	0.2521

OS: 原始样本, Original samples; PCS: 初级核心样本(取样比例为50%), Primary core samples as 50% of original samples; SCS: 次级核心样品(取样比例为20%), Secondary core samples as 20% of original samples; CS: 核心样品(取样比例为10%), Core samples as 10% of original samples; N_p : 多态位点数, N_a : polymorphic loci; N_a : 总等位基因数, Total no. of alleles; N_g : 基因型数, No. genotypes; H_s : 平均基因多样性, A verage gene diversity.

2.1.2 次级核心样本 虽然次级核心样本数只有原始样本的20%, 但其多态位点数和等位基因数仍达到原始样本的90%以上, 等位基因数和基因型数在初级核心样本的基础上又下降了10%, 仍保留原始样本的80%和75%左右。其中栽培稻的等位基因数和基因型数比野生稻减少的幅度小, 能分别达到原始样本的90%和85%左右。平均基因多样性(H_s)除粳稻的变

化较大外, 其它材料的 H_s 与原始样本相比变化较小, 仍能表现为野生稻大于栽培稻, 中国野生稻大于南亚和东南亚野生稻的趋势。

2.1.3 核心样本 由次级核心样本再压缩成核心样本, 对整个样本及野生稻和栽培稻而言, 多态位点数仍能达到原始样本的90%以上, 而具体对某组或某类材料, 如南亚和东南亚野生稻则下降幅度较大。整个样本和野生稻的等位基因数和基因型数在次级核心样本的基础上分别又下降了15%和20%, 只分别保留了原始样本的65%和55%左右。栽培稻的等位基因数和基因型数则在次级核心样本的基础上, 再下降10%和15%左右, 分别达到原始样本的80%和70%左右。

2.2 逐步聚类法与分组随机法比较

2.2.1 20% 取样比例构建的次级核心样本 从表2可知, 按照20%的取样比例, 用逐步聚类法构建的次级核心样本的多态位点数、等位基因数、基因型数及平均基因多样性等遗传多样性参数均高于分组随机法构建的次级核心样本。

表2 20% 取样比例逐步聚类取样与随机取样次级核心样品遗传多样性的比较

Table 2 Comparison of genetic diversity of secondary core samples selected by SPC and SR

		所有样品 All of samples	野生稻 (<i>O. rufipogon</i>)				栽培稻 (<i>O. sativa</i>)		
			合计 Total	南亚 South A sia	东南亚 Southeast A sia	中国 China	合计 Total	籼稻 <i>Indica</i>	粳稻 <i>Japonica</i>
Np	SPC	42	42	34	35	38	33	26	28
	SR	41	40	23	33	33	29	15	20
	SR/SPC	98%	95%	68%	94%	87%	88%	58%	71%
Na	SPC	147	141	96	94	111	97	79	82
	SR	137	132	74	91	107	90	65	67
	SR/SPC	93%	94%	77%	97%	96%	93%	82%	82%
Ng	SPC	185	175	101	98	133	107	82	83
	SR	161	156	76	94	125	90	65	67
	SR/SPC	87%	89%	75%	96%	94%	84%	79%	81%
Hs	SPC	0.3738	0.3856	0.3346	0.3212	0.3385	0.3129	0.1911	0.2301
	SR	0.3265	0.3194	0.2263	0.2837	0.2588	0.2979	0.1312	0.1691
	SR/SPC	87%	83%	68%	88%	76%	95%	69%	73%

SPC: 逐步聚类取样, Sampling progressively by clustering; SR: 随机取样, Sampling randomly; Np: 多态位点数, Na polymorphic loci; Na: 总等位基因数, Total no of alleles; Ng: 基因型数, No genotype; Hs: 平均基因多样性, A verage gene diversity.

2.2.2 10% 取样比例构建的核心样本 将两次随机取样获得的遗传多样性参数与逐步聚类法的比较结果列入表3。栽培稻中, 分组随机法构建的核心样本的遗传参数值均比逐步聚类法小。野生稻及整个样本的趋势与栽培稻相同, 但下降幅度变小。

3 讨论

核心种质的研究已日益受到国内外的重视, 我国已将水稻、小麦、大豆核心种质的研究列入国家重点基础规划项目。就水稻而言, 要将7万份材料压缩成7000份或者5000份的核心种质, 是一项难度很大的工作。构建核心种质的最终目的, 是以最少的样本数量最大程度地代表整个资源的多样性, 由此决定了构建核心种质的三项主要研究内容, 即分组原则、取样方法及遗传多样性的检测。

表3 10% 取样比例逐步聚类取样与随机取样次级核心样品遗传多样性的比较
Table 3 Comparison of genetic diversity of secondary core samples selected as 10% of original samples by SPC and SR

	所有样品 All of samples	野生稻 (<i>O. rufipogon</i>)				栽培稻 (<i>O. sativa</i>)			
		合计 Total	南亚 South Asia	东南亚 Southeast Asia	中国 China	合计 Total	籼稻 <i>Indica</i>	粳稻 <i>Japonica</i>	
N _p	SPC	41	40	27	27	32	32	23	26
	SP1	37	36	23	20	33	27	17	9
	SP1/SPC	90%	90%	85%	74%	103%	84%	74%	35%
	SR2	37	37	21	30	31	24	20	11
	SR2/SPC	90%	93%	78%	111%	97%	75%	87%	42%
N _a	SPC	120	113	77	79	91	87	69	76
	SR1	120	115	73	67	91	79	63	54
	SR1/SPC	100%	102%	95%	85%	100%	91%	91%	71%
	SR2	119	115	68	78	88	79	67	55
	SR2/SPC	99%	102%	88%	99%	97%	91%	97%	72%
N _g	SPC	137	128	78	80	96	91	69	76
	SR1	131	119	73	67	97	79	62	53
	SR1/SPC	96%	93%	94%	84%	101%	87%	90%	70%
	SR2	139	120	68	80	94	80	68	55
	SR2/SPC	101%	94%	87%	100%	98%	88%	99%	72%
H _s	SPC	0.3832	0.396	0.2854	0.294	0.3459	0.3134	0.1747	0.2521
	SR1	0.3366	0.355	0.2462	0.1939	0.3296	0.2598	0.1349	0.0845
	SR1/SPC	88%	90%	86%	66%	95%	83%	77%	34%
	SR2	0.3483	0.37	0.2235	0.2947	0.3101	0.2654	0.1854	0.1094
	SR2/SPC	91%	93%	78%	100%	90%	85%	106%	43%

SPC: 逐步聚类取样, Sampling progressively by clustering; SR1: 随机取样第一次重复, Sampling randomly 1st replication; SR2: 随机取样第二次重复, Sampling randomly 2nd replication; N_p: 多态位点数, No. polymorphic loci; N_a: 总等位基因数, Total no. of alleles; N_g: 基因型数, No. of genotype; H_s: 平均基因多样性, Average gene diversity.

Brown^[1]指出10%的资源可以代表70%的多样性。本研究采用逐步压缩构建的核心样本(10%的取样比例)的多态位点数可达到原始样本的90%以上,而等位基因数和基因型数分别为原始样本的65%和55%,没有达到70%的水平,这可能与本研究的取材有关。本研究的取材是经过筛选过的,已经具有较好的代表性,如122份野生稻来自亚洲10个国家,75份栽培稻则来自11个国家。另外本研究的197份材料相对7万份稻种资源只能算是一个小样本,如果初级与次级核心样本为2000份与5000份的大样本上再压缩成10%的核心样本结果可能不同。

核心样本的等位基因数和基因型数分别相当于原始样本的65%和55%,而次级核心样本的等位基因数和基因型数可达原始样本的80%和75%。说明在本研究中,由样本数20%压缩到10%,遗传多样性参数下降幅度明显变大(图1和图2)。对于7万份稻种资源的大样本是否存在同样的趋势,值得研究。另外,栽培稻10%的样本可包涵82%的等位基因数和73%的基因型数,而野生稻10%的样本只包涵64%的等位基因数和55%的基因型数,这可能与野生稻遗传多样性大于栽培稻有关,遗传多样性越大,压缩成核心样本时,等位基因和基因型丢失的可能性就越大。因此,对于不同的野生稻种,以及不同的栽培稻类型,如地方稻种、育成品种和杂交稻亲本,核心样本的取样比例应根据其遗传多样性的大小而有所不同,对于遗传

多样性大的样本, 取样比例可以适当提高, 反之, 可以适当降低其取样比例。

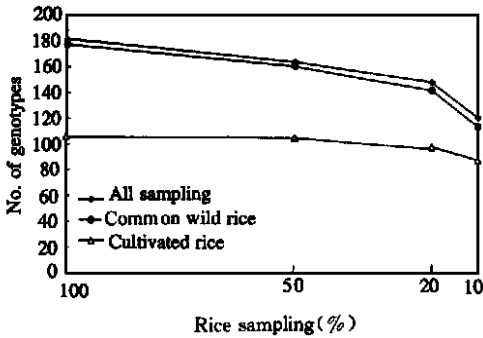


图1 取样比例与等位基因数

Fig 1 Size sampling and No. of alleles

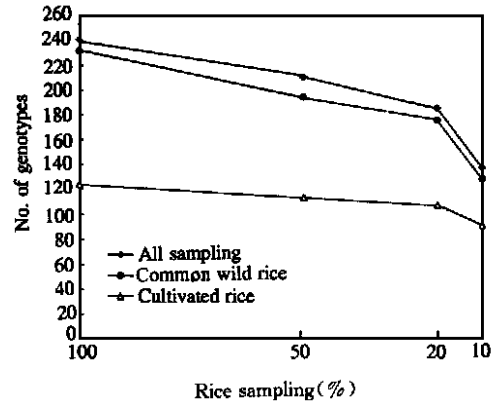


图2 取样比例与基因型数

Fig 2 Size sampling and No. of genotypes

取样方法, 将直接影响所构建的核心种质的遗传多样性。Divan^[6]等构建一年生苜蓿核心种质时, 通过11种方法的比较认为, 聚类分组方法比大随机方法构建的核心种质更能代表整个种质资源的多样性。Basigalup^[7]等也认为, 利用材料间的相似性, 采用主成分和聚类分析是很好的方法。本研究通过逐步聚类和分组随机法的比较, 无论是20%的取样比例还是10%的取样比例, 逐步聚类法构建的核心样本的遗传多样性参数, 均高于分组随机法。这是因为逐步聚类法是逐步将亲缘关系最近的材料去掉, 被保留的材料之间相似性则相对较小。

对核心种质遗传多样性检验是评价核心种质的有效性重要手段。本研究选用多态位点数、等位基因数、基因型数、平均基因多样性等4个参数进行检验。在这4个参数中, 基本不用担心多态位点数的减少, 因为即使到10%的取样比例, 仍能保持原始样本的90%。在多态位点数保持不变的条件下, 平均基因多样性只能反映等位基因间的相对频率, 不同的取样比例和方法对其影响不大。基因型数虽然在不同的取样比例中减少的幅度最大, 但只要能保持较多的等位基因, 就可能组合出不同的基因型。另外, 无论是从理论还是应用的角度考虑, 核心种质应该尽可能使整个资源的基因少丢失, 对已知的重要基因则不能丢失。因此等位基因数应该是检验遗传多样性优先考虑的参数。

参 考 文 献

- 1 Brown A H D. In: Brown AHD et al (ed) *The Use of Plant Genetic Resources*. Cambridge, England: Cambridge Univ Press, 1989a 136~ 156
- 2 孙传清, 王象坤, 吉村淳等. 中国农业科学, 1997, 30(4): 37~ 44
- 3 Rogers O S, A J Bendich. *Plant Molecular Biology Manual*. 1988 A 6: 1~ 10
- 4 Sokal R R, C D Michener. *Sci Bull Univ. Kanas*, 1958, 28: 1409~ 1439
- 5 Oka H I. *Origin of Cultivated Rice*. Tokyo: Japan Scientific Societies Press, 1988 59~ 67
- 6 Divan N, M S M Chitosh, G R Baughan. *Theor Appl Genet*, 1995, 90: 755~ 761
- 7 Basigalup D H, D K Barnes, R E Stucker. *Crop Sci*, 1995, 35: 1163~ 1168