

不结球白菜品种基于 SSR 标记的群体结构分析

秦艳梅 李 向 赵建军 王彦华 轩淑欣

(河北农业大学 园艺学院 河北 保定 071001)

摘要: 为了研究不结球白菜内部的群体结构, 从全国搜集了 78 份广泛种植的蔬菜用不结球白菜品种, 使用 SSR 标记对其进行多态性分析, 并利用 Structure 2.1 软件统计分析标记数据划分群体结构。结果表明, 初筛出的 21 对多态性 SSR 标记覆盖了白菜 10 个连锁群, 在 78 份材料中扩增共产生了 56 个等位基因变异, 平均每个 SSR 标记产生的等位基因数目为 2.7 个。Structure 分析表明, 所选的不结球白菜可分为 4 个亚类群, S1 和 S2 亚类群主要集中了来自于南北方各省份的 49 份普通白菜, 苗用大白菜主要分布在 S3 和 S4 亚类群, 乌塌菜主要分布在 S1 和 S4 亚类群, 菜心的遗传背景比较复杂; 材料的类群划分与地理来源的关系在 S2 亚类群表现的较为明显。结果表明, 不结球白菜内部存在着明显的群体结构, 为合理利用现有的种质资源指导育种和进行新基因的挖掘奠定了基础。

关键词: 不结球白菜; SSR; 群体结构; 遗传多样性

中图分类号: S634.03 文献标识码: A 文章编号: 1000-7091(2013)03-0062-05

Analysis of Population Structure of Non-heading Chinese Cabbage Based on SSR Marker

QIN Yan-mei, LI Xiang, ZHAO Jian-jun, WANG Yan-hua, XUAN Shu-xin
(College of Horticulture, Agricultural University of Hebei, Baoding 071001, China)

Abstract: To study population structure within non-heading Chinese cabbage crops, 78 non-heading Chinese cabbage varieties grown in a large planting area in China were collected. Their DNA polymorphisms were analyzed by a set of SSR markers and their population structure were clustered by Structure 2.1 software according SSR data. Out of SSR primer 120 pairs, 21 pairs originating 10 linkage groups of *Brassica rapa* were screened out with polymorphism and 56 alleles in the 78 varieties were generated. The average number of alleles per SSR locus was 2.7. The 56 markers were imported in the program structure to identify subpopulations, resulting in 4 subgroups in the 78 varieties. Subpopulations S1 and S2 mainly included 49 Pak choi varieties from different provinces in southern or northern China. Seedling Chinese cabbages were mainly distributed in subpopulation S3 and S4. Wutacai varieties were mainly distributed in subpopulation S1 and S4. The genetic backgrounds of Caixin varieties were relatively complicated. The relationship between material grouping and geographical origin was very clear in subpopulation S2. This result indicated an obvious population structure within non-heading Chinese cabbage, which established the basis for utilizing the current germplasm resources to guide breeding and discovering the new genes.

Key words: Non-heading Chinese cabbage; SSR; Population structure; DNA polymorphism

关联作图是利用现有自然群体鉴定目标性状与遗传标记或候选基因关联关系的一种新型基因定位方法^[1-2], 因为检测到的等位基因数目多、研究周期短、不需构建作图群体, 成为现今种质资源分子评价及解析研究植物数量性状基因型的主要方法之

—^[3-6]。群体结构是一个群体内存在亚群的情况。群体结构亚群的混合使整个群体所估计的 LD 强度增强, 可能导致基因多态性位点与性状的相关性并非由功能性等位基因引起, 从而提供假阳性结果^[7-9]。因此, 群体结构分析是利用自然群体进行

收稿日期: 2013-03-20

基金项目: 国家自然科学基金项目(30871713); 河北省青年科学基金项目(2010000738)

作者简介: 秦艳梅(1987-), 女, 安徽抚阳人, 在读硕士, 主要从事生物技术与蔬菜遗传育种研究。

通讯作者: 轩淑欣(1977-), 女, 河北定州人, 副研究员, 博士, 主要从事分子细胞遗传及育种应用研究。

关联作图的基础。

不结球白菜 (*Brassica campestris* ssp. *chinensis* Makino, A 基因组 2n = 20) 主要指十字花科 (Cruciferae) 芸薹属 (*Brassica*) 白菜类作物中不能形成叶球的多个栽培类型, 包括普通白菜 (又称小白菜、油菜、青菜)、乌塌菜、菜心或菜薹等, 具有适应性广、生长周期短、高产、易种、营养丰富、鲜食盐渍皆宜等特点^[10]。我国是不结球白菜的原产地之一, 遗传资源丰富, 南北方和国内外广泛的引种和杂交选育, 又形成了许多新的变异类型, 使其遗传背景非常复杂。

因此, 有必要对国内蔬菜用不结球白菜的种质资源进行分析评价, 以发掘优异基因和培育适合各生态地区的品种类型。但关于国内不结球白菜资源内部的群体结构尚无报道。

本研究选用从全国各地搜集到的不同类型的蔬菜用不结球白菜品种组成自然群体, 利用从白菜 10 条连锁群上筛选的 SSR 标记在全基因组水平上检测该群体的多态性, 对该群体的亚类群结构及与地理来源间的关系进行分析, 为不结球白菜种质资源的合理利用和品种改良提供依据。

表 1 供试不结球白菜材料名称和来源

Tab. 1 The name and origin of non-heading Chinese cabbage materials

编号 No.	品种名称 Name of varieties	类型 Type	来源 Origin	编号 No.	品种名称 Name of varieties	类型 Type	来源 Origin
1	四季快菜一号(苗用)	大白菜	北京	40	玉农鸡毛菜	普通白菜	江西
2	淮南黄心乌	普通白菜	合肥	41	华英	普通白菜	广东
3	华冠	普通白菜	日本	42	金棚夺冠	普通白菜	陕西
4	野崎 3 号	普通白菜	日本	43	矮箕青	普通白菜	河南
5	京油一号(鸿均)	普通白菜	北京	44	绿丽	普通白菜	北京
6	京油一号(绿金蓝)	普通白菜	北京	45	平油一号	普通白菜	山西
7	夏绿 2 号	普通白菜	北京	46	绿威	普通白菜	北京
8	京绿 7 号	普通白菜	北京	47	白叶四月慢	普通白菜	上海
9	京冠一号	普通白菜	北京	48	华荣青梗菜	普通白菜	广东
10	北京小白菜	普通白菜	北京	49	苏州青	普通白菜	江苏
11	五月蔓	普通白菜	北京	50	新五月慢	普通白菜	河北
12	利丰青梗白菜	普通白菜	杭州	51	上海青	普通白菜	宁夏
13	晶华	普通白菜	北京	52	四季黄秧小白菜	大白菜	河北
14	旗舰	普通白菜	北京	53	四季甜脆	大白菜	山东
15	香港青梗白菜(128)	普通白菜	杭州	54	四季小白菜	大白菜	北京
16	抗热 605 青菜(芳丰)	普通白菜	江苏	55	绿星青菜	普通白菜	南京
17	倍好	普通白菜	北京	56	四倍体矮脚黄	普通白菜	南京
18	金城	普通白菜	北京	57	绿领兔子腿矮脚黄	普通白菜	南京
19	京油 605	普通白菜	北京	58	彩芯	普通白菜	南京
20	寒绿油菜	普通白菜	天津	59	绿领金黄心	普通白菜	南京
21	碧玉	普通白菜	南京	60	绿领热优二号	普通白菜	南京
22	矮脚奶白菜	普通白菜	北京	61	绿领特矮青	普通白菜	南京
23	冬青油菜	普通白菜	北京	62	七宝青菜	普通白菜	上海
24	奶白菜	普通白菜	北京	63	特矮青菜	普通白菜	上海
25	甜脆小白菜	大白菜	北京	64	上海抗热 605	普通白菜	上海
26	百惠	普通白菜	北京	65	秋田黄心乌	普通白菜	南京
27	菊花小八叶	乌塌菜	北京	66	秋田鸡毛菜	普通白菜	南京
28	四九菜心	菜心	北京	67	秋田绿星	普通白菜	南京
29	广东超级 50 天菜心	菜心	广州	68	秋田中箕白	普通白菜	南京
30	50 天油绿菜心	菜心	广州	69	秋田四月蔓	普通白菜	南京
31	乌塌菜	乌塌菜	河北	70	金谷包心青菜	普通白菜	四川
32	台湾乌塌菜	乌塌菜	辽宁	71	改良九月鲜	菜心	湖北
33	四季全能(油菜)	普通白菜	山东	72	极早熟苔仙一号	菜心	湖北
34	夏帝	普通白菜	日本	73	31 号油青甜菜心	菜心	海南
35	臺冠	普通白菜	武汉	74	百叶塌菜	乌塌菜	湖北
36	特选京油一号	普通白菜	北京	75	绿领黑心乌	普通白菜	南京
37	清铧	普通白菜	福建	76	cc0807	大白菜	-
38	青帮油菜	普通白菜	河北	77	cc0808	大白菜	-
39	四季(杂交油菜)	普通白菜	河北	78	夏小宝(娃娃菜)	大白菜	北京

1 材料和方法

1.1 供试材料

试验材料为从国内搜集的不同地理来源的不结球白菜商业栽培品种共计 78 份,包括普通白菜 60 份、苗用大白菜 8 份、乌塌菜 4 份、菜心 6 份(表 1)。

1.2 SSR 引物

参照相关文献[11-15],共选取了分布于芸薹属 A 基因组 10 个连锁群上的 120 对 SSR 引物,引物序列来源于 <http://www.brassica.info> 和参考文献,由上海生物工程有限公司合成。

1.3 基因组 DNA 的提取与检测

基因组 DNA 提取采用 CTAB 法^[16]。

1.4 SSR-PCR 扩增及检测

PCR 反应体系总体积为 20 μL ,包括: $10\times\text{Buffer}$ (含 Mg^{2+}) 2.0 μL ,正反引物(50 ng/ μL)各 1.0 μL ,dNTPs(2.5 mmol/ μL) 0.4 μL ,Taqase(5 U/ μL) 0.2 μL ,DNA 模板(30 ng/ μL) 3.0 μL ,无菌水至 13.4 μL 。反应程序为 94 $^{\circ}\text{C}$ 预变性 5 min;94 $^{\circ}\text{C}$ 变性 1 min,60 $^{\circ}\text{C}$ 退火 30 s,此后每个循环的退火温度降低 0.5 $^{\circ}\text{C}$ 直至 55 $^{\circ}\text{C}$ (每个退火温度上进行 1 个循环),72 $^{\circ}\text{C}$ 延伸 45 s,共 10 个循环;94 $^{\circ}\text{C}$ 变性 1 min,

55 $^{\circ}\text{C}$ 退火 30 s,72 $^{\circ}\text{C}$ 延伸 45 s,25 个循环;72 $^{\circ}\text{C}$ 延伸 10 min;待温度降至 10 $^{\circ}\text{C}$ 后 4 $^{\circ}\text{C}$ 保存备用。采用 6% 的聚丙烯酰胺凝胶电泳及银染法检测扩增结果。

1.5 SSR 数据处理与分析

统计数据时,SSR 多态性定义为在扩增产物凝胶电泳带型的某一相同迁移位置上,品种某一条带的有或无,每一个多态性条带记作一个等位基因(假定在相同迁移率的条带均来自同一位点上的同一等位基因)。

参照 Zhao^[17]的方法使用 Structure 2.1 软件^[18]进行群体结构分析:扩增有带记为“1”,无带记作“0”,建立一个“0-1”二进制数据矩阵的文本文件。缺失值与不清晰的条带记作“-9”。用软件读入该文本文件,输入要分析的品种总数目、分子标记数以及有无缺失值等参数设置。在“Ancestry”选项中选择“independent allele frequency”,点击“Run Job”,保存参数设置。设置 K 值为 1~10,运行程序。进行材料的群体结构分析前,对群体的混合模型进行了相应假设即各种基因型为混合起源(Mixed ancestry)而且在群体中亚类群之间有独立的等位基因频率。不同材料根据其 SSR 等位基因频率可以按照预设的分类簇进行相应分组,亚类数目 K 设定为 1~10。

表 2 SSR 在不同栽培类群中检测到的等位基因数目

Tab. 2 Number of alleles detected in the collection and morphotypes by SSR marker

SSR 引物 SSR primers	连锁群 Linkage group	条带数 No. of alleles	苗用大白菜 Seedling Chinese cabbage	普通白菜 Pak choi	菜心 Caixin	乌塌菜 Wutacai
KBrH139G23	A ₁	5	4	5	5	3
Na12C06	A ₁	1	1	1	1	1
H009D02-3	A ₁	3	2	3	2	3
OH10D03	A ₁	4	4	4	3	1
Na14F11	A ₁	1	1	1	0	1
Na14H11	A ₂	3	2	3	3	3
Na14G02	A ₃	2	2	2	2	2
BRMS043	A ₃	1	0	1	1	1
F3HSSR2	A ₃	3	3	3	3	3
B068E07-2	A ₃	2	2	2	2	1
Na10D09	A ₄	2	2	2	1	2
Ra3H10	A ₅	4	2	4	4	3
BRMS-034	A ₅	3	2	3	3	2
Ra12D08	A ₆	2	2	2	2	2
ENA14	A ₇	3	2	3	3	2
Ra2A05b	A ₇	3	1	3	2	3
Ra2E12	A ₈	3	3	3	3	1
KBrH143K20	A ₉	4	3	4	4	2
EJU2	A ₉	3	2	3	3	2
Ra2E07	A ₁₀	3	1	3	3	3
B034N10-3	A ₁₀	1	1	1	1	0
Amount of alleles		56	42	56	51	41
Average of alleles		2.7	2.0	2.7	2.4	2.0

2 结果与分析

2.1 不结球白菜栽培群体 SSR 多态性分析

从含有 4 种不同类型的 78 份不结球白菜群体中各选择 1 份材料, 包括: 1(苗用大白菜) 27(乌塌菜) 28(菜心) 75(普通白菜), 对 120 对 SSR 引物进行初步筛选, 共筛出具有多态性且条带清晰、重复性好的 SSR 引物 21 对, 占全部引物的 17.5%, 覆盖了白菜基因组的 10 个连锁群 ($A_1 \sim A_{10}$), 每个连锁群上的 SSR 标记数目 1~5 对不等(表 2)。

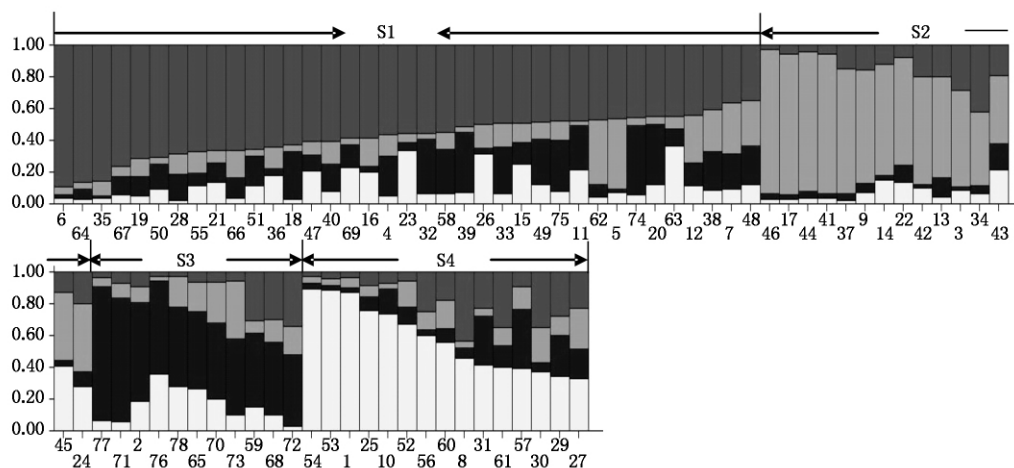
利用筛出的 21 对 SSR 引物对 78 份材料进行 SSR 分析, 共产生 56 个等位基因标记, 其中 4 对 SSR 引物只检测到 1 个等位基因, 检测到等位基因数目最多(5 个)的是位于 A_1 连锁群的 KBrH139G23 标记。等位基因数目的多样性在不同栽培类群间变化不大, 平均每个 SSR 产生的等位基因数目为 2.7 个(表 2)。

2.2 Structure 群体结构分析

使用 Structure 2.1 软件对 21 对 SSR 引物产生的 56 个标记进行群体结构分析。结果表明, 当 $K = 4$ 时, 78 份材料可分为 4 个亚类群, 从左到右分别为 S1、S2、S3、S4(图 1)。S1 亚类群最大, 包括了 37 份材料; S2 包括 15 份, S3 包括 11 份, S4 包括 15 份,

可以明显看到, 各亚类群中排列在前面的材料遗传背景整体趋于一致, 排列在后面的材料在基因型组成上则更多包含了其他亚类群白菜基因型的遗传背景。

由图 1 还可以看出, 收集的普通白菜被分为四类, 分布于 4 个亚类群中, 以 S1 和 S2 亚类群为主。S1 亚类群中的栽培类型包括了 34 份普通白菜, 还包括 1 份菜心(28) 和 2 份乌塌菜(32 和 74)。尽管每份材料的遗传背景均以 S1 亚类群所占比例为主, 但有约 40% 的材料所占遗传背景的比例不到 0.5, 排在该类群后面的 4 份材料(12, 38, 7, 48) 明显包含了 S2 和 S3 亚类群的遗传背景。S2 亚类群均为普通白菜类型, 遗传背景相对单一, 但也可以看到来自日本的 2 份材料(3 和 34) 包含了较多的 S1 亚类群的遗传背景, 基因型 45 和 24 号混合了较多的 S4 亚类群的遗传背景。5 份普通白菜与 3 份菜心和 3 份苗用大白菜分布在一起成为 S3 亚类群。S4 亚类群混合了 4 种栽培类型, 但以大白菜类型占主要遗传成份。6 份菜心分布于除 S2 的 3 个亚类群内, 除来源于湖北的 71 号材料遗传背景相对单一外, 其他材料遗传背景均比较复杂。4 份乌塌菜类型分布在 S1 和 S4 这 2 个亚类群内, 但均混合了较高比例的 S3 的遗传成份。



横坐标表示不同材料的编号, 纵坐标为不同基因型个体所占遗传背景的比例。

X-axis represents different material number, Y-axis represents the proportion of different genotype sharing the genetic background.

图 1 Structure 分析推断出的四大亚类群聚类图

Fig. 1 Analysis of the population structure in non-heading *Brassica campestris*

3 讨论

群体结构分析是目前利用自然群体进行关联作图的重要步骤之一, 其特点将直接影响性状-标记关联分析的结果, 目标群体内存在的亚群可以产生伪关联。白菜类作物变异丰富, 对其群体结构分析已有报道, 如 Zhao 等^[17]研究表明, 在白菜作物内存在 4 个主要的亚类群; 刘何^[19]研究表明, 供试的国内外 83 份白菜作物的最佳分组数是 $K = 5$, 高玉梅

等^[20]对国内 64 份白菜作物进行群体结构分析, 将 $K = 3$ 确定为最佳分组数。但是基于白菜类作物同一栽培类群内全基因组水平上的群体结构分析未见报道。本研究选用国内广泛栽培的不结球白菜栽培种为研究对象, 揭示了供试的品种类型存在明显的群体结构, 分布于 4 个亚类群, 为更好地解决群体分类对后期关联分析的影响, 降低伪关联出现的概率提供了依据。同时研究也发现, 亚类群的划分与地理来源有一定关系, 尤其在 S2 亚类群内表现得较为

明显; 由于试验材料为市场的商业品种, 个别材料的来源可能不清楚或与记载来源不符。按照遗传成份所占的比例, 排在每一个亚类群后面的材料混合程度高, 如试验选用的菜心类型不多, 但仍然反应出了这 2 个栽培类型遗传背景的复杂性。

Structure 软件分析群体结构依据的是等位基因位点的频率, 利用混合模型可以分析不同基因型间遗传信息的交流情况^[1-2, 21]。聚类分析一般依据的是不同基因型间的相似系数或遗传距离^[22-25], 不能反映材料间的遗传背景交流。但结合群体结构与分子聚类分析可以有效地帮助解释不同基因型个体的遗传关系。本研究也对研究群体进行了基于遗传距离的聚类分析(结果未列出), 发现 2 种方法获得的结果具有很高的-一致性。UPGMA 聚类分析的类群 I-1、I-2 中的材料主要分布于 Structure 结构分析中 S1 亚类群, 类群 I-3 与 S2 基本一致, 类群 II 与 S4 一致, 类群 III 与 S3 一致。个别与上述比对有差异的基因型大多为遗传背景混合程度高的材料。

参考文献:

- [1] Morton N E. Linkage disequilibrium maps and association mapping[J]. Journal of Clinical Investigation 2005, 115: 1425 - 1430.
- [2] Zhu C S, Gore M, Buckler E S *et al.* Status and prospects of association mapping in plants[J]. The Plant Genome, 2008, 1: 5 - 20.
- [3] Aranzana M, Kim S, Zhao K *et al.* Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes[J]. Plos Genetics 2005, 1(5): 60.
- [4] Kokina A, Rostoks N. Genome-wide and *m1a* locus-specific characterisation of latvian barley varieties[J]. Proceedings of the Latvian Academy of Sciences Section B 2008, 62: 103 - 109.
- [5] Xie C, Zhang S, Li M *et al.* Inferring genome ancestry and estimating molecular relatedness among 187 Chinese maize inbred lines[J]. Genetics and Genomics 2007, 34: 738 - 748.
- [6] Hasan M, Friedt W, Pons-Kühnemann J *et al.* Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *Napus*) [J]. Theoretical and Applied Genetics 2008, 116: 1035 - 1049.
- [7] Simko I, Hu J G. Population structure in cultivated lettuce and its impact on association mapping[J]. Journal of the American Society for Horticultural Science, 2008, 133(1): 61 - 68.
- [8] Pritchard J K, Stephens M, Rosenberg N A *et al.* Association mapping in structured population[J]. The American Journal of Human Genetics 2000, 67: 170 - 181.
- [9] Simko I. Population structure in cultivated lettuce and its impact on association mapping[J]. Journal of the American Society for Horticultural Science 2008, 133(1): 61 - 68.
- [10] 侯喜林. 不结球白菜育种研究新进展[J]. 南京农业大学学报 2003, 26(4): 111 - 115.
- [11] Lowe A J, Moule C, Trick M *et al.* Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species[J]. Theoretical and Applied Genetics 2004, 108: 1103 - 1112.
- [12] Piquemal J, Cinquin E, Couton F *et al.* Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers[J]. Theoretical and Applied Genetics, 2005, 111: 1514 - 1523.
- [13] Kim J S, Chung T Y, King G J *et al.* A sequence-tagged linkage map of *Brassica rapa* reveals the pattern chromosomal segmental duplications[J]. Genetics, 2006, 174: 29 - 39.
- [14] Choi S R, Teakle G R, Plaha P *et al.* The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project[J]. Theoretical and Applied Genetics 2007, 115: 777 - 792.
- [15] Soengas P, Hand P, Vicente J G *et al.* Identification of quantitative trait loci for resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica rapa* [J]. Theoretical and Applied Genetics 2007, 114: 637 - 645.
- [16] 王丽, 乔爱民, 孙一铭, 等. 菜心基因组 DNA 提取及 RAPD 反应体系的优化[J]. 西南师范大学学报: 自然科学版 2006, 31(2): 124 - 128.
- [17] Zhao J J, Maria-Joao Paulo, Diaan Jamar *et al.* Association mapping of leaf traits, flowering time and phytate content in *Brassica rapa* [J]. Genome 2007(50): 963 - 973.
- [18] Pritchard J K, Stephens M, Donnelly P. Inference of population structure from multilocus genotype data[J]. Genetics 2000, 155: 945 - 959.
- [19] 刘何. 白菜类作物种植材料的遗传多样性分析[D]. 保定: 河北农业大学 2007.
- [20] 高玉梅, 张淑江, 章时蕃, 等. 白菜类作物资源群体结构分析[J]. 中国农学通报 2009, 25(21): 332 - 334.
- [21] Falush D, Stephens M, Pritchard J K. Inference of population structure using multilocus genotype data: dominant markers and null alleles[J]. Molecular Ecology Notes, 2007, 7: 574 - 578.
- [22] 张冬玲, 张洪亮, 魏兴华, 等. 贵州栽培稻的遗传结构及其遗传多样性[J]. 科学通报 2006, 51(23): 2747 - 2754.
- [23] 刘永忠, 赵晋峰, 王高鸿, 等. 不同丝黑穗病抗性玉米自交系 ISSR 多态性分析[J]. 山西农业科学 2010, 38(7): 11 - 15.
- [24] 杨兆顺, 董海合, 吴俊强, 等. 利用 RAPD 分子标记法聚类分析糯米种质资源[J]. 天津农业科学 2005, 11(4): 48 - 49.
- [25] 刘建霞, 雷海英, 温日宇, 等. 山西省马铃薯主栽品种遗传多样性的 SSR 分析[J]. 华北农学报 2012, 27(6): 72 - 77.