

结合农艺性状和 SSR 标记构建吉林省花生初级核心种质*

牟书靓¹, 牛海龙¹, 张春宵², 刘红欣¹, 何中国^{1**}, 李玉发^{1**}

1. 吉林省农业科学院花生研究所, 公主岭 136100; 2. 吉林省农业科学院作物资源研究所, 公主岭 136100

摘要:以 260 份花生材料为试材,采用不同的遗传距离构建吉林省花生初级核心种质。在 50%的取样比例下,采用数量性状参数均值差异百分率、方差差异百分率、极差符合率、变异系数变化率 4 个指标评价不同方法构建的初级核心种质的可行性和有效性。最终选出一种合适的方法构建花生初级核心种质,并利用等位变异数、有效等位变异数、Shannon-Weaver 指数、Nei 期望杂合度等 SSR 标记相关参数进行验证。研究结果表明:基于欧式距离,采用优先取样法结合最短距离法进行聚类分析,构建包含 128 个样品花生核心种质。采用 SSR 标记参数进行验证,表明 128 份初级核心种质可以代表原种质 80% 的遗传信息,较好地代表了原种质的遗传多样性。

关键词:花生;核心种质;农艺性状;SSR 标记;吉林省

中图分类号: S565.2 **文献标识码:** A **文章编号:** 1000-5684(2017)04-0392-11

DOI: 10.13327/j.jjlau.2017.3325

引文格式: 牟书靓,牛海龙,张春宵,等.结合农艺性状和 SSR 标记构建吉林省花生初级核心种质[J].吉林农业大学学报,2017,39(4):392-402.

Construction of Peanut Primary Core Germplasm in Jilin Province by Agronomic Traits and SSR Markers

MU Shujing¹, NIU Hailong¹, ZHANG Chunxiao², LIU Hongxin¹, HE Zhongguo^{1**}, LI Yufa^{1**}

1. Peanut Research Institute, Jilin Academy of Agricultural Sciences, Gongzhuling 136100, China;
2. Crop Resources Institute, Jilin Academy of Agricultural Sciences, Gongzhuling 136100, China

Abstract: Using 260 peanut test materials, the primary peanut core germplasm in Jilin province was constructed by different genetic distances. In 50% sampling ratio, feasibility and validity of constructing primary core germplasm with different methods were evaluated by using quantitative trait parameters including mean difference percentage, variance difference percentage, range coincidence rate and variation ratio of variation coefficient. Finally, the appropriate method was used to construct the primary peanut core germplasm. The related parameters of SSR markers were used to verify the above results, which included allelic variance, effective number of allelic variation, Shannon-Weaver index, Nei expected heterozygosity, etc. The results showed that Euclidean distance, the preferred sampling method and the shortest distance method were applied in cluster analysis, and the peanut core germplasm of 128 accessions was constructed. Via the validation of the related parameters of SSR markers, 128 accessions represented 80% genetics information of the primary germplasm and could very well represent the genetic diversity of the original germplasm.

* 基金项目: 吉林省农业科技创新工程重大项目(6215350101), 吉林省科技发展计划项目(20150307028NY)

作者简介: 牟书靓,女,硕士,从事花生资源抗旱鉴定研究。

收稿日期: 2016-08-04

** 通信作者

Key words: peanut; core germplasm; agronomic trait; SSR marker; Jilin province

花生起源于南美洲的热带、亚热带地区^[1]。目前,世界上生产花生的国家 100 多个。其中,印度的花生种植面积最大,但其总产量却居于第 2 位;中国花生种植面积仅次于印度,但年均总产量却居于首位^[2]。因此,中国的花生产业在世界花生生产中占有重要的地位^[3]。我国东北三省的大部分地区以及河北燕山东段以北的地区构成东北花生产区,产区土质多为丘陵沙地和风沙地,适于花生生长。吉林省具有得天独厚的条件,光照充足,病虫害较少,特别是严重影响花生品质的黄曲霉显著低于其他地区,是我国出口优质花生的主要产地^[4-5]。自“九五”以来,吉林省的花生生产一直处于上升趋势,其品质和产量也逐步提高;“十五”期间,全省花生种植面积已经达到 8.4 万 hm^2 ,到“十一五”期间,其种植面积已经扩大到 11.9 万 hm^2 。随着种植面积的增加,花生的产量也备受关注,挖掘更好的种质资源是当下解决花生高产、稳产最有效的途径^[6]。

种质资源的收集、整理和保存是选育新品种的物质基础。核心种质的构建,有助于简化种质资源的评价,发掘并保存重要性状。核心种质就是以最少数量的遗传资源最大限度地保存整个资源群体的遗传多样性。在小麦^[7-8]、水稻^[9]、大麦^[10]、木豆^[11]、红花^[12]、玉米^[13-17]、豌豆^[18]、大豆^[19-20]等作物上建立核心种质,通过植物学性状鉴定,表明核心种质都能代表各物种种质资源 80% 以上的遗传多样性。目前,对核心种质的研究已经从植物学性状研究转向植物学性状和分子标记相结合的研究,尤其是基因多样性分析,为新基因发掘创造了条件。核心种质的遗传多样性及品种间亲缘关系的研究可为核心种质的有效利用、作物遗传基础的拓宽、杂交亲本的选配提供分子依据。

1 材料与方法

1.1 材料

本试验所使用的 260 份花生材料均由吉林省农业科学院花生研究所提供。

1.2 农艺性状测量方法

2014 年 5 月 18 日,在吉林省农业科学院公主岭花生研究所试验地种植 260 份花生材料,每

穴 2 粒,行长 4 m,行距 0.65 m,株距 0.14 m。每个材料 2 行,成熟期时每个材料随机选择 5 株,测量主茎倒三叶叶长、叶宽,每株测量 1 叶;在收获前 1 周同样每个材料随机取 5 株,测量主茎高、第一侧枝长、总分枝、单株结果数;收获后,在实验室内进行考种,测量单株生产力、百果重、百仁重、荚果长、荚果宽、粒长、粒宽,5 次重复,最后取其各性状平均值作为原始数据。

2014 年 6 月,采集嫩叶放入自封袋中,于 $-80\text{ }^\circ\text{C}$ 冻存,用于 SSR 数据分析。

1.3 DNA 的提取与浓度检测

本试验采用改进的 CTAB 法提取花生基因组 DNA,利用 NanoDrop 2000C 超微量分光光度计进行检测,其 $\text{OD}_{260}/\text{OD}_{280}$ 值为 1.8~2.0,适合于 SSR 体系扩增,将其放入 $-20\text{ }^\circ\text{C}$ 冰箱保存备用。

1.4 SSR 引物的筛选

本试验主要从国际热带半干旱地区作物研究所(ICRISAT)、美国塔斯基吉大学(Tuskegee University)、中国农业科学院搜集并选取 476 对 SSR 引物,由北京鼎国昌盛生物技术有限公司、长春库美生物科技有限公司进行合成。

1.5 变性聚丙烯酰胺凝胶电泳和银染

SSR 扩增产物要在 6% 的变性聚丙烯酰胺凝胶电泳上进行分离,按照 70 W 恒功率下电泳约 45 min,然后按染色、漂洗、显影、定影等步骤进行显色。

1.6 数据的统计分析方法

采用经典数量遗传学分析软件 QGA Station1.0,结合 DPS 7.05 对花生数量性状进行分析。选择 2 种遗传距离(D1.欧氏距离;D2.马氏距离)、8 种系统聚类方法(C1.最短距离法;C2.最长距离法;C3.中间距离法;C4.重心法;C5.类平均法;C6.可变类平均法;C7.可变法;C8.离差平方和法)和 3 种取样方法(S1.随机取样法:从聚类图中每组随机选择 1 个株系进入下一轮聚类,如果组内只有 1 个株系,则该株系直接进入下一轮聚类分析。S2.优先取样法:在聚类图最低分类水平的 2 个株系中,优先选择具有最大或最小性状表型值的株系进入下一轮聚类,如果 2 个株系都有最大或最小性状表型值,那么 2 个株系都被选择进入下一轮聚类,如果组内只有 1 个遗传材料,则

该材料直接进入下一轮聚类分析^[21]。S3. 偏离度取样法:在聚类图最低分类水平的2个或2个以上株系中,选择具有较大或最大偏离度的株系进入下一轮聚类,如果组内只有1个遗传材料,则该材料直接进入下一轮聚类分析)进行遗传分析。

SSR 标记数据采用 NTSYS2.0 和 MEGA5.04 对其进行处理。(1)根据 SSR 标记扩增结果,在同样的迁移位置,1 表示有带,0 表示无带,用 9 代表数据缺失,构建数据库。(2)利用简单相配系数(Simple matching coefficient, SM)计算自交系遗传相似系数(Genetic similarity, GS)。公式为 $GS = m / (m + n)$ 。其中 m 代表基因型间共有带的数目; n 代表基因型间差异带的数目。(3)根据遗传相似系数矩阵和中心化(Double-center)处理,处理后的自交系 SSR 相似系数矩阵进行聚类分析,采用 UPGMA(Unweighted Pair Group Method Arithmetic Average)方法,建立树状图。将矩阵 X 的每一元素减掉相应的变量(即列),计算平均值,中心化数据为数值所得的矩阵。本研究采用的是 Dcenter 模块(NTSYpce-2.11 软件中)将 GS 矩阵中心化处理。(4)根据遗传相似系数做主坐标分析。(5)采用 NTSYS 软件,以简单相配系数(SM)计算遗传距离(GD)矩阵,按 MEGA5.04 软件中的 Phylogeny 模块,构建系统发育树。(6)某个引物所扩增获得的全部等位基因数,即为总等位变异数。有效等位基因数是纯合度的倒数,反映等位基因间的相互影响,是衡量基因纯合度的另一指标。每个 SSR 位点的多态性信息量(Polymorphism Information Content, 简称 PIC)按公式 $PIC = 1 - \sum f_i^2$, 其中 f_i 为 i 位点的等位基因频率。Shannon-weaver 多样性指数是按照公式 $H = -\sum p_i \times \ln p_i$ 计算,其中 p_i 为某个 SSR 位点的第 i 个等位变异出现的次数占该位点全部等位变异出现次数的百分比。Nei 期望杂合度是按照公式 $nHe = 1 - \sum p_i^2$ 计算,其中 p_i 是第 i 个等位基因的频率, n 是总等位变异数。

本试验对核心种质的评价参数有 4 种:均值差异百分率(Mean difference percentage, MD)、方差差异百分率(Variance difference percentage, VD)、极差符合率(Coincidence rate of range, CR)、变异系数变化率(Changeable rate of coefficient of variation, VR)。利用未标准化的群体计

算以上参数(即用标准化的群体构建核心种质,根据构建结果从未标准化的群体中找到对应的核心材料,计算各个评价参数的值,计算公式参照文献^[21]),然后根据计算结果对其进行评价。只有当群体同时满足均值差异百分率 < 20%、极差符合率 > 80%时,才能表示这个核心种质能够代表原种质资源的遗传多样性。而且,如果均值差异百分率越小,方差差异百分率、极差符合率和变异系数变化率越大,则这个核心种质就越能代表原群体的遗传多样性^[21]。

2 结果与分析

2.1 花生种质资源遗传多样性分析

基于已建立的花生 SSR 标记鉴定技术体系和筛选确立的 40 对核心 SSR 引物,构建 260 份花生试材的 DNA 指纹图谱库。结果表明,40 对 SSR 核心引物在 260 份花生材料间共检测出 139 个等位基因变异,每对 SSR 引物扩增出的等位基因数 2~7 个,平均 3.56 个;多态性信息量(PIC)变化为 0.122 5~0.750 0,平均 0.512 3。利用 40 对 SSR 标记的 139 个等位基因变异,对 260 份花生材料进行 UPGMA 聚类分析(表 1),将遗传相似系数矩阵进行中心化处理,按 UPGMA 法进行聚类分析,花生供试材料划分为 A、B、C、D 4 个类群。遗传相似系数变化范围 0.357 1~0.986 1,利用筛选出的 40 对引物能够区分全部材料。

使用 MEGA4 软件,基于 Nei 遗传距离(GD)对 260 份花生材料进行系统发育树的重建。NJ、ME 和 UPGMA 法进化树构建的结果表明,260 份花生材料在遗传进化关系上可以划分为 2 个种质类群,划分结果与 UPGMA 聚类结果一致。

2.2 种质资源农艺性状变异分析及核心种质构建方法的选择

试验材料各性状的变异系数和变异幅度,见表 2。由表 2 可知,各性状均有较大差异,不同性状间的变异程度也不同。其中,单株生产力、单株结果数的变异程度比较大,其变异系数 > 20%;而叶长、荚果宽、粒宽、生育期的变异程度比较小,其变异系数 < 10%。从各性状的方差分析来看,各个性状在试验株系间的差异均到达极显著水平。从各性状的极差来看,不同材料性状间的差异均较大。

表 1 260 份花生材料的聚类结果

Table 1. Clustering results of 260 peanut accessions

类群	材料名称	数量
A	HX1、吉花 07-8、花育 26、C-20、科濮、品 16、鲁花 15、华实 9616、I2、华实 9618、鲁花 14、C-22、J-14、C-49、C-9、日本 2 号、R99-5、C-45、京白 2、环 10-5、C-40、C-53、C-46、漂花 4807、豫花 9327、C-50、郑花、开府 23、锦 2010、P06-9、P03-9、冀 9814、远杂 9847、P06-1、濮花 23、锦 9203-608、誉宇 3 号、高密 1 号、403-11、锦 9806-F1、无名 12 号、无名 18 号、花育 33、湘花 2008、吉花 07-9、C-28、冀花 8、环 10-2、I7、无名 13 号、C-25、C-29、C-42、阜花 11 号、SD-5、C-10、黑 1、C-37、无名 16 号、冀花 9、锦花 14、花育 37、2 号、吉花 1069、吉花 09-9、豫花 9719、冀 0211、花育 40、1 号、吉花 4 号	70
B	无名 5 号、开 17-15、V04-9、徐 9505、花育 23、花育 27、C-55、C-56、环 10-3、小日本、无名 15 号、无名 17 号、无名 1 号、开农 49、J-1、J-17、J-19、豫花 9717、吉花 08-36、C-16、J-18、花育 24、J-2、豫 9633、豫紫 1、环 10-8、J-20、C-48、豫花 9503、J-15、C-47、开农 37、C-41、J-8、I5、吉花 07-1、吉花 07-3、科富 2 号、红粒王、金花 7 号、无名 3 号、铁花 2 号、锦花 7 号、锦 9909、J-6、锦 2015、开花 21、付花 11、冀油 4 号、吉扶 4、J-13、豫花 10、J-21、金花 13 号、吉花 07-11、C-18、唐油 4 号、I3、吉花 07-2、铁花 3 号、铁花 1 号、莲花 609、吉花 07-4、远育 16-8、J-7、402-3、402-6、402-7、J-3、J-4、J-11、莲玉 6、C-51、环 10-5、花选 15、吉花 07-5、花育 22、鲁花 11、花育 25、长岭引	80
C	远杂 9805、C-54、远杂 9307、无名 4 号、远杂 9102、豫花 23、C-32、无名 20 号、花小宝、无名 19 号、冀花 7、无名 21 号、青花 6、R10-20、吉花 08-35、濮科花 28、濮科花 5、漂花 5、冀 0607-17、C-59、花育 16、小白沙、铁花 195、扶花 1 号、扶花 3 号、扶花 201、扶花 202、白城引、无名 6 号、花生花生 3、多粒粉、环 10-6、锦 9804、濮花 24、E12	35
D	吉花 07-7、阜花 13、阜花 10 号、锦 9818、锦 9826、锦 9938、驻 20150、冀 9606、中花 4 号、广州珍珠红、锦 2006、冀 002-2-9、李修海引、锦花 10、铁花 226、花生花生 2、皖花 2、粤油 45、无名 2 号、无名 8 号、吉花 09-24、无名 9 号、无名 10 号、白沙 1016、锦 2001、远杂 2-3-3、鲁花 12、阜 9658、吉扶 2、多粒白、湘 B、阜花 16、淮花 9、无名 14 号、吉扶 3、SD-3、C-31、C-58、C-61、阜花 14、J-5、J10、辽宁红粒、8525、J-12、C-62、冀 02-6、双辽红粒王、吉扶 201、2102、无名 24 号、无名 7 号、无名 11 号、莲花 209、大粒后代、誉宇 1 号、皖花 4、唐花 11、3023、双花 1 号、吉花 08-41、兰娜、莱农 06、花育 28、403-2、403-10、丰花 1 号、花选 39、C-43、锦 K-9704、濮花 22、誉宇 2 号、徐 9318、科富 3 号、R03-3	75

表 2 供试材料各性状的参数变异

Table 2. Parameter variation of different traits for the tested materials

性状	变异幅度	均值	极差	方差	标准差	变异系数/%
主茎高	18.80~55.40	34.97	36.60	32.24	5.68	16.24
侧枝长	23.40~58.20	38.02	34.80	28.03	5.29	13.92
总分枝数	3.80~12.80	7.18	9.00	1.60	1.26	17.59
单株生产力	8.40~54.45	27.73	46.05	42.06	6.49	23.39
单株结果数	10.60~44.00	25.39	33.40	33.97	5.83	22.95
百果重	55.47~188.58	123.41	133.11	400.11	20.00	16.21
百仁重	23.95~77.79	55.39	53.84	81.55	9.03	16.30
叶长	4.41~7.08	5.49	2.67	0.26	0.51	9.29
叶宽	2.16~3.60	2.80	1.44	0.09	0.30	10.66
荚果长	22.76~45.42	32.71	22.66	18.28	4.28	13.07
荚果宽	10.52~17.18	13.97	6.66	1.26	1.12	8.03
粒长	10.46~19.36	14.61	8.90	4.10	2.02	13.85
粒宽	5.44~10.24	7.64	4.80	0.49	0.70	9.18
生育期	124.00~132.00	129.33	8.00	2.32	1.52	1.18

鉴于吉林省农业科学院所收集的花生材料范围较窄和份数较少,而且仅开展一年一点的表型调查,为慎重起见本研究采用 50% 的取样比例,构建初级核心种质,并在 50% 的取样比例下采用不同的组合抽取 128 个初级核心种质(表 3)。由表 3 可知,在以 50% 的取样比例,采用随机取样法

结合最短距离法构建的 2 个核心种质中,采用欧氏距离和马氏距离的均值差异百分率均<8%、极差符合率均>84%,这表明采用 2 种距离都能很好地保存原种质的遗传信息。但是,所构建核心种质的方差差异百分率和变异系数变化率欧氏距离均高于马氏距离。因此,在花生核心种质构建中,

采用欧氏距离比采用马氏距离更好。

表 3 核心种质与原种质性状的差异百分率

Table 3. Percentage of trait differences between the core collection and the initial collection

%

遗传距离	取样方法	聚类方法	核心种质	MD	VD	CR	VR
	随机取样法(S1)	最短距离法(C1)	D1S1C1	0	64.29	93.78	122.66
		最长距离法(C2)	D1S1C2	0	7.14	87.15	107.03
		中间距离法(C3)	D1S1C3	0	0.00	85.99	107.03
		重心法(C4)	D1S1C4	0	14.29	89.57	107.98
		类平均法(C5)	D1S1C5	0	7.14	93.23	110.55
		可变类平均法(C6)	D1S1C6	0	7.14	92.11	112.01
		可变法(C7)	D1S1C7	0	7.14	89.97	110.28
		离差平方和法(C8)	D1S1C8	0	21.43	91.63	109.57
欧氏距离(D1)	优先取样法(S2)	最短距离法(C1)	D1S2C1	0	85.71	100.00	130.98
		最长距离法(C2)	D1S2C2	0	42.86	100.00	122.18
		中间距离法(C3)	D1S2C3	0	50.00	100.00	119.01
		重心法(C4)	D1S2C4	7.14	64.29	100.00	124.08
		类平均法(C5)	D1S2C5	0	42.86	100.00	121.07
		可变类平均法(C6)	D1S2C6	0	42.86	100.00	120.40
		可变法(C7)	D1S2C7	0	35.71	100.00	120.99
		离差平方和法(C8)	D1S2C8	7.14	35.71	100.00	120.59
偏离度取样法(S3)		最短距离法(C1)	D1S3C1	0	92.86	98.84	128.09
		最长距离法(C2)	D1S3C2	0	35.71	96.73	121.12
		中间距离法(C3)	D1S3C3	0	14.29	95.66	113.22
		重心法(C4)	D1S3C4	0	35.71	96.61	118.05
		类平均法(C5)	D1S3C5	0	35.71	98.94	119.85
		可变类平均法(C6)	D1S3C6	0	35.71	98.39	120.18
		可变法(C7)	D1S3C7	0	50.00	98.86	121.17
		离差平方和法(C8)	D1S3C8	0	64.29	98.88	122.65
马氏距离(D2)	随机取样法(S1)	最短距离法(C1)	D2S1C1	0	28.57	90.12	118.01
		最长距离法(C2)	D2S1C2	0	0.00	89.62	106.43
		中间距离法(C3)	D2S1C3	0	7.14	87.15	108.11
		重心法(C4)	D2S1C4	0	7.14	88.22	111.60
		类平均法(C5)	D2S1C5	0	0.00	84.66	103.79
		可变类平均法(C6)	D2S1C6	0	0.00	85.90	105.77
		可变法(C7)	D2S1C7	0	0.00	87.66	105.41
		离差平方和法(C8)	D2S1C8	0	7.14	89.66	107.95
马氏距离(D2)	优先取样法(S2)	最短距离法(C1)	D2S2C1	0	85.71	100.00	128.99
		最长距离法(C2)	D2S2C2	0	50.00	100.00	123.80
		中间距离法(C3)	D2S2C3	7.14	42.86	100.00	119.84
		重心法(C4)	D2S2C4	0	64.29	100.00	126.74
		类平均法(C5)	D2S2C5	0	57.14	100.00	123.85
		可变类平均法(C6)	D2S2C6	0	35.71	100.00	121.16
		可变法(C7)	D2S2C7	0	35.71	100.00	119.43
		离差平方和法(C8)	D2S2C8	0	64.29	100.00	122.70
偏离度取样法(S3)		最短距离法(C1)	D2S3C1	0	78.57	98.69	128.97
		最长距离法(C2)	D2S3C2	0	50.00	94.38	121.79
		中间距离法(C3)	D2S3C3	0	21.43	94.89	119.71
		重心法(C4)	D2S3C4	0	64.29	93.82	122.36
		类平均法(C5)	D2S3C5	0	50.00	93.96	121.75
		可变类平均法(C6)	D2S3C6	0	28.57	92.76	121.11
		可变法(C7)	D2S3C7	0	42.86	93.01	120.70
		离差平方和法(C8)	D2S3C8	0	42.86	93.87	122.14

由表 3 可知,在 50%的取样比例,采用欧氏距离结合最短距离法构建的 3 个核心种质中,采用随机取样法、优先取样法和偏离度取样法的均值差异百分率均<8%、极差符合率均>84%,表明采用这 3 种取样方法都能很好地保存原种质的遗传信息。而且,由于欧氏距离结合优先取样法构建的核心种质的极差符合率达到了 100%,不仅保存了原种质的所有变异幅度,同时又具有相对较高的方差差异百分率和变异系数变化率,因此本试验选用优先取样法。

对于欧氏距离优先取样法,在 50%的取样比例下,结合 8 种聚类方法所构建的 8 个核心种质中的均值差异百分率均<8%,各个核心种质间的方差差异百分率有较大的差异,极差符合率均为 100%,变异系数变化率>119%,说明每个核心种质均较好地保存了原种质的遗传变异,在 8 种聚类方法下比较不同核心种质间 4 个代表性的评价参数发现,利用最短距离法所构建的核心种质具有最高的方差差异百分率(85.71%)和最高的变异系数变化率(130.98%),因此采用欧氏距离优先取样法逐步聚类构建核心种质时,采用 8 种聚类方法都可以得到较好的聚类效果,其中最短距离法略优于其他 7 种聚类方法。在马氏距离的比较中也发现,最短距离法的聚类效果好于其他 7 种聚类方法。

表 4 原种质和核心种质主成分分析的特征值及累计贡献率

Table 4. Eigen value and cumulative contribution for the initial collection and the core collection

主成分	原种质			核心种质		
	特征值	贡献率/%	累积贡献率/%	特征值	贡献率/%	累积贡献率/%
1	3.455	24.677	24.677	4.104	29.312	29.312
2	2.182	15.583	40.26	2.805	20.034	49.345
3	2.028	14.483	54.743	1.862	13.299	62.644
4	1.572	11.231	65.974	1.573	11.233	73.877
5	1.191	8.510	74.484	1.225	8.751	82.628

2.3.2 基于 SSR 变异构建的核心种质 利用 40 对 SSR 标记的 139 个等位基因变异,对 128 份花生材料进行 UPGMA 聚类分析(表 7),将遗传相似系数矩阵进行中心化处理,按 UPGMA 法进行聚类分析,将供试花生材料同样划分为 A、B、C、D 4 个类群,并且这 4 个类群的材料(表 6)与原种质的 4 个类群所划分的材料基本相符。

综上所述,在 50%的取样比例下,构建的 48 个花生核心种质中,利用欧氏距离,采用优先取样法进行逐步聚类,结合最短距离法构建的核心种质具有最大的遗传变异量,该结合方法是最适合花生核心种质构建的方法。以 D1S2C1-50 方法筛选出的核心种质 128 份,较好地保存了原种质的遗传信息。

2.3 核心种质的确定

2.3.1 基于形态特性变异构建的核心种质 由表 4 可知,利用主成分分析对所构建的初级核心种质进行确认,比较可以看出,原种质和核心种质的主成分分析结果的特征值、贡献率和累计贡献率非常相近。选入特征值>1 的 5 个主成分,其中,原种质的第 5 个主成分的特征值为 1.191,累计贡献率为 74.484%;而所构建的核心种质的第 5 个主成分的特征值为 1.225,累计贡献率已达到 82.628%;一次所构建的核心种质能较好地保存原种质的遗传信息。

比较原种质和核心种质的相关系数(表 5)发现,核心种质与随机抽取的核心种质相比,所构建的核心种质性状间的相关系数与原种质性状间的相关系数更接近,并且还能很好地保持原种质的相关性。这表明所构建的初级核心种质能更好地保存原种质性状间的遗传和变异。

对 128 份初级核心种质进行分析,结果表明,40 对 SSR 引物共检测出 139 个等位基因变异,每对 SSR 引物扩增出的等位基因数 2~7 个,平均 3.49 个;多态性信息量(PIC)为 0.749 9~0.085 4,平均 0.506 5(表 7)。这与原种质检测出的结果非常接近,可见,这 128 份初级核心种质从等位基因的变异方面可代表原种质。

表5 原种质和核心种质性状间的相关系数

Table 5. Correlation coefficients among traits in the core collection and the initial collection

项目	主茎高	侧枝长	总分枝数	单株生产力	单株结果数	百果重	百仁重	叶长	叶宽	荚果长	荚果宽	粒长	粒宽	生育期
	1 ^a													
主茎高	1 ^b													
	1 ^c													
侧枝长	0.969**	1												
	0.989**	1												
	0.982**	1												
总分枝数	-0.125	-0.079	1											
	-0.167	-0.161	1											
	-0.215	-0.217	1											
单株生产力	-0.125	-0.143*	0.105	1										
	0.018	0.005	0.264	1										
	-0.119	-0.177	0.095	1										
单株结果数	-0.032	-0.056	0.058	0.678**	1									
	-0.014	-0.063	0.270	0.693**	1									
	-0.023	-0.075	0.128	0.683**	1									
百果重	-0.039	-0.069	-0.022	0.493**	-0.089	1								
	0.204	0.184	-0.034	0.559**	0.033	1								
	-0.246	-0.264	-0.017	0.652**	-0.014	1								
百仁重	-0.103	-0.130	0.118	0.509**	-0.016	0.795**	1							
	0.054	0.020	0.022	0.588**	0.076	0.863**	1							
	-0.330	0	0.103	0.650**	0.068	0.809**	1							

续表 5
Continued table 5

项目	主茎高	侧枝长	总分枝数	单株生产力	单株结果数	百果重	百仁重	叶长	叶宽	荚果长	荚果宽	粒长	粒宽	生育期
叶长	0.123	0.132	-0.178**	0.116	-0.006	0.179**	0.191**	1						
	0.232	0.234	-0.351	0.055	-0.143	0.304	0.356	1						
	0.248	0.193	-0.480**	0.152	0.122	0.189	0.058	1						
叶宽	0.037	0.049	-0.155*	0.077	0.153*	-0.004	0.028	0.708**	1					
	0.157	0.156	-0.375	-0.016	-0.031	0.166	0.210	0.861**	1					
	0.124	0.082	-0.410*	0.125	0.222	0.092	-0.122	0.757**	1					
荚果长	0.034	0.031	0.082	0.185**	-0.255**	0.512**	0.511**	0.152*	-0.180**	1				
	0.429	0.441*	0.107	0.427	-0.089	0.556**	0.532*	0.257	0.155	1				
	0.177	0.190	-0.014	0.226	-0.288	0.497**	0.495**	0.044	-0.126	1				
荚果宽	-0.110	-0.125	0.159*	0.212**	0.014	0.262**	0.363**	0.094	0.058	0.490**	1			
	0.114	0.061	0.413	0.400	0.256	0.299	0.452*	0.248	0.195	0.527*	1			
	-0.278	-0.295	0.219	0.178	-0.130	0.395*	0.497**	0.096	0.015	0.427*	1			
粒长	-0.031	-0.060	0.212**	0.152*	-0.217**	0.410**	0.509**	0.012	-0.304**	0.841**	0.505**	1		
	-0.103	-0.115	0.462*	0.523*	0.098	0.469*	0.654**	-0.041	-0.246	0.620**	0.592**	1		
	-0.207	-0.216	0.351	0.281	-0.073	0.412*	0.532**	-0.332	-0.393*	0.673**	0.552**	1		
粒宽	-0.112	-0.117	0.019	0.055	0.197**	-0.075	-0.043	-0.181**	-0.033	-0.090	0.274**	0.013	1	
	0.151	0.128	0.229	-0.055	0.227	-0.176	-0.042	0.069	-0.018	-0.100	0.452*	0.039	1	
	-0.274	-0.260	0.060	-0.166	-0.023	-0.128	-0.030	-0.201	-0.075	-0.254	0.230	-0.046	1	
生育期	-0.048	-0.090	-0.001	-0.062	-0.090	0.068	0.180**	0.144*	0.063	0.088	0.175**	0.180**	-0.102	1
	-0.153	-0.186	0.011	0.184	0.189	0.182	0.287	0.112	0.038	0.058	0.216	0.309	0.096	1
	-0.493**	-0.594**	0.139	0.408*	0.305	0.377*	0.483**	0.161	0.171	-0.003	0.347	0.287	0.031	1

注：“a”为原性状间的相关系数；“b”为完全随机取样所获得核心种质性状间的相关系数；“c”为欧氏距离最短聚类结合优先取样法构建的花生核心种质(DIS2C1-50)性状相关系数；“*”和“**”分别表示相关达到0.05和0.01显著水平

表 6 128 份花生材料的聚类结果

Table 6. Clustering results of 128 peanut materials

类群	材料名称	数量
D	吉花 07-8、小白沙、花育 16、花育 25、锦 9203-608、J-21、吉扶 4、锦 2010、P06-9、冀 9814、远杂 9847、濮花 23、C-40、京白 2、环 10-7、郑花、吉花 07-9、花育 40、环 10-2、无名 5 号、C-22、鲁花 15、科濮、J2、花育 33、湘花 2008、莱农 06、花育 28、花选 39、豫花 9327、冀 0211、C-37、冀花 9、锦花 14、花育 37	35
A	吉花 07-1、403-2、403-10、唐花 11、17、黑 1、铁花 2 号、锦 2015、J-6、科富 2 号、开花 21、开农 37、花育 23、花育 27、环 10-3、小日本、无名 17 号、13、阜花 11 号、SD-5、吉花 07-2、铁花 3 号、铁花 1 号、莲花 609、吉花 07-4、远育 16-8、J-7、402-7、J-3、莲玉 6、花育 22、402-3、吉花 07-11、环 10-5、锦花 7 号、J-1、J-17、403-11、豫紫 1、豫花 9503	40
B	锦 9806-F1、锦 9804、濮花 24、C-59、E12、扶花 201、白城引、多粒粉、无名 16 号、J-2、远杂 9307、无名 4 号、远杂 9102、豫花 23、冀花 7、开府 23、吉花 08-36、红粒王	18
C	濮花 5、吉花 07-7、阜花 13、阜花 10 号、广州珍珠红、锦 9818、锦 9826、锦 9938、冀 9606、李修海引、锦花 10、R03-3、无名 2 号、无名 8 号、无名 10 号、锦 2001、远杂 2-3-3、鲁花 12、SD-3、阜花 16、淮花 9、冀 02-6、阜花 14、双花 1 号、吉花 08-41、濮花 22、誉宇 2 号、科富 3 号、C-25、双辽红粒王、无名 24 号、无名 7 号、花小宝、莲花 209、誉宇 1 号、	35

表 7 40 对核心引物信息

Table 7. Information of 40 core primers

引物名称	总等位变异数	有效等位变异数	Shannon-Weaver 指数	Nei 期望杂合度	多态性信息量
7G2	4	1.853 5	0.839 8	0.460 5	0.460 5
PM36	6	2.946 5	1.297 6	0.660 6	0.660 6
2A6	4	2.057 5	0.811 9	0.514 0	0.514 0
18C5F	4	2.199 9	0.970 4	0.545 4	0.545 4
13E 9	2	1.825 1	0.644 4	0.452 1	0.452 1
14A07	2	1.830 1	0.646 0	0.453 6	0.453 6
2E 6	6	2.950 8	1.284 0	0.661 1	0.661 1
2F5	2	1.152 9	0.257 3	0.132 7	0.132 6
2G03	3	2.908 5	1.083 4	0.656 2	0.656 2
2A5	3	2.680 6	1.042 4	0.626 9	0.626 9
GNB353	2	1.244 0	0.347 0	0.196 2	0.196 1
GNB38	2	1.307 2	0.397 6	0.235 0	0.235 0
GNB127	2	1.579 0	0.553 1	0.366 7	0.366 7
GNB100	2	1.205 3	0.311 7	0.170 4	0.170 3
GNB278	3	2.226 5	0.870 2	0.550 9	0.550 9
GNB387	3	2.757 4	1.050 5	0.637 3	0.637 3
GNB523	5	3.372 7	1.343 3	0.703 5	0.703 5
GNB873	3	2.539 9	1.013 9	0.606 3	0.606 3
GNB989	4	2.232 9	0.903 4	0.552 2	0.552 2
GNB1121	3	1.093 5	0.211 0	0.085 5	0.085 4
GNB519	4	2.986 4	1.195 9	0.665 1	0.665 1
GNB580	4	2.549 6	1.050 8	0.607 8	0.607 8
GNB620	3	2.815 9	1.067 3	0.644 9	0.644 9
GNB716	5	2.767 3	1.172 3	0.638 6	0.638 6
GNB733	5	1.290 3	0.534 2	0.225 0	0.225 1
GNB837	5	2.346 6	1.000 5	0.573 9	0.573 9
GNB840	4	3.523 8	1.320 7	0.716 2	0.716 2
GNB703	3	2.445 8	0.970 8	0.591 1	0.591 1
GNB1027	6	1.647 4	0.805 4	0.393 0	0.393 0
GNB619	3	2.611 0	1.021 3	0.617 0	0.617 0
GNB649	3	2.227 6	0.870 0	0.551 1	0.551 1
GNB686	4	2.564 9	1.071 8	0.610 1	0.610 1

续表 7

Continued table 7

引物名称	总等位变异数	有效等位变异数	Shannon-Weaver 指数	Nei 期望杂合度	多态性信息量
GNB876	2	1.846 5	0.651 0	0.458 4	0.458 5
GNB902	2	1.553 2	0.541 5	0.356 2	0.356 1
GNB1021	4	3.997 8	1.386 0	0.749 9	0.749 9
GNB181	3	2.812 7	1.066 9	0.644 5	0.644 4
GNB324	4	1.193 6	0.375 3	0.162 2	0.162 1
GNB613	3	2.230 1	0.929 7	0.551 6	0.551 6
GNB274	4	3.707 3	1.344 7	0.730 3	0.730 3
SP9-4	7	4.283 2	1.692 5	0.523 1	0.766 5

对聚类结果的亲缘关系比较,其结果见表 8。从 260 份材料中初步筛选出 128 份核心种质,经过显著性的 T 测验后,其差异均不显著,则表明

所构建的花生初级核心种质可以代表原种质 80% 的遗传信息。

表 8 花生核心种质和原种质的符合率及 T 测验Table 8. T -testing and comparison of peanut core and initial collections

遗传信息	总等位变异数	有效等位变异数	Shannon-Weaver 指数	Nei 期望杂合度
原种质(260 份)	3.56	2.300	0.89	0.51
核心种质(128 份)	3.49	2.280	0.88	0.51
t 值	0.28	0.098	0.22	0.14

3 讨论

目前,为了更全面地体现群体的遗传结构,一定要有不同的数据,这些数据不仅仅局限于数量性状,更要有能够代表整体遗传变异的分子方面的数据。为了计算遗传距离,Bar-Hen^[22]等在 1995 年首次提利用二阶段分析策略对玉米自交进行分析,其实质在于利用分子标记信息计算各材料之间的遗传距离,如果得到的值高于之前设定好的阈值,则个体间就存在遗传差异,再对有差异的进行表型数据分析,这种分阶段的分析数据并不是实际的信息整合。Noirot^[23]等在 1996 年为了剔除多余因素对主成分分析所造成的干扰,只对分析所得到的 18 个主成分中的 6 个主成分进行遗传距离的计算。Islam^[24]等在 2004 年利用数量性状和质量性状的整合进行主成分分析,实质上就是对数量性状进行转换,最后统一进行分析。Liu^[25]等在 2002 年对 151 份美国大麦核心种质的同工酶数据资料进行聚类分析和主坐标分析。Xu^[26]等在 2006 年对 232 1 份大豆的分子标记数据和数量性状的数据进行主成分分析,构建核心种质。Wang 等^[27]在 2007 年将数量性状数据和分子标记数据进行整合,计算所整合

数据的混合遗传距离。张春雨等^[28]在 2009 年采用位点优先取样法进行取样,通过 SM、Jaccard 或 Nei & Li 遗传距离进行多次聚类构建新疆野苹果核心种质。

截至目前,未见有利用分子标记结合数量性状构建吉林省花生核心种质的相关文献报道。本研究利用混合遗传距离整理数量性状数据,并与 SSR 标记相结合构建花生初级核心种质。结果表明,利用数量性状数据结合 SSR 标记数据构建的花生初级核心种质既能代表原种质的表型性状又能代表基因型性状。

利用 14 个数量性状以 260 份花生为研究材料构建花生形态数据库,在未标准化的数量性状数据构建花生初级核心种质的过程中,欧式距离优于马氏距离,最短距离法优于最长距离法、中间距离法、重心法、类平均法、可变类平均法、可变量和离差平方和法,优先取样法优于偏离度取样法和随机取样法。在 50% 的取样比例下,利用欧式距离,采用优先取样法结合最短距离法进行聚类分析,选出 128 个样品构建花生核心种质。

利用主成分分析对所构建的初级核心种质进行确认,结果发现原种质和核心种质的主成分分析结果的特征值、贡献率和累计贡献率非常相近。

选入特征值>1的5个主成分,其中,原种质的第5个主成分的特征值为1.191,累计贡献率为74.484%;而所构建的核心种质的第5个主成分的特征值为1.225,累计贡献率已达到82.628%;比较原种质和初级核心种质性状间的相关系数发现,初级核心种质能较好地保存原种质性状间的遗传和变异。同时,对花生初级核心种质利用SSR分子标记技术对其进行遗传多样性分析,40对SSR引物共检测出139个等位基因变异,每对SSR引物扩增出的等位基因数2~7个,平均3.49个;多态性信息量(PIC)为0.749 9~0.085 4,平均0.506 5。按UPGMA法进行聚类分析,同样划分为A、B、C、D 4个类群,并且这4个类群的材料与原种质的4个类群所划分的材料基本相符。使用MEGA4软件对128份花生材料进行系统发育树的重建,划分结果与UPGMA聚类结果一致。结果表明,利用数量性状数据结合SSR分子标记数据构建的花生初级核心种质既能代表原种质的表型性状又能代表基因型性状。

参考文献:

- [1] 周雪松,赵谋明.我国花生食品产业现状与发展趋势[J].食品与发酵工业,2004,30(6):84-89.
- [2] FAO. Statistical Database. Food and Agriculture Organization (FAO) of the United Nations, Rome[EB/OL](2008)[2016-08-04]. <http://www.fao.org>.
- [3] 凤桐,高华援,赵叶明,等.吉林省花生生产现状与发展优势[J].吉林农业科学,2010,35(1):23-25,27.
- [4] 丁小霞,李培武,白艺珍,等.中国花生黄曲霉毒素风险评估中膳食暴露非参数概率评估方法[J].中国油料作物学报,2011,33(4):402-408.
- [5] 陈冉,李培武,马飞,等.花生黄曲霉毒素污染臭氧脱毒技术研究[J].中国油料作物学报,2013,35(1):92-96.
- [6] 白雪,王柏涛.吉林省农科院花生课题研究技术见成效[N].中国食品报(第A5版),2013-05-24.
- [7] Wang J, Sun J Z, Liu D C, et al. Analysis of Pina and Pinb alleles in the micro-core collections of Chinese wheat germplasm by Eco-tilling and identification of a novel Pinb allele[J]. Journal of Cereal Science, 2008, 48: 836-842.
- [8] 吴旭江,程凯,臧淑江,等.普通小麦株高的遗传分析[J].吉林农业大学学报,2016,38(3):274-280.
- [9] Wang J C, Hu J, Zhang C F, et al. Assessment on evaluating parameters of rice core collections constructed by genotypic values and molecular marker information[J]. Rice Science, 2007, 14(2): 101-110.
- [10] Liu F, Sun G L, Salomon B, et al. Characterization of genetic diversity in core collection accessions of wild barley, *Hordeum vulgare* ssp. *Spontaneum*[J]. Hereditas, 2002, 136: 67-73.
- [11] Reddy L J, Upadhyaya H D, Gowda C L L, et al. Development of core collection in pigeonpea [*Cajanuscajan* (L.) Millspaugh] using geographic and qualitative morphological descriptors[J]. Genetic Resources and Crop Evolution, 2005, 52: 1049-1056.
- [12] Dwivedi S L, Upadhyaya H D, Hegde D M. Development of core collection using geographic information and morphological descriptors in safflower (*Carthamus tinctorius* L.) germplasm[J]. Genetic Resources and Crop Evolution, 2005, 52: 821-830.
- [13] Li Y, Shi Y S, Cao Y S, et al. Establishment of a core collection for maize germplasm preserved in Chinese National GenBank using geographic distribution and characterization data[J]. Genetic Resources and Crop Evolution, 2004, 51: 845-852.
- [14] 才源,李凯,姜涛,等.玉米秃尖性状的遗传模型分析及SSR标记[J].吉林农业大学学报,2015,37(1):6-13.
- [15] 孟强,戴冬青,杨利,等.糯玉米种质SSR标记遗传及品质性状分析[J].吉林农业大学学报,2014,36(3):258-264.
- [16] 田晓艳,刘强,曲静,等.9个不同耐密性自交系的SSR标记分析[J].吉林农业大学学报,2013,35(5):511-515.
- [17] 姜涛,李凯,才源,等.玉米自交系SSR标记技术的优化[J].吉林农业大学学报,2014,36(6):653-659.
- [18] 宗绪晓,关建平,王述民,等.国外栽培豌豆遗传多样性分析及核心种质构建[J].作物学报,2008,34(9):1518-1528.
- [19] Ma Y S, Wang W H, Wang L X, et al. Genetic diversity of soybean and the establishment of a core collection focused on resistance to soybean cyst nematode[J]. Journal of Integrative Plant Biology, 2006, 48(6): 722-731.
- [20] Wang L X, Guan Y, Guan R X, et al. Establishment of Chinese soybean (*Glycine max*) core collections with agronomic traits and SSR markers[J]. Euphytica, 2006, 151: 215-223.
- [21] 刘遵春,张春雨,张艳敏,等.利用数量性状构建新疆野苹果核心种质的方法[J].中国农业科学,2010,43(2):358-370.
- [22] Bars-Hen A, Charcosset A, Bourgoin M, et al. Relationship between genetic markers and morphological traits in maize inbred lines collection[J]. Euphytica, 1995, 84: 145-154.
- [23] Noirot M, Hamon S, Anthony F. The Principal component scoring: A new method of constituting a core collection using quantitative data[J]. Genetic Resources and Crop Evolution, 1996, 43: 1-6.
- [24] Islam F M A, Beebe S, Muffoz M, et al. Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool[J]. Theoretical and Applied Genetics, 2004, 108: 243-252.
- [25] Liu Fang, Sun Guo, Salomon B, et al. Characterization of genetic diversity in core collection accessions of wild barley, *Hordeum vulgare* ssp. *spontaneum* [J]. Hereditas, 2002, 136: 67-73.
- [26] Xu Haiming, Hu Jin, Zhu Jun, et al. Sampling a core collection of Island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits[J]. Genetic Resources and Crop Evolution, 2006, 53: 515-521.
- [27] Wang Jiancheng, Hu Jin, Xu Haiming. Strategy on constructing core collections by least distance stepwise sampling[J]. Theoretical and Applied Genetics, 2007, 115(1): 1-8.
- [28] 张春雨,陈学森,张艳敏,等.采用分子标记构建新疆野苹果核心种质的方法[J].中国农业科学,2009,42(2):597-604.

(责任编辑:赵立华)