

基于农艺性状指标的山西高粱地方品种核心种质构建

李 萌^{1,2}, 秦慧彬¹, 王宇楠¹, 穆志新¹, 杜慧玲²

(¹ 山西农业大学农业基因资源研究中心 / 农业部黄土高原作物基因资源与种质创制重点实验室 / 杂粮种质资源发掘与
遗传改良山西省重点实验室, 太原 030031; ² 山西农业大学农学院, 太谷 030801)

摘要: 利用 1285 份山西省高粱地方品种 18 个农艺性状的历史数据, 通过比较不同取样方法、取样比例和聚类方法组合的构建方法, 确定了“多次聚类偏离度取样法+15% 取样比例+欧氏距离+最长距离法”为山西省高粱地方核心种质构建的方法。192 份初选核心种质和所有样本的均值差异百分率、方差差异百分率、极差符合率和变异系数变化率分别为 0、83.33%、97.45% 和 119.63%。同时, 在此基础上补充选择 6 个具有特殊性状但未选入的种质资源, 最终确定 198 份高粱资源组成山西省高粱地方品种核心种质, 取样量为 15.4%。经核心种质和所有样本不同性状的均值 *t* 测验、极值和标准差比较、遗传多样性指数的 *t* 值检验以及主成分分析, 最终得出核心种质的 198 份高粱资源能够代表山西省高粱地方品种的遗传多样性, 为山西省高粱地方品种评价和利用提供了优先样本。

关键词: 高粱; 地方品种; 农艺性状; 核心种质

A Core Collection of Sorghum Landraces Formed by Taking Use of Agronomic Traits in Shanxi Province

LI Meng^{1,2}, QIN Hui-bin¹, WANG Yu-nan¹, MU Zhi-xin¹, DU Hui-ling²

(¹ Center for Agricultural Genetic Resources Research, Shanxi Agricultural University/Key Laboratory of Crop Gene Resources and Germplasm Enhancement on Loess Plateau, Ministry of Agriculture /Shanxi Key Laboratory of Genetic Resources and Genetic Improvement of Minor Crops, Taiyuan 030031; ² College of Agronomy, Shanxi Agricultural University, Taigu 030801)

Abstract: This study attempted to build up a core collection of 1285 sorghum germplasm accessions at Shanxi Gene Bank. These accessions were historically investigated for the phenotypic variations 18 agronomic traits, and these datasets were subjected for tests using the combinations on sampling methods, sampling proportion and clustering methods. A pipeline combining multiple cluster deviation sampling method, 15% sampling proportion, euclidean distance and complete method was used. By pairwise comparison of phenotypic datasets in core collection (192 accessions) to all samples, the mean difference percentage (MD) was 0, the variance difference percentage (VD) was 83.33%, the coincidence rate of range (CR) was 97.45%, and the variable rate of coefficient of variation (VR) was 119.63%. Furthermore, eight accessions with particular characters were incorporated. This core collection including 198 landraces has been proposed, followed by tests using t-test, extreme value, standard deviation, the t-test of Shannon-Weaver diversity index, as well as principal component. As a result, 198 sorghum landraces are representative on the diversity of sorghum landraces hosted by Shanxi Gene Bank, and this collection might be useful in practical uses for sorghum breeding.

Key words: sorghum; landraces; agronomic traits; core collection

收稿日期: 2020-06-10 修回日期: 2020-06-22 网络出版日期: 2020-07-08

URL: <http://doi.org/10.13430/j.cnki.jpgr.20200610001>

第一作者研究方向为高粱种质资源评价, E-mail: nkypzslm@163.com

通信作者: 穆志新, 研究方向为作物种质资源, E-mail: muzx2008@sina.com

杜慧玲, 研究方向为化学调控及农药毒理研究, E-mail: duhuiling66@163.com

基金项目: 山西省农业科学院农业科技创新研究 (YCX2020YQ41)

Foundation project: Agricultural Science and Technology Innovation Research of SAAS (YCX2020YQ41)

植物种质资源的收集、保存和研究一直以来都是科学家们选育高产、优良、抗逆新品种的基础,丰富的种质资源为科学研究提供了大量的材料,但繁多的种质资源又会给保存、鉴定等工作带来巨大的挑战,建立一种种质资源分析处理的方法是非常有必要的^[1]。Frankel 等^[2]最早提出核心种质(Core collection)的概念。迄今为止,马铃薯^[3]、水稻^[4]、大豆^[5]、小麦^[6-7]、玉米^[8]、棉花^[9]、豌豆^[10]、茶^[11]等植物均进行了核心种质的构建,国外学者也进行了高粱核心种质构建研究;Grenier 等^[12]采用对数取样方法选取 225 份高粱构建了核心种质,其中入选比例为 10%;Tariq 等^[13]根据 SSR 标记结果选择了 107 份资源构建了高粱的核心种质;Dahlberg 等^[14]对保存在美国国家植物种质体系(NPGS)的高粱进行了核心种质构建,全球 77 个国家的高粱资源按 10% 的比例随机选择构建核心种质。

高粱(*Sorghum bicolor*(L.) Moench)抗旱、抗涝、耐盐碱、耐瘠薄,具有广泛的适应性和较强的抗逆能力。山西作为高粱生产种植的中心地带,多样的地形和气候形成了丰富多彩的高粱种质资源。由于近年来种质资源收集数目的增加,给山西省高粱种质资源的保存、研究和利用带来了挑战,因此本研究对山西省种质资源库编目入库的 1285 份山西高粱地方品种进行核心种质构建,为山西省高粱的有效保存、遗传研究以及新品种的选育提供依据。

1 材料与方法

1.1 试验材料

供试高粱材料 1285 份,为山西省种质库编目入库的山西省地方高粱种质资源,来源于山西省 11 个市的 100 个区/县(表 1)。

表 1 山西高粱地方品种的来源信息

Table 1 The collection sites of sorghum landraces in Shanxi province

来源地 Origin	所属区/县 Subordinate district /county	数量 Number
大同 Datong	新荣区、平城区、云冈区、云州区、灵丘、浑源、天镇、阳高、广灵、左云	133
朔州 Shuozhou	朔城、平鲁、山阴、怀仁、应县	81
忻州 Xinzhou	忻府、原平、保德、代县、静乐、定襄、河曲、繁峙、五台、五寨、宁武、岢岚、偏关、神池	300
吕梁 Lvliang	汾阳、孝义、离石、临县、文水、石楼、兴县、交城、交口、岚县、柳林、中阳	134
太原 Taiyuan	阳曲、娄烦、清徐	76
阳泉 Yangquan	平定、盂县	20
晋中 Jinzhong	榆次、太谷、介休、祁县、榆社、和顺、昔阳、左权、灵石、平遥、寿阳	223
临汾 Linfen	侯马、霍州、曲沃、翼城、襄汾、洪洞、古县、安泽、浮山、吉县、乡宁、隰县、永和、蒲县、大宁	94
长治 Changzhi	潞城、屯留、襄垣、武乡、沁县、沁源、黎城、壶关、平顺、长子、武乡	136
晋城 Jincheng	城区、高平、沁水、阳城、陵川	42
运城 Yuncheng	盐湖、永济、稷山、绛县、临猗、垣曲、万荣、闻喜、夏县、新绛、平陆、芮城	46

1.2 方法

1.2.1 数据整理 1285 份高粱地方品种农艺性状数据来源于山西省农业科学院农作物品种资源研究所高粱种质资源数据库。18 个农艺性状中包括芽鞘色、幼苗叶色、主脉色、穗型、穗形、颖壳包被度、颖壳色、粒色、分蘖性 9 个质量性状;参照《高粱种质资源描述规范和数据标准》^[15]对质量性状进行赋值,同时将代码数除以代码总数,转换成[0, 1]数值。9 个数量性状包括株高、茎粗、主穗长度、主穗柄长度、单穗粒重、千粒重、全生育期、角质率、着壳率。

1.2.2 核心种质的构建方法、评价参数 应用 QGA2.0 软件构建核心种质^[16],从 3 种取样方法

(多次聚类随机取样法、多次聚类优先取样法、多次聚类偏离度取样法)、6 种取样比例(5%、10%、15%、20%、25%、30%)、8 种聚类方法(最短距离法、可变法、可变类平均法、类平均法、离差平方和法、中间距离法、重心法、最长距离法)和 2 种遗传距离(欧氏距离、马氏距离)4 个层次,共构建了 288 个初选核心种质,通过均值差异百分率、方差差异百分率、极差符合率和变异系数变化率 4 个参数评价 288 个初选核心种质子集,确定最佳取样方法。均值差异百分率小于 20%,同时极差符合率参数大于 80% 时,核心种质的代表性较好,且均值差异百分率越小,极差符合率、方差差异百分率和变异系数变化率越大,核心种质的代表性越强。

1.2.3 核心种质的验证 采用 t 检验进行核心种质和所有样本的均值差异显著性分析; 通过 Shannon-Wiener's 多样性指数 (H') 进行遗传多样性评价, 计算各性状的 H' :

$$H' = -\sum p_i \ln p_i \quad (1)$$

其中, H' 表示某性状的遗传多样性指数; p_i 表示某一性状第 i 个代码出现的频率;

通过 t 值^[17] 进行核心种质和所有样本的 Shannon-Wiener's 多样性指数比较, 检验核心种质的代表性。计算各性状的 t 值:

$$\begin{aligned} \text{Var}(H') &= \frac{\sum (P_j \ln P_j)^2 - (\sum P_j \ln P_j)^2}{N} + \frac{n-1}{2N^2}, \\ m &= \frac{[\text{Var}(H'_1) + \text{Var}(H'_2)]^2}{\frac{\text{Var}(H'_1)^2}{N_1} + \frac{\text{Var}(H'_2)^2}{N_2}}, \\ t &= \frac{(H'_1 - H'_2)}{\sqrt{\text{Var}(H'_1) + \text{Var}(H'_2)}} \end{aligned} \quad (2)$$

其中, H'_1 、 H'_2 分别表示所有样本和核心种质某一性状的多样性指数, P_j 表示某一性状第 j 个代码出现的频率, N_1 、 N_2 表示所有样本和核心种质的样本数, $\text{Var}(H'_1)$ 、 $\text{Var}(H'_2)$ 为方差, m 为自由度, t 为检验 t 值。

2 结果与分析

2.1 核心种质构建方法的确定

2.1.1 取样方法比较 构建的 288 组初选核心种质的评价参数见表 2。基于均值差异百分率参数小于 20% 的评价标准, 各取样方法中均有均值差异百分率参数大于 20% 的核心种质子集, 多次聚类随机取样法和多次聚类偏离度取样法的最大均值差异百分率参数均为 27.78%, 小于多次聚类优先取样法的最大均值差异百分率参数 (33.33%), 且多次聚类随机取样法的均值差异百分率总平均值最小 (3.36%), 其次是多次聚类偏离度取样法 (4.69%), 多次聚类优先取样法 (5.04%)。因此, 在均值差异百分率评价参数上, 多次聚类随机取样法和多次聚类偏离度取样法优于多次聚类优先取样法。基于极差符合率大于 80% 的评价标准, 构建的各组初选核心种质的极差符合率均大于 80%, 且 3 种取样方法的平均极差符合率均超过 90%。基于方差差异百分率和变异系数变化率越大越好的评价原则, 构建的各组初选核心种质的变异系数变化率均超过 100%, 保留了原样本种质的遗传变异。在极差符合率、方差差异百分率和变异系数变化率评价参数上, 多次聚类偏离度取样法的均优于多次聚类随机取样法。综合

表 2 288 个初选核心种质的评价参数

Table 2 Evaluation parameter of 288 primary core collection subset

(%)

取样比例 Sample ratio	遗传距离 Genetic distance	聚类方法 Cluster method	多次聚类随机取样法 Multiple cluster random sampling method				多次聚类优先取样法 Multiple cluster preferred sampling method				多次聚类偏离度取样法 Multiple cluster deviation sampling method			
			均值差	方差差	极差	变异系	均值差	方差差	极差符	变异系	均值差	方差差	极差符	变异系
			异百分	异百分	符合率	数变化	异百分	异百分	合率	数变化	异百分	异百分	合率	数变化
			率 MD	率 VD	CR	率 VR	率 MD	率 VD	CR	率 VR	率 MD	率 VD	CR	率 VR
5	欧氏距离	最短距离法	16.67	72.22	93.35	127.93	22.22	61.11	100.00	128.28	27.78	66.66	95.53	136.20
		最长距离法	0	27.78	88.29	113.38	5.56	61.11	100.00	121.24	0	72.22	95.52	133.34
		中间距离法	11.11	5.56	83.44	106.40	5.56	50.00	100.00	119.39	0	61.11	93.83	122.77
		重心法	5.56	33.33	89.85	109.94	16.67	55.56	100.00	122.94	11.11	44.44	93.82	118.44
		类平均法	0	38.89	89.38	111.10	0	50.00	100.00	119.31	5.56	61.11	93.62	130.67
		可变类平均法	0	27.78	87.75	112.82	0	55.56	100.00	122.28	5.56	61.11	93.80	129.49
		可变法	0	22.22	90.19	112.43	5.56	55.56	100.00	126.40	0	72.22	93.57	130.52
		离差平方和法	0	50.00	91.98	113.55	11.11	50.00	100.00	122.71	0	61.11	94.15	130.07
	马氏距离	最短距离法	27.78	61.11	92.46	123.31	22.22	61.11	100.00	128.52	27.78	72.22	93.54	137.98
		最长距离法	5.56	38.89	88.42	114.30	5.56	61.11	100.00	121.24	5.56	77.78	93.82	129.99
		中间距离法	5.56	27.78	87.76	110.43	0	55.56	100.00	123.56	27.78	55.56	90.79	130.27
		重心法	5.56	50.00	91.70	120.15	16.67	66.67	100.00	128.57	22.22	72.22	92.07	132.11
		类平均法	0	33.33	88.00	113.98	5.56	61.11	100.00	124.82	0	72.22	92.69	132.65
		可变类平均法	0	27.78	86.76	111.72	0	61.11	100.00	123.68	5.56	66.67	94.64	132.93

表 2(续)

取样 比例 Sample ratio	遗传距离 Genetic distance	聚类方法 Cluster method	多次聚类随机取样法				多次聚类优先取样法				多次聚类偏离度取样法			
			Multiple cluster random sampling				Multiple cluster preferred sampling				Multiple cluster deviation			
			method				method				sampling method			
			均值差 异百分 率 MD	方差差 异百分 率 VD	极差 符合率 CR	变异系 数变化 率 VR	均值差 异百分 率 MD	方差差 异百分 率 VD	极差符 合率 CR	变异系 数变化 率 VR	均值差 异百分 率 MD	方差差 异百分 率 VD	极差符 合率 CR	变异系 数变化 率 VR
10	欧氏距离	可变法	0	22.22	84.75	109.88	0	61.11	100.00	121.06	5.56	66.67	94.01	132.64
		离差平方和法	0	22.22	92.13	112.54	5.56	50.00	100.00	122.15	11.11	72.22	93.94	133.54
		最短距离法	27.78	61.11	93.79	118.36	16.67	66.67	100.00	121.26	22.22	77.78	97.54	127.54
		最长距离法	0	33.33	91.52	110.90	0	50.00	100.00	115.35	0	66.67	95.92	124.24
		中间距离法	5.56	11.11	89.40	105.89	5.56	50.00	100.00	110.61	0	61.11	94.67	118.24
		重心法	5.56	22.22	92.34	104.44	5.56	38.89	100.00	108.08	0	38.89	94.56	114.42
		类平均法	0	33.33	92.27	108.85	0	50.00	100.00	114.70	0	66.67	95.63	121.01
		可变类平均法	0	33.33	90.03	109.20	0	44.44	100.00	114.03	0	77.78	95.24	122.82
	马氏距离	可变法	0	33.33	91.31	109.07	0	50.00	100.00	117.90	0	72.22	94.91	121.66
		离差平方和法	0	50.00	93.20	110.65	0	55.56	100.00	113.55	5.56	72.22	95.15	122.80
		最短距离法	27.78	55.56	96.23	117.43	33.33	66.67	100.00	122.25	22.22	77.78	97.83	129.36
		最长距离法	5.56	27.78	89.35	110.35	0	44.44	100.00	114.87	0	72.22	96.29	121.70
		中间距离法	0	27.78	89.81	107.69	0	55.56	100.00	113.70	5.56	61.11	92.98	122.33
		重心法	0	33.33	92.37	108.01	22.22	55.56	100.00	115.49	5.56	66.67	96.09	121.00
		类平均法	0	27.78	89.03	109.88	5.56	55.56	100.00	116.79	0	72.22	94.86	123.22
		可变类平均法	0	11.11	90.79	107.39	5.56	61.11	100.00	115.54	0	72.22	95.11	122.82
15	欧氏距离	可变法	0	27.78	87.79	107.73	0	38.89	100.00	113.90	5.56	66.67	95.41	122.81
		离差平方和法	0	27.78	92.92	110.82	11.11	50.00	100.00	113.81	5.56	66.67	95.38	123.62
		最短距离法	22.22	55.56	93.86	113.42	16.67	66.67	100.00	116.13	22.22	77.78	98.21	122.59
		最长距离法	0	33.33	92.74	109.05	0	44.44	100.00	112.12	0	83.33	97.45	119.63
		中间距离法	0	11.11	91.94	103.79	0	38.89	100.00	107.58	0	55.56	95.45	113.81
		重心法	0	22.22	93.71	103.67	0	38.89	100.00	106.07	0	44.44	97.04	112.29
		类平均法	0	33.33	93.31	107.04	0	38.89	100.00	109.54	0	72.22	95.72	117.71
		可变类平均法	0	27.78	90.03	106.77	0	33.33	100.00	109.76	0	72.22	95.24	117.92
	马氏距离	可变法	5.56	16.67	92.86	105.69	5.56	55.56	100.00	113.23	0	72.22	96.03	118.22
		离差平方和法	0	33.33	94.24	108.07	0	50.00	100.00	110.31	0	77.78	96.11	118.22
		最短距离法	22.22	44.44	96.78	112.81	22.22	55.56	100.00	116.04	27.78	77.78	98.50	123.08
		最长距离法	5.56	38.89	91.89	108.44	5.56	55.56	100.00	111.34	0	72.22	96.29	118.27
		中间距离法	0	22.22	90.83	106.76	11.11	44.44	100.00	109.65	5.56	61.11	94.66	117.44
		重心法	5.56	27.78	92.68	106.50	16.67	44.44	100.00	111.34	0	61.11	96.09	117.23
		类平均法	0	27.78	92.40	107.59	0	44.44	100.00	110.94	0	72.22	96.25	117.26
		可变类平均法	0	11.11	91.65	105.66	5.56	50.00	100.00	111.19	0	72.22	95.48	118.00
20	欧氏距离	可变法	0	27.78	91.75	107.04	0	44.44	100.00	111.41	5.56	83.33	96.45	119.41
		离差平方和法	0	33.33	93.28	109.52	5.56	50.00	100.00	111.36	5.56	66.67	96.81	118.41
		最短距离法	22.22	44.44	95.84	110.88	22.22	61.11	100.00	111.82	22.22	72.22	99.24	118.93
		最长距离法	0	38.89	96.10	107.38	0	50.00	100.00	108.92	0	72.22	98.86	115.57
		中间距离法	0	22.22	94.95	103.92	0	16.67	100.00	105.33	0	55.56	96.29	111.90
		重心法	0	22.22	93.82	103.11	0	27.78	100.00	105.53	0	50.00	97.98	111.51
		类平均法	0	16.67	93.33	105.88	0	38.89	100.00	107.75	0	72.22	96.61	113.68
		可变类平均法	0	27.78	95.83	105.76	0	33.33	100.00	106.86	0	83.33	96.58	114.58
		可变法	0	27.78	93.84	105.45	0	55.56	100.00	110.10	0	83.33	96.58	114.33
		离差平方和法	0	33.33	94.24	106.49	5.56	44.44	100.00	108.02	0	83.33	97.12	115.01

表 2(续)

取样 比例 Sample ratio	遗传距离 Genetic distance	聚类方法 Cluster method	多次聚类随机取样法				多次聚类优先取样法				多次聚类偏离度取样法			
			Multiple cluster random sampling method				Multiple cluster preferred sampling method				Multiple cluster deviation sampling method			
			均值差 异百分 率 MD	方差差 异百分 率 VD	极差 符合率 CR	变异系 数变化 率 VR	均值差 异百分 率 MD	方差差 异百分 率 VD	极差符 合率 CR	变异系 数变化 率 VR	均值差 异百分 率 MD	方差差 异百分 率 VD	极差符 合率 CR	变异系 数变化 率 VR
25	马氏距离	最短距离法	27.78	50.00	97.05	111.36	33.33	66.67	100.00	113.39	22.22	72.22	99.51	119.23
		最长距离法	5.56	33.33	95.69	107.30	0	44.44	100.00	109.22	0	77.78	96.90	115.76
		中间距离法	0	16.67	94.02	105.68	0	33.33	100.00	106.79	5.56	66.67	94.67	115.01
		重心法	0	22.22	92.76	104.29	16.67	38.89	100.00	108.82	0	61.11	96.88	114.47
		类平均法	0	5.56	93.84	104.92	0	44.44	100.00	109.57	0	72.22	97.82	114.29
		可变类平均法	0	22.22	92.68	105.41	0	33.33	100.00	108.07	0	77.78	97.20	116.06
		可变法	0	38.89	94.27	106.83	0	44.44	100.00	109.58	11.11	83.33	96.79	116.73
		离差平方和法	0	38.89	94.71	108.36	5.56	38.89	100.00	108.15	11.11	72.22	97.92	116.81
	欧氏距离	最短距离法	5.56	38.89	97.48	108.39	22.22	55.56	100.00	109.09	11.11	72.22	99.24	115.69
		最长距离法	0	22.22	96.49	105.16	0	44.44	100.00	107.24	0	66.67	98.86	112.82
		中间距离法	0	11.11	96.80	102.89	0	22.22	100.00	104.64	0	55.56	96.29	109.76
		重心法	0	16.67	94.49	103.39	0	27.78	100.00	104.18	0	50.00	98.78	110.10
		类平均法	0	27.78	94.28	106.29	0	44.44	100.00	107.03	0	72.22	97.45	112.88
		可变类平均法	0	0.00	96.35	103.29	0	33.33	100.00	106.89	0	72.22	96.83	110.82
		可变法	0	5.56	93.92	102.68	0	50.00	100.00	108.87	0	72.22	96.83	110.81
		离差平方和法	0	16.67	94.41	104.13	0	44.44	100.00	107.49	0	72.22	97.38	111.24
30	马氏距离	最短距离法	16.67	44.44	97.19	108.03	16.67	55.56	100.00	110.79	27.78	72.22	99.51	115.90
		最长距离法	5.56	16.67	95.69	104.99	5.56	55.56	100.00	108.15	0	61.11	96.90	111.87
		中间距离法	0	11.11	94.31	104.91	0	27.78	100.00	105.46	5.56	66.67	96.77	112.57
		重心法	0	16.67	94.38	104.31	5.56	38.89	100.00	106.70	0	66.67	97.50	112.06
		类平均法	0	27.78	97.00	106.33	0	44.44	100.00	108.27	0	77.78	97.98	112.72
		可变类平均法	0	5.56	92.68	102.62	0	33.33	100.00	106.74	0	61.11	97.20	111.60
		可变法	0	5.56	94.27	104.65	0	44.44	100.00	107.96	0	61.11	97.04	112.84
		离差平方和法	0	33.33	94.71	106.69	0	38.89	100.00	106.91	0	61.11	97.92	112.52
	欧氏距离	最短距离法	5.56	44.44	97.65	107.13	16.67	61.11	100.00	107.18	5.56	66.67	99.24	112.17
		最长距离法	0	22.22	97.38	105.38	0	38.89	100.00	106.32	0	66.67	99.10	111.55
		中间距离法	0	11.11	96.80	102.67	0	22.22	100.00	104.84	0	55.56	97.16	109.19
		重心法	0	5.56	94.49	101.99	0	22.22	100.00	104.27	0	44.44	98.77	108.35
		类平均法	0	11.11	94.28	103.51	0	27.78	100.00	105.32	0	61.11	97.45	110.09
		可变类平均法	0	0	96.35	103.29	0	33.33	100.00	105.42	5.56	66.67	99.03	111.22
		可变法	0	16.67	96.71	104.33	0	38.89	100.00	107.59	5.56	66.67	99.03	111.31
		离差平方和法	0	16.67	94.41	104.13	0	44.44	100.00	106.59	0	66.67	99.03	111.05
	马氏距离	最短距离法	16.67	44.44	97.62	107.05	11.11	61.11	100.00	109.15	0	72.22	99.51	113.05
		最长距离法	5.56	27.78	97.45	105.64	0	44.44	100.00	107.06	0	72.22	98.30	111.76
		中间距离法	0	11.11	94.39	102.67	0	11.11	100.00	104.75	5.56	66.67	97.53	110.99
		重心法	0	16.67	95.99	103.38	5.56	11.11	100.00	104.37	0	66.67	98.12	111.14
		类平均法	0	16.67	94.54	104.91	0	38.89	100.00	106.70	0	61.11	98.26	110.09
		可变类平均法	0	11.11	92.76	104.24	0	38.89	100.00	106.27	0	66.67	99.32	111.85
		可变法	0	22.22	94.27	104.93	0	33.33	100.00	106.65	0	66.67	98.37	112.21
		离差平方和法	0	38.89	97.84	106.72	0	33.33	100.00	106.37	0	72.22	98.38	112.37

MD: Mean difference percentage, VD: Variance difference percentage, CR: Range coincidence rate, VR: Variation coefficient change rate

比较 3 种取样方法下的代表性评价参数,认为多次聚类偏离度取样法为最佳取样方法,既保持了初选核心种质与所有样本的性状一致性和变异范围,又有效地降低了遗传冗余度。

2.1.2 取样比例比较 在多次聚类偏离度取样法下,随着取样比例的增加,极差符合率均值逐渐升高,而变异系数变化率则逐步降低,方差差异百分率均值在 15% 和 20% 的取样比例下较大;综合比较各取样规模下均值差异百分率均值的变化后,在 15% 取样规模下,均值差异百分率均值较小,初选核心种质以较少的样本保持了同原始样本的一致性,据此认为 15% 的取样比例最为合适。

参照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,核心种质代表性越好的原则,在多次聚类偏离度取样法和 15% 取样规模下,采用“多次聚类偏离度取样法+15% 取样比例+欧氏距离+最长距离法”构建的初选核心种质均值差异百分率、方差差异百分率、极差符合

率和变异系数变化率分别为 0、83.33%、97.45% 和 119.63%,明显优于其他构建组合。

2.2 山西省高粱地方品种核心种质

根据获得的最佳取样策略将选出的 192 份初选核心种质同所有样本进行对比,补充 6 个具有特殊性状但未选入的种质资源,最终确定 198 份高粱资源组成山西省高粱地方品种核心种质,取样量为 15.4%。

2.3 核心种质的评价

2.3.1 所有样本和核心种质均值、极值、标准差比较分析 核心种质与所有样本 18 个表型性状的均值经 t 检验(表 3),差异均未达到显著水平,从均值上反映了总体与核心种质的差异不显著;由表 4 看出,所有样本和核心种质群体中的最大值和最小值均一致,说明核心种质保留了总体中的极值种质;核心种质各性状的标准差大于或等于总体,说明核心种质较所有样本异质性更强。综合比较分析得出核心种质对所有样本的表型变异特征具有较强的代表性。

表 3 所有样本与核心种质 18 个表型性状均值比较

Table 3 Pairwise comparison on 18 phenotypic characters between all samples and core collection

性状 Traits	所有样本平均值 Average of all samples	核心种质平均值 Average of core collection	t 值 t value	P 值 P value
株高 (cm) PH	262.11	258.76	0.58	0.56
茎粗 (cm) SD	1.39	1.40	-0.11	0.91
主穗长度 (cm) MPL	26.32	26.77	-0.57	0.57
主穗柄长度 (cm) PL	43.84	43.57	0.29	0.78
单穗粒重 (g) GWPS	52.12	50.97	0.61	0.54
千粒重 (g) TGW	24.84	24.50	0.73	0.47
全生育期 (d) PD	111.55	110.94	0.60	0.55
角质率 (%) EC	48.19	49.37	-0.49	0.62
着壳率 (%) GSP	22.40	23.21	-0.41	0.68
粒色 GCOL	0.65	0.62	1.61	0.11
穗形 PS	0.43	0.47	-1.42	0.16
颖壳色 GC	0.61	0.61	0.12	0.91
穗型 PT	0.56	0.57	-0.64	0.52
颖壳包被度 GCOV	0.64	0.65	-0.89	0.38
分蘖性 TI	0.44	0.44	0.00	1.00
主脉色 MVC	0.38	0.40	-1.64	0.10
芽鞘色 CC	0.89	0.87	1.52	0.13
幼苗叶色 LCS	0.59	0.57	0.94	0.35

PH: Plant height, SD: Stem diameter, MPL: Main panicle length, PL: Peduncle length, GWPS: Grain weight per spike, TGW: Thousand grain weight, PD: Period of duration, EC: Endosperm corneous, GSP: Glume seed percentage, GCOL: Grain color, PS: Panicle shape, GC: Glume color, PT: Panicle type, GCOV: Grain covering, TI: Tillering, MVC: Medium vein color, CC: Coleoptile color, LCS: Leaf color of seedling. The same as below

2.3.2 所有样本和核心种质多样性分析 通过比较所有样本与核心种质的 18 个农艺性状的多样性指数,能够从表型水平上了解核心种质的资源类型及

其与总体的差异。核心种质的茎粗、主穗长度、主穗柄长度、单穗粒重、千粒重、全生育期、角质率、着壳率以及幼苗叶色的多样性指数低于所有样本的多

表 4 所有样本与核心种质性状极值和标准差比较
Table 4 Contrast of extreme value and standard deviation between all samples and core collection

性状 Traits	所有样本 All samples			核心种质 Core collection		
	最大值 Max.	最小值 Min.	标准差 SD	最大值 Max.	最小值 Min.	标准差 SD
株高 (cm) PH	519.90	76.00	60.05	519.90	76.00	76.46
茎粗 (cm) SD	2.31	0.84	0.18	2.31	0.84	0.25
主穗长度 (cm) MPL	79.50	8.20	8.26	79.50	8.20	10.55
主穗柄长度 (cm) PL	93.80	2.90	9.67	93.80	2.90	12.65
单穗粒重 (g) GWPS	135.40	10.20	19.37	135.40	10.20	25.29
千粒重 (g) TGW	53.60	9.30	4.77	53.60	9.30	6.28
全生育期 (d) PD	155.00	84.00	11.59	155.00	84.00	13.38
角质率 (%) EC	100.00	0	28.22	100.00	0	31.31
着壳率 (%) GSP	100.00	0	25.29	100.00	0	27.77
粒色 GCOL	1.00	0.13	0.18	1.00	0.13	0.21
穗形 PS	1.00	0.13	0.36	1.00	0.13	0.38
颖壳色 GC	1.00	0.14	0.18	1.00	0.14	0.23
穗型 PT	1.00	0.20	0.29	1.00	0.20	0.30
颖壳包被度 GCOV	1.00	0.20	0.15	1.00	0.20	0.18
分蘖性 TI	1.00	0.25	0.15	1.00	0.25	0.17
主脉色 MVC	1.00	0.33	0.15	1.00	0.33	0.18
芽鞘色 CC	1.00	0.33	0.21	1.00	0.33	0.23
幼苗叶色 LCS	1.00	0.33	0.32	1.00	0.33	0.32

样性指数,株高、粒色、穗形、颖壳色、穗型、颖壳包被度、分蘖性、主脉色和芽鞘色核心种质的多样性指数高于所有样本,其中分蘖性的多样性指数差异显著,颖壳色和颖壳包被度的多样性指数差异极显著(表 5),表明核心种质保留了更多的表型变异特征,具有相对较好的代表性。

表 5 基本样本和核心种质多样性指数的比较
Table 5 Comparison of diversity index of core collection and all samples

性状 Traits	所有样本多样性指数 Diversity index of all samples	核心种质多样性指数 Diversity index of core collection	t 值 t value	P 值 P value
株高 PH	2.0707	2.0858	-0.3635	0.7165
茎粗 SD	2.0661	2.0227	0.8591	0.3911
主穗长度 MPL	1.9122	1.8808	0.6162	0.5383
主穗柄长度 PL	2.0333	2.0099	0.4781	0.6330
单穗粒重 GWPS	2.0665	2.0163	1.3048	0.1931
千粒重 TGW	2.0554	2.0509	0.1017	0.9191
全生育期 PD	2.0687	2.0002	1.7639	0.0788
角质率 EC	2.0217	1.9660	1.9108	0.0572
着壳率 GSP	1.5038	1.4701	0.5332	0.5943
粒色 GCOL	1.2500	1.3475	-1.3357	0.1828
穗形 PS	1.4440	1.4848	-0.6016	0.5479
颖壳色 GC	1.0844	1.3530	-3.8496	0.0001**
穗型 PT	1.5639	1.5832	-0.9139	0.3616
颖壳包被度 GCOV	1.0854	1.2830	-4.2001	0.0000**
分蘖性 TI	0.8771	0.9744	-1.9894	0.0477*
主脉色 MVC	0.3909	0.4998	-1.6054	0.1097
芽鞘色 CC	0.7023	0.7826	-1.3768	0.1698
幼苗叶色 LCS	0.6666	0.6478	0.7616	0.4471

*: 差异达到 0.05 显著水平, **: 差异达到 0.01 极显著水平
* shows that difference is significant at 0.05 level, ** shows that difference is significant at 0.01 level

2.3.3 所有样本和核心种质的主成分分析 所有样本和核心种质的表型性状经均值化后进行主成分分析(表6),结果显示二者的特征根、贡献率和累计贡献率均比较接近。特征根以1为基准,所有样本有

5个主成分,核心种质有6个,且核心种质第1个主成分的特征值、贡献率以及前6个主成分的累计贡献率都大于所有样本的对应值,说明核心种质能有效克服所有样本的遗传冗余,提高群体代表性。

表6 所有样本和核心种质的主成分分析

Table 6 The principal component analysis of all samples and core collection

主成分 Component	所有样本 All samples			核心种质 Core collection		
	特征根	贡献率(%)	累计贡献率(%)	特征根	贡献率(%)	累计贡献率(%)
	Eigen value	Variance	Cumulative proportion	Eigen value	Variance	Cumulative proportion
1	3.406	18.924	18.924	3.428	19.043	19.043
2	2.965	16.470	35.394	2.868	15.933	34.976
3	1.731	9.616	45.011	1.684	9.357	44.333
4	1.559	8.660	53.671	1.621	9.005	53.337
5	1.071	5.952	59.623	1.097	6.095	59.433
6	0.946	5.253	64.876	1.041	5.784	65.217

3 讨论

3.1 核心种质构建方法的选择

核心种质的构建可以利用表型性状^[18]、等位酶(同工酶)^[19]、DNA^[20]等不同水平的遗传变异数据进行研究,依据表型性状的遗传多样性进行核心种质构建既方便快捷又经济,徐宁等^[21]根据地理来源分组,利用12项农艺性状和3项营养品质性状数据进行聚类分析构建了小豆核心种质;代攀虹等^[22]根据品种类型分组,利用21个表型性状聚类分析构建了陆地棉核心种质;李洪果等^[23-24]利用不分组的方法,从取样比例、取样方法、遗传距离和系统聚类4个方面构建了杜仲资源的核心种质。

山西是高粱的种植大省,分为春播早熟区和春播晚熟区,多样的生态环境丰富了高粱品种遗传的多样性,但不同地理居群之间存在互相的种质渗透^[25],因此,本研究采用不分组的核心种质构建方法,以高粱18个农艺表型性状对材料进行聚类分组,所构建的初选核心种质与所有样本的均值差异百分率为0、方差差异百分率为83.33%、极差符合率为97.45%、变异系数变化率为119.63%,说明运用此方法构建核心种质较好的保留了所有样本的表型变异。

3.2 核心种质评价手段的确定

核心种质是通过多方优先评价,在众多的资源中筛选出的具有代表性、有研究利用价值的优良种质,因此选择最优的评价手段极其重要^[26-27]。李自超等^[28]利用方差、平均值、变幅、标准差等检测连续性指标,运用Shannon多样性信息指数等检测离散型指标,创造性地提出了评价核心种质遗传多样性

的代表度。郝晓鹏等^[29]利用表型保留比、表型方差、变异系数等进行山西普通菜豆的核心种质评价。本文构建的核心种质与所有样本各性状的均值和多样性指数进行了显著性检验,同时,利用主成分分析对所构建核心种质进行确认,减少了性状间相关造成的遗传冗余,从遗传变异和遗传结构上评价了所得核心种质的代表性和实用性,更有说服力。

参考文献

- [1] 胡晋,徐海明,朱军. 基因型值多次聚类法构建作物种质资源核心库. 生物数学学报, 2000, 15(1): 103-109
Hu J, Xu H M, Zhu J. Constructing core collection of crop germplasm by multiple clusters based on genotypic values. Journal of Biomath Matics, 2000, 15(1): 103-109
- [2] Frankel O H, Brown A H D. Plant genetic resources today: a critical appraisal//Holden J H W, Williams J T. Crop genetic resources: Conservation and Evaluation. London: George Allen and Unwin Limitid, 1984: 249-257
- [3] Chandra S, Huaman Z, Krishna S H, Ortiz R. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data simulation study. Theoretical and Applied Genetics, 2002, 104: 1325-1334
- [4] Li Z C, Zhang H L, Zeng Y W, Yang Z Y, Shen S Q, Sun C Q, Wang X K. Studies on sampling schemes for the establishment of core collection of rice landraces in Yunnan, China. Genetic Resources and Crop Evolution, 2002, 49: 67-74
- [5] Zhao L M, Dong Y S, Liu B, Hao S, Wang K J, Li X H. Establishment of a core collection for Chinese annual wild soybean (*Glycine soja*). Chinese Science Bulletin, 2005, 50(10): 989-996
- [6] Balfourier F, Roussel V, Strelchenko P, Exbrayat-Vinson F, Sourdille P, Boutet G, Koenig J, Ravel C, Mitrofanova O, Beckert M, Charmet G. A worldwide bread wheat core collection arrayed in a 384-well plate. Theoretical and Applied Genetics, 2007, 114(7): 1265-1275
- [7] 董玉琛,曹永生,张学勇,刘三才,王兰芬,游光霞,庞斌双,李

- 立会,贾继增. 中国普通小麦初选核心种质的产生. 植物遗传资源学报, 2003, 4(1): 1-8
- Dong Y C, Cao Y S, Zhang X Y, Liu S C, Wang L F, You G X, Pang B S, Li L H, Jia J Z. Establishment of candidate core collection in chinese common wheat germplasm. 2003, 4(1): 1-8
- [8] Malvar R A, Butron R A, Alvarez A, Padilla G, Carrea M E, Revilla P, Ordás A. Yield performance of the european union maize landrace core collection under multiple corn borer infestations. Crop Protection, 2007, 26: 775-781
- [9] Xu H, Mei Y, Hu J, Zhu J, Gong P. Sampling a core collection of island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits. Genetic Resources and Crop Evolution, 2007, 53(3): 515-521
- [10] Porter L D, Hoheisel G, Coffman V A. Resistance of peas to *Sclerotinia sclerotiorum* in the *Pisum* core collection. Plant Pathology, 2009, 58: 52-60
- [11] Wang X C, Chen L, Yang Y J. Establishment of core collection for Chinese tea germplasm based on cultivated region grouping and phenotypic data. Frontiers of Agriculture in China, 2011, 5(3): 344-350
- [12] Grenier C, Bramel-Cox P J, Noirot M, Prasada Rao K E, Hamon P. Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs non-random sampling procedures. Theoretical and Applied Genetics, 2000, 101: 190-196
- [13] Tariq S, Hisato O, Makoto K, Okuno K. Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. Genetic Resources and Crop Evolution, 2009, 56: 809-827
- [14] Dahlberg J A, Burke J J, Rosenow D T. Development of a sorghum core collection: refinement and evaluation of a subset from sudan. Economic Botany, 2004, 58: 556-567
- [15] 陆平. 高粱种质资源描述规范和数据标准. 北京: 中国农业出版社, 2006
- Lu P. Descriptors and Data Standard for Sorghum [*Sorghum bicolor* (L.) Moench]. Beijing: Chinese Agricultural Publishers, 2006
- [16] 徐海明, 胡晋, 朱军. 构建作物种质资源核心库的一种有效抽样方法. 作物学报, 2000, 26(2): 157-162
- Xu H M, Hu J, Zhu J. An efficient method of sampling core collection from crop germplasm. Acta Agronomica Sinica, 2000, 26(2): 157-162
- [17] 胡兴雨, 王纶, 张宗文, 陆平, 张红生. 中国黍稷核心种质的构建. 中国农业科学, 2008, 41(11): 3489-3502
- Hu X Y, Wang L, Zhang Z W, Lu P, Zhang H S. Establishment of broomcorn millet core collection in China. Scientia Agricultura Sinica, 2008, 41(11): 3489-3502
- [18] Malosetti M, Abadie T. Sampling strategy to develop a core collection of Uruguayan maize landraces based on morphological traits. Genetic Resources and Crop Evolution, 2001, 48: 381-390
- [19] Geraci A, Chèvre A, Divaret I, Eber F, Raimondo F. Isozyme analysis of genetic diversity in wild Sicilian populations of *Brassica* sect. *Brassica* in view of genetic resources management. Genetic Resources and Crop Evolution, 2004, 51(2): 137-146
- [20] Fan Q L, Zou P, Xian S. Constructing a core collection for maize (*Zea mays* L.) landrace from wuling mountain region in China. Agricultural Sciences in China, 2008, 7(12): 1423-1432
- [21] 徐宁, 程须珍, 王素华, 王丽侠, 赵丹. 以地理来源分组和利用表型数据构建中国小豆核心种质. 作物学报, 2008, 34(8): 1366-1373
- Xu N, Cheng X Z, Wang S H, Wang L X, Zhao D. Establishment of an adzuki bean (*Vigna angularis*) core collection based on geographical distribution and phenotypic data in China. Acta Agronomica Sinica, 2008, 34(8): 1366-1373
- [22] 代攀虹, 孙君灵, 贾银华, 杜雄明, 王溢. 利用表型数据构建陆地棉核心种质. 植物遗传资源学报, 2016, 17(6): 961-968
- Dai P H, Sun J L, Jia Y H, Du X M, Wang M. Construction of core collection of upland cotton based on phenotypic data. Journal of Plant Genetic Resources, 2016, 17(6): 961-968
- [23] 李洪果, 杜庆鑫, 王淋, 杜红岩, 陈锡民. 利用表型数据构建杜仲雌株核心种质. 分子植物育种, 2017, 15(12): 5197-5209
- Li H G, Du Q X, Wang L, Du H Y, Chen X M. Establishment of female core collection of *Eucommia ulmoides* oliv. Based on phenotypic characters. Molecular Plant Breeding, 2017, 15(12): 5197-5209
- [24] 李洪果, 杜红岩, 贾宏炎, 湛红辉, 许基煌, 杜庆鑫, 王璐. 利用表型性状构建杜仲雄性资源核心种质. 分子植物育种, 2018, 16(2): 591-601
- Li H G, Du H Y, Jia H Y, Chen H H, Xu J H, Du Q X, Wang L. Establishment of male core collection of *eucommia ulmoides* based on phenotypic traits. Molecular Plant Breeding, 2018, 16(2): 591-601
- [25] 李萌, 秦慧彬, 王宇楠, 闫建俊, 穆志新. 基于 SSR 标记的山西高粱地方品种遗传多样性和遗传结构分析. 分子植物育种, 2020, 18(18): 6068-6076
- Li M, Qin H B, Wang Y N, Yan J J, Mu Z X. Genetic diversity and structure of sorghum landrace germplasm in Shanxi based on SSR markers. Molecular Plant Breeding, 2020, 18(18): 6068-6076
- [26] 张春雨. 新疆野苹果 (*Malus sieversii*) 群体遗传结构与核心种质构建方法. 泰安: 山东农业大学, 2008
- Zhang C Y. Population genetic structure and method of constructing core collection for *Malus sieversii*. Tai'an: Shandong Agricultural University, 2008
- [27] Odong T L, Jansen J, Eeuwijk F A, Van Hintum T J L. Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. Theoretical and Applied Genetics, 2013, 126(2): 289-305
- [28] 李自超, 张洪亮, 孙传清, 王象坤. 植物遗传资源核心种质研究现状与展望. 中国农业大学学报, 1999, 4(5): 51-62
- Li Z C, Zhang H L, Sun C Q, Wang X K. Status and prospects of core collection in plant germplasm resource. Journal of China Agricultural University, 1999, 4(5): 51-62
- [29] 郝晓鹏, 王燕, 田翔, 郜欣, 畅建武. 基于农艺性状的山西普通菜豆初级核心种质构建. 植物遗传资源学报, 2016, 17(5): 815-823
- Hao X P, Wang Y, Tian X, Hao X, Chang J W. Construction of primary core collection of common bean (*Phaseolus vulgaris* L.) based on agronomic traits in Shanxi province. Journal of Plant Genetic Resources, 2016, 17(5): 815-823