

整合农艺性状和分子标记数据构建新疆野苹果核心种质

刘遵春^{1,2}, 刘大亮¹, 崔美¹, 李敏¹, 焦其庆¹, 高利平¹, 陈学森^{1,*}

(¹ 山东农业大学作物生物学国家重点实验室, 山东泰安 271018; ² 河南科技学院园艺园林学院, 河南新乡 453003)

摘要: 以新疆野苹果 [*Malus sieversii* (Ledeb.) Roem.] 60 份初级核心种质为试材, 采用不同遗传距离构建新疆野苹果核心种质, 利用数量性状参数均值差异百分率 (*MD*)、方差差异百分率 (*VD*)、极差符合率 (*CR*)、变异系数变化率 (*VR*) 和分子标记参数多态位点百分率 (*p*)、平均有效等位基因数 (*M_{Ne}*)、平均多态信息含量 (*M_{PI}*) 和平均 Shannon's 信息指数 (*M_I*) 等 8 个指标评价不同方法构建核心种质的优劣, 用选出的合适方法构建新疆野苹果核心种质。研究结果表明: 利用混合遗传距离 (*D_{mix}*) 构建的核心种质整合了两类不同数据, 优于单独使用农艺性状表型值数据或分子标记数据构建的核心种质, 其 *VD* 为 50%, *CR* 为 96.56%, *VR* 为 117.98%, *p* 为 97.67%, *M_{Ne}* 为 97.15%, *M_{PI}* 为 101.88%, *M_I* 为 100.44%; 利用 15 个农艺性状检测表明, 所构建的 42 份新疆野苹果核心种质保留了 300 份原始种质 93% 以上的农艺性状, 很好地代表了原始种质的遗传多样性。

关键词: 新疆野苹果; 核心种质; 农艺性状; 分子标记

中图分类号: S 661.1

文献标识码: A

文章编号: 0513-353X (2012) 06-1045-10

Combining Agronomic Traits and Molecular Marker Data for Constructing *Malus sieversii* Core Collection

LIU Zun-chun^{1,2}, LIU Da-liang¹, CUI Mei¹, LI Min¹, JIAO Qi-qing¹, GAO Li-ping¹, and CHEN Xue-sen^{1,*}

(¹ State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an, Shandong 271018, China; ² College of Horticulture Landscape Architecture, Henan Institution of Science and Technology, Xinxiang, He'nan 453003, China)

Abstract: To construct *Malus sieversii* (Ledeb.) Roem. core collection, based on 15 agronomic traits data and molecular marker from *M. sieversii* primary core collection, a series of subsets were sampled at 50% proportion by different genetic distances, respectively. The genetic variation among collections was compared by evaluating the *MD* (mean difference percentage), *VD* (variance difference percentage), *CR* (coincidence rate of range), *VR* (changeable rate of coefficient of variation) of agronomic traits and the *p* (percentage of polymorphic loci), *M_{Ne}* (average effective number of alleles), *M_I* (average Shannon's information index) and *M_{PI}* (average polymorphism information content) of molecular marker information. The results showed that the core collection constructed on the basis of a combination

收稿日期: 2012-03-13; **修回日期:** 2012-05-25

基金项目: 国家自然科学基金项目 (30871679); 国家“863”计划重点项目 (2006AA100108); 山东省农业良种工程项目

* 通信作者 Author for correspondence (E-mail: chenxs@sda.edu.cn; Tel: 0538-8249338)

of data for agrinomic traits and molecular marker data is more representative than that constructed on the basis of data from agrinomic traits or molecular marker data alone. The value of VD , CR , VR , p , M_{Ne} , M_{PIC} and M_I is 50%, 96.56%, 117.98%, 97.67%, 97.15%, 101.88% and 100.44%, respectively. The mixed genetic distance (D_{mix}) is more preferable to construct core collection. The core collection was evaluated by using 15 agrinomic traits, and the data showed that it retained more than 93% agrinomic traits and could highly represent the genetic diversity of the initial collection of 300 *Malus sieversii* accessions.

Key words: *Malus sieversii*; core collection; agronomic trait; molecular marker

种质资源的收集和保存是培育高质量新品种的物质基础。中国是世界苹果遗传多样性中心和起源中心。新疆野苹果 [*Malus sieversii* (Ledeb.) Roem.] 是现代苹果遗传育种的宝贵基因库, 但数量庞大的新疆野苹果资源也为如何合理保存、评价、研究及利用带来了很大的负担。目前核心种质已经成为国际遗传资源研究的热点, 并已在许多一年生作物上得到广泛应用。对于多年生木本植物的核心种质构建研究, 目前在苹果 (张春雨 等, 2009)、桃 (李天红, 2005; 李银霞 等, 2006, 2007)、枣 (白瑞霞, 2008)、果梅 (高志红 等, 2005)、柚类 (刘勇 等, 2006) 和山葡萄 (刘闯萍 等, 2008; 吴子龙 等, 2011) 等果树上有一些报道。

核心种质构建所用的数据包括形态农艺性状数据和分子标记数据两种类型。形态农艺性状数据是核心种质构建的主要数据类型, 它能最直观地反映种质特性, 也是检测其他类型数据真实性的重要参照。李银霞等 (2006) 以国家种质资源圃 (北京) 编目的 558 份桃品种的 18 项形态学和农艺学性状为基本数据构建了桃的初级核心种质。白瑞霞 (2008) 以 170 份样品的 23 项数量性状和 11 项质量性状作为基本数据构建了枣的核心种质。刘闯萍等 (2008) 采用来源省份分组—花型分组, 再以果实性状为主的数据进行聚类构建了山葡萄的核心种质。

由于形态农艺性状数据易受环境和人为因素的影响, 存在一定的局限性。而随着分子标记的出现及其技术的日臻完善, 分子标记数据成为果树核心种质研究中重要的数据来源之一。刘勇等 (2006) 根据 678 个 SSR 和 AFLP 分子标记聚类结果, 对 110 份柚类资源采用逐级压缩法构建了柚类的核心种质。张春雨等 (2009) 以 109 个新疆野苹果实生株系的 128 个 SSR 位点为材料, 研究了新疆野苹果核心种质的构建方法。但分子标记不等同于基因, 不能反映控制性状变异基因的差异, 其仍存在一定的局限性。随着研究的进一步深入发现, 整合形态农艺性状和分子标记等各类数据能更准确地检测个体间的遗传差异, 提高核心种质遗传代表性。高志红等 (2005) 运用形态特征、农艺性状并参考同工酶、RAPD 和 SSR 的结果, 构建了果梅的核心种质。李银霞等 (2007) 利用果实性状和 SSR 数据对桃初级核心种质进一步聚类取样, 最终抽取 45 份桃核心种质。

笔者曾利用 15 个形态农艺性状数据对新疆伊犁地区巩留县莫合乡的 300 份新疆野苹果原始种质进行初级核心种质构建研究, 获得了 60 份初级核心样品 (刘遵春 等, 2010)。本研究中以这 60 份新疆野苹果初级核心种质为材料, 利用形态农艺性状数据和 SSR 分子标记数据对其进一步聚类压缩, 确定构建新疆野苹果核心种质的最佳取样方案, 依此方案构建了新疆野苹果核心种质并对其代表性进行评价。

1 材料与方法

1.1 材料

本试验材料来自新疆伊犁地区的巩留县莫合乡的新疆野苹果林, 选取立地条件基本一致并具有

代表性的 300 个株系, 并进行 GPS 定位。

于 2006—2007 年连续两年对新疆野苹果实生株系的叶片、花朵和果实进行采集, 并在山东农业大学果树生物学实验室对 15 个性状进行了测定。于 4 月下旬每株随机采集 20 朵完全开放的花朵, 立地用游标卡尺测量花冠宽度、花柄长度、花瓣长度和花瓣宽度, 8 月下旬—9 月上旬每株随机采集 50 片发育正常的叶片和 20 个果实, 空运回实验室, 对其性状进行测定。第 2 年在相同株系上采集作为重复。

用于 SSR 分析的材料于 2007 年 4 月进行采集, 分别取其幼叶立即放入装有干燥硅胶的自封塑料袋中干燥, 备用 (何天明 等, 2004)。

1.2 新疆野苹果表型农艺形态性状分析

根据冯涛等 (2006) 的方法对叶片长度、叶片宽度、叶片厚度、叶柄长度、花冠宽度、花柄长度、花瓣长度、花瓣宽度、果实总糖、可溶性固形物 (TSS) 含量、单果质量、果实纵径、果实横径、果形指数以及钙元素含量进行测定。

1.3 新疆野苹果 SSR 分子标记分析

采用改良 CTAB 法提取新疆野苹果基因组 DNA (Doyle & Doyle, 1990), SSR-PCR 扩增及 SSR 引物信息参考 Zhang 等 (2007) 的方法。

1.4 构建核心种质

1.4.1 混合遗传距离的计算

参考 Wang 等 (2007) 的方法, 提出了农艺性状表型值遗传距离 (D_p) 的计算方法。第 i 个样品与第 j 个样品之间的农艺性状表型值遗传距离 (D_{pij}) 可用下列公式计算:

$$D_{pij} = \frac{1}{m} \sum_{k=1}^m \frac{|p_{ki} - p_{kj}|}{R_k},$$

其中 p_{ki} 和 p_{kj} 表示第 k 个性状下的两个种质的农艺性状表型值, R_k 表示第 k 个性状下的所有农艺性状表型值的极差, m 表示性状总数。

对于分子标记, 提出了分子标记遗传距离 (D_m) 的计算方法。第 i 个样品与第 j 个样品之间的分子标记遗传距离 (D_{mij}) 可表示为:

$$D_{mij} = \frac{1}{n} \sum_{k=1}^n S_k,$$

其中 S_k 表示两个样品在第 k 个分子标记位点上的表现状况, 即当两个样品在第 k 个分子标记位点上表现一致, 则 $S_k = 0$, 否则 $S_k = 1$ 。 n 表示分子标记位点总数。因此, 第 i 个样品与第 j 个样品之间的混合遗传距离 (D_{mixij}) 为: $D_{mixij} = D_{pij} + D_{mij}$ 。

1.4.2 遗传距离的筛选

本试验以 50% 取样比例, 参照 Hu 等 (2000) 的逐步聚类随机取样法, 使用不加权类平均聚类法 (unweighted pair-group average method) 结合一系列遗传距离构建核心种质。

3 种遗传距离用于表型值数据, 包括欧氏距离 (Euclidean distance, Euclid)、马氏距离 (Mahalanobis distance, Mahal) 和提出的农艺性状表型值遗传距离 (D_p)。

4 种遗传距离用于分子标记数据, 包括 Sokal-Michener 距离、Nei-Li 距离、Jaccard 距离和提出的分子标记遗传距离 (D_m); 1 种用于表型值和分子标记的混合遗传距离 (D_{mix})。

通过综合比较从中筛选出最佳的遗传距离。

1.4.3 取样比例的筛选

采用以上筛选出的最佳遗传距离,按照9种不同的取样比例(10%、20%、30%、40%、50%、60%、70%、80%和90%)从60株中抽取9个核心种质,根据各自的遗传多样性保留量,从中筛选出最佳取样比例。

1.4.4 核心种质的评价

共选择8个评价参数,包括前人(Hu et al., 2000)提出和本研究中提出的。评价数量性状的参数为: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)和变异系数变化率(changeable rate of coefficient of variation, VR)。评价分子标记信息的参数为:多态位点百分率(percentage of polymorphic loci, p)、平均有效等位基因数(average effective number of alleles, M_{Ne})、平均多态信息含量(average polymorphism information content, M_{PIC})和平均 Shannon's 信息指数(average Shannon's information index, M_I)。

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参照刘遵春等(2010)的计算公式计算。所有评价数量性状参数的计算均是基于未标准化的群体,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。

1.4.5 核心种质的确认

采用最佳构建策略构建新疆野苹果核心种质,并采用主成分分析法对构建的核心种质进行确认。核心种质和初级核心种质的分布情况由最初的两个主坐标轴确定。

1.4.6 核心种质对原始种质代表性的检验

核心种质的有效性检验是通过比较300份原始种质和核心种质15个数量性状,即总糖、可溶性固形物、单果质量、果实横径、果实纵径、果形指数、钙含量、叶片宽度、叶片长度、叶片厚度、叶柄长度、花冠宽度、花柄长度、花瓣长度和花瓣宽度的变异幅度、极差、平均值、标准误差、变异系数以及保留比例,检验核心种质的代表性。

1.5 数据处理

数据标准化、遗传距离的计算、聚类分析、核心种质的构建和核心种质评价参数的计算均在MATLAB环境中编程实现。

2 结果与分析

2.1 农艺性状表型值遗传距离与常用遗传距离的比较

在50%取样比例下,基于表型值采用两种常用遗传距离Euclid和Mahal,以及本研究中提出的 D_p 遗传距离构建的核心种质,结果表明其 MD 均为0,而 CR 均大于80%(表1)。因此,采用3种遗传距离构建的核心种质都能够代表原始群体的遗传多样性。

采用 D_p 遗传距离构建的核心种质具有最大值的 VD 、 CR 和 VR ,与Mahal和Euclid均达到显著水平。在参数 p 、 M_{Ne} 、 M_{PIC} 和 M_I 上,采用 D_p 遗传距离和Mahal距离构建的核心种质差异不显著,与Euclid距离构建的核心种质数值差异达到显著水平。因此, D_p 遗传距离适用于构建的核心种质,其代表性好于Mahal和Euclid两种遗传距离。

表 1 50%取样比例下基于表型值数据 3 种遗传距离构建核心子集 8 个参数的比较
Table 1 Comparison of three genetic distances for phenotypic values within eight evaluating parameters at sampling percentages of 50%

遗传距离种类 Genetic distance type	MD (%)	VD (%)	CR (%)	VR (%)	p (%)	M _{Ne} (%)	M _{PIC} (%)	M _I (%)
Euclid	0	30 c	87.06 c	99.74 c	94.07 b	95.85 a	90.3 b	85.9 b
Mahal	0	50 b	95.67 b	115.69 b	94.81 a	95.10 b	95.6 a	93.5 a
D _p	0	60 a	97.12 a	120.31 a	95.06 a	95.25 b	96.8 a	95.9 a

注：MD (%)：核心种质与初级核心种质的均值差异百分率；VD (%)：核心种质与初级核心种质的方差差异百分率；CR (%)：核心种质与初级核心种质的极差符合率；VR (%)：核心种质与初级核心种质的变异系数变化率；p (%)，M_{Ne} (%)，M_{PIC} (%) 和 M_I (%) 为 p，M_{Ne}，M_{PIC} 和 M_I 相对于初级核心种质的百分率，不同字母表示差异达到显著水平 (P < 0.05)。下同。

Note: MD(%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for mean difference percentage of traits; VD(%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for variance difference percentage of traits; CR (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for coincidence rate of range of traits; VR (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for changeable rate of the coefficient of variation of traits; p (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for percentage of polymorphic loci; M_{Ne} (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for average effective number of alleles; M_{PIC} (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for average polymorphism information content; M_I (%). Percentage of significant difference (α = 0.05) between core collection and the primary collection for average Shannon's information index. Data followed by different letters are significantly different (P < 0.05). The same below.

2.2 分子标记遗传距离与常用遗传距离的比较

在 50%取样比例下，对于 3 种分子标记数据遗传距离 Jaccard、Nei-Li、Sokal-Michener 和本研究中提出的 D_m 遗传距离构建的核心种质，其 MD 均小于 20%，而 CR 均大于 80% (表 2)。因此，所有 4 种遗传距离构建的核心种质都能够代表原始群体的遗传多样性。

与 Jaccard、Nei-Li、Sokal-Michener 遗传距离构建的核心种质相比较，由 D_m 遗传距离构建的核心种质具有相同的 MD 值；其 VD 值大于 Nei-Li 遗传距离，差异显著；其 CR、VR、p、M_{Ne}、M_{PIC} 和 M_I 值大于 Jaccard 和 Nei-Li 遗传距离，差异显著，小于 Sokal-Michener 遗传距离，在评价参数 VR 和 M_{Ne} 上差异达到显著水平。

因此，本研究中提出的 D_m 遗传距离适宜于构建核心种质的研究，其代表性次于 Sokal-Michener 遗传距离，好于 Jaccard 和 Nei-Li 两种遗传距离。

表 2 50%取样比例下基于分子标记数据 4 种遗传距离构建核心种质 8 个参数的比较
Table 2 Comparison of four genetic distances for molecular marker information within eight evaluating parameters at sampling percentages of 50%

遗传距离种类 Genetic distance type	MD (%)	VD (%)	CR (%)	VR (%)	p (%)	M _{Ne} (%)	M _{PIC} (%)	M _I (%)
Jaccard	0	20 a	81.36 b	92.14 c	97.39 b	100.30 c	101.75 b	98.31 b
Nei-Li	0	10 b	81.16 b	92.07 c	97.32 b	100.25 c	100.10 c	96.50 c
Sokal-Michener	0	20 a	82.12 a	93.31 a	98.31 a	100.88 a	102.43 a	101.68 a
D _m	0	20 a	82.02 a	92.45 b	97.58 a	100.48 b	102.03 a	100.93 a

2.3 农艺性状表型值遗传距离 (D_p)、分子标记遗传距离 (D_m) 和混合遗传距离 (D_{mix}) 的比较

在 50%取样比例下，通过 3 种遗传距离所构建核心种质的 MD 均为 0，CR 均大于 80%，说明 3 种遗传距离构建的核心种质都能够代表初级核心种质 (表 3)。由 D_{mix} 混合遗传距离构建的核心种质由于整合了两类不同数据，与以上两种核心种质相比较，在表型值评价参数 VD、CR 和 VR 上与由

D_m 遗传距离构建的核心种质差异显著, 与 D_p 遗传距离差异不显著, 在分子标记信息评价参数 p 、 M_{Ne} 、 M_{PIC} 和 M_I 上与由 D_p 遗传距离构建的核心种质差异显著, 与 D_m 遗传距离差异不显著。由 D_p 遗传距离或 D_m 遗传距离构建的核心种质均在一类数据评价参数上表现良好, 而在另一类数据上表现较差, 差异达到显著水平。因此, 采用本研究的整合方法构建核心种质比单独使用表型值或分子标记构建的核心种质更加可靠。

表 3 50%取样比例下基于表型值遗传距离、分子标记遗传距离和混合遗传距离在 8 个参数上的比较
Table 3 Comparison of the phenotypic values distance, molecular marker distance, and mixed distance within eight evaluating parameters at sampling percentages of 50%

遗传距离种类 Genetic distance type	MD (%)	VD (%)	CR (%)	VR (%)	p (%)	M_{Ne} (%)	M_{PIC} (%)	M_I (%)
D_p	0	60 a	97.12 a	120.31 a	95.83 b	95.06 b	96.80 b	95.90 b
D_m	0	20 b	82.02 b	92.45 b	98.12 a	97.58 a	102.03 a	100.93 a
D_{mix}	0	50 a	96.56 a	117.98 a	97.67 a	97.15 a	101.88 a	100.44 a

2.4 取样比例的确定

采用混合遗传距离, 利用逐步聚类随机取样法取样, 按照 10%、20%、30%、40%、50%、60%、70%、80% 和 90% 的比例分别从 60 份初级核心种质中抽取 6、12、18、24、30、36、42、48 和 54 份, 构建了 9 个核心种质, 分别记为 CorePM1、CorePM2、CorePM3、CorePM4、CorePM5、CorePM6、CorePM7、CorePM8 和 CorePM9 (表 4)。

从表 4 可知, 随着抽样比例的降低, 各核心种质的 CR 、 p 和 M_{Ne} 逐渐减小, 在 20% 取样比例时, CR 值已低于 80%。相反随着抽样比例的降低, 各核心种质的 MD 、 VD 和 VR 逐渐增大, 在 20% 取样比例时, MD 值已经高于 20%。随着抽样比例的降低, M_{PIC} 和 M_I 参数值呈现出先增大后减小的趋势, 在 70% 取样比例时, 两个参数值均达到最大, 分别为 103.052% 和 102.058%。因此, 20% ~ 70% 是抽取核心种质的适宜比例, 针对到初级核心种质的样品数量, 70% 作为抽取核心种质的最适宜取样比例, 此时 CR 和 M_{PIC} 已分别达到初级核心种质的 98.4% 和 103.052%, 可以认为很好地保存了初级核心种质的遗传多样性。核心种质库 CorePM7 共有 42 份材料组成, 其余 18 份剩余材料可作为新疆野苹果的备用种质。

表 4 9 种取样比例下核心种质与初级核心种质在 8 个评价参数上的差异
Table 4 Comparison of the mixed distance within eight evaluating parameters at 9 sampling percentages

取样比例 Sampling percentage	代号 Code	MD (%)	VD (%)	CR (%)	VR (%)	p (%)	M_{Ne} (%)	M_{PIC} (%)	M_I (%)
90%	CorePM9	0	10	98.40	105.91	99.57	99.18	97.40	100.39
80%	CorePM8	0	30	98.40	109.32	99.13	98.67	100.84	101.42
70%	CorePM7	0	30	98.40	113.53	98.83	98.09	103.05	102.06
60%	CorePM6	0	40	96.40	116.65	98.56	97.82	102.80	101.82
50%	CorePM5	0	50	95.12	117.98	97.67	97.15	101.88	100.44
40%	CorePM4	10	50	86.86	118.89	94.67	95.73	98.39	98.69
30%	CorePM3	10	50	80.44	120.31	93.19	93.78	94.92	95.22
20%	CorePM2	20	70	77.64	126.04	90.28	91.52	92.33	92.28
10%	CorePM1	40	70	62.50	128.75	83.14	88.67	83.02	83.06

2.5 核心种质对初级核心种质代表性的确认

利用主成分分析法对所构建核心种质库的代表性进行确认 (图 1)。

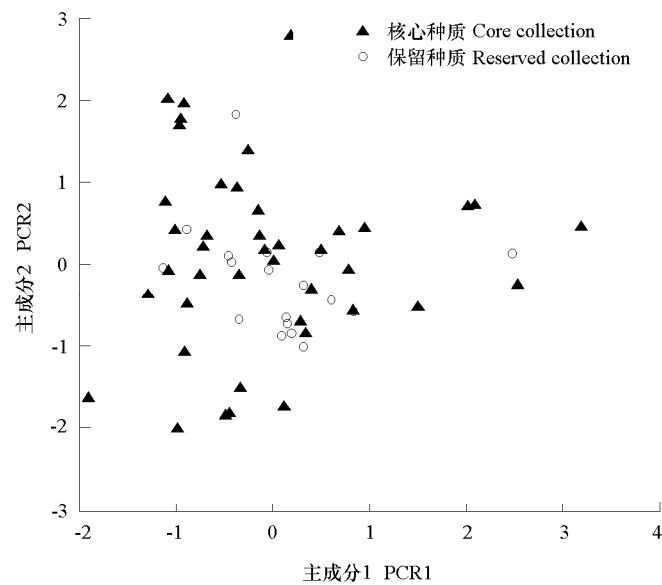


图 1 70%取样比例下核心种质和保留种质的主成分分析图

Fig. 1 Principal component plots for the core and reserved collections at 70% sampling proportion

图 1 清楚地显示出第一、二主成分分布大致代表了样品的分布，第一、二主成分对初级核心种质遗传变异的贡献率累计达到 69%。另一方面，样品分布的几何特征同时表明，核心种质库（CorePM7）的株系分布仍保存了初级核心种质分布的几何形状和特征，并且具有极值的外围个体均被选入到核心库中，遗传相似性高的株系中仅有一个株系入选为核心种质，进一步降低了核心种质的冗余，同时又确保了核心种质的代表性。由此可推断，核心种质库（CorePM7）很好地保存了初级核心种质的遗传多样性和结构，确保了核心种质库的实用性。

2.6 核心种质对原始种质代表性的评价

利用 15 个数量性状分别计算所构建的核心种质与 300 份原始种质的变异幅度、极差、平均值、标准误差、变异系数和保留比例，以检验核心种质对原始种质的代表性（表 5）。

表 5 核心种质的评价

Table 5 Evaluation of core collection

性状 Trait	原始群体 Original germplasms					核心种质 Core collection					保留比例 / % Reserved ratio
	变异幅度 Variation range	极差 Range	平均值 Mean	标准误差 Standard deviation	变异系 数 / % Variation coefficient	变异幅度 Variation range	极差 Range	平均值 Mean	标准误差 Standard deviation	变异系 数 / % Variation coefficient	
总糖 / (mmol · kg ⁻¹) Total acid	56.70 ~ 543.30	486.60	240.20	8.40	34.98	59.10 ~ 507.60	448.50	236.30	7.80	33.02	93.40
TSS / %	7.40 ~ 17.07	9.67	11.65	1.50	12.88	9.77 ~ 17.76	7.99	12.15	1.78	14.67	93.12
单果质量 / g Fruit weight	6.79 ~ 81.20	74.41	15.77	7.12	45.08	6.75 ~ 77.49	70.75	22.83	16.46	72.09	99.73
果实横径 / cm Vertical length	1.97 ~ 4.81	2.84	3.00	0.43	14.17	2.42 ~ 4.62	2.21	3.33	0.57	17.17	86.67

续表 5

性状 Trait	原始群体 Original germplasms					核心种质 Core collection					保留比例 / % Reserved ratio
	变异幅度 Variation range	极差 Range	平均值 Mean	标准误差 Standard deviation	变异系 数 / % Variation coefficient	变异幅度 Variation range	极差 Range	平均值 Mean	标准误差 Standard deviation	变异系 数 / % Variation coefficient	
果实纵径 / cm Transverse length	2.33 ~ 6.04	3.71	3.51	0.47	13.45	2.91 ~ 5.90	2.99	3.84	0.68	17.78	85.19
果形指数 Shape index	0.72 ~ 1.33	0.62	0.86	0.06	6.89	0.76 ~ 1.17	0.41	0.87	0.07	8.31	100.00
钙含量 / (mg · kg ⁻¹) Ca content	20.93 ~ 184.70	163.74	67.60	2.60	38.40	31.80 ~ 161.00	129.20	69.60	2.77	39.76	86.25
叶片厚度 / cm Leaf thickness	0.008 ~ 0.020	0.019	0.014	0.003	19.43	0.010~ 0.019	0.018	0.018	0.009	22.13	91.53
叶片长度 / cm Leaf length	4.95 ~ 8.54	3.59	6.30	0.82	13.00	4.95 ~ 8.33	3.38	6.75	0.97	15.40	96.31
叶片宽度 / cm Leaf width	2.68 ~ 5.03	2.35	3.72	0.46	12.41	2.82 ~ 5.03	2.21	3.94	0.55	13.61	92.65
叶柄长度 / cm Leaf stalk length	1.65 ~ 3.43	1.78	2.52	0.39	15.27	1.73 ~ 3.23	1.50	2.72	0.45	16.57	90.37
花冠宽度 / cm Anadem width	1.19 ~ 3.01	1.92	2.10	0.44	20.75	1.21 ~ 3.01	2.80	2.48	0.68	22.35	95.69
花柄长度 / cm Flower stalk length	3.81 ~ 5.54	1.73	4.54	0.42	9.32	3.93 ~ 5.34	1.41	5.24	0.47	9.85	97.63
花瓣长度 / cm Petal length	1.97 ~ 2.81	0.96	1.43	0.20	8.78	2.07 ~ 2.67	0.60	2.43	0.45	9.38	92.75
花瓣宽度 / cm Petal width	1.32 ~ 2.06	0.74	1.56	0.19	11.20	1.43 ~ 2.01	0.58	1.75	0.23	12.85	94.02

结果表明，核心种质除总糖的变异系数略低于原始种质外，其他性状的变异系数均高于原始种质，表明核心种质具有很好的异质性，在一定程度上剔除了遗传重复。

不同性状的保留比例不同，果形指数的保留比例最高为 100%，其次为单果质量、花柄长度、叶片长度、花冠宽度、花瓣宽度、总糖含量、可溶性固形物含量、花瓣长度、叶片宽度、叶片厚度和叶柄长度，分别为 99.73%、97.63%、96.31%、95.69%、94.02%、93.40%、93.12%、92.75%、92.65%、91.53%和 90.37%，果实横径、果实纵径和钙含量的保留比例较低，但仍高于 85%。15 个数量性状的平均保留比例为 93.02%，说明所构建的核心种质尽管有一些极值材料的丢失，但核心种质库（CorePM7）保留了原始种质 93% 以上的变异类型。可以认为通过以上方法所构建的核心种质能够很好地代表原始种质。

3 讨论

在种质资源研究中，要想较为全面的反映出群体的遗传多样性，必须综合各种不同类型的性状数据。这些数据不仅包含呈连续型变异的数量性状数据，也包含呈离散型变异的质量性状或分子标记数据。这两类数据具有不同的变异模式，各自所反映的遗传多样性信息量各不相同，遗传距离的计算方法及距离的变异范围也存在较大差异，因此对这两类数据不能只进行简单的合并。为充分利用这两类不同数据的遗传信息，Bar-Hen 等（1995）首次提出采用二阶段分析策略分析玉米自交系

材料间的遗传关系,其本质是首先采用分子标记信息计算自交系材料间的遗传距离,当遗传距离超过预先确定的阈值时则认为两者存在遗传差异。基于此再对这些存在遗传差异的材料用表型数据进行分析。由于它每一分析阶段仅用到一种类型的数据信息,因此,这种二阶段分析策略并不是真正意义上的信息整合。Noirot 等(1996)先对数量性状进行主成分分析,从 18 个主成分中仅选择了 6 个主成分计算个体间的遗传距离,这有利于排除一些未知因素对分析的干扰。Islam 等(2004)采用主成分分析整合数量性状和质量性状进行分析,其本质上是根据数量性状的均值和方差将数量性状转换成质量性状,然后进行统一分析。

由于受费用、工作量等因素的影响,直接对分子标记数据进行主成分分析构建核心种质的报道相对较少,分子标记数据主要用来研究核心种质的多样性、群体结构及冗余材料的鉴别等。Liu 等(2002)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 等(2006)对分子标记和数量性状数据分别进行主成分分析,将各个体的主成分值进行标准化,选出累积变异达到 85%的主成分计算个体间的欧式距离,构建了 2 321 份大豆种质资源的核心种质。Wang 和 Hu(2007)提出了整合数量性状数据和分子标记数据的混合遗传距离,该方法排除了难以处理共显性标记和不同性状量纲影响的问题。

果树核心种质构建研究起步较晚,许多理论和技术还不完善,大多学者意识到要综合运用多种数据构建核心种质的必要性,但终未见具体运用方法研究的报道。本研究首次尝试运用混合遗传距离整合数量性状数据和 SSR 分子标记数据来构建新疆野苹果核心种质。结果表明:本研究中提出的 D_p 遗传距离、 D_m 遗传距离和 D_{mix} 遗传距离均适用于核心种质的构建,但单一使用遗传距离 D_p 或遗传距离 D_m 构建的核心种质仅在一类数据评价参数上表现良好,而在另一类数据上表现不理想;而以混合遗传距离 D_{mix} 构建的核心种质由于其整合了两类不同数据信息,与以上两种核心种质相比较,其在表型值评价参数 VD 、 CR 和 VR 和分子标记信息评价参数 p 、 M_{Ne} 、 M_{PIC} 和 M_I 上均表现出较好的效果。因此,采用混合遗传距离 D_{mix} 构建核心种质比单独使用表型值或分子标记构建的核心种质更加可靠。

References

- Bai Rui-xia. 2008. Studies on genetic diversity and core collection construction of *Ziziphus jujuba* germplasm resources using AFLP and SRAP marker [Ph. D. Dissertation]. Baoding: Agricultural University of Hebei. (in Chinese)
- 白瑞霞. 2008. 枣种质资源遗传多样性的分子评价及其核心种质的构建 [博士论文]. 保定: 河北农业大学.
- Bar-Hen A, Charcosset A, Bourgoin M, Guiard J. 1995. Relationship between genetic markers and morphological traits in maize inbred lines collection. *Euphytica*, 84: 145 - 154.
- Doyle J J, Doyle J L. 1990. Isolation of plant DNA from fresh tissue. *Focus*, 12: 13 - 15.
- Feng Tao, Zhang Hong, Chen Xue-sen, Zhang Yan-min, He Tian-ming, Feng Jian-rong, Xu Zheng. 2006. Genetic diversity of fruit morphological traits and content of mineral element in *Malus sieversii* (Ldb.) Roem. and its elite seedings. *Journal of Plant Genetic Resources*, 7 (3): 270 - 276. (in Chinese)
- 冯涛, 张红, 陈学森, 张艳敏, 何天明, 冯建荣, 许正. 2006. 新疆野苹果果实形态与矿质元素含量多样性以及特异性状单株. *植物遗传资源学报*, 7 (3): 270 - 276.
- Gao Zhi-hong, Zhang Zhen, Han Zhen-hai, Fang Jing-gui. 2005. Development and evaluation of core collection of Japanese apricot germplasms in China. *Scientia Agricultura Sinica*, 38 (2): 363 - 368. (in Chinese)
- 高志红, 章镇, 韩振海, 房经贵. 2005. 中国梅核核心种质的构建与检测. *中国农业科学*, 38 (2): 363 - 368.
- He Tian-ming, Chen Xue-sen, Wu Yan. 2004. Preparation of DNA from silica gel dried leaves of Rosaceae. *Journal of Shihezi University: Natural Science*, 22 (4): 316 - 319. (in Chinese)
- 何天明, 陈学森, 吴燕. 2004. 从蔷薇科果树硅胶干燥叶片中制备 DNA. *石河子大学学报: 自然科学版*, 22 (4): 316 - 319.

- Hu J, Zhu J, Xu H M. 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theoretical and Applied Genetics*, 101: 264 – 268.
- Islam F M A, Beebe S, Muffoz M, Tohme J, Redden R J, Basford K E. 2004. Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool. *Theoretical and Applied Genetics*, 108: 243 – 252.
- Li Tian-hong. 2005. Studies on the core collection of peach [*P. persica* (L.) Batsch.] genetic resources. *Chinese Agricultural Science Bulletin*, 21 (8): 296 – 298. (in Chinese)
- 李天红. 2005. 桃遗传资源核心种质的研究. *中国农学通报*, 21 (8): 296 – 298.
- Li Yin-xia, An Li-jun, Jiang Quan, Zhao Jian-bo, Li Tian-hong. 2007. Establishment and evaluation of the core collection of peach [*Prunus persica* (L.) Batsch.] cultivars. *Journal of China Agricultural University*, 12 (5): 22 – 28. (in Chinese)
- 李银霞, 安丽君, 姜 全, 赵剑波, 李天红. 2007. 桃 [*Prunus persica* (L.) Batsch.] 品种核心种质的构建与评价. *中国农业大学学报*, 12 (5): 22 – 28.
- Li Yin-xia, Gao Qi-jie, Li Tian-hong. 2006. Sampling strategy based on fruit characteristics for a primary core collection of peach cultivars. *Journal of Fruit Science*, 23 (3): 359 – 364. (in Chinese)
- 李银霞, 高其洁, 李天红. 2006. 基于果实相关性状的桃品种初级核心种质取样策略研究. *果树学报*, 23 (3): 359 – 364.
- Liu Chuang-ping, Wang Jun, Shen Yu-jie, Lu Wen-peng. 2008. Preliminary construction of core collection in *Vitis amurensis*. *Journal of Plant Genetic Resources*, 9 (3): 372 – 374. (in Chinese)
- 刘闯萍, 王 军, 沈育杰, 路文鹏. 2008. 山葡萄 (*Vitis amurensis*) 资源核心种质的初步构建. *植物遗传资源学报*, 9 (3): 372 – 374.
- Liu Fang, Sun Guo, Salomon B, Bothmer R V. 2002. Characterization of genetic diversity in core collection accessions of wild barley, *Hordeum vulgare* ssp. *spontaneum*. *Hereditas*, 136: 67 – 73.
- Liu Yong, Sun Zhong-hai, Liu De-chun, Wu Bo, Zhou Qun. 2006. Screening the core collection of pomelo germplasm based on molecular marker. *Journal of Fruit Science*, 23 (3): 339 – 345. (in Chinese)
- 刘 勇, 孙中海, 刘德春, 吴 波, 周 群. 2006. 利用分子标记技术选择柚类核心种质资源. *果树学报*, 23 (3): 339 – 345.
- Liu Zun-chun, Zhang Chun-yu, Zhang Yan-min, Zhang Xiao-yan, Wu Chuan-jin, Wang Hai-bo, Shi Jun, Chen Xue-sen. 2010. Study on method of constructing core collection based on quantitative traits for *Malus sieversii*. *Scientia Agricultura Sinica*, 43 (2): 358 – 370. (in Chinese)
- 刘遵春, 张春雨, 张艳敏, 张小燕, 吴传金, 王海波, 石 俊, 陈学森. 2010. 利用数量性状构建新疆野苹果核心种质方法研究. *中国农业科学*, 43 (2): 358 – 370.
- Noirot M, Hamon S, Anthony F. 1996. The Principal component scoring: A new method of constiutting a core collection using quantitative data. *Genetic Resoueres and Crop Evolution*, 43: 1 – 6.
- Wang Jian-cheng, Hu Jin. 2007. Assessment of different genetic distances in constructing cotton core subset by genotypic values. *Journal of Zhejiang University*, 9 (5): 356 – 362.
- Wang Jian-cheng, Hu Jin, Xu Hai-ming. 2007. Strategy on constructing core collections by least distance stepwise sampling. *Theoretical and Applied Genetics*, 115 (1): 1 – 8.
- Wu Zi-long, Wang Jun, Shen Yu-jie, Lu Wen-peng. 2011. Pilot study on establishment of core germplasm of *Vitis amurensis* by different grouping method. *Northern Horticulture*, (7): 117 – 122. (in Chinese)
- 吴子龙, 王 军, 沈育杰, 路文鹏. 2011. 不同分组取样方法初步构建山葡萄核心种质的研究. *北方园艺*, (7): 117 – 122.
- Xu Hai-ming, Hu Jin, Zhu Jun, Gong P. 2006. Sampling a core collection of Island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits. *Genetic Resources and Crop Evolution*, 53: 515 – 521.
- Zhang Chun-yun, Chen Xue-sen, He Tian-ming, Feng Tao, Yuan Zhao-he. 2007. Genetic structure of *Malus sieversii* population from Xinjiang, China, revealed by SSR markers. *Journal of Genetics and Genomics*, 34 (10): 947 – 955.
- Zhang Chun-yu, Chen Xue-sen, Zhang Yan-min, Yuan Zhao-he, Liu Zun-chun, Wang Yan-ling, Lin Qun. 2009. A method for constructing core collection of *Malus sieversii* using molecular markers. *Scientia Agricultura Sinica*, 42 (2): 597 – 604. (in Chinese)
- 张春雨, 陈学森, 张艳敏, 苑兆和, 刘遵春, 王延龄, 林 群. 2009. 采用分子标记构建新疆野苹果核心种质的方法. *中国农业科学*, 42 (2): 597 – 604.