

DOI: 10.11829/j.issn.1001-0629.2019-0435

王英芳, 张业猛, 刘德梅, 沈迎芳, 王海庆. 基于转录组测序的青藏扁蓿豆 EST-SSR 标记开发与验证. 草业科学, 2020, 37(4): 718-727.

WANG Y F, ZHANG Y M, LIU D M, SHEN Y F, WANG H Q. Development and verification of EST-SSR markers in *Medicago archiducis-nicolai* by transcriptome sequencing. Pratacultural Science, 2020, 37(4): 718-727.



基于转录组测序的青藏扁蓿豆 EST-SSR 标记开发与验证

王英芳^{1,2}, 张业猛^{1,2}, 刘德梅^{1,3}, 沈迎芳^{1,4}, 王海庆^{1,3}

(1. 中国科学院高原生物适应与进化重点实验室 / 中国科学院西北高原生物研究所, 青海 西宁 810001;

2. 中国科学院大学, 北京 100049; 3. 青海省作物分子育种重点实验室, 青海 西宁 810001;

4. 中国科学院三江源国家公园研究院, 青海 西宁 810001)

摘要: 青藏扁蓿豆 (*Medicago archiducis-nicolai*) 广泛分布于青藏高原及其毗邻地区, 对高海拔极端环境具有极强的适应性。本研究利用高通量转录组测序数据, 对青藏扁蓿豆的 EST-SSR 标记进行了开发。结果发现, 随机选取的 477 对引物中, 350 对可在青藏扁蓿豆中有效扩增, 其中 346 对引物在扁蓿豆 (*M. ruthenica*) 中可实现跨物种扩增。利用其中 64 对独立遗传且符合哈迪-温伯格平衡的标记引物对青藏扁蓿豆和扁蓿豆野生群体的遗传多样性和群体遗传结构分析发现, 两个物种存在明显的种间和种内的遗传分化, 且提示种内群体间的遗传分化与地理隔离或环境选择有关。本研究所开发的 EST-SSR 标记为今后进一步评价青藏扁蓿豆和扁蓿豆的遗传多样性以及解释其适应极端环境的遗传机制提供了重要的研究工具。

关键词: 青藏扁蓿豆; 转录组测序; EST-SSR 标记; 遗传多样性; 群体遗传结构

文献标志码: A 文章编号: 1001-0629(2020)04-0718-10

Development and verification of EST-SSR markers in *Medicago archiducis-nicolai* by transcriptome sequencing

WANG Yingfang^{1,2}, ZHANG Yemeng^{1,2}, LIU Demei^{1,3}, SHEN Yingfang^{1,4}, WANG Haiqing^{1,3}

(1. Key Laboratory of Adaptation and Evolution of Plateau Biota / Northwest Institute of Plateau Biology,

Chinese Academy of Sciences, Xining 810001, Qinghai, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Key Laboratory of Crop Molecular Breeding of Qinghai Province, Xining 810001, Qinghai, China;

4. Institute of Three-River-Source National Park, Chinese Academy of Sciences, Xining 810001, Qinghai, China)

Abstract: *Medicago archiducis-nicolai*, which is endemic to the Qinghai-Tibet Plateau and its adjacent regions, is a species of *Platycarpae* in the genus *Medicago*. This species is characterized by its superior adaptability to extremely harsh environments. In this study, we developed EST-SSR markers of *M. archiducis-nicolai* using transcriptome sequencing. Polymerase chain reactions verified 350 of 477 randomly chosen EST-SSR primer sets in *M. archiducis-nicolai*. Of the verified primer sets, 346 were successfully amplified in *M. ruthenica*, indicating high transferability of the EST-SSR primers to *M. ruthenica*. The genetic diversity and structure of *M. archiducis-nicolai* and *M. ruthenica* populations were characterized

收稿日期: 2019-08-26 接受日期: 2019-10-14

基金项目: 国家自然科学基金 (31770365); 青海省应用基础研究项目 (2017-ZJ-784); 青海省创新平台建设项目 (2017-ZJ-Y14)

第一作者: 王英芳 (1994-), 女, 青海大通人, 在读硕士生, 主要从事豆科牧草资源利用与评价。E-mail: w5297159868@163.com

通信作者: 王海庆 (1971-), 男, 青海湟中人, 副研究员, 博士, 主要从事豆科牧草资源评价与抗逆改良。E-mail: wanghq@nwpb.cas.cn

<http://cykx.lzu.edu.cn>

using 64 non-genetically linked EST-SSR loci that did not deviate from Hardy-Weinberg equilibrium. Results show significant inter- and intra-species genetic differentiation of the two species, implying geographic distance and environmental selection influenced intra-species genetic differentiation. The EST-SSR markers developed in this study provide useful tools for detailed characterization of the genetic diversity of *M. archiducis-nicolai* and *M. ruthenica* and reveal the mechanisms underlying their adaptation to harsh environments.

Keywords: *Medicago archiducis-nicolai*; transcriptomic sequencing; EST-SSR markers; genetic diversity; population genetic structure

Corresponding author: WANG Haiqing E-mail: wanghq@nwipb.cas.cn

青藏扁蓿豆 (*Medicago archiducis-nicolai*) 和扁蓿豆 (*M. ruthenica*) 为苜蓿属多年生二倍体物种, 是矩荚苜蓿族的核心种^[1]。青藏扁蓿豆分布于青藏高原及其毗邻的高海拔高寒地区^[2-3]; 扁蓿豆则广泛分布于北半球的西伯利亚、蒙古以及我国北方高纬度地区^[4-5]。由于对干旱、寒冷以及盐碱等不利环境具有极强的适应性, 扁蓿豆被认为是具有驯化潜力、有望在紫花苜蓿 (*Medicago sativa*) 等苜蓿属栽培种无法越冬地区栽培利用的优质多年生豆科牧草资源^[6]。相对于扁蓿豆, 青藏扁蓿豆作为青藏高原土著物种, 其生长环境更为严酷, 对极端环境具有更强的适应性。

群体遗传多样性评价是植物种质资源利用和保护研究的重要内容。到目前为止, 对扁蓿豆的遗传多样性在形态学、生化和分子水平上已有较多的报道^[7-11], 但研究对象多集中于来自我国内蒙古的扁蓿豆群体, 而对青藏扁蓿豆遗传多样性评价以及群体遗传结构分析研究则只有少数报道^[3, 12]。利用核基因组标记和叶绿体基因组标记对青藏高原及其周边青藏扁蓿豆和扁蓿豆的野生群体分析发现, 由于海拔差异引起的环境异质性可能导致了上述两个物种间或种内群体间的适应性分化^[3]。揭示青藏扁蓿豆适应青藏高原复杂极端环境的机制, 将为其人工驯化栽培, 以及苜蓿属栽培种的抗逆性改良提供重要的参考价值。因此, 有必要利用 SSR (simple sequence repeat) 等分子标记, 进一步对青藏高原复杂多样性环境下青藏扁蓿豆的群体遗传多样性和群体遗传结构特点进行分析。

SSR 也称为微卫星 DNA (microsatellite DNA), 存在于基因组基因编码区或非编码区, 由 1~6 个碱基为单位构成串联重复序列^[13-14]。作为最常用的分子标记类型, SSR 被广泛运用于遗传作图、标记辅

助选择和群体遗传学研究等领域。对尚无基因组序列信息物种的基因组 SSR 标记开发, 因涉及基因组文库构建和大量的常规测序操作, 存在成本高和费时等限制因素^[15-16]。但是, 研究人员通过对迄今为止数量一直不断增加的基于转录本测序的 EST (expressing sequence tag) 序列信息挖掘分析, 开发 EST-SSR 标记是一条省时且成本低廉的途径^[15, 17]。此外, 由于属内物种间基因组转录区序列间存在较高的保守性, EST-SSR 标记引物可在物种间实现高效转移, 也为 EST-SSR 标记在尚缺少 EST 序列信息的物种上的利用带来了方便^[18-19]。但与之同时也发现, 在物种间高效转移的 EST-SSR 标记位点通常存在多态性低的问题^[20]。因此, 在遗传多样性研究、高密度遗传图谱构建以及作物育种中, 开发针对特定物种的 EST-SSR 标记仍然十分必要。

过去十年中, 高通量测序技术的兴起与发展, 极大地提高了低成本高效率获得转录本序列信息的能力, 为规模化开发特定物种的 EST-SSR 标记提供了条件。本研究利用 RNA-seq 技术, 对青藏扁蓿豆 EST-SSR 标记进行了开发和验证, 发现开发的标记在扁蓿豆中具有高效转移性。利用所开发的引物对青藏扁蓿豆和扁蓿豆野生群体的遗传多样性和群体遗传结构初步分析发现, 上述两个种间和种内群体间存在显著的遗传分化。本研究获得的 EST-SSR 标记和初步分析结果, 为今后进一步分析青藏扁蓿豆的群体遗传特性以及揭示上述两个物种间和物种内的适应性分化机制提供了重要的工具和有价值的信息。

1 材料与方法

1.1 材料和培养条件

用于 RNA-seq 测序的青藏扁蓿豆植株采集自青海

省同德县(海拔 3 234 m, 100°38'52" E, 35°15'46" N)天然草地。将同一植株的茎段在蛭石基质中扦插, 在温度 21 °C、光 16 h/暗 8 h 人工气候室培养, 获得遗传背景一致的无性系幼苗后, 转移到相同光照时间、温度为 4 °C 的光照培养箱低温处理。取低温处理 0 和 24 h 幼苗地上部分液氮速冻保存, 进行后续总 RNA 提取和测序文库构建。用于 EST-SSR 引物验证、多态性筛选以及遗传多样性分析的青藏扁蓿豆和扁蓿豆群体采集地点和地理信息如表 1 所列。

1.2 RNA-seq 测序与转录本组装

用 TRIzol 试剂 (Life Technologies, Carlsbad, CA, USA) 提取总 RNA, 每个样品取 3 μg 总 RNA 用 NEB Next Ultra RNA Library Prep Kit for Illumina 试剂盒 (NEB, Beverly, MA, USA) 构建 RNA-seq 测序文库, 在 Illumina HiSeq 2500 测序仪进行双端 reads 测序。对测序原始数据 (raw reads) 过滤剔除文库中污染的接头序列和未知核苷酸数大于 10% 或碱基质量 ≤ 10 ($Q \leq 10$) 的 reads。将两个样品过滤后的 reads 合并后用 Trinity 软件^[21] 进行转录本组装。组装获得的转录本 (transcript) 经过聚类后, 从中选取最长的转录本 (unigene) 进行 SSR 位点检测。

1.3 EST-SSR 位点鉴定与引物设计

EST-SSR 位点鉴定由 MISA 程序 (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) 完成, 设定 SSR 位点侧翼序列长度在 150 bp 以上, 一碱基、二碱基、三碱基、四碱基、五碱基和六碱基重复单元的最小重复次数分别为 12、6、6、5、5 和 4 次。对经过检测含有 SSR 位点的 unigene, 使用 Primer3 ([\[sourceforge.net/projects/primer3\]\(http://sourceforge.net/projects/primer3\)\) 软件设计引物, 引物设计参数: 引物长度 18 ~ 28 nt, 23 nt 为最优; 扩增片段长度 100 ~ 300 bp; 退火温度 55 ~ 65 °C, 60 °C 为最优温度。](http://</p>
</div>
<div data-bbox=)

1.4 EST-SSR 引物验证和多态性筛选

对采集的 3 个青藏扁蓿豆和 3 个扁蓿豆群体的共计 94 个个体 (表 1), 用 CTAB 法^[22] 提取基因组 DNA。分别从西宁西山 (XNXS) 青藏扁蓿豆群体和湟中西堡 (HZXB) 扁蓿豆群体中随机取 5 个个体的 10 ng DNA 样品等量混合为模板, 利用随机挑取合成的 477 对 EST-SSR 引物进行 PCR 扩增验证, rTaq DNA 聚合酶为大连宝生物 (Takara) 产品。扩增反应条件: 94 °C 预变性 4 min; 94 °C 变性 30 s, 60 °C 退火 30 s, 72 °C 延伸 30 s, 共计 5 个循环; 而后 94 °C 变性 30 s, 55 °C 退火 30 s, 72 °C 延伸 30 s, 共计 30 个循环; 最后 72 °C 延伸 10 min。

扩增产物经过 2% 琼脂糖凝胶电泳进行扩增有效性验证后, 进一步利用 6% 聚丙烯酰胺凝胶电泳进行初步多态性筛选。将经过上述检测获得的扩增条带符合预期分子量范围, 且具有潜在多态性的标记反向引物 5' 端用 HEX、6-FAM 或 TAMARA 荧光基团进行标记, 对 3 个青藏扁蓿豆和 3 个扁蓿豆野生群体共计 94 个个体进行扩增, 扩增产物送上海生工利用 ABI3130 毛细管电泳系统进行检测, 用 GeneMapper 2.2.0 (Applied Biosystems, Foster City, CA, USA) 软件读取扩增条带分子量, 并进行进一步人工校正。

1.5 群体遗传多样性和遗传结构分析

各 EST-SSR 位点等位基因数目 (N_A)、位点丰富

表 1 青藏扁蓿豆和扁蓿豆野生群体采样点信息

Table 1 *Medicago archiducis-nicolai* and *M. ruthenica* population sources

物种 Species	群体代码 Population code	地点 Location	经度 Latitude (E)	纬度 Longitude (N)	海拔 Altitude/m	样本 Size
青藏扁蓿豆 <i>Medicago archiducis-nicolai</i>	XNXS	青海西宁西山 Xishan, Xining, Qinghai	101°44'17.5"	36°37'17.5"	2 462	16
	RYS	青海日月山 Riyueshan, Qinghai	101°7'11.7"	36°26'50.6"	3 311	16
	YSMC	青海玉树 Yushu, Qinghai,	96°50'51.2"	33°27'27.8"	4 023	16
扁蓿豆 <i>M. ruthenica</i>	HZLZ	青海湟中 Huangzhong, Qinghai	101°31'37"	36°24'34"	2 422	16
	HZXB	青海湟中 Huangzhong, Qinghai	101°36'44.8"	36°34'36.2"	2 455	14
	ZN	甘肃卓尼 Zhuoni, Gansu	103°29'41"	34°32'38"	2 689	16

<http://cykx.lzu.edu.cn>

度 (A_R)、群体间遗传分化系数 (F_{ST})、个体相对于总群体和各个群体的近交系数 (F_{IT} 和 F_{IS}) 用 FSTAT 2.9.3.2 软件 (<http://www2.unil.ch/popgen/softwares/fstat.html>) 估计; 多态性信息含量 (PIC) 用 Cervus 3.0 程序计算^[23]。利用 GenALEx 6.41^[24-25] 计算每个群体的有效等位基因数目 (N_E)、观察和期望杂合度 (H_O and H_E) 以及固定系数 (F)。对各个位点在每个群体中的哈迪-温伯格平衡 (Hardy-Weinberg equilibrium) 和连锁不平衡 (genotypic linkage disequilibrium) 用 GENEPOP 4.2^[26] 软件进行, 并予以 Bonferroni 校正 ($P < 0.05$)。

群体主成分分析和贝叶斯聚类 (Bayesian clustering) 分别用 GenALEx 6.4.3 和 STRUCTURE 2.3.4 程序^[27] 进行。STRUCTURE 软件分析选用混合模型 (admixture model) 并对群体间的位点频率进行校正, K 值为 1 ~ 6。对每个 K 值进行 10 次运算, 每次运算 burn-in 和 MCMC repetition 参数分别为 100 000 和 30 000。最优 K 值由 ΔK 方法^[28] 和 $\ln P(K)$ 曲线拐点来确定。POWERMARKER 3.25^[29] 用于计算群体间 Nei's 遗传距离和构建基于此的 UPGMA (unweighted pair group method with arithmetic) 聚类树。利用 ARLEQUIN 3.5^[30] 用于 AMOVA 分析。

2 结果与分析

2.1 转录组测序、转录本组装

对青藏扁蓿豆幼苗低温处理 0 和 24 h 两个样品经转录组测序获得原始数据 (raw reads), 经过滤去除低质量的 reads 后, 每个样品最终获得的 Q20

碱基比例不低于 98% 的有效数据 (clean reads) 在 5.28×10^7 条。上述两个样品 clean reads 测序结果已提交至 NCBI 数据库, 收录编号分别为 SRR8797417 和 SRR8797418。在此基础上, 将两个样品的 clean reads 文件合并后, 经过从头组装、转录本 (transcript) 聚类, 从各个聚类转录本中选取最长转录本作为代表性转录本 (unigene), 最终获得了 99 635 条 unigene 序列^[31]。

2.2 EST-SSR 位点的检测与引物设计

通过 MISA 软件对 99 635 条 unigene 的检测, 发现 2 887 条 unigene 中存在侧翼序列长度在 150 bp 以上的 3 253 个潜在的 SSR 位点。其中 242 个 unigene 含有一个以上 SSR 位点, 但位点之间间隔序列小于 150 bp, 为复合 SSR (complex SSR) 位点; 其余的 unigene 仅含有一个 SSR 位点, 或一个以上且间隔序列大于 150 bp 的位点, 为简单 SSR (simple SSR) 位点。获得的 SSR 中以三碱基重复单元比例最高 (42.61%), 其次为单碱基重复序列 (30.25%) (表 2)。根据上述结果, 利用 SSR 位点的侧翼序列, 设计获得了 3 080 对 EST-SSR 引物, 涉及 2 829 条 unigene, 其中 2 605、198、25 和 1 条 unigene 中分别含有 1、2、3 和 4 个简单 EST-SSR 引物扩增位点, 另外还从含有复合 SSR 位点的 98 条 unigene 中获得了 98 对复合 SSR 标记引物 (表 2)。本研究所获得的 unigene 序列、EST-SSR 位点信息、引物序列和本文后续分析获得的遗传参数数据集 (data sets) 均已提交到 Mendeley 数据库^[31] (<https://data.mendeley.com/>)。

表 2 青藏扁蓿豆代表性转录本中 EST-SSR 位点检测和引物设计统计
Table 2 Summary of detection and primer designing of EST-SSR loci in *Medicago ruthenica* unigenes

重复单元类型 Repeat type	数量 Number	频率 Frequency/%	引物对数 Number of primer pairs
单碱基 Mono-nucleotide	984	30.25	931
二碱基 Di-nucleotide	595	18.29	530
三碱基 Tri-nucleotide	1 386	42.61	1 275
四碱基 Tetra-nucleotide	71	2.18	58
五碱基 Penta-nucleotide	101	3.10	88
六碱基 Hexa-nucleotide	116	3.57	100
复合位点 Compound loci	107	3.29	98
合计 Total	3 253	100.00	3 080

2.3 EST-SSR 标记引物验证与多态性筛选

为了验证所计 ESR-SSR 引物的有效性, 从中随机选取合成 477 对引物^[31], 以来自西宁西山 (XNXS) 的青藏扁蓿豆群体中 5 个植株基因组 DNA 等量混合为模板进行了扩增。对扩增产物的琼脂糖凝胶电泳结果表明, 其中有 350 对引物能够扩增获得与预期分子量接近的扩增条带^[31]。上述 350 对引物对湟中西堡 (HZXB) 扁蓿豆群体 5 个植株等量混合 DNA 模板扩增, 发现有 346 对引物能够获得预期扩增条带^[31]。此外, 有 28 对 (5.87%) 引物在扁蓿豆和青藏扁蓿豆中的扩增片段长度在 400~1 500 bp^[31], 可能与扩增区段存在内含子有关。

进一步对扩增产物聚丙烯酰胺凝胶电泳分析发现, 165 对和 114 对引物分别在青藏扁蓿豆和扁蓿豆种内个体间存在潜在的多态性; 另外有 44 对引物的扩增产物虽不存在种内个体间的多态性, 但具有种间多态性^[31]。在此基础上, 对其中 100 对引物的反向引物 5'端用荧光素标记后, 通过对 3 个青藏扁蓿豆和 3 个扁蓿豆共计 94 个个体 DNA 模板进行扩增, 毛细管电泳检测确认, 最终获得了 75 对条带清晰可辨且具有多态性的引物^[31]。该 75 对引物在 6 个群体中共计检测到了 588 个等位变异, 每个位点平均等位基因数 (N_A) 为 7.71 ± 4.63 , 其中 51 个位点 (68%) 均有较高的多态性变异 ($PIC \geq 0.5$)^[31]。各位点丰富度 (A_R) 在 1.311~8.909, 平均为 3.706; 观察杂合度 (H_O) 在 0.010~0.83, 平均为 0.455; 期望杂合度 (H_E) 在 0.030~0.837, 平均为 0.461; F_{IS} 在 -0.586~0.093, 平均为 0.049; F_{IT} 在 -0.262~0.853, 平均为 0.302; F_{ST} 在 -0.010~0.794, 平均为 0.200 (表 3)^[31]。哈迪-温伯格平衡检验结合 Bonferroni 校正 ($P < 0.05$) 发现 9 个位点表现为杂合子过量或杂合子缺失^[31]。进一步对符合哈迪-温伯格平衡的 66 个位点的连锁不平衡检验发现, 两对位点之间 (Man184 和 Man251, Man228 和 Man257) 存在遗传连锁^[31]。为此将 Man257 和 Man228 位点剔除后, 最终保留 64 个独立遗传且不存在相互遗传连锁的位点^[31] 进行后续群体遗传多样性和遗传结构分析。

2.4 群体遗传多样性、遗传结构与分化

利用上述 64 个独立遗传且不存在相互遗传连

表 3 基于 6 个群体的 75 个 EST-SSR 位点的遗传参数
Table 3 Summary of genetic parameters for 75 EST-SSR loci in six populations

参数 Parameters	平均值 Mean \pm SE	范围 Range
丰富度 A_R	3.706 \pm 1.586	1.311~8.909
多态性信息含量 PIC	0.582 \pm 0.210	0.031~0.918
观察杂合度 H_O	0.455 \pm 0.013	0.010~0.830
期望杂合度 H_E	0.461 \pm 0.011	0.030~0.837
群体内近交系数 F_{IS}	0.049 \pm 0.017	-0.586~0.693
群体间遗传分化系数 F_{ST}	0.200 \pm 0.015	-0.010~0.794
总近交系数 F_{IT}	0.302 \pm 0.021	-0.262~0.853

A_R , allele richness; PIC, polymorphism information content; H_O , observed heterozygosity; H_E , expected heterozygosity; F_{IS} , inbreeding coefficient within populations; F_{ST} , genetic differentiation index among populations; F_{IT} , inbreeding coefficients with individuals; similarly for Table 4.

锁的 EST-SSR 位点数据, 对 3 个青藏扁蓿豆和 3 个扁蓿豆野生群体分析发现, 各群体的平均等位基因数 (N_A) 在 3.047 (YSMC) ~ 4.688 (ZN), 平均为 3.823; 有效等位基因数 (N_E) 在 2.204 (RYS) ~ 2.741 (ZN), 平均为 2.330; 观察杂合度 (H_O) 在 0.437 (RYS) ~ 0.487 (ZN); 期望杂合度 (H_E) 在 0.428 (XNXS) ~ 0.528 (ZN) (表 4)。扁蓿豆较青藏扁蓿豆群体均有相对较高的遗传多样性, 且以来自甘肃卓尼 (ZN) 群体最高。各群体的固定系数 (F 值) 在 -0.064 (YSMC) ~ 0.089 (ZN), 平均为 0.010 (表 4), 因此各群体内几乎不存在近交 (Self-crossing)。

运用 STRUCTURE 软件的贝叶斯聚类方法, 可将 6 个群体归为和形态学分类相一致的两个种, 但存在明显的青藏扁蓿豆向扁蓿豆群体的基因渗入 (图 1A, B)。主成分分析结果显示, 占总变异 67.62% 的第一主成分 (PC_1) 将 6 个野生群体区分为与 STUCUTRE 结果一致的两类, 占总变异 19.07% 的第二主成分 (PC_2) 将青藏扁蓿豆的 YSMC 群体和另外两个群体区分开来 (图 1C), 这表明, 青藏扁蓿豆可能存在地理隔离相关的遗传分化, 基于 Nei's 遗传距离的 UPGMA 结果进一步支持了上述推断 (图 1D)。基于上述结果的分级 AMOVA 分析结果显示, 64 个位点解释的遗传变异主要存在个体内 (62.95%, $F_{IT} = 0.370 5$, $P = 0.000 0$), 其余依次为种间 (24.56%, $F_{CT} = 0.245 6$, $P = 0.000 0$)、种内群体间

表 4 基于 64 个 EST-SSR 位点的群体遗传参数
Table 4 Genetic diversity parameters of populations revealed by 64 EST-SSR loci

群体代码 Population code	等位基因数 N_A	有效等位基因数 N_E	位点丰富度 A_R	观察杂合度 H_O	期望杂合度 H_E	固定系数 F
XNXS	3.438 ± 0.232	2.227 ± 0.156	3.370 ± 1.781	0.444 ± 0.034	0.428 ± 0.032	-0.047 ± 0.020
RYS	3.484 ± 0.208	2.204 ± 0.145	3.410 ± 1.600	0.437 ± 0.030	0.445 ± 0.028	0.005 ± 0.030
YSMC	3.047 ± 0.146	2.173 ± 0.123	3.005 ± 1.373	0.473 ± 0.036	0.442 ± 0.031	-0.064 ± 0.027
HZLZ	4.328 ± 0.271	2.452 ± 0.162	4.173 ± 2.036	0.481 ± 0.033	0.495 ± 0.028	0.063 ± 0.033
HZXB	3.953 ± 0.244	2.184 ± 0.128	3.953 ± 1.955	0.454 ± 0.031	0.456 ± 0.027	-0.001 ± 0.030
ZN	4.688 ± 0.322	2.741 ± 0.204	4.532 ± 2.421	0.487 ± 0.031	0.528 ± 0.028	0.089 ± 0.029
平均 Mean ± SE	3.823 ± 0.104	2.330 ± 0.064	3.741 ± 1.950	0.463 ± 0.013	0.481 ± 0.012	0.010 ± 0.012

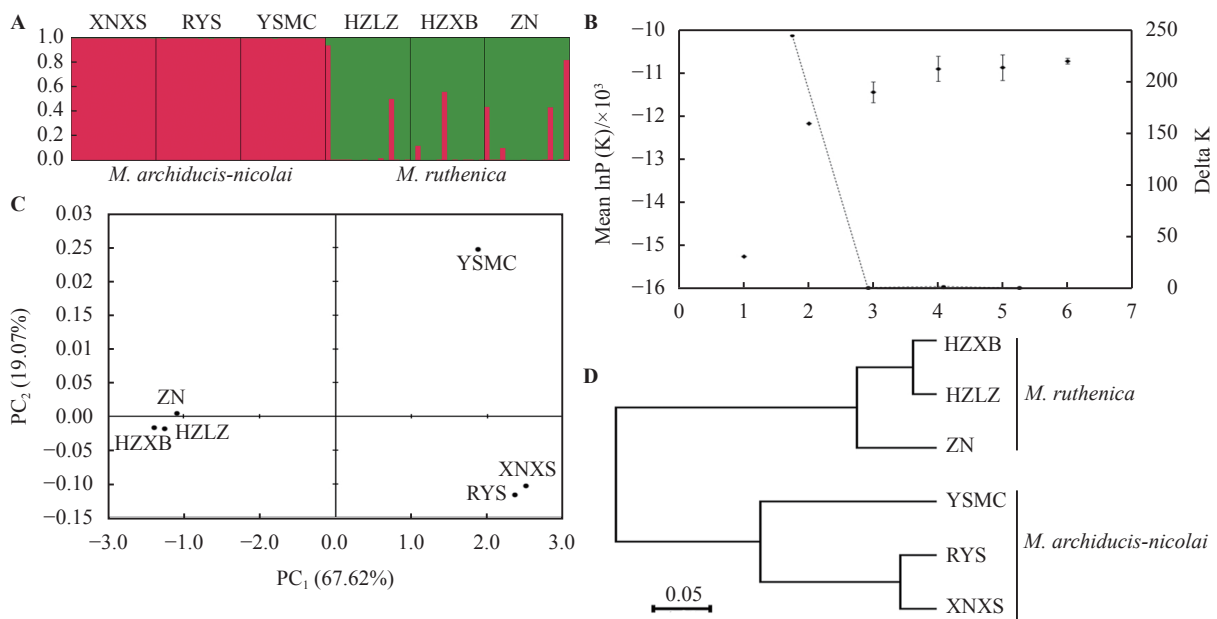


图 1 群体遗传结构分析

Figure 1 Population genetic structure of six populations

A. STRUCTURE 群体分组； B. STRUCTURE 最佳分组 (K 值) 确定； C. 主成分分析； D. 基于 Nei's 遗传距离的 UPGMA 群体分组。XNXS, RYS, YSMC, HZLZ, HZXB, ZN 表示群体代码，详如表 1 所列；下表同。

A. Individuals and populations assigned to the STRUCTURE-inferred genetic clusters for K = 2. B. Plot of lnP(K) and ΔK from STRUCTURE analysis, which shows optimal clustering of the six populations for K = 2. C. Principal component analysis. D. Unweighted pair group method with averages dendrogram based on Nei's genetic distance. XNXS, RYS, YSMC, HZLZ, HZXB, and ZN indicate population codes of the six populations as listed in Table 1; similarly for the following tables.

(9.91%, $F_{SC} = 0.1314$, $P = 0.0000$) 和群体内个体间 (2.58%, $F_{IS} = 0.0394$, $P = 0.0000$) (表 5)。

对群体间遗传分化分析显示，除了扁蓿豆 HZXB 和 HZLZ 群体间存在显著 ($P < 0.05$) 的遗传分化外，其余群体间均存在极显著 ($P < 0.01$) 的遗传分化，不同种的群体间的遗传分化系数高于种内群体间，而且种内地理距离较远的群体间的遗传分化明显高于地理距离较近的群体间的遗传分

化 (表 6)。

3 讨论与结论

属内种间基因组转录区序列上相对较高的保守型为 EST-SSR 标记的跨物种转移利用提供了可能^[18, 20]。有报道表明，苜蓿属内物种间 ESR-SSR 标记的转移效率可在 70% 以上^[32-34]，但同时也发现具有物种间可转移性的 EST-SSR 标记的多态性较

表5 青藏扁蓊豆和扁蓊豆群体分级分子变异分析

Table 5 Hierarchical analysis of molecular variance (AMOVA) in *M. archiducis-nicolai* and *M. ruthenica* populations

变异来源 Source of variation	自由度 Degrees of freedom	平方和 Sum of squares	变异组成 Variance components	变异百分比 Variation percentage/%	固定指数 Fixation index
种间 Among species	1	632.257	5.779 51 Va	24.56	$F_{CT} = 0.245 6$ ($P = 0.000 0$)
种内群体间 Among populations within species	4	356.244	2.332 77 Vb	9.91	$F_{SC} = 0.131 4$ ($P = 0.000 0$)
群体内个体间 Among individuals within populations	88	1 410.647	0.608 13 Vc	2.58	$F_{IS} = 0.039 4$ ($P = 0.000 0$)
个体内 Within individuals	94	1 392.500	14.010 64 Vd	62.95	$F_{IT} = 0.370 5$ ($P = 0.000 0$)
总变异 Total	187	3 791.628	23.534 24		

表6 群体间遗传分化系数 (F_{ST})Table 6 Index of genetic differentiation (F_{ST}) between populations

群体代码 Population code	XNXS	RYS	YSMC	HZLZ	HZXB	ZN
XNXS	0.000	0.061**	0.231**	0.352**	0.377**	0.318**
RYS		0.000	0.240**	0.339**	0.365**	0.307**
YSMC			0.000	0.351**	0.377**	0.321**
HZLZ				0.000	0.014*	0.093**
HZXB					0.000	0.106**
ZN						0.000

“*”和“**”分别表示群体间存在显著($P < 0.05$)和极显著($P < 0.01$)在遗传分化。

“*” and “**” indicate significant differentiation between populations at 0.05 and 0.01 levels, respectively.

低^[20]。此外,由于不同物种间扩增目标区段可能存在的细微变异,会导致扩增的假阳性或物种特异性区段扩增的失败。本研究利用 Illumina 高通量测序技术,对青藏扁蓊豆的 EST-SSR 位点进行了鉴定,发现其中以三碱基重复单位主要类型,其次为单碱基重复类型。最近对扁蓊豆的 EST-SSR 位点检测表明主要为二碱基重组单元,三碱基重复单元次之^[33]。造成上述差异的原因可能与本研究在 SSR 位点鉴定时剔除了侧翼序列小于 150 bp 的位点,使得转录本的开放读码框得到了相对富集有关。本研究中开发的青藏扁蓊豆的 EST-SSR 标记引物在扁蓊豆中扩增发现存在较高的可转移性;而利用上述引物在与之不同族(Section)的蒺藜苜蓿(*M. truncatula*)基因组中的查询则发现 350 对引物中仅有 137 对引物存在潜在的扩增靶点或区段^[31]。上述结果表明,苜蓿属内不同族的物种间转录区仍然存在一定程度的变异。因此,为了获得高密度多态性位点,针对苜蓿属特定的物种或者

族,开发相应的 EST-SSR 仍然十分必要。本研究中开发的青藏扁蓊豆和最近报道开发的扁蓊豆中的 EST-SSR 标记^[32],将会为今后解释矩荚苜蓿组内物种的群体遗传多样性和遗传结构提供有用的研究工具。

青藏扁蓊豆和扁蓊豆为异花授粉植物,异花授粉植物群体一般被认为存在较高水平的遗传多样性^[8, 35-38]。此前,利用 SSR 标记位点的期望杂合度对扁蓊豆野生群体的遗传多样性分析结果不尽一致^[7, 8, 32]。本研究中利用 64 个不存在偏离哈迪-温伯格平衡且独立遗传的 EST-SSR 位点,对位于青藏高原及其毗邻地区的青藏扁蓊豆和扁蓊豆野生群体的初步分析显示其平均期望杂合度为 0.481,与最近报道的基于 EST-SSR 标记的扁蓊豆群体的分析结果接近^[32]。上述结果上的差异可能与分析群体大小、群体瓶颈效应、奠基者效应或者人类活动的干扰等因素有关。此外,与来自基因间区的 SSR 标记相比,EST-SSR 标记多态性较低,也可能对

物种遗传多样评价结果产生影响。

本研究中的遗传结构分析结果得到了与笔者此前利用核基因组和叶绿体基因组标记^[3,12]相一致的结果,均可以将青藏扁蓿豆和扁蓿豆归为和形态学鉴定结果相一致的两个物种,此外也发现,存在青藏扁蓿豆向扁蓿豆群体的基因渗入。但是本研究中主成分分析、UPGMA 聚类以及群体间遗传分化比较的结果提示上述两个物种种内群体间遗传分化可能与地理距离之间存在相关性,而此前的研究则没有发现相关性^[3]。因此,EST-SSR 标记较此前使用的核基因组和叶绿体基因组标记具有更高的灵敏性。上述研究结果为今后解释青藏高原地理和气候环境对上述两个物种遗传多样性和遗传结构的影响提供了重要的信息。

EST-SSR 标记的多态性是基因组转录区段变异水平的反应,通过对转录区序列在核苷酸和蛋白质数据库中的查询可以获得含有 SSR 位点的转录

本可能的功能。采用遗传位点 F_{ST} 离群检验和环境信息相关联的方法,已被用于物种在环境选择下的变异规律以及引起局部适应的环境因子筛选研究^[39-43]。对分子标记位点所在基因组区段潜在功能的了解,将会为解析物种适应环境的机制提供更为丰富和详细的信息^[43-44]。与扁蓿豆相比,青藏扁蓿豆分布于海拔较高的青藏高原地区,因此青藏扁蓿豆较扁蓿豆对青藏高原极端低温环境具有更强的适应性。青藏高原在历史上曾经历了隆升和大规模的气候波动^[45-46]。青藏扁蓿豆和扁蓿豆在上述过程中也随之经历了局部环境下的存活、历史冰纪终期或后期的局部扩增^[3]。因此,进一步鉴定与历史气候波动以及复杂地貌环境导致的局部气候异质性相关的 EST-SSR 位点的遗传分化程度,将不仅有助于揭示上述植物物种适应青藏高原极端环境的遗传机制,同时也会将有助于苜蓿属栽培种抗逆性改良。

参考文献 References:

- [1] SMALL E, JOMPHE M. A synopsis of the genus *Medicago* (Leguminosae). *Canadian Journal of Botany*, 1989, 67(11): 3260-3294.
- [2] JIN D M, MA J J, MA W H, LIANG C Z, SHI Y, HE J S. Legumes in Chinese natural grasslands: Species, biomass, and distribution. *Rangeland Ecology & Management*, 2013, 66(6): 648-656.
- [3] WU X P, LIU D M, GULZAR K, SHEN Y F, WANG H Q. Population genetic structure and demographic history of *Medicago ruthenica* (Fabaceae) on the Qinghai-Tibetan Plateau based on nuclear ITS and chloroplast markers. *Biochemical Systematics and Ecology*, 2016, 69: 204-212.
- [4] CAMPBELL T A, BAO G, XIA Z L. Agronomic evaluation of *Medicago ruthenica* collected in Inner Mongolia. *Crop Science*, 1997, 37(2): 599-604.
- [5] BALABAEV G A. Yellow lucernes of Siberia, *Medicago ruthenica* (L.) Lebd. and *M. platycarpus* (L.) Lebd. *Bulletin of Applied Botany of Genetic and Plant Breeding Service*, 1934, 7: 13-123.
- [6] 乌云飞, 玉柱, 石凤翎. 直立型扁蓿豆的选育及其生物学特性的研究. *内蒙古草业*, 1993(Z1): 51-53.
WU Y F, YU Z, SHI F L. Breeding and biological characteristics of a cultivar of *Medicago ruthenica* (L.) Sojak cv. 'Zhilixing'. *Inner Mongolia Prataculture*, 1993(Z1): 51-53.
- [7] LI H Y, LI Z Y, CAI L Y, SHI W G, MI F G, SHI F L. Analysis of genetic diversity of *Ruthenia Medic* (*Medicago ruthenica* (L.) Trautv.) in Inner Mongolia using ISSR and SSR markers. *Genetic Resources and Crop Evolution*, 2013, 60(5): 1687-1694.
- [8] YAN J, CHU H J, WANG H C, LI J Q, SANG T. Population genetic structure of two *Medicago* species shaped by distinct life form, mating system and seed dispersal. *Annals of Botany*, 2009, 103(6): 825-834.
- [9] 李鸿雁, 李志勇, 米福贵, 卢新雄, 师文贵. 利用微卫星标记鉴定扁蓿豆种质资源. *华北农学报*, 2008, 23(3): 67-71.
LI H Y, LI Z Y, MI F G, LU X X, SHI W G. Appraisal of *Medicago ruthenica* germplasms by SSR markers. *Acta Agriculturae Boreali-Sinica*, 2008, 23(3): 67-71.
- [10] 曹喆, 谷安琳, 赵来喜, 刘昊, 戎郁萍. 扁蓿豆遗传多样性的等位酶研究. *西北植物学报*, 2008, 28(10): 2005-2010.
CAO Z, GU A L, ZHAO L X, LIU H, RONG Y P. Genetic variation of allozymes in *Medicago ruthenica*. *Acta Botanica Boreali-Occidentalia Sinica*, 2008, 28(10): 2005-2010.

- [11] 乌日娅, 雍世鹏, 包贵平. 扁蓿豆生态生物学特性的比较研究. 中国草地学报, 1994(2): 1-7.
WU R Y, YONG S P, BAO G P. A Comparative study on ecobiological characteristics of *Melilotoides ruthenica*. Chinese Journal of Grassland, 1994(2): 1-7.
- [12] 吴小培, 沈迎芳, 王海庆. 基于 trnL-trnF 序列的扁蓿豆和青藏扁蓿豆遗传多样性及其群体遗传结构分析. 草业科学, 2016, 33(6): 1136-1146.
WU X P, SHEN Y F, WANG H Q. Analysis of genetic diversity and population genetic structure of *Medicago archiducis-nicolai* and *Medicago ruthenica* populations based on cpDNA trnL-trnF sequences. Pratacultural Science, 2016, 33(6): 1136-1146.
- [13] TOTH G, GASPARI Z, JURKA J. Microsatellites in different eukaryotic genomes: Survey and analysis. Genome Research, 2000, 10(7): 967-981.
- [14] POWELL W, MACHRAY G C, PROVAN J. Polymorphism revealed by simple sequence repeats. Trends in Plant Science, 1996, 1(7): 215-222.
- [15] SQUIRRELL J, HOLLINGSWORTH P M, WOODHEAD M, RUSSEL J, LOWE A J, GIBBY M, POWELL W. How much effort is required to isolate nuclear microsatellites from plants? Molecular Ecology, 2003, 12(6): 1339-1348.
- [16] ZANE L, BARGELLONI L, PATARNELLO T. Strategies for microsatellite isolation: A review. Molecular Ecology, 2002, 11(1): 1-16.
- [17] SAHA M C, MIAN M A R, EUJAYL I, AWONITZER J C, WANG L J, MAY G D. Tall fescue EST-SSR markers with transferability across several grass species. Theoretical and Applied Genetics, 2004, 109(4): 783-791.
- [18] ELLIS J R, BURKE J M. EST-SSRs as a resource for population genetic analyses. Heredity, 2007, 99(2): 125-132.
- [19] VARSHNEY R K, GRANER A, SORRELLS M E. Genic microsatellite markers in plants: features and applications. Trends in Biotechnology, 2005, 23(1): 48-55.
- [20] PONCET V, RONDEAU M, TRANCHANT C, CAYREL A, HAMON S, DE KOCHKO A, HAMON P. SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. Molecular Genetics and Genomics, 2006, 276(5): 436-449.
- [21] GRABHERR M G, HAAS B J, YASSOUR M, LEVIN J Z, THOMPSON D A, AMIT I, ADICONIS X FAN L, RAYCHOWDHURY R, ZENG Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 2011, 29(7): 644-652.
- [22] DOYLE J J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, 1987, 19: 11-15.
- [23] KALINOWSKI S T, TAPER M L, MARSHALL T C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, 2007, 16(5): 1099-1106.
- [24] PEAKALL R, SMOUSE P E. GenA1Ex 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics, 2012, 28(19): 2537-2539.
- [25] PEAKALL R, AND SMOUSE P E. GenA1Ex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 2006, 6(1): 288-295.
- [26] ROUSSET F. Genepop'007: A complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources, 2008, 8(1): 103-106.
- [27] FALUSH D, STEPHENS M, PRITCHARD J K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics, 2003, 164(4): 1567-1587.
- [28] EVANNO G, REGNAUT S, GOUDET J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 2005, 14(8): 2611-2620.
- [29] LIU K J, MUSE S V. PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics, 2005, 21(9): 2128-2129.
- [30] EXCOFFIER L, LISCHER H E L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 2010, 10(3): 564-567.
- [31] WANG Y F, ZHANG Y M, LIU D M, SHEN Y F, WANG H Q. Data for: EST-SSR marker development in *Medicago archiducis-nicolai* and transferability to *Medicago ruthenica*. Mendeley Data, 2019, V1, doi: 10.17632/m86xxy2gkt.1

- [32] LI J, LI H Y, CHI E H, HUANG F, LIU L, DING Z Y, SHI W G, MI F G, LI Z Y. Development of simple sequence repeat (SSR) markers in *Medicago ruthenica* and their application for evaluating outcrossing fertility under open-pollination conditions. *Molecular Breeding*, 2018, 38(12): 142-148.
- [33] WANG Z, YU G H, SHI B B, WANG X M, QIANG H P, GAO H W. Development and characterization of simple sequence repeat (SSR) markers based on RNA-Sequencing of *Medicago sativa* and *in silico* mapping onto the *M. truncatula* genome. *PLoS One*, 2014, 9(3): e92029.
- [34] EUJAYL I, SLEDGE M K, WANG L, MAY G D, CHEKHOVSKIY K, ZWONITZER J C, MIAN M A R. *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. *Theoretical and Applied Genetics*, 2004, 108(3): 414-422.
- [35] ELLWOOD S R, D'SOUZA N K, KAMPHUIS L G, BURGESS T I, NAIR R M, OLIVER R P. SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theoretical and Applied Genetics*, 2006, 112(5): 977-983.
- [36] STENOIEN H K, FENSTER C B, TONTERI A, SAVOLAINEN O. Genetic variability in natural populations of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, 2005, 14(1): 137-148.
- [37] FLAJOULOT S, RONFORT J, BAUDOUIN P, BARRE P, HUGUET T, HUYGHE C, JULOER B. Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a breeding program, using SSR markers. *Theoretical and Applied Genetics*, 2005, 111(7): 1420-1429.
- [38] CLAUSS M J, COBBAN H, MITCHELL-OLDS T. Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). *Molecular Ecology*, 2002, 11(3): 591-601.
- [39] SONG Z J, ZHANG M M, LI F G, WENG Q J, ZHOU C P, LI M, LI J, HUAN H H, MO X Y, GAN S M. Genome scans for divergent selection in natural populations of the widespread hardwood species *Eucalyptus grandis* (Myrtaceae) using microsatellites. *Scientific Reports*, 2016, 6: 34941.
- [40] LIND-RIEHL J F, SULLIVAN A R, GAILING O. Evidence for selection on a CONSTANS-like gene between two red oak species. *Annals of Botany*, 2014, 113(6): 967-975.
- [41] LI C, SUN Y, HUANG H W, CANNON C H. Footprints of divergent selection in natural populations of *Castanopsis fargesii* (Fagaceae). *Heredity*, 2014, 113(6): 533-541.
- [42] SUGAI K, SETSUKO S, NAGAMITSU T, MURAKAMI N, KATO H, YOSHIMARU H. Genetic differentiation in *Elaeocarpus photiniifolia* (Elaeocarpaceae) associated with geographic distribution and habitat variation in the Bonin (Ogasawara) Islands. *Journal of Plant Research*, 2013, 126(6): 763-774.
- [43] BRADBURY D, SMITHSON A, KRAUSS S L. Signatures of diversifying selection at EST-SSR loci and association with climate in natural *Eucalyptus* populations. *Molecular Ecology*, 2013, 22(20): 5112-5129.
- [44] STRASBURG J L, SHERMAN N A, NATASHA A, WRIGHT K M, MOYLE L C, WILLIS J H, RIESEBERG L H. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2012, 367: 364-373.
- [45] 施雅风, 李吉均, 李炳元. 青藏高原晚新生代隆升与环境变化. 广州: 广东科技出版社, 1998.
SHI Y F, LI J J, LI B Y. Uplift and Environmental Changes of Qinghai-xizang (Tibetan) Plateau in the Cenozoic. Guangzhou: Guangdong Science & Technology Press, 1998.
- [46] ZHENG D. The System of Physico-Geographical Regions of the Qinghai-Xizang (Tibet) Plateau. *Science in China Series D-Earth Sciences*, 1996, 39(4): 410-417.

(责任编辑 王芳)