

基于刺槐转录组测序的 SSR 引物开发及不同来源无性系种质的遗传多样性分析

毛秀红

(山东林业科学研究院 山东济南 250000)

摘要: 刺槐 (*Robinia pseudoacacia* L.) 为豆科 (Leguminosae) 刺槐属 (*Robinia*) 阔叶乔木, 耐干旱、耐瘠薄、耐轻度盐碱、抗污染能力强、自然更新能力强, 是荒山造林和水土保持先锋树种, 且具有观赏、材用、蜜源和饲用价值。刺槐作为外来物种, 在我国已经栽培了 140 年, 栽培范围自南京、青岛逐渐扩大至 27 个省(市、自治区)。目前国内外关于刺槐生态特性、适应性、无性繁殖和有性生殖方面的研究较多, 而遗传信息相对匮乏, 遗传学基础研究比较薄弱。中国各地区栽培的刺槐无性系来源记载不清, 遗传关系不明, 遗传多样性状况信息甚少, 同物异名现象时有发生, 急需系统开展刺槐无性系遗传多样性分析和指纹图谱构建, 为下一步开展新品种选育、遗传改良和品种鉴定打好基础。本研究得到的结果如下:

(1) 开展了刺槐转录组测序和基因功能注释工作, 获得 12.2 Gb 刺槐花序和叶片转录组数据, 得到 141 948 个 unigenes, 平均长度 1 209 bp, N50 为 1 997 bp。将这些 Unigenes 注释到 Nt、Swiss-Prot、KOG 和 KO 这 4 个数据库, 分别有 104 239 (73.43%)、67 194 (47.34%)、60 814 (43.15%) 和 81 439 (57.37%) 个 Unigenes 获得功能注释。挖掘 SSR 位点 35 423 个, 平均每 4.84 kb 出现一个 SSR 位点, SSR 检出率为 24.95%。分析了刺槐 SSR 的分布规律, 起始位点主要集中于 unigene 的 0~2 000 bp 之间, SSR 序列长度主要集中在 12~20 bp 之间, 平均 16.5 bp。二核苷酸、三核苷酸和五核苷酸为优势重复类型, 二、三核苷酸重复基元的主要类型分别为 GA/CT 和 GGC/CCG。刺槐只有二、三核苷酸重复微卫星丰度和重复基元长度呈显著负相关, 相关系数为-0.5003186。

(2) 开发了 25 882 个 EST-SSR 分子标记, 引物设计成功率为 73.07%。随机挑选 30 对进行了引物有效性和通用性研究, 扩增成功率为 80%, 其中有 9 对引物多态性较高, 都位于 UTR 区域, 对 385 个无性系的分子鉴定率为 95.32%。9 对引物能够在 3 个刺槐属物种中通用, 并且具有明显的多态性。

(3) 8 对 SSR 引物在 383 个刺槐无性系中共扩增出 68 个等位变异, 每个位点的等位变异数为 4~16 个, 平均每对引物扩增出 8.5 个等位变异; 不同位点的多态性信息量 PIC (polymorphism informative content) 值变化范围很大, 在 0.146~0.898 之间, 平均 PIC 值为 0.528。

(4) 8 对 SSR 引物在 9 个地域共 318 个样本中扩增出 66 条多态性条带, 每个位点的等位变异数为 4~16 个, 平均每对引物扩增出 8.25 个等位变异; 9 个地域的平均等位基因数是 5.617, 其中国内 8 个地域的平均值是 5.597, 山东 6.00; 实际观察杂合度 H_o 在 0.479 (辽宁)~0.555 (山东) 之间, 平均值为 0.519, 其中山东的 H_o 值略高于国内 8 个地域的平均值 0.514; 平均期望杂合度 H_e 在 0.509 (辽宁)~0.553 (美国) 之间, 平均 H_e 值为 0.533, 其中山东的 H_e 值为 0.548, 略高于国内 8 个地域的平均值 0.533。9 个地域的 H_e 值均大于 0.5, 表明遗传多样性均比较丰富。其中山东和山西 H_o/H_e 大于 1, 表明这两个地域的无性系杂合度较高。以上数据表明山东的刺槐无性系遗传多样性水平略高于国内遗传多样性水平, 拥有比较丰富的遗传多样性。山东、甘肃等有私有等位基因, 杂交育种时可以选择作为亲本; 国内不同省份之间可以进行适当引种, 以增加地域内的遗传多样性。9 个地域 318 个无性系之间的遗传变异分析表明只有 2% 来源于地域间, 8% 来自于无性系间, 90% 来自于无性系内。不同地域主坐标分析和亲缘关系分析结果基本一致。山东、美国、辽宁、河南和山西无性系亲缘关系比较近, 聚为一大类群; 甘肃和河北亲缘关系较近, 聚为一大类群; 陕西和北京亲缘关系较近, 聚为一大类群。山东和美国的无性系遗传距离最小为 0.011, 说明这两个地域的无性系遗传分化差异较小。

(5) 8 对 SSR 引物在山东省 49 个无性系中共扩增出 51 个等位变异, 每个位点的等位变异数为 2~15

个,平均每个位点扩增出 6.375 个;不同位点的 PIC 值变化范围很大,在 0.092~0.879 之间,平均为 0.5098。聚类分析结果表明无性系分组与现有栽培区没有明显的规律。引物 Rply109 和 rops16 对 49 份无性系的分子鉴定率为 91.84%,可作为指纹图谱构建、分子鉴定的高效分子标记。

(6) 建立了以 SSR 标记为技术基础的刺槐无性系分子鉴定体系。利用 18 对引物可以区分 378 份刺槐属植物无性系,分子鉴定率为 96.36%。利用 18 对引物对收集到的国内外 383 份刺槐种质资源进行 NJ 分析,结果表明所有无性系并没有严格按照不同来源地区聚集到一起,但是又有一定的规律,大部分来源相同或相近的种质聚到一起。相同种源无性系没有聚到一起主要是由于互相引种和遗传变异造成的。

(7) 研究结果表明荧光 SSR 毛细管电泳检测技术是一种简便、快速、高效的初步鉴定植物是否是多倍体的方法。利用引物 RP13 检测出 293 个无性系有三至六个等位变异,占 76.1%,确定该引物为刺槐指纹图谱构建、分子鉴定、尤其是多倍体初步鉴定的高效分子标记。

(8) 本研究在利用 18 对荧光 SSR 引物构建指纹图谱时,发现 385 个无性系中有 343 个在 1 个位点或者多个位点(最多 7 个位点)出现了 3~6 个等位变异,占 89.09%。

关键词: 刺槐; SSR 分子标记; 遗传多样性; 指纹图谱; 聚类分析