

杨会肖, 刘天颐, 徐 斌, 等. 火炬松生长和松脂性状相关候选基因的单核苷酸多样性分析[J]. 华南农业大学学报, 2016, 37(1): 75-81.

火炬松生长和松脂性状相关候选基因的单核苷酸多样性分析

杨会肖^{1,2}, 刘天颐², 徐 斌¹, 刘纯鑫², 王金榜³, 黄少伟²

(1 广东省林业科学研究院, 广东 广州 510520; 2 广东省森林植物种质创新与利用重点实验室/华南农业大学林学与风景园林学院, 广东 广州 510642; 3 英德市林业科学研究所, 广东 英德 513000)

摘要:【目的】对火炬松 *Pinus taeda* 生长和松脂产量相关功能基因单核苷酸多态性 (Single nucleotide polymorphism, SNP) 位点进行筛选与分析, 为分子标记辅助育种提供技术基础。【方法】利用生物信息学对火炬松 cDNA 文库进行整理、比对、剪接、注释, 挑选出与生长素、赤霉素、细胞分裂素及萜烯合成相关的非重复序列基因 (Unigene) 作为候选基因片段, 利用 MEGA5.0 和 DnaSP4.0 软件对火炬松 36 个单株的 8 个候选基因片段进行序列比对和分析。【结果】所测序列总长为 5 177 bp, 检测到 184 个 SNP 位点, 平均 36.9 bp 的基因序列中出现 1 个 SNP 位点, 其中 123 个为非同义突变、61 个为同义突变 SNP 位点。核苷酸多态性 π_a 和 θ_w 分别为 0.020 和 0.016。对 8 个候选基因片段内 SNP 位点进行的连锁不平衡分析显示, 随着核苷酸序列长度的延伸, SNP 位点的连锁不平衡在基因内部迅速衰退 ($R^2 \leq 0.2$)。【结论】在火炬松中, 基于候选基因内 SNP 位点间的连锁不平衡作图是可行的。

关键词: 火炬松; 候选基因; 单核苷酸多态性; 连锁不平衡
中图分类号: S722.3 **文献标志码:** A **文章编号:** 1001-411X(2016)01-0075-07

Analysis of single nucleotide polymorphisms within candidate genes involved in growth and resin biosynthesis of loblolly pine

YANG Huixiao^{1,2}, LIU Tianyi², XU Bin¹, LIU Chunxin², WANG Jinbang³, HUANG Shaowei²

(1 Guangdong Academy of Forestry, Guangzhou 510520, China; 2 Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm/College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, China; 3 Yingde Institute of Forestry, Yingde 513000, China)

Abstract:【Objective】Single nucleotide polymorphisms (SNPs) within candidate genes involved in growth and resin biosynthesis of loblolly pine, *Pinus taeda*, were screened and analyzed, preparing technical basis for later marker-assisted breeding.【Method】The candidate unigenes related to auxin, gibberellins, cytokinins or pinene synthesis were selected from loblolly pine cDNA library using bioinformatics. The sequences of eight candidate genes from 36 loblolly pine individuals were aligned and compared using MEGA5.0 and DnaSP4.0 softwares.【Result】The total length of measured sequences was 5 177 bp. There were 184 SNPs to be detected, and in average there was one SNP site in every 36.9 bp sequence. Among them, 123 were nonsynonymous SNPs and 61 were silent SNPs. The levels of nucleotide diversity

收稿日期:2015-04-07 优先出版时间:2015-12-07
优先出版网址:<http://www.cnki.net/kcms/detail/44.1110.s.20151207.1133.028.html>
作者简介:杨会肖 (1981—), 女, 助理研究员, 博士研究生, E-mail: hxyang@sinogaf.cn; 通信作者: 黄少伟 (1964—), 男, 教授, 博士, E-mail: shwhuang@scau.edu.cn
基金项目: 广东省林业科技创新专项资金 (2011KJ CX013-02); “十二五”农村领域国家科技计划专题 (2012BAD01B0203); 联合国发展计划署援建的火炬松改良种子园项目 (CPR/91/153, 1992-1996)

for sequenced regions represented by π_a and θ_w were 0.020 and 0.016, respectively. The linkage disequilibrium (LD) of SNPs in the eight candidate genes declined rapidly ($R^2 \leq 0.2$) with the nucleotide sequences increased in length. 【Conclusion】 LD mapping of SNPs in the five candidate genes are useful for loblolly pine breeding.

Key words: loblolly pine; candidate gene; single nucleotide polymorphism (SNP); linkage disequilibrium (LD)

林木的重要性状大多属于数量性状,由许多基因和环境共同作用来调控其生长与发育。传统的数量遗传学不能检测控制数量性状的基因的数量、基因在染色体上的位置、每个位点的贡献大小以及基因间的关系。在数量性状的多基因体系中,尽管可能存在效应较大的主效基因,但多数基因对目标性状可能只表现微效作用,而且表型受环境影响很大,因而采用传统的育种途径对这些性状进行选择有很大的局限性^[1-2]。近年来,连锁分析和连锁不平衡作图已成为许多树种检测复杂性状的手段。关联分析是对控制数量性状的多基因系统进行遗传剖析的有效手段,而关联作图则是植物遗传学研究最重要的工具之一^[3]。对某一植物种而言,筛选出的单核苷酸多态性 (Single nucleotide polymorphism, SNP) 位点在进行关联分析之前,有必要对目标基因的群体遗传学特征如核苷酸多样性、基因进化动力及连锁不平衡特征进行分析,以便为 SNP 基因型分析位点的确定提供参考^[4]。

火炬松 *Pinus taeda* 材性偏软,具有粗糙的质感,适用于船舶楼宇及一般建筑。生长素 (IAA)、细胞分裂素与赤霉素在树木生长过程中起到重要的调节作用。火炬松含有较丰富的松脂,可为制药和化工等行业提供原料。当松树受到甲虫危害时,能在伤口处产生组成型或诱导型的松脂^[5],松脂是一种复杂的混合物,它包含单萜烯 (松节油),二萜类树脂酸 (松香) 和少量的倍半萜烯,而松节油的主要成分为蒎烯^[6]。目前,关于松树 SNP 标记的研究较多,主要集中在生长^[7]、材性^[8-9]、代谢^[10]、抗旱抗性^[4,11] 等方面,鲜见火炬松激素和蒎烯相关候选基因 SNP 标记研究的报道。本研究以生长和松脂产量为目标,从引自美国的火炬松基因资源试验林中选出 36 个单株为材料,并选择 8 个相应的候选基因片段进行研究。通过 PCR 扩增、纯化回收后连接于 T 载体上,对目的片段的 SNP 位点进行了筛选,分析了目标基因核苷酸多样性,连锁不平衡特征。研究结果为后

<http://xuebao.scau.edu.cn>

续 SNP 标记与表型性状的关联分析提供参考。

1 材料与方法

1.1 植物材料

通过美国北卡罗莱纳州立大学林木改良工业协作组织引进的 258 个不同改良程度的火炬松家系,营造了基因资源林,以生长 (树高和胸径) 和松脂产量为目标性状,利用 ASReml 软件估算了 258 个家系的生长和松脂产量育种值,选取能够最大程度地反映基因资源变异水平的 36 个家系,每家系 1 株,共 36 个单株 (36 个基因型),采集嫩芽或 1 年生新鲜松针, -80 ℃ 冰箱中保存备用。

1.2 基因组 DNA 提取

取出保存的松针或嫩芽,称取 0.3 g,放入 2 mL 离心管中,加入小钢珠,利用 Qiagen 的 MM301 磨样仪将其磨碎,期间用液氮浸泡数次。基因组 DNA 的提取按照新型植物基因组 DNA 提取试剂盒 (天根生化科技有限公司) 描述的方法进行。

1.3 引物设计和 PCR 扩增

从公共数据库 (GenBank 和 <http://pinetree.ccgb.umn.edu>) 下载火炬松 cDNA 文库的所有记录,约 32 万条,利用生物信息学方法进行序列的比对、剪接、组装、注释,获得非重复序列基因 (Unigene)。选取与生长、松脂蒎烯合成相关的 8 个候选基因片段,采用 Primer6 和 Oligo7 软件设计和评估正反向 PCR 扩增引物^[12]。

以选取的 36 个单株样本为材料,对 8 个候选基因片段进行 PCR 扩增。PCR 反应体系为 25 μL,含 30 ~ 60 ng 模板 DNA,每种 dNTP 0.2 mmol · L⁻¹,正、反向引物各 0.2 mmol · L⁻¹, 1 × PCR Buffer, 1 U *Taq* 聚合酶和体积分数为 1% 的 DMSO。采用 PTC-200 型 PCR 仪运行 PCR 反应,反应程序为: 94 ℃ 预变性 5 min 后进入循环;每个循环 94 ℃ 变性 30 s, 45 ~ 62 ℃ 退火 30 s, 72 ℃ 延伸 60 s, 反应共 36 个循环; 72 ℃ 延伸 7 min。

1.4 PCR 产物的克隆、测序及计算机分析

依据所选的 8 个候选基因片段的核苷酸序列设

计基因特异的引物,以 36 个样本的总 DNA 为模板进行 PCR 扩增,将扩增得到的目的基因片段纯化回收后连接于 pMD19-T 载体上。连接产物转化大肠埃希菌 *Escherichia coli* DH5 α ,筛选阳性克隆,送华大基因公司测序。应用 DNAMAN 6.0 软件和 CLUSTALX 软件包推导出氨基酸序列并在 NCBI 进行检索分析,初步判断候选序列位于编码区还是非编码区。将去除载体后的序列采用 BioEdit 的 ClustalW2 程序进行联配分析,获得基因片段的一致序列,结合每条序列测序图,在 BioEdit 中确定 SNP 位点并做记录。

1.5 基因的 SNP 多样性分析和连锁不平衡检测

利用 MEGA5.0 和 DnaSP4.0 软件对每个基因的 36 个序列进行比对分析,计算 SNP 频率、颠换和转换的 SNP 数量,估算各基因的核苷酸多样性水平^[13]。分别用参数 θ_w 和 π_a 表示核苷酸多样性水平,其中 θ_w 的估算以序列间每位点差异核苷酸的平均数为基础^[14], π_a 的估算以分离位点的数量(S)为

基础^[15]。单倍型多样性(H_e)按照 Nei^[14] 的公式计算。除以上定义外,核苷酸多样性的描述还包括:静默位点多样性(θ_{wsil})、非同义突变多样性(π_{nsyn})和同义突变多样性(π_{syn})。将所有位点依次对应的物理距离(bp)及其 R^2 进行组合,利用非线性回归模型估算 8 个候选基因片段整体水平上的连锁不平衡(LD)延伸。

2 结果与分析

2.1 候选基因片段的序列分析

8 个候选基因片段序列分析结果见表 1。每一候选基因片段的扩增长度为 431 ~ 1 007 bp(不包括插入或缺失位点),8 个候选基因片段的测序总长度为 5 177 bp,其中,编码区的序列长度为 3 322 bp,非编码区的序列长度为 1 855 bp(5'-UTR,3'-UTR 和内含子)。共有 24 个样本发生插入或缺失突变,总的插入片段长度为 18 bp,缺失片段长度为 80 bp。

表 1 火炬松生长和松脂性状候选基因片段序列分析
Tab.1 Sequence analysis of candidate genes involved in growth and resin traits of *Pinus taeda*

候选基因片段	基因功能	所获序列长度/bp					
		总长度	5'非编码区	外显子	内含子	3'非编码区	插入(缺失)长度
<i>Pt30-1</i>	α -蒎烯合成	767		568	199		5(37)
<i>Pt30-2</i>	α -蒎烯合成	1 007		472	535		6(10)
<i>Pt30-3</i>	α -蒎烯合成	922		529	393		3(25)
<i>Pt30-4</i>	α -蒎烯合成	464		309		155	4(8)
<i>CKX7_ARATH</i>	细胞分裂素	431		431	0		0
<i>ARR17_ARATH</i>	细胞分裂素	453		453	0		0
<i>COGT2_ARATH</i>	细胞分裂素	527		413	114		0
<i>SSGP_VOLCA</i>	生长素	606	134	147	325		0
总计		5 177	134	3 322	1 566	155	18(80)

2.2 SNP 位点筛选

SNP 位点筛选和分析结果见表 2,对 8 个候选基因片段的序列进行整合后共检测到 184 个 SNP 位点,平均 36.9 bp 的基因序列中出现 1 个 SNP 位点。*SSGP_VOLCA* 基因片段上检测了 41 个 SNP 位点,平

均 14.8 bp 的基因序列中出现 1 个 SNP 位点,而在 *Pt30-2* 基因片段上检测了 18 个 SNP 位点,平均 55.9 bp 的基因序列中出现 1 个 SNP 位点。非同义突变的位点有 123 个,静默突变(同义突变)有 61 个 SNP 位点。

表 2 火炬松生长和松脂候选基因片段序列核苷酸变异
Tab.2 Summary of SNPs in candidate genes in *Pinus teada* for growth and resin traits

候选基因片段	长度/bp	SNP 数量	SNP 频率/(个 $\cdot$ bp ⁻¹)	非同义 SNP 数量	静默 SNP 数量
<i>Pt30-1</i>	767	14	1: 54.8	10	4
<i>Pt30-2</i>	1 007	18	1: 55.9	11	7
<i>Pt30-3</i>	922	56	1: 16.5	38	18
<i>Pt30-4</i>	464	9	1: 51.6	8	1
<i>CKX7_ARATH</i>	431	9	1: 47.9	7	2
<i>ARR17_ARATH</i>	453	16	1: 28.3	6	10
<i>COGT2_ARATH</i>	527	21	1: 25.1	14	7
<i>SSGP_VOLCA</i>	606	41	1: 14.8	29	12
总计	5 177	184	1: 36.9 ¹⁾	123	61

1) 平均值。

2.3 核苷酸多样性

采用 MEGA5.0 和 DnaSP4.0 软件对各候选基因 36 个样本序列进行分析,计算 8 个候选基因的核苷酸多样性系数(表 3)。8 个候选基因平均核苷酸多态性 π_a 和 θ_w 分别为 0.020 和 0.016,处于较高水平,其中核苷酸多态性 π_a 的变化范围介于 0.006 ~ 0.034,核苷酸多态性 θ_w 的变化范围介于 0.003 ~ 0.030。这比先前报道的相同树种其他性状或不同树种同类性状相关基因的核苷酸多态性高(表 4)。有报道显示,火炬松的材性和生长性状相关基因的核苷酸多态性平均为 0.005 9^[4,8,16],是本研究的 0.3 倍。Pot 等^[17]报道,海岸松 *P. pinaster* 和辐射松 *P. radiata* 材性方面的核苷酸多态性是 0.002 41 和 0.001 86,分别是本研究的 0.12 和 0.09 倍。核苷酸多态性的研究也在其他针叶树种中有所报道,例如樟子松 *P. sylvestris* var. *mongolica*^[18-19],花旗松 *Pseudotsuga menziesii*^[9,16,20] 和柳杉 *Cryptomeria japonica*^[21]。

在 8 个候选基因片段中,具有最高核苷酸多态性 π_a 和 θ_w 的基因片段分别是 *ARR17_ARATH* (π_a =

0.034) 和 *SSGP_VOLCA* (θ_w = 0.030),具有最低核苷酸多态性 π_a 和 θ_w 的基因片段是 *Pt30-4* 和 *CKX7_ARATH* (π_a = 0.006, θ_w = 0.003)。静默位点的平均核苷酸多样性(θ_{wsil})为 0.022,该值与非同义突变位点多样性(π_{nsyn} = 0.014)基本相同。本研究与同一树种的其他性状如抗旱性、材性和抗性的候选基因相比,核苷酸多样性 θ_w 高出其他相关研究 3.2 和 4.0 倍^[4,8]。基因片段 *COGT2_ARATH* 的非同义突变多样性(π_{nsyn})是基因片段 *CKX7_ARATH* 的 14.5 倍,非同义突变位点多样性在 0.029 ~ 0.002 之间变化(表 3),这表明基因间在适应性进化历史中变化较大^[22]。基因片段 *Pt30-2*、*Pt30-4* 和 *CKX7_ARATH* 的非同义分歧很低,即 0.002 ~ 0.004,表明这些基因具有很强的保守核苷酸序列。8 个候选基因片段中的 2 个基因片段(*Pt30-1* 和 *Pt30-4*)的 K_a/K_s 比值大于 1,说明这些基因受正向选择的影响较大。其余基因片段的 K_a/K_s 比值都小于 1,说明他们受负向选择的影响较大,核苷酸序列相对保守。

表 3 火炬松生长和松脂性状候选基因片段核苷酸多样性分析¹⁾

Tab.3 Estimates of nucleotide diversities of candidate genes for growth and resin traits of *Pinus taeda*

候选基因片段	核苷酸多样性			同义/非同义/静默位点核苷酸多样性							单倍型多样性		R_m
	S	θ_w	π_a	π_{syn}	π_{nsyn}	π_{nsyn}/π_{syn}	θ_{wsil}	K_s	K_a	K_a/K_s	N_h	H_e (SD)	
<i>Pt30-1</i>	14	0.012	0.023	0.011	0.012	1.061	0.012	0.010	0.011	1.061	13	0.989 ± 0.031	1
<i>Pt30-2</i>	18	0.005	0.007	0.009	0.004	0.464	0.009	0.008	0.004	0.464	6	0.952 ± 0.096	0
<i>Pt30-3</i>	56	0.023	0.021	0.028	0.022	0.768	0.028	0.026	0.020	0.768	11	1.000 ± 0.039	1
<i>Pt30-4</i>	9	0.003	0.006	0.002	0.004	2.385	0.002	0.001	0.003	2.385	7	0.773 ± 0.128	0
<i>CKX7_ARATH</i>	9	0.003	0.006	0.005	0.002	0.370	0.003	0.005	0.002	0.370	7	0.625 ± 0.139	0
<i>ARR17_ARATH</i>	16	0.022	0.034	0.055	0.013	0.229	0.055	0.053	0.012	0.229	21	0.975 ± 0.021	2
<i>COGT2_ARATH</i>	21	0.029	0.030	0.030	0.029	0.952	0.030	0.027	0.026	0.952	11	0.985 ± 0.040	4
<i>SSGP_VOLCA</i>	41	0.030	0.032	0.040	0.028	0.693	0.040	0.037	0.026	0.693	14	0.971 ± 0.024	7
平均值	23	0.016	0.020	0.023	0.014	0.865	0.022	0.021	0.013	0.865	11.25	0.909 ± 0.065	1.875

1) S 为分离位点的数量, θ_w 、 π_a 为总核苷酸多样性, π_{syn} 、 π_{nsyn} 分别为同义和非同义突变多样性, θ_{wsil} 为静默位点多样性, K_s 为编码区同义突变平均数, K_a 为编码区非同义突变平均数, N_h 为单倍型个数, H_e 为单倍型多样性, R_m 为历史重组事件最小数。

单倍型的平均数(N_h)和单倍型多样性(H_e)平均值分别为 11.25 和 0.909。在 8 个候选基因片段中发现核苷酸数和核苷酸多样性有较大的变化,其中单倍型个数的变化范围介于 6 ~ 21,单倍型多样性变化范围介于 0.625 ~ 1.000。采用四配子检测进行最小历史重组事件(R_m)估算,结果表明 8 个候选基因片段的 R_m 为 1 ~ 7。

2.4 连锁不平衡和重组

将 8 个候选基因片段的位点间连锁不平衡数据(R^2)汇总后进行综合分析,期望从总体上了解火炬

松基因组中连锁不平衡情况。对 R^2 值的非线性回归分析结果见图 1,随着候选基因片段核苷酸序列长度的增加,SNPs 的连锁不平衡程度迅速削弱,当长度达 400 bp 左右时, $R^2 < 0.2$,连锁不平衡迅速衰退。这一结果表明,在火炬松中,SNP 的连锁不平衡在候选基因片段内部就已衰退。这与其他针叶树种的报道一致^[8-9]。连锁不平衡衰退率在不同的植物中的变化较大,玉米 *Zea mays* 的候选基因片段长度为 1 500 bp,当在 200 bp 左右时, R^2 降低到 0.2 以下;拟南芥 *Arabidopsis thaliana* 的候选基因片段长度为

表 4 不同性状的候选基因核苷酸多样性在其他林木中的应用¹⁾

Tab. 4 Nucleotide diversities reported for candidate genes involved in different traits of various forest species								
树种	DNA 部位	基因数量	基因总长度/bp	θ_w	π_{nsyn}	π_{syn}	H_e	参考文献
火炬松 <i>Pinus taeda</i>	针叶	5	5 177	0.016	0.014	0.023	0.909	本研究
	单倍体的大配子体	18	10 116	0.005	0.002	0.009	0.680	[4]
	单倍体的大配子体	19	18 027	0.004	0.001	0.006		[8]
	单倍体的大配子体	28		0.005	0.002	0.007		[16]
海岸松 <i>Pinus pinaster</i>	单倍体的大配子体	10	4 746	0.002	0.000		0.425	[17]
辐射松 <i>Pinus radiata</i>	单倍体的大配子体	10	4 746	0.002	0.000		0.376	[17]
樟子松 <i>Pinus sylvestris</i> var. <i>mongolica</i>	单倍体的大配子体	2	4 136	0.001	0.000			[18]
	单倍体的大配子体	1	2 043	0.001	0.000		0.359	[19]
花旗松 <i>Pseudotsuga menziesii</i>	单倍体的大配子体	18	15 183	0.007	0.002	0.011	0.820	[9]
	单倍体的大配子体	12		0.009	0.005	0.013		[16]
	单倍体的大配子体	121	59 173	0.004	0.002	0.008		[20]
欧洲山杨 <i>Populus tremula</i>	叶	5	6 188	0.011	0.006	0.016		[22]
黑杨 <i>Populus nigra</i>	叶	9	8 056	0.007	0.005	0.009	0.931	[23]
	叶	9	10 135	0.004	0.002	0.005	0.894	[23]
柳杉 <i>Cryptomeria japonica</i>	萌发后的大配子体	10		0.002			0.486	[21]
日本扁柏 <i>Chamaecyparis obtusa</i>	萌发后的大配子体	10		0.002			0.453	[21]
美国水松 <i>Taxodium distichum</i>	单倍体的大配子体	10	8 623	0.003		0.007		[24]

1) θ_w 为总核苷酸多态性, π_{nsyn} 和 π_{syn} 分别为非同义和同义突变多样性, H_e 为单倍型多样性。

100 000 bp, 当在 50 000 bp 左右时, R^2 降低到 0.2 以下^[25]。

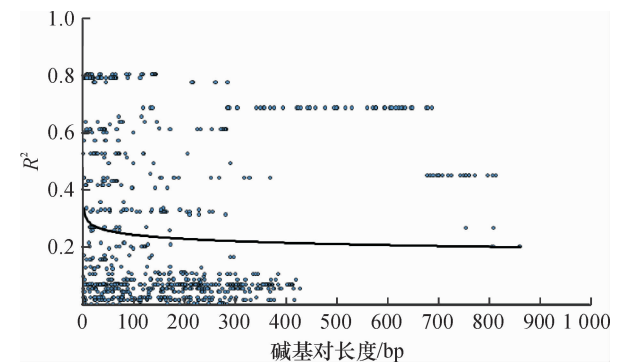


图 1 8 个候选基因片段总体连锁不平衡分析的 R^2 随碱基对长度衰减情况

Fig.1 Scatter plot of distances in base pairs and linkage disequilibrium estimates (R^2) between SNP sites in eight candidate genes in *Pinus taeda*

3 讨论与结论

火炬松是南方商品林的主要树种之一,适宜在粤北和粤东北低山种植。火炬松的遗传改良时间较长,目前正处在第 2 代向第 3 代过渡阶段。在火炬松第 1、2 代育种过程中,生长量获得显著提高,但忽略了对产脂量的选择。为了适应林业生产对松树多用途的需求,本研究以高产脂和速生为目标,选择了 8 个候选基因片段,其中 4 个基因片段与松脂合成有

关,4 个基因片段与火炬松生长(包括生长素和细胞分裂素)有关,将 8 个候选基因片段进行 PCR 扩增测序,共得到 5 177 bp 的基因片段序列,其中编码区的序列长度为 3 322 bp,非编码区的序列长度为 1 855 bp。非同义突变 SNP 位点(123)多于静默 SNP 位点(61)。8 个候选基因片段的 SNP 筛选分析研究表明,编码区的 SNP 数量高于非编码区,与其他松树的 SNP 分析结果一致,表明火炬松的编码区比其他部位具有更高的保守性。但是不同候选基因的 SNP 频率不一样,说明各个基因的保守性不同,可能是由于在进化过程中受到的选择压力不同所致。

8 个候选基因片段的平均核苷酸多样性 π_a 和 θ_w 分别为 0.020 和 0.016,与先前报道的核苷酸多样性相比,本研究 8 个候选基因片段的核苷酸多样性较高,这可能是因为材料具独立的变迁历史和地理分布,不同的遗传基础造成了核苷酸多样性的差异,今后应更大范围取样以便更全面地了解该树种的遗传多样性水平。另一部分原因可能是本研究所用的材料为针叶部位,这个部位的基因组比大配子体复杂。8 个候选基因片段的单倍型个数(N_h)和单倍型多样性(H_e)平均值分别为 11.25 和 0.909,单倍型多样性的平均标准误为 0.065, R_m 为 1~7。本文检测到了较高程度的核苷酸多样性,表明利用 SNP 标记对火炬松进行分子标记辅助选择(MAS)需要较多的

SNP 标记。

本研究发现,随着 LD 衰退距离的增加,总体趋势基本平稳,但火炬松 8 个候选基因片段总体的 LD 水平稍高,这可能是由于所选样品数量较少。Kim 等^[26]提出当个体遗传背景差异较小时,较大的样本量一般具有较低的 LD 水平。大量研究表明林木具有较低的 LD 水平,适宜基于候选基因片段的连锁不平衡作图研究。在杨树中,基于 SNP 的连锁不平衡检测发现,杨属 *Populus* 近缘种的候选基因片段内 LD 衰退距离为 300 ~ 1 700 bp 不等^[22,27-29]。Thumma 等^[30-31]发现亮果桉 *Eucalyptus nitens* LD 水平在基因 *CCR* 和 *COBRA-like* 内已衰退至不明显。在针叶树种中,基于 SNP 标记的 LD 水平在几千个碱基长度内已衰退至不明显^[8-10,32],这些研究结果与本研究的 LD 检测基本一致,揭示了基于候选基因片段的关联遗传研究是可行的。本研究进行的连锁不平衡分析仅限于候选基因片段内部,火炬松基因间隔区域、不同染色体及全基因组水平的 LD 需要更多的标记去验证。随着火炬松全基因组测序、拼接、注释完成及重测序研究的开展,该物种基因组水平的 LD 衰退趋势将会得到更深入全面的检测。

参考文献:

- [1] 黄少伟,钟伟华,陈炳铨. 火炬松半同胞子代配合选择的遗传增益估算[J]. 林业科学,2006,42(4):33-37.
- [2] 汤在祥,陈志军,王学枫,等. 标记辅助选择育种中 QTL 基因型的多点联合推断[J]. 分子植物育种,2006,4(2):293-298.
- [3] SIMKO I, HAYNES K G, JONES R W. Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers[J]. Genetics, 2006,173(4):2237-2245.
- [4] GONZA'LEZ-MARTI'NEZ S C, ERSON E, BROWN G R, et al. DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L[J]. Genetics, 2006, 172(3):1915-1926.
- [5] ROBERDS J H, STROM B L, HAIN F P, et al. Estimates of genetic parameters for oleoresin and growth traits in juvenile loblolly pine[J]. Can J Forest Res, 2003, 33(12):2469-2476.
- [6] PHILLIPS M A, WILDUNG M R, WILLIAMS D C, et al. cDNA isolation, functional expression, and characterization of (+)- α -pinene synthase from loblolly pine (*Pinus taeda*): Stereocontrol in pinene biosynthesis [J]. Arch Biochem Biophys, 2003, 411(2):267-276.
- [7] CUMBIEL W P, ECKERT A, WEGRZYN J, et al. Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. [J]. Heredity, 2011, 107(1):105-114.
- [8] BROWN G R, GILL G P, KUNTZ R J, et al. Nucleotide diversity and linkage disequilibrium in loblolly pine [J]. Proc Natl Acad Sci USA, 2004, 101(42):15255-15260.
- [9] KRUTOVSKY K V, NEALE D B. Nucleotide diversity and linkage disequilibrium in cold-hardiness and wood quality-related candidate genes in *Douglas fir* [J]. Genetics, 2005, 171(4):2029-2041.
- [10] GONZA'LEZ-MARTNEZ S C, WHEELER N C, ERSOZ E, et al. Association genetics in *Pinus taeda* L.: I: Wood property traits [J]. Tree Genet Genom, 2007, 175(1):399-409.
- [11] ECKERTA J, WEGRZYN J L, CUMBIE W P, et al. Association genetics of the loblolly pine (*Pinus taeda*, Pinaceae) metabolome [J]. New Phytol, 2012, 193(4):890-902.
- [12] 杨会肖. 火炬松基因资源的遗传分析及 SNP 位点筛选 [D]. 广州:华南农业大学,2013.
- [13] ROZAS J, SA'NCHEZ-DEL-BARRIO J C, MESSEGUER X, et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods [J]. Bioinformatics, 2003, 19(18):2496-2497.
- [14] NEI M. Molecular evolutionary genetics [M]. New York: Columbia University Press, 1987.
- [15] WATTERSON G A. On the number of segregating sites in genetical models without recombination [J]. Theor Popul Biol, 1975, 7(2):256-276.
- [16] NEALE D B, SAVOLAINEN O. Association genetics of complex traits in conifers [J]. Trends Plant Sci, 2004, 9(7):325-330.
- [17] POT D, MCMILLAN L, ECHT C, et al. Nucleotide variation in genes involved in wood formation in two pine species [J]. New Phytol, 2005, 167(1):101-112.
- [18] GARCÍA-GIL M R, MIKKONEN M, SAVOLAINEN O. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris* [J]. Mol Ecol, 2003, 12(5):1195-1206.
- [19] DVORNYK V, SIRVIÖ A, MIKKONEN M, et al. Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestris* [J]. Mol Biol Evol, 2002, 19(2):179-188.
- [20] ECKERT A J, BOWER A D, WEGRZYN J L, et al. Association genetics of coastal douglas fir (*Pseudotsugamenziesii* var. *menziesii*, Pinaceae): I: Cold-hardiness related traits [J]. Genetics, 2009, 182(4):1289-1302.
- [21] KADO T, MATSUMOTO A, UJINO-IHARA T, et al. A-

mounts and patterns of nucleotide variation within and between two Japanese conifers, sugi (*Cryptomeria japonica*) and hinoki (*Chamaecyparis obtusa*) (Cupressaceae *sensu lato*) [J]. *Tree Genet Genom*, 2008, 4(1): 133-141.

[22] INGVARSSON P K. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae) [J]. *Genetics*, 2005, 169(2): 945-953.

[23] 褚延广. 欧洲黑杨(*Populus nigra* L.)水、光资源高效利用相关单核苷酸多态性(SNP)研究 [D]. 北京: 中国林业科学研究院, 2009.

[24] KUSUMI J, LI Z D, KADO T, et al. Multilocus patterns of nucleotide polymorphism and demographic change in *Taxodium distichum* (Cupressaceae) in the lower Mississippi River alluvial valley [J]. *Am J Bot*, 2010, 97(11): 1848-1857.

[25] RAFALSKI A, MORGANTE M. Corn and humans: Recombination and linkage disequilibrium in two genomes of similar size [J]. *Trends Genet*, 2004, 20(2): 103-111.

[26] KIM S, PLAGNOL V, HU T T, et al. Recombination and linkage disequilibrium in *Arabidopsis thaliana* [J]. *Nat Genet*, 2007, 39(8): 1151-1155.

[27] INGVARSSON P K, GARCIA M V, LUQUEZ V, et al. Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae) [J]. *Genetics*, 2008, 178(4): 2217-2226.

[28] WĘGRZYN J L, ECKERT A J, CHOI M, et al. Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem [J]. *New Phytol*, 2010, 188(2): 515-532.

[29] XU B, YANG X, LI B, et al. Isolation, expression and single nucleotide polymorphisms analysis of cellulose synthase gene (*PtCesA4*) from *Populus tomentosa* [J]. *Scientia Silvae Sinicae*, 2009, 45(5): 1-10.

[30] THUMMA B R, MATHESON B A, ZHANG D Q, et al. Identification of a *Cis*-acting regulatory polymorphism in a eucalypt *COBRA*-like gene affecting cellulose content [J]. *Genetics*, 2009, 183(3): 1153-1164.

[31] THUMMA B R, NOLAN M F, EVANS R, et al. Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in micro fibril angle in *Eucalyptus* spp. [J]. *Genetics*, 2005, 171(3): 1257-1265.

[32] DVOMYK V, SIRVIO A, MIKKONEN M, et al. Low nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris* [J]. *Mol Biol Evol*, 2002, 19: 179-199.

【责任编辑 李晓卉】