

# 马尾松和黄山松两个核基因位点的群体遗传多样性和种间分化

张利锐 彭艳玲 任广朋 周永锋 李忠虎 刘建全\*

兰州大学生命科学学院, 干旱与草地生态教育部重点实验室分子生态研究所, 兰州 730000

**摘要** 利用两个核基因座位 $C3H$ 和 $GI$ , 对重叠分布于中国东南部的两个松属(*Pinus*)物种马尾松(*P. massoniana*)和黄山松(*P. hwangshanensis*)的22个群体88个个体进行了遗传多样性和种间分化模式研究。在这两个核基因座位上, 两种植物都表现出较低的核苷酸多样性水平(马尾松 $\pi_{sil} = 0.001\ 71$ ; 黄山松 $\pi_{sil} = 0.003\ 40$ ), 但是马尾松要显著低于黄山松; 在种内分化水平上, 马尾松的种内遗传分化也明显低于黄山松(马尾松 $F_{ST} = 0.059$ ; 黄山松 $F_{ST} = 0.339$ )。这可能是由于黄山松的海拔分布高于马尾松, 而高海拔分布使黄山松的分布区域更加片段化, 促使其形成较高的种内遗传多样性和遗传分化。分子变异分析(AMOVA)发现, 两物种基于两个核基因座位的种间差异为48.86%, 而 $GI$ 基因座位上的种间差异明显高于 $C3H$ 座位( $GI$ : 77.24%,  $C3H$ : 20.48%), 同时, 基因谱系显示两物种的共享单倍型仅在 $C3H$ 座位上存在。结合这两个基因的功能, 推测 $GI$ 基因可能在物种形成过程中受到了一定的选择压力, 因为 $GI$ 基因参与调控植物的开花时间, 而 $C3H$ 与木质素表达水平的调控有关。不同的选择压力使得 $GI$ 的进化速度相对较快, 从而加速了黄山松和马尾松的物种分化。

**关键词** 核苷酸多样性, 黄山松, 马尾松, 群体遗传结构, 物种分化

## Population genetic diversity and species divergence of *Pinus massoniana* and *P. hwangshanensis* at two nucleotide loci

ZHANG Li-Rui, PENG Yan-Ling, REN Guang-Peng, ZHOU Yong-Feng, LI Zhong-Hu, and LIU Jian-Quan\*

Institute of Molecular Ecology, Key Laboratory of Arid and Grassland Ecology of Ministry of Education, School of Life Sciences, Lanzhou University, Lanzhou 730000, China

### Abstract

**Aims** Our objective was to examine population genetic diversity and species divergence of two pine species (*Pinus massoniana* and *P. hwangshanensis*) with overlapping distributions in southeastern China.

**Methods** We estimated levels of nucleotide diversity and an analysis of molecular variance (AMOVA) of the two species based on sequence data of two nuclear loci for 88 individuals from 22 populations of the two species across most of their range of distribution. Gene genealogies of each locus were constructed by coalescent simulations using the Median-Joining model.

**Important findings** Both species exhibit low levels of nucleotide diversity at two nucleotide loci, and the level of silent nucleotide diversity is two times higher in *P. hwangshanensis* ( $\pi_{sil} = 0.003\ 40$ ) than that in *P. massoniana* ( $\pi_{sil} = 0.001\ 71$ ). The population differentiation ( $F_{ST}$ ) is also significantly different between the two species (*P. massoniana*, 0.059; *P. hwangshanensis*, 0.339,  $p < 0.05$ ). These genetic differences in the population structure of the two species may result from their differences in distribution and habitat preference. Hierarchical AMOVA revealed that the average of variance components between species is 48.86% based on the two loci and is significantly higher at locus  $GI$  (77.24%) than locus  $C3H$  (20.48%). In addition, shared haplotypes were only observed in  $C3H$  rather than  $GI$ . Thus we speculate that  $GI$  (control of flowering time) may have experienced speciation-related selection, which further accelerated its lineage-sorting divergence between the species.

**Key words** nucleotide diversity, *Pinus hwangshanensis*, *P. massoniana*, population genetics structure, species divergence

遗传结构是一个物种或种内群体在空间上的遗传学组成及分布状况, 即种内不同群体之间和同

一群体内不同个体间的遗传变异(胡新生, 2002), 群体遗传结构差异是遗传多样性的一种重要体现。生

物在长期进化的过程中, 很多因素(包括突变、交配系统、遗传漂变、基因流动以及自然选择等)都会影响群体内和群体间的核苷酸多样性水平, 继而影响物种的遗传多样性(Arunyawat *et al.*, 2007), 因此, 对物种遗传多样性的研究可用于推断过去的种群历史和选择事件。比较近缘物种之间的遗传结构不仅可以研究生境差异对物种群体遗传结构的影响, 更为重要的是, 可研究不同基因座位在物种分化中的谱系筛选状况(Rieseberg *et al.*, 2007)。由于谱系筛选不完全和杂交渗入, 近缘物种在一些基因座位上可能共享相同的基因型, 而在另外一些座位上可能呈现物种之间完全的遗传分化(Belahbib *et al.*, 2001; Mallet, 2007)。因而, 研究物种间不同基因位点的分化程度和谱系筛选状况对理解物种形成过程和寻找相关的物种形成基因具有重要意义(Strasburg & Rieseberg, 2008; Li *et al.*, 2010b; Zheng & Ge, 2010)。物种间完全的生殖隔离可能会导致物种间整个基因组水平的同时分化(Mayr, 1942), 但是物种形成过程中往往并非是这样状况, 而是在种间分化的过程中伴随有基因流存在(Orr *et al.*, 2004)。无论在何种情况下, 与物种形成相关或者连锁的基因可能在物种分化过程中因为选择作用的影响而具有较快的进化速度, 从而使种间的谱系筛选速度更快或者分化更为明显(Wu, 2001)。目前仅有少数与植物物种生殖隔离相关的基因被报道, 如: 控制细胞质雄性不育杂种不育基因(*CMS*) (Chase, 2007)、授粉媒介变化相关花色的*myb*型转录因子(Hoballah *et al.*, 2007)以及杂交坏死相关基因(Bombliès & Weigel, 2007)。它们的突变会导致新物种的形成, 从而也可能使物种在突变基因上的遗传变异更具有物种特异性。但是, 对于一些参与物种形成或者物种形成过程中受到强烈选择的基因(如影响开花时间的基因)在物种谱系筛选中的分化, 目前还了解较少。

随着DNA测序技术的发展, 基于核基因对松柏类植物遗传多样性和物种分化的研究相继出现(Kado *et al.*, 2003; Ma *et al.*, 2006; Li *et al.*, 2010a), 但是这些研究结果都表明松柏类植物的群体遗传结构与以往研究比较成熟的模式植物拟南芥(*Arabidopsis thaliana*)和玉米(*Zea mays*)等不同(Wright & Gaut, 2005)。松柏类植物较大的有效群体大小、较长的世代时间以及基因渐渗, 致使松柏类

植物种间具有大量共享的单倍型以及低水平的种间遗传分化。因此, 推断此类植物的进化历史和物种形成过程的相关工作极具挑战性(Willyard *et al.*, 2009; Chen *et al.*, 2010)。我们这里选择两个可能在物种形成过程中具有不同作用的基因来研究它们在两种松属植物中的谱系筛选和物种分化。

马尾松(*Pinus massoniana*)和黄山松(*P. hwangshanensis*)是近乎重叠地分布于中国东南部的两个松科近缘物种, 两物种在形态学、细胞学以及木材解剖学上都不同, 并且在其分布区的环境和空间上显示明显的分离。曾认为黄山松应作为台湾松(*P. taiwanensis*)的异名(Fu *et al.*, 1999), 但是染色体和分子研究的证据表明, 它与台湾松和琉球松(*P. luchuensis*)存在明显区别, 并且亲缘关系较远(与马尾松的亲缘关系更近), 应作为一个独立的物种(邹惠渝等, 2003; Chiang *et al.*, 2006)。经过多年的野外调查, 发现黄山松分布于海拔1 000 m以上, 而马尾松分布于700 m以下, 在700–1 000 m马尾松和黄山松交接处普遍存在两者的天然杂交群体, 这些杂交群体的属性已经分别在形态水平和分子水平上得到验证(邢有华等, 1992; 罗世家等, 2001)。这两个物种在生态位的主要差别在于海拔, 随着海拔的升高, 环境因子如氧分压、空气温度和湿度、土壤温度和湿度、日照和紫外线强度都随之改变。这些环境因子与植物生长和适合度密切相关, 它们的改变可引起植物表型及适合度的分化, 并保持物种的特异性特征(Li *et al.*, 2010b)。马尾松和黄山松二者不存在完全的谱系筛选和生殖隔离(Li *et al.*, 2010b; Zhou *et al.*, 2010), 是研究不同核基因在物种间遗传分化与谱系筛选的较好模式。我们选用了分布于中国东南部不同海拔的马尾松和黄山松10个同域分布的群体(黄山松和马尾松各5个)以及12个邻域分布群体(7个马尾松群体, 5个黄山松群体)。选择的两个核基因分别是*C3H*和*GI*。*C3H*是与植物木质素形成有关的基因, 对木质素的合成具有下调作用(Ralph *et al.*, 2006)。*GI*参与植物开花时间的调控, 该基因位点突变导致开花时间存在显著差异(Huq *et al.*, 2000)。本研究的目的是: 1)基于两个核基因位点估算黄山松和马尾松的核苷酸多态性以及这些核苷酸多态性呈现的群体遗传结构, 并与其他松柏类植物进行比较, 进而推测形成这些模式的可能的进化动力; 2)回答是否因为这两个基因在物种形成

过程中的功能不同, 导致了它们在物种间的谱系筛选和遗传分化存在显著差异。

## 1 材料和方法

### 1.1 群体取样

本实验选取了来自12个马尾松群体和10个黄山松群体共88个个体的种子, 其中包括同域分布的马尾松和黄山松群体各5个, 邻域分布的7个马尾松群体和5个黄山松群体, 每个群体取样2到4个不等, 见图1A, 采集的种子置于4 °C冰箱储存备用。

### 1.2 胚乳总DNA的提取、扩增以及测序

将收集的种子浸泡于蒸馏水中, 在4 °C放置过

夜后, 剥皮取得胚乳, 再利用优化的CTAB法提取胚乳的总DNA (Doyle & Doyle, 1987)。马尾松和黄山松的胚乳与其卵细胞一样, 都为单倍体, 因此, 直接扩增胚乳单拷贝基因就可获得不同的等位基因。我们对马尾松和黄山松的两个单拷贝核基因座位 *C3H* 和 *GI* 进行了测序分析。利用引物设计软件 Primers (Dumolin-Lapègue *et al.*, 1997) 设计了两个核基因座位的引物序列, 详细信息见表1。

测序反应和测序纯化反应使用ABI公司提供的 Big Dye Terminator (Applied Biosystems, Foster City, CA, American) 试剂和实验手册, 测序在ABI 3130x1 PCR 自动DNA测序仪 (Applied Biosystems, Foster City, CA, American) 上进行。测得序列利用MEGA v. 3.1 (Kumar *et al.*, 2004) 和CodonCode Aligner v. 3.0 (www.codoncode.com/aligner/download.htm) 进行比对和人工校正两次, 得到不同的DNA序列类型, 即不同的单倍型。

### 1.3 数据分析

#### 1.3.1 核苷酸多态性

利用DnaSP v. 5.0 (Rozas *et al.*, 2003) 计算了两个物种的核苷酸多态性参数: 分离位点数目  $S$ 、Watterson参数  $\theta_w$  (Watterson, 1975)、核苷酸多样性  $\pi$  (Tajima, 1983) 以及单倍型数目  $h$  和单倍型多样性  $H_d$  (Nei, 1987)。采用Hudson和Kaplan (1985) 的四配子检验法, 计算了物种内发生的最小历史重组事件的数目 ( $R_m$ )。

#### 1.3.2 中性检测

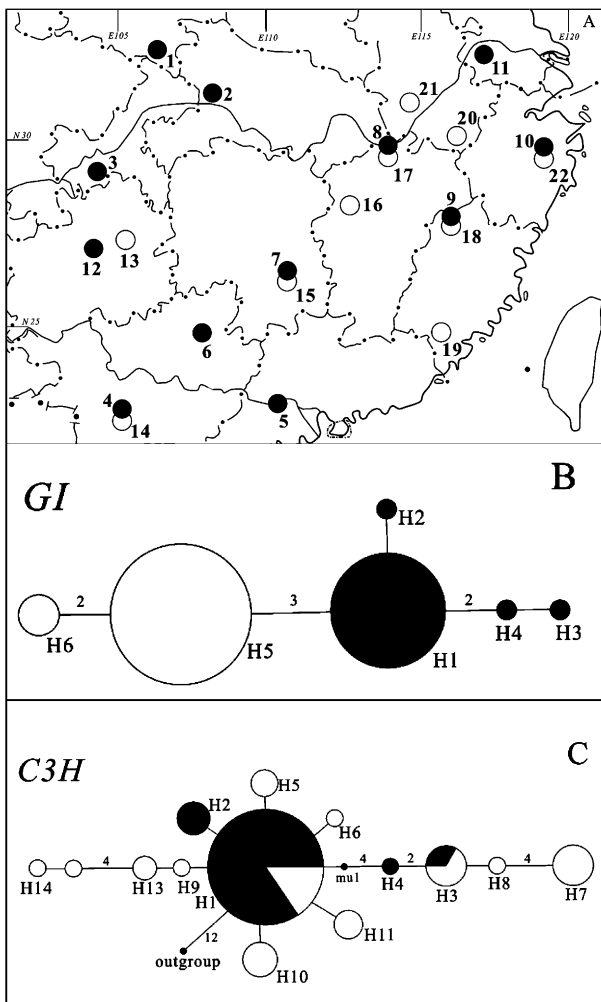
利用DnaSP v. 5.0 (Rozas *et al.*, 2003) 分别估算了两个物种每个座位的中性检验统计量Tajima's  $D$  (Tajima, 1989)、Fu和Li's  $D$  和  $F$  统计 (Fu & Li, 1993) 参数, 以判断这些座位是否符合中性进化模式。

#### 1.3.3 遗传结构

我们用Wright的固定指数  $F_{ST}$  (Wright, 1951) 检测每个物种群体内的遗传差异程度。  $F_{ST}$  的值用ARLEQUIN version 3.1 (http://cmpg.unibe.ch/software/arlequin3/) 软件包中的AMOVA (analysis of molecular variance) (Excoffier *et al.*, 2005) 分析方法运算得到, 重复10 000次以取得显著性统计参数。

#### 1.3.4 基因谱系

基因谱系反映了群体等位基因间的遗传关系, 它受自然选择、有效群体大小、群体亚结构等诸多因素影响, 可以从基因水平上反映群体进化历史等



审图号: GS(2011)239号

- 黄山松 *P. hwangshanensis*
- 马尾松 *P. massoniana*

图1 马尾松和黄山松群体采样地点分布(A)和核基因座位 *C3H* 和 *GI* 的基因谱系(B, C)。

Fig. 1 Sampling localities of *Pinus massoniana* and *P. hwangshanensis* (A) and gene genealogies of nuclear loci *C3H* and *GI* (B, C).

表 1 核基因座位 *C3H* 和 *GI* 描述

Table 1 Description of the *C3H* and *GI* locus

基因座位 Gene locus	可能的功能 Putative function	PCR 引物序列和退火温度 PCR primers and annealing temperature	长度 Aligned length	编码区 Coding region	非编码区 Non-coding region
<i>C3H</i>	下调木质素 Down-regulation on lignin	F: 5'-CAITTCCTGCGCCCTAAGAT-3' R: 5'-TTGCCAGCCTTGTGAGA-3' 60 °C	827 bp	外显子 Exon (1-93)H(492-827)	内含子 Intron 94-491
<i>GI</i>	控制开花时间 Control flowering time	F: 5'-CACAGTTCACACAGCAAC-3' R: 5'-ACAAAGCGGTTCAGTGGCTC-3' 53 °C	522 bp	1-522	

表 2 马尾松和黄山松核基因座位 *C3H* 和 *GI* 核苷酸多样性和中性检测

Table 2 Nucleotide diversity and neutrality tests over the *C3H* and *GI* locus in *Pinus massoniana* and *P. hwangshanensis*

基因座位 Gene locus	<i>n</i>	<i>S</i>	$\pi$ ( $\times 10^{-3}$ )	$\theta_w$ ( $\times 10^{-3}$ )	<i>h</i>	<i>H<sub>d</sub></i>	$\pi_{sil}$ ( $\times 10^{-3}$ )	$\pi_a$ ( $\times 10^{-3}$ )	$\pi_s$ ( $\times 10^{-3}$ )	<i>R<sub>m</sub></i>	<i>D</i>	<i>D</i> *	<i>F</i> *
马尾松 <i>P. massoniana</i>													
<i>C3H</i>	45.0	7.0	1.01	2.00	3.0	0.129	1.28	0.64	0.94	0	-1.335	1.212	0.515
<i>GI</i>	44.0	4.0	0.75	1.76	4.0	0.251	2.15	0.33	2.15	0	-1.339	-0.090	-0.555
平均值 Average	44.5	5.5	0.88	1.88	3.5	0.190	1.71	0.48	1.55	0	-1.337	1.122	-0.140
黄山松 <i>P. hwangshanensis</i>													
<i>C3H</i>	34.0	14.0	4.89	4.31	12.0	0.898	5.68	3.75	3.84	0	0.560	1.018	1.032
<i>GI</i>	29.0	3.0	0.76	1.46	2.0	0.133	1.11	0.66	1.11	0	-1.135	0.959	0.418
平均值 Average	31.5	8.5	2.83	2.89	7.0	0.516	3.40	2.21	2.48	0	-0.575	0.059	0.614

*H<sub>d</sub>*, Nei 单倍型多样性;  $\pi$ , 总核苷酸多样性;  $\pi_a$ , 非同义位点核苷酸多样性;  $\pi_{sil}$ , 沉默位点核苷酸多样性;  $\pi_s$ , 同义位点核苷酸多样性; *h*, 单倍型数; *S*, 分离位点数; *R<sub>m</sub>*, 最小重组数;  $\theta_w$ , Watterson 参数; *D*, *D*\*

和 *F*\* 分别为中性检验的 Tajima's *D*, Fu & Li's *D* 和 Fu & Li's *F*. *H<sub>d</sub>*, Nei's haplotype diversity;  $\pi$ , nucleotide diversity across all loci;  $\pi_a$ , nucleotide diversity at nonsynonymous sites;  $\pi_{sil}$ , nucleotide diversity at synonymous sites; *h*, number of haplotypes; *S*, number of segregating sites; *R<sub>m</sub>*, the minimum number of recombination events;  $\theta_w$ , Watterson's parameter; *D*, *D*\* and *F*\* are Tajima's *D*, Fu & Li's *D*, and Fu & Li's *F* of neutrality tests, respectively.

相关信息(Schaal *et al.*, 1998)。因此我们应用Network 4.5 (Bandelt *et al.*, 1999)软件, 选择Median-joining模型通过溯祖模拟构建了两物种两个座位的基因谱系, 并经过手工调整将分离位点作为独立的进化单元, 同时删除了所有的插入缺失。

## 2 结果

### 2.1 核苷酸多态性

本研究涉及两个核基因座位(*C3H*, *GI*), 我们平均在每个位点上选取了马尾松和黄山松的76个雌配子体(取自76个不同的个体), 共测序得到了152条序列: *C3H* 79条, *GI* 73条。测得的序列总长1 349 bp, 其中951 bp为编码区, 398 bp位于非编码区。在检测到的23个变异位点中, 16个来自*C3H* (2个插入缺失), 7个来自*GI*。

我们利用DnaSP v. 5.0估算了两个物种的种内遗传多样性(表2), 分析时剔除了插入缺失位点。马尾松和黄山松的种内多态性位点分别为11个和17个。由表2可知, 马尾松的全部位点 $\pi$ 、沉默位点 $\pi_{\text{sil}}$ 、以及非同义位点 $\pi_a$ 多态性平均值均明显低于黄山松(马尾松:  $\pi = 0.000\ 88$ ,  $\pi_{\text{sil}} = 0.001\ 71$ ,  $\pi_a = 0.000\ 48$ ; 黄山松:  $\pi = 0.002\ 83$ ,  $\pi_{\text{sil}} = 0.003\ 40$ ,  $\pi_a = 0.002\ 21$ )。两物种的核苷酸多样性差异主要源于*C3H*基因, 该座位在两物种中总共检测到14个变异位点, 马尾松只有7个, 而黄山松有14个。另外, 马尾松基于两个核基因的单倍型数目 $h$ 和单倍型多样性 $H_d$ 的平均值也明显低于黄山松(马尾松 $h = 3.5$ ,  $H_d = 0.190$ ; 黄山松 $h = 7.0$ ,  $H_d = 0.516$ ), 主要差异也源于*C3H*基因, 与上述核苷酸多样性的结果一样。

### 2.2 中性检测

我们从物种水平上对两个核基因座位的进化模式进行了Tajima's  $D$ 、Fu和Li's  $D$ 和Fu和Li's  $F$ 中

性检测(表2)。结果显示, 虽然马尾松在两个座位都检测到负的Tajima's  $D$ 值(*C3H*:  $-1.334$ ; *GI*:  $-1.339$ ), 黄山松的*GI*座位也检测到负的Tajima's  $D$ 值( $-1.135$ ), 但是差异都不显著(表2)。所以两物种两个核基因座位的进化都不拒绝中性假设。

### 2.3 群体遗传结构

两个物种核基因座位的AMOVA结果见表3。马尾松和黄山松群体的遗传分化系数 $F_{\text{ST}}$ 分别为0.059和0.339 ( $p < 0.01$ )。两物种的种间差异有48.86%来自种间, 11.35%来自种内群体间, 39.74%来自群体内, 其中核基因座位*C3H*和*GI*对两物种种间差异的贡献分别为20.48%和77.24%。

### 2.4 基因谱系分析

两个座位的基因谱系结构明显不同(图1B, 1C)。*C3H*座位在两个物种79个个体中总共得到了14种单倍型, 其中, 马尾松的单倍型有4种(H1-H4), 主要以H1为主, 其余10种来自黄山松(H5-H14)。单倍型的频率都比较低, 两个物种共有的单倍型为H1和H3。*GI*座位在两个物种76个个体中共获得了6种单倍型, 其中马尾松有4种(H1-H4), 主要以H1为主; 黄山松有2种(H5和H6), 主要以H5为主, 两个物种没有共有单倍型。

## 3 讨论

### 3.1 核苷酸多样性

由于沉默位点的核苷酸多态性基本符合中性进化, 故能很好地反应物种的遗传多样性水平(Tajima, 1989)。我们计算了马尾松和黄山松两个核基因沉默位点的核苷酸多态性, 结果显示两个物种的两个基因座位均存在低水平的多样性, 马尾松的平均沉默位点多样性显著低于黄山松(马尾松 $\pi_{\text{sil}} = 0.001\ 71$ , 黄山松 $\pi_{\text{sil}} = 0.003\ 4$ ,  $p < 0.05$ )。与以前报

表3 马尾松和黄山松核基因座位*C3H*和*GI*变异的分子方差(AMOVA)分析(10 000次重复)

Table 3 Analysis of molecular variance (AMOVA) of nuclear loci *C3H* and *GI* variation in *Pinus massoniana* and *P. hwangshanensis* (10 000 permutations)

基因座位 Gene locus	变异来源 Source of variation (%)			遗传分化系数 Coefficient of gene differentiation ( $F_{\text{ST}}$ )	
	种间 Among species	种内群体间 Among populations within species	群体内 Within population	马尾松 <i>P. massoniana</i>	黄山松 <i>P. hwangshanensis</i>
<i>C3H</i>	20.48	15.81	63.71	0.071	0.210*
<i>GI</i>	77.24	6.99	15.77	0.047	0.467**
平均值 Average	48.86	11.35	39.74	0.059	0.339

\*,  $p < 0.05$ ; \*\*,  $p < 0.01$ .

道的基于多个核基因研究松柏类树木物种的平均多态性值相比,除黄山松的多态性高于欧洲赤松(*Pinus sylvestris*,  $\pi_{\text{sil}} = 0.0019$ , Garcia-Gil *et al.*, 2003)和银杉(*Cathaya argyrophylla*,  $\pi_{\text{sil}} = 0.0029$ , Wang & Ge, 2006)外,本研究中两个松属物种的多态性均低于其他大多数松柏类植物,如高山松(*Pinus densata*,  $\pi_{\text{sil}} = 0.0122$ )、云南松(*P. yunnanensis*,  $\pi_{\text{sil}} = 0.0184$ )、油松(*P. tabulaeformis*,  $\pi_{\text{sil}} = 0.0119$ ) (Ma *et al.*, 2006)、祁连圆柏(*Juniperus przewalskii*,  $\pi_{\text{sil}} = 0.00292$ , Li *et al.*, 2011)、火炬松(*P. taeda*,  $\pi_{\text{sil}} = 0.0064$ , Brown *et al.*, 2004)和挪威云杉(*Picea abies*,  $\pi_{\text{sil}} = 0.0039$ , Heuertz *et al.*, 2006)等。总之,我们的研究结果说明,马尾松和黄山松基于两个核基因座位的核苷酸多态性水平较低,支持松柏类植物具有低水平核苷酸多态性的观点(Wang & Ge, 2006)。

影响物种核苷酸多样性水平的因素很多,如研究物种取样数量的差异、所检测的基因片段、基因突变、自然选择、交配系统以及群体动态历史等。马尾松和黄山松较低的核苷酸多态性水平可能与松柏类植物较长的生活史以及低的突变率有关(Brown *et al.*, 2004)。同时,重组率也可影响物种的核苷酸多样性,我们研究的两个核基因座位的最小重组数都为0,因此低水平的重组可能对两个物种低的核苷酸多态性有一定贡献。

由于马尾松和黄山松的分布模式不同,马尾松分布于海拔700 m以下,而黄山松分布于1 000 m以上。这种分布模式使得黄山松趋向于片段化分布,而片段化分布是影响整个物种变异水平很重要的因素(Wakeley, 2001),黄山松的片段化分布使其具有高于马尾松的核苷酸多样性。另外,它们的群体动态历史,如两个物种分化后是否经历过瓶颈效应? 群体扩张等也会影响它们的核苷酸多样性水平的差异(Abbott & Comes, 2004; Akey *et al.*, 2004)。

中性检测显示, Tajima's *D*, Fu和Li's *D*和Fu和Li's *F*的值有正有负,而且差异都不显著,所以两个物种两个核基因座位的进化都没有显著偏离中性模型,因此我们选用的两个核基因位点未检测到是否经历过自然选择作用。

### 3.2 种内和种间遗传分化

我们利用AMOVA分析了两个物种核基因座位的群体遗传结构,估算了遗传分化系数 $F_{\text{ST}}$ 。结果显示:马尾松的种内遗传分化明显低于黄山松(马尾

松 $F_{\text{ST}} = 0.059$ ; 黄山松 $F_{\text{ST}} = 0.339$ ),与其他松柏类植物核基因得到的值相比,马尾松处于相对较低的种内分化水平,而黄山松具有较高的种内分化水平(Ledig, 2000; Li *et al.*, 2010a; Ren *et al.*, 2010; Li *et al.*, 2011)。松柏类植物是风媒远交多年生裸子植物,群体间的差异较小,遗传变异主要存在于群体内(Bouillé & Bousquet, 2005),从而形成了较低的种内遗传分化。然而,物种在片段化及有限的分布区内将趋向于形成高水平的种内遗传分化(Li *et al.*, 2010a)。相对于马尾松来说,黄山松生长于较高海拔而呈现片段化的分布,这可能会导致黄山松具有较高水平的遗传分化。

我们对马尾松和黄山松进行的AMOVA分析结果显示,两个物种基于两个核基因座位的种间差异为48.86%。近年来松柏类植物物种形成的研究表明,近期分化的物种间存在着不完全的谱系筛选(incomplete lineage sorting),短的分化时间以及物种形成后二次接触时相互间的渐渗杂交(hybridization)使物种间存在大量的共享多态性和较少的固定性差异(Willyard *et al.*, 2009; Chen *et al.*, 2010; Li *et al.*, 2010b)。对马尾松和黄山松这两个松科近缘物种来说,它们重叠分布于中国东南部,在形态学、细胞学以及木材解剖学上都不同,且存在普遍天然杂交群体(邢有华等, 1992; 罗世家, 2001),这些可能也导致它们在分子水平上的不完全分化。

另外,我们还可以看出, *GI*座位的种间差异明显高于*C3H* (*C3H*: 20.48%, *GI*: 77.24%),同时,基因谱系结果还显示,在*C3H*基因上具有共享单倍型,而在*GI*上没有(图1B, 1C)。基因谱系分析表明,两个物种在*C3H*座位上共有的单倍型为H1和H3(图1C),其中共享单倍型H1直接与外类群火炬松序列相连,并且位于network的中央,而且单倍型频率最高,此单倍型可能是较古老的单倍型,由不完全的谱系筛选所导致的物种间共享,而共享单倍型H3更靠近network的外部分支,则有可能是由两个物种分化之后再杂交造成的。但是,两种单倍型在两个物种之间的分布是祖先型共享还是杂交渗入导致的,仍需要进一步研究。*GI*基因在network图上显示没有共享的单倍型。这些研究结果表明,两个基因在两个物种之间的遗传分化和谱系筛选存在显著的区别。*C3H*是与植物木质素形成相关的基因,对木质素的合成具有下调作用(Ralph *et al.*, 2006),显然,该基

因与物种形成或者受物种形成的影响较少; *GI*参与调控植物的开花时间(Huq *et al.*, 2000), 而开花时间的差异是许多近缘物种生殖隔离或者物种分化的前提(Orr *et al.*, 2004)。这种功能上的差异显然与它们在物种之间的分化式样密切相关。但是, 我们对*GI*基因在两个物种中的进化历史进行了检验, 未发现选择痕迹(表2)。这有可能是因为我们研究的该基因长度不够或其他原因造成的。另外, 该基因在所研究的两个物种中是物种形成的相关或者连锁基因, 还是物种分化后导致它们加速分化的基因, 仍然不清楚。总之, 我们的研究表明, 近缘物种间的不同基因在物种分化过程中, 由于选择或者功能不同, 可能导致相应的种间基因流和遗传分化不同, 从而也为近年来逐渐兴起的邻域物种形成的基因观点提供了证据(Wu, 2001; Orr *et al.*, 2004)。然而, 我们的结果只是基于两个核基因座位研究得出的, 还需要更多基因座位的研究和数据的计算来支持。

**致谢** 国家自然科学基金(30725004)资助。感谢王召锋博士和毛康珊博士等为本文提出宝贵意见。

## 参考文献

- Abbott RJ, Comes HP (2004). Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (purple saxifrage). *New Phytologist*, 161, 211–224.
- Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L (2004). Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biology*, 2, 1591–1599.
- Arunyawat U, Stephan W, Städler T (2007). Using multilocus sequence data to assess population structure, natural selection, and linkage disequilibrium in wild tomatoes. *Molecular Biology and Evolution*, 24, 2310–2322.
- Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37–48.
- Belahbib N, Pemonge MH, Ouassou A, Sbay H, Kremer A, Petit RJ (2001). Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, 10, 2003–2012.
- Bomblies K, Weigel D (2007). Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics*, 8, 382–393.
- Bouillé M, Bousquet J (2005). Trans-species shared polymorphism at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. *American Journal of Botany*, 92, 63–73.
- Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004). Nucleotide diversity and linkage disequilibrium in Loblolly pine. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 15255–15260.
- Chase CD (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. *Trends in Genetics*, 23, 81–90.
- Chen J, Källman T, Gyllenstrand N, Lascoux M (2010). New insights on the speciation history and nucleotide diversity of three boreal spruce species and a Tertiary relict. *Heredity*, 104, 3–14.
- Chiang YC, Hung KH, Schaal BA, Ge XJ, Tsu TW, Chiang TY (2006). Contrasting phylogeographical patterns between mainland and island taxa of the *Pinus luchuensis* complex. *Molecular Ecology*, 15, 765–779.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11–15.
- Dumolin-Lapègue S, Pemonge MH, Petit RJ (1997). An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, 6, 393–397.
- Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47–50.
- Fu YX, Li WH (1993). Statistical tests of neutrality of mutations. *Genetics*, 133, 693–709.
- Fu LG, Li N, Elias TS, Mill RR (1999). Pinaceae. In: Wu ZY, Raven PH eds. *Flora of China*, Volume 4. Science Press, Beijing. 17.
- García-Gil MR, Mikkonen M, Savolainen O (2003). Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Molecular Ecology*, 12, 1195–1206.
- Heuertz M, de Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006). Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce (*Picea abies* (L.) Karst). *Genetics*, 174, 2095–2105.
- Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell’Olivio A, Arnold ML, Kuhlemeier C (2007). Single gene-mediated shift in pollinator attraction in *Petunia*. *The Plant Cell*, 19, 779–790.
- Huq E, Tepperman JM, Quail PH (2000). *GIGANTEA* is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 9789–9794.
- Hu XS (胡新生) (2002). A review on understanding the genetic structure of population. *Scientia Silvae Sinicae* (林业科学), 38, 119–128. (in Chinese with English abstract)
- Hudson RR, Kaplan NL (1985). Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, 111, 147–164.
- Kado T, Yoshimaru H, Tsumura Y, Tachida H (2003). DNA variation in a conifer, *Cryptomeria japonica* (Cupressaceae sensu lato). *Genetics*, 164, 1547–1599.
- Kumar S, Tamura K, Nei M (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, 5,

- 150–163.
- Ledig FT (2000). Founder effects and the genetic structure of Coulter pine. *Journal of Heredity*, 91, 307–315.
- Li Y, Stocks M, Hemmilä S, Källman T, Zhu HT, Zhou YF, Chen J, Liu JQ, Lascoux M (2010a). Demographic histories of four spruce (*Picea*) species of the Qinghai-Tibetan Plateau and neighboring areas inferred from multiple nuclear loci. *Molecular Biology and Evolution*, 27, 1001–1014.
- Li SX, Chen Y, Gao HD, Yin TM (2010b). Potential chromosomal introgression barriers revealed by linkage analysis in a hybrid of *Pinus massoniana* and *P. hwangshanensis*. *BMC Plant Biology*, 10, doi: 10.1186/1471-2229-10-37.
- Li ZH, Zhang Q, Liu JQ, Källman T, Lascoux M (2011). The Pleistocene demography of an alpine juniper of the Qinghai-Tibetan Plateau: *tabula rasa*, cryptic refugia or something else? *Journal of Biogeography*, 38, 31–43.
- Luo SJ (罗世家), Zou HY (邹惠渝), Liang SW (梁师文) (2001). Study on the introgressive hybridization between *Pinus hwangshanensis* and *P. massoniana*. *Scientia Silvae Sinicae* (林业科学), 6, 118–122. (in Chinese with English abstract)
- Ma XF, Szmidi AE, Wang XR (2006). Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci. *Molecular Biology and Evolution*, 23, 807–816.
- Mallet J (2007). Hybrid speciation. *Nature*, 446, 279–283.
- Mayr E (1942). *Systematics and the Origin of Species*. Columbia University Press, New York, USA.
- Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, USA.
- Orr HA, Masly JP, Presgraves DC (2004). Speciation genes. *Current Opinion in Genetics and Development*, 14, 675–679.
- Ralph J, Akiyama T, Kim H, Lu F, Schatz PF, Marita JM, Ralph SA, Reddy MSS, Chen F, Dixon R (2006). Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *The Journal of Biological Chemistry*, 281, 8843–8853.
- Ren GP, Abbott R, Zhou YF, Zhang LR, Peng YL, Liu JQ (2010). Genetic divergence, range expansion and homoploid hybrid speciation among four pine species in northeast China. *Molecular Ecology*. (under review)
- Rieseberg LH, Kim SC, Randell RA, Whitney KD, Gross BL, Lexer C, Clay K (2007). Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica*, 129, 149–165.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 19, 2496–2497.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998). Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, 7, 465–474.
- Strasburg JL, Rieseberg LH (2008). Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—large effective population sizes and rates of long-term gene flow. *Evolution*, 62, 1936–1950.
- Tajima F (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, 105, 437–460.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.
- Wakeley J (2001). The coalescent in an island model of population subdivision with variation among demes. *Theoretical Population Biology*, 59, 133–144.
- Wang HW, Ge S (2006). Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Molecular Ecology*, 15, 4109–4122.
- Watterson GA (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7, 256–276.
- Willyard A, Cronn R, Liston A (2009). Reticulate evolution and incomplete lineage sorting among the Ponderosa pines. *Molecular Phylogenetics and Evolution*, 52, 498–511.
- Wright S (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323–354.
- Wright SI, Gaut BS (2005). Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution*, 22, 506–519.
- Wu CI (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, 14, 851–865.
- Xing YH (邢有华), Fang YX (方永鑫), Wu GR (吴根荣) (1992). The preliminary study of natural hybridizations between *Pinus hwangshanensis* and *P. massoniana* in Dabie Mountain of Anhui Province. *Anhui Forestry Science and Technology* (安徽林业科技), 4, 5–9. (in Chinese)
- Zheng XM, Ge S (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology*, 19, 2439–2454.
- Zhou YF, Abbott RJ, Jiang ZY, Du FK, Milne RI, Liu JQ (2010). Gene flow and species delimitation: a case study of two pine species with overlapping distributions in southeast China. *Evolution*, 64, 2342–2352.
- Zou YH (邹惠渝), Li Z (黎志), Luo SJ (罗世家), Yi YM (易咏梅), Wang JR (王静茹) (2003). Study on taxonomy of *Pinus hwangshanensis*. *Bulletin of Botanical Research* (植物研究), 23, 278–279. (in Chinese with English abstract)

责任编辑: 李庆军 实习编辑: 黄祥忠