

# 应用单拷贝核基因标记研究我国欧洲山杨自然居群的遗传多样性

杜淑辉<sup>1,2</sup>, 王兆山<sup>1</sup>, 保尔江·阿布都哈米提<sup>3</sup>, 邢世岩<sup>2\*</sup>, 张建国<sup>1\*</sup>

(1. 国家林业局林木培育重点实验室, 中国林业科学研究院林业研究所, 北京 100091; 2. 山东农业大学林学院, 山东 泰安 271000;  
3. 新疆阿勒泰地区林业科学研究所, 新疆 阿勒泰 836500)

**摘要:** [目的] 对分布于我国新疆地区欧洲山杨自然居群的遗传多样性与遗传分化进行研究。 [方法] 对 4 个欧洲山杨自然居群共 72 个个体的 6 个单拷贝核基因标记 (sing-copy nuclear markers) 进行扩增与测序, 并计算相应的遗传多样性和遗传分化参数。 [结果] 表明: 比对后 6 个基因的长度在 459 bp 到 747 bp 之间, 平均多态核苷酸位点个数为 14, 遗传多样性指数  $\pi$  和  $\theta_w$  分别达到了 0.004 8 和 0.004 4, 基因多样性指数为 0.73, 以上结果表明欧洲山杨表现出较高的遗传多样性水平。 Tajima 中性检验结果均不显著, 表明所用 6 个单拷贝核基因均符合中性进化假设。 AMOVA 分析表明 89.72% 的遗传变异存在于居群内, 居群间的平均遗传分化指数  $F_{st}$  为 0.10, 居群间平均基因流 (Nm) 为 3.235。 [结论] 高度异交, 高碱基突变率及超长距离的基因流机制能够很好的解释我国新疆地区欧洲山杨表现出的较高的遗传多样性水平与较低水平的遗传分化。

**关键词:** 新疆; 欧洲山杨; 单拷贝核基因标记; 遗传多样性; 遗传分化

中图分类号: S718.46

文献标识码: A

## A Study on Genetic Diversity of *Populus tremula* in China by Single-copy Nuclear Markers

DU Shu-hui<sup>1,2</sup>, WANG Zhao-shan<sup>1</sup>, BAOERJIANG Abuduhamiti<sup>3</sup>, XING Shi-yan<sup>2</sup>, ZHANG Jian-guo<sup>1</sup>

(1. Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China; 2. College of Forestry, Shandong Agricultural University, Tai'an 271000, Shandong, China;  
3. Altay Research Institute of Forestry, Altay 836500, Xinjiang, China)

**Abstract:** [Objective] This paper aims at investigating the genetic diversity and genetic differentiation of *Populus tremula* distributed in Xinjiang, China. [Method] Six single-copy nuclear gene markers of 72 *P. tremula* individuals from 4 populations in Xinjiang were amplified and sequenced. The parameters of genetic diversity and genetic differentiation were calculated. [Result] After aligning, the length of the six genes ranged between 459 bp and 747 bp and the average number of polymorphism sites was 14. Diversity parameters,  $\pi$  and  $\theta_w$ , reached 0.004 8 and 0.004 4, respectively. The mean value of gene diversity was 0.73. These indicated high level of genetic diversity of *P. tremula*. The results of Neutral tests using Tajima's D were not significant, which indicated all the six genes did not violate from neutrality. The results of AMOVA analysis showed that 89.72% of the total variance existed within populations. The average genetic differentiation parameter  $F_{st}$  and gene flow among populations reached 0.10 and 3.235, respectively. [Conclusion] High level of outcrossing rate and nucleotide mutation rate as well as long distance dispersal seed contributed to the high level of genetic diversity and low genetic differentiation in *P. tremula* in

收稿日期: 2015-01-05

基金项目: 国家自然科学基金(31470665)、北京市优秀博士学位论文指导教师科技项目、江苏省高等教育联合创新计划

作者简介: 杜淑辉(1985—), 男, 山东泰安人, 山东农业大学林学博士后科研流动站, 博士, 从事林木种群遗传学研究。

\* 通讯作者: 邢世岩 xingsy@sdau.edu.cn, 张建国 zhangjg@caf.ac.cn

Xinjiang, China.

**Key words:** Xinjiang; *Populus tremula*; sing-copy nuclear markers; genetic diversity; genetic differentiation

欧洲山杨(*Populus tremula* L.)是杨柳科(Salicaceae)杨属(*Populus*)白杨派(section *Populus*)代表物种之一,为落叶植物,广泛分布于西伯利亚,高加索及欧洲地区,而在我国只分布于新疆阿尔泰、塔城,天山东部北坡至西部伊犁山区。欧洲山杨染色体数目  $2n = 38$ ,高度异交,根系发达,主要靠根蘖繁殖,也可靠种子繁殖,花粉和种子扩散能力强<sup>[1]</sup>。近年来,针对分布于欧洲地区的欧洲山杨遗传多样性及遗传分化的研究不断增多。Ingvarsson<sup>[2-3]</sup>利用77个核基因位点对分布于欧洲地区的14个欧洲山杨自然居群进行了研究,结果表明欧洲山杨表现出高水平的遗传多样性。Fussi等<sup>[4]</sup>利用叶绿体标记的研究也得到了类似的结果。然而,目前国内外对分布于我国新疆地区欧洲山杨自然居群遗传多样性与遗传分化的研究还未见报道,对欧洲山杨自然居群遗传多样性的研究,一方面可以为合理构建核心种质,制定种质资源保存和利用策略及资源的可持续利用提供依据,另一方面也可以为种质创新奠定基础。

单拷贝核基因(single-copy nuclear genes)由于易于扩增测序、多态核苷酸位点多、不存在基因复

制、谱系分选等优点<sup>[5]</sup>,现已广泛应用于植物种群遗传学的研究<sup>[6-7]</sup>。毛果杨(*P. trichocarpa*)全基因组数据的公布<sup>[8]</sup>,使得利用生物信息学的方法开发杨属内单拷贝核基因标记成为可能<sup>[9]</sup>。因此,本研究利用单拷贝核基因标记对分布于我国新疆地区的4个欧洲山杨自然居群进行分析,旨在揭示其遗传多样性与遗传分化水平,为欧洲山杨引种、育种策略的科学制定和种质资源的有效保护及利用提供理论依据。

## 1 材料与方法

### 1.1 试验材料

2011年5月份对新疆地区欧洲山杨进行了采样,总共采集4个居群72个个体的样品,具体居群编号、采样地点、样品个数等信息见表1。采样时选取居群内生长期良好,无病虫害的单株。采样时尽量覆盖整个居群,每个居群随机选择样品10~20个左右,相邻株间至少相距100m左右。每个个体采集新鲜叶片保存于变色硅胶中,带回实验室用于DNA的提取。

表1 欧洲山杨自然居群采样信息

居群	地点	样品个数	经纬度	海拔/m
ES	新疆维吾尔自治区哈巴河县 Habahe Xinjiang	17	48°22' N, 85°56' E	744
SS	新疆维吾尔自治区布尔津县 Buerjin Xinjiang	23	47°42' N, 86°50' E	472
TS	新疆维吾尔自治区天山 Tianshan Xinjiang	20	43°47' N, 87°37' E	895
AS	新疆维吾尔自治区阿勒泰县 Aletai Xinjiang	12	47°58' N, 88°11' E	1 305

### 1.2 试验方法

采用改良的CTAB法提取DNA<sup>[10]</sup>。所提DNA经1.5%琼脂糖凝胶电泳检测,并用分光光度法定量后,于-20℃保存备用。6个单拷贝核基因标记的详细信息见表2。PCR扩增采用30 μL的反应体系:50 ng DNA, Mg<sup>2+</sup> 2.0 mmol · L<sup>-1</sup>、dNTP 0.24 mmol · L<sup>-1</sup>、引物0.24 μmol · L<sup>-1</sup>、Taq酶1.5 U。所有试剂均为北京天根试剂有限公司生产。PCR扩增程序为:94℃预变性4 min;94℃变性30 s,退火90 s(退火温度随引物而定,见表2),72℃延伸90 s,30个循环;72℃延伸10 min;10℃保存。扩增产物用1.8%琼脂糖凝胶电泳检测,G-BOX紫外凝胶成像

系统观察,拍照记录。扩增产物采用DNA纯化试剂盒(Amersham Pharmacia Biotech, Piscataway, USA)纯化后利用ABI 3730 DNA analyzer (Applied Biosystems, Foster City, California, USA)直接测序。扩增和测序采用相同引物。对于个别个体直接测序结果出现套峰的标记位点,采用克隆测序。PCR产物经纯化后连接至pGEM-T easy Vector System II (Promega, Madison, WI, USA),每个样品随机选取6~12个阳性克隆使用通用引物SP6(5'-ATTTAGGTGACACTATAG 3')和T7(5'-TAATACGACTC ACTATAGGG 3')进行双向测序。

表2 引物序列及退火温度

基因	引物序列 (5'-3')	染色体位置	退火温度/°C	参考文献
<i>DSH3</i>	F:TCTGCTTCCACTTCTTGC R:CATACTCTCCATTGTCC	VI	55	
<i>DSH5</i>	F:TGGCAGAATCACCAGACCCTC R:CCAATTTAGCATCTTCAGCCTCAT	XII	59	
<i>DSH6</i>	F:GCCTCCTGATTATTATGC R:TATTACAAGCCCTTCCAG	XV	54	[9]
<i>DSH7</i>	F:TGTCCAGAAACGCATCC R:CAAACCTTACCACCCCA	XVI	58	
<i>DSH12</i>	F:CACCACATCCCGCTTTCTCTCTTC- ACTT R:TAAACCCAGGAGGCAAAAACAG- CACCAG	II	57	
<i>DSH21</i>	F:CATGCTTATGAAGGTGTGGGCTT R:TGCAAACATCTCACTGTGACTG	XVII	53	

### 1.3 数据统计与分析

序列数据采用 CLUSTALX<sup>[11]</sup> 进行比对并利用 BioEdit<sup>[12]</sup> 进行手工校正,输出用于后续分析的 FSTAT 格式数据文件。利用 Dnasp 5.10<sup>[13]</sup> 对每个基因的 FSTAT 数据文件进行 PHASE 运算(运算次数 10 000 次,预热 100 次)并计算多态核苷酸位点数目( $S$ )、基因多态性( $H_d$ )<sup>[14]</sup>、居群间基因流( $Nm$ )<sup>[15]</sup>和 Tajima's D 中性检验(显著性检验由 1 000 次联合模拟计算得到)<sup>[16]</sup>。遗传多样性水平使用一对序列间每碱基平均差异值  $\pi$ <sup>[17]</sup> 和基于多

态核苷酸位点数目  $S$  的 Watterson's 参数  $\theta_w$ <sup>[18]</sup> 来衡量,以上两个参数也由 Dnasp 5.10 计算得到。全部计算完成后,由 Dnasp 5.10 导出软件 Arlequin 3.5.1.3<sup>[19]</sup>需要的 ARF 文件格式并利用 Arlequin 3.5.1.3 中的分子方差分析(AMOVA)计算在每个基因上物种内居群间以及居群内差异对总变异的贡献率(显著性检验利用 1000 次重复模拟得到)和居群间的遗传分化指数  $F_{st}$ <sup>[15]</sup>,显著性检验同样利用 1 000 次重复模拟得到。

## 2 结果与分析

### 2.1 欧洲山杨自然居群遗传多样性

经过扩增测序,得到了欧洲山杨 4 个自然居群 72 个个体在 6 个单拷贝核基因的全部序列。比对后序列长度在 459 bp 到 747 bp 之间。在 6 个核基因中,多态核苷酸位点数目从 11 个到 17 个不等,基因多样性从 0.57 到 0.85,平均值为 0.73,各居群间基因流的平均值达到了 3.235。多样性指数  $\pi$  和  $\theta_w$  的计算结果表明,*DSH5* 的  $\pi$  值最高,为 0.007 2,*DSH7* 的  $\pi$  值最低,为 0.002 3,6 个基因  $\pi$  的平均值达到了 0.004 8;*DSH6* 的  $\theta_w$  值最高为 0.005 7,而 *DSH21* 的  $\theta_w$  最低,为 0.003 9,6 个基因  $\theta_w$  的平均值为 0.004 4。Tajima's D 检验结果均不显著( $P > 0.05$ ),表明所选基因均符合中性进化假设(表 3)。

表3 各位点遗传多样性计算结果

基因	长度 L	多态位点数目 S	核苷酸多样性 $\pi$	核苷酸多样性 $\theta_w$	Tajima's D D	基因多样性 $H_d$	基因流 Nm
<i>DSH3</i>	602	13	0.005 4	0.003 9	1.01	0.85	1.66
<i>DSH5</i>	459	11	0.007 2	0.004 4	1.62	0.70	0.53
<i>DSH6</i>	542	17	0.005 9	0.005 7	0.105	0.85	6.37
<i>DSH7</i>	529	14	0.002 3	0.004 8	-1.365	0.57	4.81
<i>DSH12</i>	548	12	0.005 2	0.004 0	0.799	0.62	0.56
<i>DSH21</i>	747	16	0.002 6	0.003 9	-0.849	0.81	5.48
平均值	571	14	0.004 8	0.004 4	-	0.73	3.235

### 2.2 欧洲山杨自然居群间遗传分化

AMOVA 分析结果表明,在 6 个核基因位点欧洲山杨自然居群间和居群内都存在显著的遗传变异( $P < 0.01$ ),其中居群间变异占总变异的百分比从 5.03% 到 16.34% 不等,平均值为 10.28%,而居群

内变异占总变异的百分比在 83.66% 到 97.12% 之间,平均值为 89.72%。遗传分化指数  $F_{st}$  计算结果均显著( $P < 0.01$ ),最大值为 0.16(*DSH3*),最小值为 0.029(*DSH6*),平均值为 0.10(表 4)。

表4 遗传分化计算结果

变异来源	<i>DSH3</i>	<i>DSH5</i>	<i>DSH6</i>	<i>DSH7</i>	<i>DSH12</i>	<i>DSH21</i>	平均值
居群间	16.34	15.51	2.88	5.03	9.79	5.44	10.28
居群内	83.66	84.49	97.12	94.97	90.21	94.56	89.72
$F_{st}$	0.16	0.16	0.029	0.050	0.098	0.054	0.10

### 3 讨论

#### 3.1 欧洲山杨自然居群遗传多样性

尽管不同基因标记间遗传多样性有一定差异,分布于我国新疆地区的欧洲山杨在6个核基因上表现出较高的遗传多样性( $\pi = 0.0023 \sim 0.0072$ , 平均值为0.0048,  $\theta_w = 0.0039 \sim 0.0057$ , 平均值为0.0044)(表3),比其他同纬度分布的木本植物如挪威云杉(*Picea abies*,  $\pi = 0.0039$ )<sup>[20]</sup>和同属的香脂杨(*P. balsamifera*,  $\pi = 0.00375$ ,  $\theta_w = 0.004$ )、毛果杨(*P. trichocarpa*,  $\pi = 0.0029$ ,  $\theta_w = 0.0020$ )<sup>[21]</sup>, 窄叶杨(*P. angustifolia*,  $\pi = 0.0024$ )<sup>[22]</sup>等物种的遗传多样性要高,而与其近缘种中国山杨(*P. davidiana*,  $\pi = 0.0044$ )和美洲山杨(*P. tremuloides*,  $\pi = 0.0035$ )<sup>[23]</sup>表现出相似的遗传多样性水平。而比分布于欧洲地区的欧洲山杨的遗传多样性水平略低( $\pi = 0.007 \sim 0.0120$ ,  $\theta_w = 0.005 \sim 0.0129$ )<sup>[2-3]</sup>,这主要是由于 Ingvarsson<sup>[2-3]</sup>的研究中所使用的基因受到了平衡选择(balancing selection)的影响,而平衡选择可使受影响的基因区域保持较高的遗传多样性水平<sup>[6]</sup>。

影响物种遗传多样性水平的因素很多,包括取样的代表性、碱基突变速率、繁育系统及自然选择等<sup>[24]</sup>。本研究取样代表性问题我们给予了充分的考虑,所取样品遍及欧洲山杨在我国新疆地区的分布范围,因此应排除取样代表性对研究结果的影响。自然选择会对遗传多样性产生显著的影响<sup>[24]</sup>,根据本研究中 Tajima's *D* 检测结果表明所有位点均符合中性进化模型,因此,自然选择作用也不是导致欧洲山杨表现出较高的遗传多样性水平的因素。杨属物种为高度异交植物,异交能够增加物种的有效群体大小和有效重组率,从而增加物种的遗传多样性水平<sup>[25]</sup>。另据 Richardson 等<sup>[26]</sup>估计,多年生木本植物核基因的突变率大约在  $3.2 \times 10^{-10}$  到  $5.7 \times 10^{-10}$ /年/同义碱基之间,而杨属核基因突变率却高达  $2.5 \times 10^{-9}$ /年/同义碱基<sup>[8]</sup>。因此高度异交和高碱基突变率能够很好的解释欧洲山杨表现出的较高的遗传多样性水平。

#### 3.2 欧洲山杨自然居群的遗传分化

杨属物种为异交风媒植物,并且种子和花粉的扩散能力很强,因此杨属物种内居群间遗传分化应该相对略低<sup>[27]</sup>。如表4所示,在6个核基因中,欧洲山杨不同居群间遗传分化指数  $F_{st}$  平均值为0.10,

要略低于杨属不同物种内居群间平均遗传分化水平( $F_{st} = 0.11$ )<sup>[28]</sup>。AMOVA 分析结果表明居群内的变异占总变异的绝大部分(89.72%),只有一小部分变异存在于居群间(10.28%)。针对分布于欧洲地区的欧洲山杨的研究也得到了类似的结果<sup>[4]</sup>,并且这种情况也普遍存在于其他杨属物种中<sup>[22]</sup>。张建国<sup>[29]</sup>通过研究欧洲山杨的克隆多样性发现欧洲山杨具有超长距离的基因流机制(种子飞散特性)。本研究中也发现在6个核基因上居群间基因流平均值达到了3.235。超长距离的基因流机制能够很好的解释本研究中欧洲山杨自然居群间表现出的遗传分化水平。

#### 参考文献:

- [1] 中国科学院中国植物志编辑委员会[M].《中国植物志》北京:科学出版社,1984,20(2):12.
- [2] 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.
- [3] Ingvarsson P K. Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula* [J]. Genetics, 2008, 180(1): 329-340.
- [4] Fussi B, Lexer C, Heinze B. Phylogeography of *Populus alba* (L.) and *Populus tremula* (L.) in Central Europe: secondary contact and hybridisation during recolonisation from disconnected refugia [J]. Tree Genetics & Genomes, 2010, 6(3): 439-450.
- [5] Wendel J F, Doyle J J. Phylogenetic incongruence: window into genome history and molecular evolution [J]. Molecular Systematics of Plants. London: Springer. 1998: 265-296.
- [6] Zhang LB, Ge S. Multilocus analysis of nucleotide variation and speciation in *Oryza officinalis* and its close relatives [J]. Molecular Biology and Evolution, 2007, 24(3): 769-783.
- [7] Wang Z, Du S, Dayanandan S, et al. Phylogeny reconstruction and hybrid analysis of *Populus* (Salicaceae) based on nucleotide sequences of multiple single-copy nuclear genes and plastid fragments [J]. PloS One, 2014, 9(8): e103645.
- [8] Tuskan G A, Difazio S, Jansson S, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray) [J]. Science, 2006, 313(5793): 1596-1604.
- [9] Du S H, Wang Z H, Zhang J G. A novel set of 15 single-copy nuclear DNA markers for genetic studies of Salicaceae [J]. Genetics and Molecular Research, 2015, 13(3): 4911-4917.
- [10] 李际红,邢世岩,王聪聪,等.银杏基因组 DNA 甲基化修饰位点的 MSAP 分析 [J]. 园艺学报, 2011, 38(8): 1429-1436.
- [11] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. Nucleic Acids Research, 1997, 25(24): 4876-4882.

- [12] Hall TA 1999. BioEdit: a user – friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. 95 – 98 [Z].
- [13] Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data [J]. *Bioinformatics*, 2009, 25 (11): 1451 – 1452.
- [14] Nei M. *Molecular evolutionary genetics*[M]. New York: Columbia University Press, 1987.
- [15] Nei M. Evolution of human races at the gene level, *Human genetics, part A: The unfolding genome*[M]. New York: Alan R. Liss, 1982: 167 – 181.
- [16] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism [J]. *Genetics*, 1989, 123 (3): 585 – 595.
- [17] Nei M, Li W H. Mathematical model for studying genetic variation in terms of restriction endonucleases [J]. *Proceedings of the National Academy of Sciences*, 1979, 76(10): 5269 – 5273.
- [18] Watterson G A. On the number of segregating sites in genetical models without recombination [J]. *Theoretical Population Biology*, 1975, 7(2): 256 – 276.
- [19] Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis [J]. *Evolutionary Bioinformatics*, 2005, 1: 47.
- [20] Heuertz M, De Paoli E, Källman T, *et al.* Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst] [J]. *Genetics*, 2006, 174(4): 2095 – 2105.
- [21] Ismail M, Soolanayakanahally R Y, Ingvarsson P K, *et al.* Comparative nucleotide diversity across North American and European *Populus* species [J]. *Journal of Molecular Evolution*, 2012, 74: 257 – 272.
- [22] DiFazio S P, Slavov G T, Joshi C P. *Populus*: a premier pioneer system for plant genomics. *Genetics, genomics and breeding of poplar* [M]. Lebanon: Science Publishers, Inc. 2011: 1 – 28.
- [23] 杜淑辉. 杨属分子系统发育及三个山杨物种的生物地理学研究 [D]. 中国林业科学研究院博士学位论文, 2014.
- [24] Wright S I, Gaut B S. Molecular population genetics and the search for adaptive evolution in plants [J]. *Molecular Biology and Evolution*, 2005, 22(3): 506 – 519.
- [25] Charlesworth D. Effects of inbreeding on the genetic diversity of populations [J]. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 2003, 358(1434): 1051 – 1070.
- [26] Richardson J E, Pennington R T, Pennington T D, *et al.* Rapid diversification of a species-rich genus of neotropical rain forest trees [J]. *Science*, 2001, 293(5538): 2242 – 2245.
- [27] Hamrick J, Godt M. Effects of life history traits on genetic diversity in plant species [J]. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 1996, 351(1345): 1291 – 1298.
- [28] Petit R J, Aguinalalde I, de Beaulieu J L, *et al.* Glacial refugia: hotspots but not melting pots of genetic diversity [J]. *Science*, 2003, 300(5625): 1563.
- [29] 张建国. 新疆杨属植物研究 [M]. 北京: 科学出版社, 2014: 92 – 93.

(责任编辑:张 研)