

马尾松表达序列标签多态性初步分析

尹佟明 李 东 陈 颖 黄敏仁 王明麻

(南京林业大学林木遗传与基因工程开放研究实验室 南京 210037)

关键词: 单链构象多态性(SSCP), 表达序列标签(EST), 马尾松, 遗传图谱

中图分类号: S718.46; Q943.2 文献标识码: A 文章编号: 1001-7488(2004)06-0176-05

Preliminary Detection of Polymorphisms of Expressed Sequence Tag
in *Pinus massoniana*

Yin Tongming Li Dong Chen Ying Huang Minren Wang Mingxiu

(The Key Laboratory of Forest Genetics and Gene Engineering, Nanjing Forestry University Nanjing 210037)

Abstract: We reported our preliminary study on mapping of the expressed sequence tag (EST) using single-stranded DNA conformation polymorphism (SSCP) on polymerase chain reaction (PCR) product. Five ESTs primer pairs, developed from nuclear genes including CAD (cinnamyl-alcohol dehydrogenase), CHS (chalcone synthase), NIR (nitrite reductase), ARA554 (cDNA expressed in differentiating xylem in *Pinus taeda*) and GS (glutamine synthetase), were selected to optimize the experimental protocol and generate mapping markers in the megagametophytes from a single tree of *Pinus massoniana*. Our ultimate goal was to use SSCP approach to construct a transcriptional map for comparative mapping studies in *P. massoniana*. The efficiency to construct a transcriptional map in *P. massoniana* based on the published EST primer pairs derived from other pine species critically depended on the successful amplification of EST fragments and the ratio of the heterozygous loci revealed in *P. massoniana*. In this study, 3 (60%) out of 5 tested EST primer pairs were succeeded in amplification, however, only 1 (20%) gene was heterozygous in the tested tree. The ratio of heterozygous loci detected in this study was similar to that revealed by anonymous marker in *P. taeda*. Therefore, a low efficiency would be expected if the map would be constructed using single pedigree. The segregation ratio of loci revealed by primer pairs would be higher if multiple pedigrees would be used. We proposed that consensus mapping approach based on multiple-pedigree should be used for EST mapping and therefore to increase the efficiency of EST mapping.

Key words: Single-strand conformation polymorphism (SSCP), Expressed sequence tag (EST), *Pinus massoniana*, Linkage map

遗传图谱是现代分子数量遗传学研究的一个基本平台,在过去的十几年中,林木遗传学家在几十个树种中构建了分子标记连锁图谱,并进行了一些重要数量性状的 QIL 分析。(Bradshaw *et al.*, 1994; Cervera *et al.*, 2001; Devey *et al.*, 1994; Echt *et al.*, 1997; Grattapaglia *et al.*, 1994; Mukai *et al.*, 1995; Nelson *et al.*, 1993; Plomion *et al.*, 1995; Remington *et al.*, 1999; Wu *et al.*, 2000; Yin *et al.*, 2003)。可是林木中大部分的遗传图谱是利用匿名的显性标记构建的,如 RAPD 标记和 AFLP 标记等。近几年来才有一些在松树 (Temesgen *et al.*, 2001) 和杨树 (Cervera *et al.*, 2001; Yin *et al.*, 2004) 中利用高保守性的共显性标记进行遗传图谱构建的报道,尽管利用随机匿名标记可以快速建成某一个杂交组合的遗传图谱,但所获得的信息是零散的,而且往往具有杂交组合特异性,因而无法进行基因组比较 (Yin *et al.*, 2003)。

林木是一个处在初级改良阶段的物种,其遗传改良程序与作物及牲畜等以品系为改良对象的育种程序有所不同,杂交组合特异性的遗传信息在实际的育种程序应用中有很大的局限性。要确定一个检测到的 QIL 能否应用到实际改良程序中,首先需要测定它在其他杂交组合中是否稳定。因此在实际改良过程中,人们感兴趣的是那些在不同遗传背景下都表达的“通用”QIL (Lerceteau *et al.*, 2000)。然而在林木中由于遗传组成的复杂性及缺少有效的标记系统,很难将遗传背景的效应分离出来。而且利用随机匿名标记很难在天

收稿日期: 2003-11-20。

基金项目: 国家自然科学基金(30200224, 30230300)和霍英东基金(81024)资助。

然群体水平上找到与某一 QTL 显著相关的分子标记 (Strauss *et al.*, 1992)。在林木天然群体进化历程中,由于标记与 QTL 间重组的不断发生,在不同世代及不同个体间,标记与 QTL 间的相关可能会丧失。因此分子标记辅助选择育种在林木实际育种程序中应用的可操作性越来越受到林木经典育种学家的质疑。

随着林木基因组研究的进展,人们具有了解决上述问题的方法。可以利用与性状表达有关的基因本身发展分子标记,这样就可以避免在群体水平上由于标记与目的基因间连锁不平衡消失带来的问题,从而使在林木群体水平上开展标记辅助选择具有可行性 (Strauss *et al.*, 1992)。目前大规模的 EST (Expressed Sequence Tag, 表达序列标签) 测序计划已在松树 (<http://www.cbc.umn.edu/ResearchProjects/Pine/index.html>) (Temesgen *et al.*, 2001) 和杨树 (<http://www.biochem.kth.se/PopulusDB>) (Sterky *et al.*, 1998) 中展开。EST 标记可以作为图谱比较及在不同群体中获得的 QTL 信息比较的桥梁。为进行有效的 EST 图谱构建,需要高效且成本低廉的检测技术。EST 标记的开发主要是基于导致点突变的单核苷酸多态性 (single nucleotide polymorphism, SNP)。目前有 6 种主要的方法可用于点突变检测:单链构象多态性 (single-strand conformation polymorphism, SSCP) 电泳技术、DNA 化学切割 (chemical cleavage, CCM)、RNA 酶非匹配切割 (RNase mismatch cleavage)、利用水溶性碳二酰亚胺 (carbodiimide, CDI) 与 DNA 杂合链 (DNA heteroduplexes) 反应、序列测定和变性凝胶梯度电泳 (denaturing gradient gel electrophoresis, DGGE) (Andersen *et al.*, 1998; Bui *et al.*, 2002; Cotton, 1993; Economou *et al.*, 1992; Ellis *et al.*, 1998; Fauda *et al.*, 2000; Grange *et al.*, 1990; Liu *et al.*, 1999; Marshall *et al.*, 1997; Hayashi, 1991; Krizkova *et al.*, 1994)。在这 6 种技术中,尽管 SSCP 及 DGGE 检测技术不能获得准确的突变碱基及位置信息,但这 2 种技术都是分离野生型和突变型 DNA 及 RNA PCR 产物的非常敏感的方法,对区分含有单核苷酸突变的等位基因十分有效,而且都适合于大量样品的分析,因此是 EST 作图的常用分析方法。目前已有全自动化仪器用于完成这 2 种分析,但是仪器价格较为昂贵。由于 DGGE 分析中灌制梯度凝胶人工操作较难,因而 SSCP 就成了更简单和成本低的技术选择。本文的目的之一是对针叶树分析的 SSCP 检测实验程序进行优化,同时对 EST 标记在松属不同树种间的通用性进行检测。在此基础上,对马尾松转录图谱构建的更有效策略进行了探讨。

1 材料和方法

1.1 试验材料和 DNA 提取 试验所用材料马尾松 (*Pinus massoniana*) 来自于安徽黄山,利用单株树上的针叶和种子提取 DNA (由南京林业大学南方林木种子中心李淑娴同志提供)。其中一份 DNA 模板是由二倍体针叶组织中提取的,另有 24 份 DNA 样品是从种子的单倍体胚乳组织中提取的。DNA 的提取采用了植物 DNA 提取试剂盒 (DNeasy Plant Mini Kit, Qiagen) 并按试剂盒提供的实验程序提取,从叶和种子中提取的 DNA 最后分别溶解在 50 μ L 和 100 μ L TE 缓冲液中。

1.2 引物的选用 选用了 5 个在其他松属树种中开发的 EST 引物 (Plomion *et al.*, 1999),这些 EST 对应的基因功能已知或其功能可以根据与其他物种基因的序列相似性推知。这几个基因 (表 1) 分别是:苯丙烯醇脱氢酶 (cinnamyl-alcohol dehydrogenase, AD) 基因,该基因在木素合成代谢过程中起重要作用¹⁾;查尔酮合酶 (chalcone synthase, CHS) 基因,该基因编码类黄酮和联苯乙烯生物合成过程中的关键酶;亚硝酸盐降解酶 (nitrite reductase, NIR) 基因,该基因编码 2 种亚硝酸盐降解酶中的一种酶,其功能是将亚硝酸盐降解为氨盐 (Neininger *et al.*, 1994); ARA554 为在火炬松 (*P. taeda*) 木质部中差异表达的 cDNA,这个 cDNA 的序列与编码阿拉伯半乳糖苷酶的基因有很高的相似性 (Loopstra *et al.*, 1995);谷氨酰胺合成酶 (glutamine synthetase, GS) 基因,这个基因编码的酶催化一个 ATP 依赖型的反应,可将氨基基团和谷氨酸合成谷氨酰胺 (Plomion *et al.*, 1999)。

1.3 PCR 扩增 PCR 扩增反应中利用了 20~50 ng 的 DNA 模板,反应总体积为 15 μ L,反应体系中含有终浓度为 0.1 pmol 的正、反向引物,1 \times PCR 反应缓冲液 (Gibco BRL),100~300 μ mol 的 MgCl₂ (浓度随不同基因变化),200 μ mol dNTPs 和 1.0 U Taq DNA 聚合酶 (Gibco BRL)。扩增反应在 PE-9700 热循环仪上进行,程序如下:94 变性 3 min,然后进行 20 个梯度变温循环,梯度变温循环中的第一个循环的退火温度比最终确定的

1) McKay J J. A mutation in lignin biosynthesis in loblolly pine: genetic, molecular and biochemical analyses. Ph. D. Thesis, North Carolina State University, Raleigh, NC, 1996:144

特异退火温度高 2.0 ,在随后的梯度变温循环中,每一循环退火温度降低 0.1 ,20 个循环后,在恒定的特异退火温度下再进行 10 个循环的反应。这些循环的退火时间为 45 s,变性条件为 94 30 s,延伸条件为 72 1 min 30 s。最终特异退火温度的测试范围在 45 ~ 65 之间。循环结束后,最后在 72 下保温 10 min。在 SSCP 分析前,利用 1.0 μ L PCR 产物在 1.5 %琼脂糖凝胶上进行电泳并进行溴化乙锭染色检测。

表 1 所用引物及针对不同基因的 PCR 扩增反应优化条件

Tab.1 Description of primers and optimal PCR conditions used for specific amplification of different genes

基因在 Gene Bank 中的检索号 GeneAccession No. in Gene Bank	引物序列 ** Primer sequences	扩增片段期望长度 * Expected Length/bp	退火温度 Annealing temperature/	镁离子终浓度 Final MgCl ₂ concentration/ μ mol L ⁻¹	在被检样品中的基因型 Heterozygosity in the tested tree
CAD Z37991	F CGAATCCTGTGAA GTGGIT R CTCGCTACCAATCTCTGTT	870 Genomic DNA	55	250	Homozygous
CHS X60754	F ACTCCCCCTAATGCGGTTGA R CTTGGCTGCGGCTTCTTTC	370 Genomic DNA	62	150	Heterozygous
NIR X74949	F AAGCAAAGGCCCTGAAAAAT R CCTGATGGCTCCAAA GTG	303 cDNA	54	150	Homozygous
ARA554 U09554	F CAGCAATGACGATGGTTTIA R GCCTGGGGACGAGAAC	306 cDNA			failed
GS AJ00119	F AAGTTGGGCTGCACGT R TTGGAAGCTGGCCTGC	600 cDNA			failed

* 表示根据 Gene Bank 中的对应序列扩增出的片段期望长度,并标明序列是来自基因组 DNA,还是来自 cDNA。 ** F, 正向引物; R, 反向引物。 * Indicated the length between the priming sites of the corresponding sequences in gene bank and also whether a genomic clone or a cDNA clone was used to design the primer pairs. ** F, forward primer; R, reverse primer.

1.4 SSCP 分析 SSCP 分析参照 Plomion 等(1999)的方法进行并做了适当调整,将 1 μ L PCR 反应产物与 10 μ L 上样缓冲液(95 %去离子甲酰胺, 10 mmol 氢氧化钠, 0.05 %二甲苯, 0.05 %溴酚蓝)混合,然后在 95 下变性 5 min 并迅速在冰上冷却。单链片段在非变性聚丙烯酰胺凝胶上分离[0.5 \times MDE (Mutation Detection Enhancement gel, BMA)]。电泳在 220 V 电压下、0.6 \times TBE 缓冲液中持续 16 ~ 24 h(随片段大小而异),并对不同的电泳温度进行测试。电泳结束后,凝胶的银染检测按以下步骤进行:1)在含有 10 %乙醇和 1 %乙酸的固定液中固定 2 次,每次 7 min;2)用去离子水清洗 3 次;3)在 0.15 %的硝酸银中染色 30 ~ 60 min;4)用去离子水清洗 1 次;5)在含有 1.5 g 氢氧化钠、1 %甲醛的 100 mL 显色液中显色,这一步骤大约需要 5 ~ 15 min;6)用去离子水清洗 2 次并将凝胶存放在含有 6 %甘油、40 %乙醇及 54 %去离子水的固定液中。最后利用佳能 G2 数码相机进行凝胶拍摄。

1.5 数据收集及分析 通过 SSCP 电泳片段在二倍体针叶和单倍体胚乳组织中的比对,分别进行了等位基因的记录,利用 χ^2 检验对观测到的等位基因的分离是否符合孟德尔比进行了检测。理论上,由于变性不完全及某些片段可能有相同的电泳迁移率,一般在单倍体组织中可能观测到 1 ~ 3 条 SSCP 片段,而在二倍体组织中可能观测到 2 ~ 6 条片段。与群体分析不同的是,对家系材料进行分析时,可以根据等位基因在家系子代中的分离情况从复杂的带型中清晰地识别等位基因。

2 结果与讨论

在对本研究选择的测试基因分析中,根据 Plomion 等(1999)提供的实验条件,没能获得理想的扩增产物。因此首先对 PCR 扩增反应的条件进行了优化,主要对模板浓度、镁离子浓度、退火温度及反应缓冲液的成份进行了测试。针对每一基因的最优反应条件列在表 1 中。在 PCR 反应程序中加入了梯度变温循环以提高扩增反应的特异性。

在这 5 个选出的基因中,ARA554 和 GS 在任何尝试的反应条件下,都不能在马尾松中获得扩增产物。因此认为在马尾松中,这 2 个基因在引物的接合域可能发生了碱基突变。由于 EST 引物序列在种内不同个体间是高度保守的,所以扩增失败不应是取样导致的。因此要在马尾松中分析这 2 个基因,需要根据基因不同位置的序列重新设计引物。其他 3 个基因都在马尾松中成功地扩增。扩增片段的大小与期望片段大小相近(图 1)。因此,与基因来源物种相比,马尾松中对应基因没有因插入或缺失而导致长度的明显变化。

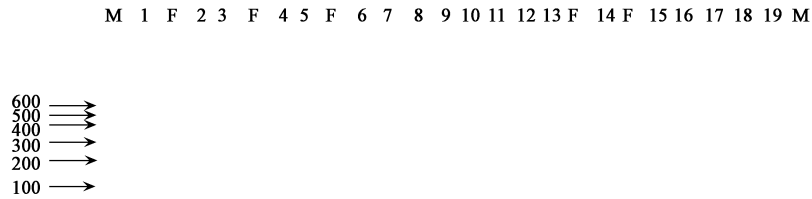


图 1 CHS 基因 PCR 产物的琼脂糖凝胶检测

Fig.1 Detection of the PCR products of CHS in agarose gel

CHS PCR 扩增产物的大小约为 370 bp,与期望片段大小相似,在琼脂糖凝胶上未检测到多态性。M 为标准分子量;F 为扩增失败的样品;1 为从针叶组织提取的 DNA 模板;2~19 为从种子胚乳组织中提取的 DNA 模板;带箭头的数字 100~600 是以碱基数为单位的对应片段的分子量。The size of PCR product of CHS is estimated at about 370 bp,approximately in the similar size as expected and no polymorphism could be identified. M is the molecular weight standard;F stands for the sample failed in amplification;1 is the DNA extracted from needles of the seeding tree;2~19 are DNAs isolated from the megagametophytes of seeds from the same tree;100~600 with arrows are the molecular standard in base pairs.

通过 SSCP 分析发现,CAD 和 NIR 基因在分析的个体中纯合,因此在种子胚乳组织中不发生分离。CHS 在 SSCP 带谱上,在二倍体组织中共产生了 4 条谱带,这 4 条谱带分别代表了 2 个杂合等位基因的 4 条 DNA 单链。这 4 条单链在胚乳组织中形成 2 组分离的谱带,分别对应 2 个不同的等位基因(图 2)。² 分析发现,该基因的分离比率符合孟德尔分离比 ($\chi^2 = 0.22, < 3.84, p = 0.05$)。因此,在检测的 5 个基因中有 3 个(60%)引物在马尾松中通用,扩增出的基因位点中,只有一个(20%)引物位点杂合。这与基于随机标记在火炬松基因组中检测到的位点杂合度相近(Remington *et al.*, 1999)。

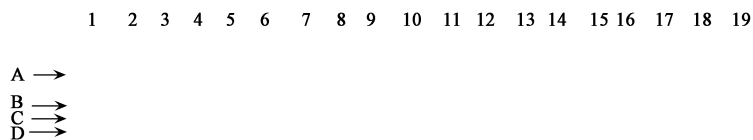


图 2 CHS 基因 PCR 产物的 SSCP 分析

Fig.2 SSCP analysis of the PCR products of CHS

1~19 为图 1 中对应样品的 PCR 产物;A、B、C 和 D 为该基因的 2 个等位基因对应的 4 条 DNA 单链。A 和 B 为等位基因 1 的 2 条单链,C 和 D 为等位基因 2 的 2 条单链。CHS 在检测个体的二倍体组织中是杂合的,因而在单倍体胚乳组织中发生分离。1~19 are the PCR products as indicated in fig.1. A、B、C and D are the bands corresponding to each single strand of the two alleles of this gene. A and B are the single strand of allele 1,C and D are the single strand of allele 2. The CHS gene is heterozygous in the diploid tissue of the tested tree,thus segregates in its haploid megagametophytes.

在马尾松中测试了数量有限的在其他松树中开发的 EST 引物,尽管本研究的数据不足以提供一个可靠的统计参数,但是本研究结果显示:在松属不同树种间能够通用且在单株树中为杂合的 EST 位点比例是比较低的。对于通用的引物是否可以定位,主要取决于它在作图个体中是否杂合。由于只有杂合的位点才能用于作图,因此可以利用多家系构建共祖先图谱的策略进行转录图谱的构建,因为在某一个体中纯合的基因位点,在另一个体中可能为杂合。这样随着作图家系的增多,能够检测到分离的杂合位点比例也会相应增加,种内不同个体间基因在染色体上的排列具有高度的同线性顺序,因此可以利用这些基因产生的标记构建出某一树种的共祖先图谱,将这一物种作为一个大遗传系统进行研究(Yin *et al.*, 2004)。

3 结论

利用编码核 DNA 的序列开发 PCR 标记为转录图谱的构建提供了一个可行的方法(Plomion *et al.*, 1999)。如果引物的序列是根据功能已知的基因设计而来,那么很快就会完成松属转录图谱的构建。在本文提到的突变检测技术中,SSCP 技术操作简单,成本低廉且检测突变的灵敏度高。本研究建立了在松属中进行 SSCP 分析的实验室手工操作的优化程序,同时为正在进行的马尾松转录图谱的构建提供了一个初步的研究结果。根据本研究获得的信息,采用多家系作图策略进行 EST 转录图谱构建比常用的单家系作图策略更为有效。

参 考 文 献

- Andersen M, Larsen P S, Kanters L A *et al.* Mutation detection by cleavase in combination with capillary electrophoresis analysis: Application to mutations causing hypertrophic cardiomyopathy and long-QT syndrome. *Molecular Diagnosis*, 1998, 3:105 - 111
- Bradshaw H D, Villar M, Watson B D *et al.* Molecular genetics of growth and development in *Populus 3*: A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theoretical and Applied Genetics*, 1994, 89:167 - 178
- Bui R G H, Rees C T, Lambrinakos K *et al.* Site-selective reactions of imperfectly matched DNA with small chemical molecules: applications in mutation detection. *Bioorganic Chemistry*, 2002, 30:216 - 232
- Cevera M T, Storme V, Ivens B *et al.* Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and Microsatellite markers. *Genetics*, 2001, 158:787 - 809
- Cotton R G H. Current methods of mutation detection. *Mutation Research*, 1993, 285:125 - 144
- Devey M E, Fiddler T A, Liu B H *et al.* An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. *Theoretical and Applied Genetic*, 1994, 88:273 - 278
- Echt C S, Nelson C D. Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). *Theoretical and Applied Genetics*, 1997, 94:1031 - 1037
- Economou E P, Kazazian H H Jr, Antonarakis S E. Detection of mutations in the factor 8 gene using single-stranded conformational polymorphism (SSCP). *Genomics*, 1992, 5:874 - 879
- Ellis R G H, Humphrey T P, Smith K E *et al.* Chemical cleavage of mismatch: a new look at an established method/recent developments. *Hum Mutat*, 1998, 11: 345 - 353
- Faudoa R, Xue Z, Lee F *et al.* Detection of novel NF2 mutations by an RNA mismatch cleavage method. *Hum Mutat*, 2000, 15:474 - 478
- Grange J C, Gottesman D K, Lewis G S *et al.* Detection of point mutations in type-1 collagen by RNase digestion of RNA hybrids. *Nucleic Acids Res*, 1990, 18: 4227 - 4236
- Grattapaglia D, Sederoff R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics*, 1994, 137:1121 - 1137
- Hayashi K. PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods App*, 1991, 1:34 - 38
- Krizkova J, Sakthivel L, Olowe R *et al.* Human SP-A: genotype and single-strand conformation polymorphism analysis. *Am J Physiol*, 1994, 266:519 - 527
- Liu Z J, Karsi A, Dunham R A. Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. *Marine Biotech*, 1999, 1:437 - 447
- Loopstra C A, Sederoff R. Xylem-specific gene expression in loblolly pine. *Plant Mol Biol*, 1995, 27:277 - 291
- Marshall M, Heisler D J, Lyamichev H *et al.* Determination of hepatitis C virus genotypes in the United States by cleavase fragment length polymorphism analysis. *Journal of Clinical Microbiology*, 1997, 35:3156 - 3162
- Mukai Y, Suyama Y, Tsumura Y *et al.* A linkage map for sugi (*Cryptomeria japonica*) based on RFLP, RAPD, and isozyme loci. *Theoretical and Applied Genetics*, 1995, 90:835 - 840
- Neininger A, Seith B, Hoch B *et al.* Gene expression of nitrite reductase in Scots pine (*Pinus sylvestris* L.) as affected by light and nitrate. *Plant Mol Biol*, 1994, 25:449 - 457
- Nelson C D, Nance W L, Doudrick R L. A partial genetic linkage map of slash pine (*Pinus elliottii* Engelm. var. *elliottii*) based on random amplified polymorphic DNAs. *Theoretical and Applied Genetics*, 1993, 87:145 - 151
- Plomion C, Bahman N, Durel C E *et al.* Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Heredity*, 1995, 74:661 - 668
- Plomion C, Hurme P, Frigerio J M *et al.* Developing SSCP markers in two *Pinus* species. *Molecular Breeding*, 1999, 5:21 - 31
- Remington D L, Wheten R W, Liu B H *et al.* Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theoretical and Applied Genetic*, 1999, 98:1279 - 1292
- Sterky F, Regan S, Karlsson J *et al.* Gene discovery in the wood-forming tissues of poplar: Analysis of 5692 expressed sequence tags. *PNAS*, 1998, 95:13330 - 13335
- Strauss S H, Lande R, Namkoong G. Limitations of molecular-marker-aided selection in forest tree breeding. *Can J For Res*, 1992, 22:1050 - 1061
- Temesgen B, Brown G R, Harry D E *et al.* Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda* L.). *Theor Appl Genet*, 2001, 102:664 - 675
- Wu R L, Han H F, Hu J J *et al.* An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. *Theor Appl Genet*, 2000, 100:1249 - 1256
- Yin T M, Wang X R, Andersson B *et al.* Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. *Theor and appl Genet*, 2003, 106:1075 - 1083
- Yin T M, Difazio S P, Gunter L E *et al.* Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. *Theor and appl Genet*, 2004, 109 (3): 451 - 463