

## 杉木纤维素合成酶类似蛋白基因 *ClCsID1* 的克隆 及其生物信息学分析

彭沙沙, 童再康, 黄华宏, 周厚君, 时 剑, 林二培

(浙江农林大学 林业与生物技术学院, 浙江 临安 311300)

**摘要:** 纤维素合成酶类似蛋白(CSL)作为一种重要的膜蛋白与纤维素合成酶具有相类似的蛋白结构, 都含有 D, D, D, QXXRW 保守区。利用其他物种中的 *CesA* 基因的保守序列设计简并引物, 采用反转录-聚合酶链式反应(RT-PCR)结合 cDNA 末端快速扩增技术(RACE)成功地从杉木 *Cunninghamia lanceolata* 中扩增出一个含有完整阅读框架的 cDNA 序列, 长度为 4 150 bp, 开放阅读框架为 3 396 bp, 编码 1 132 个氨基酸并含有保守的 D, D, D, QXXRW 天冬氨酸残基序列。应用美国国家生物技术信息中心(NCBI)数据库对获得的基因进行序列分析比对, 发现它属于 *CsID* 基因家族, 命名为 *ClCsID1* 基因。多重序列比对结果分析表明, 该蛋白与来自颤杨 *Populus tremuloides*, 水稻 *Oryza sativa*, 拟南芥 *Arabidopsis thaliana* 等的同源基因相似性高达 71% 以上。利用生物软件对其进行生物学分析, 为进一步研究其在植物纤化中的功能奠定基础。图 5 参 10

**关键词:** 林木育种学; 杉木; 纤维素合成酶类似蛋白; 克隆; 表达

中图分类号: S722.1; Q785 文献标志码: A 文章编号: 2095-0756(2012)01-0001-06

## Cloning and bioinformatic analysis of the cellulose synthase-like protein gene, *ClCsID1* from *Cunninghamia lanceolata*

PENG Sha-sha, TONG Zai-kang, HUANG Hua-hong, ZHOU Hou-jun, SHI Jian, LIN Er-pei

(School of Forestry and Biotechnology, Zhejiang A & F University, Lin'an 311300, Zhejiang, China)

**Abstract:** The protein structure of the cellulose synthase-like protein (CSL) was similar to cellulose synthase (*CesA*), including the conservative sequence D,D,D,QXXRW. One full-length cDNA of the cellulose synthase-like protein D (*CsID*) gene was cloned by reverse transcriptase (RT)-polymerase chain reaction (PCR) with 5', 3' rapid amplification of cDNA ends (RACE) methods using degenerate primers designed from the homologous sequences of the *CesA* genes. A multiple comparison sequence analysis was conducted concurrently with bioinformatic methods to analyze the obtained sequence. Results of the sequence analysis showed that this cDNA was 4 150 bp in length and contained a single open reading frame encoding a protein of 1 132 amino acids. The multiple comparison sequence analysis showed that the deduced amino acid sequence shared high similarity (over 71%) with the *ClCsID* genes from *Populus tremuloides*, *Oryza sativa*, and *Arabidopsis thaliana*. This work will help lay an important foundation for further molecular studies with cellulose synthesis of plants. [Ch, 5 fig. 10 ref.]

**Key words:** forest tree breeding; *Cunninghamia lanceolata*; cellulose synthase-like (CSL) protein; clone; expression

纤维素是植物细胞壁的主要成分, 自然界中大约有 1 800 亿 t·a<sup>-1</sup> 的纤维素产物产生, 具有较高的经

收稿日期: 2011-02-10; 修回日期: 2011-06-20

基金项目: 浙江省重大科技专项(2008C02004-1); 浙江省自然科学基金资助项目(Y3090510); 浙江省林业厅资助项目(08A10)

作者简介: 彭沙沙, 从事林木遗传育种研究。E-mail: shashabingyu1129@163.com。通信作者: 童再康, 教授, 博士, 从事林木遗传育种研究。E-mail: zktong@zafu.edu.cn

济价值<sup>[1]</sup>。研究表明：纤维素的生物合成是纤维素合成基因、非纤维素多糖合成基因以及结构蛋白基因等在细胞发育调控下进行协同表达、相互作用的过程。纤维素合成酶作为植物纤维素合成途径中的关键酶影响着纤维的产量和品质，探明纤维素合成酶的作用机制及其在纤维合成过程中的功能，对于材性品质改良至关重要。纤维素合成酶类似蛋白作为一种重要的膜蛋白，与纤维素合成酶具有相似的蛋白结构，都含有 D, D, D, QXXRW 保守区，与棉花 *Gossypium hirsutum* 纤维素合成酶相比，最明显的差异在于后者具有额外的特异序列。在拟南芥 *Arabidopsis thaliana*<sup>[2-4]</sup>，水稻 *Oryza sativa*<sup>[5-6]</sup>，烟草 *Nicotiana glauca*<sup>[7]</sup>，美洲山杨 *Populus tremuloides*<sup>[8]</sup>等多种植物种已发现，纤维素合成酶类似蛋白基因有 A, B, C, D, E, F, G, H 等 8 个庞大的基因家族，也证实了几个类纤维合成酶(Csl)直接参与纤维的生物合成，但对其功能的了解不十分清楚。本研究选用杉木 *Cunninghamia lanceolata* 为材料来研究纤维素合成酶类似蛋白基因，克隆了 1 个 *CsID* 基因，并对其进行生物信息学分析，为该基因的功能研究奠定基础，同时也为利用 *CsID* 基因对杉木进行遗传改良奠定基础。

## 1 材料与方法

### 1.1 试验材料

杉木取自浙江农林大学遗传学科苗圃地(29°56'N, 118°51'E)，取其茎叶作为 mRNA 的提取材料。

### 1.2 试剂

克隆用的大肠埃希菌 *Escherichia coli* 5 $\alpha$  为本研究小组实验室保存；PureLink™ Plant RNA Reagen 试剂盒(Invitrogen)；cDNA 合成试剂盒(Clontech)；T4 DNA 连接酶和克隆载体 pMD-T19 均为 TaKaRa 公司生产；AxyPrep DNA 凝胶回收试剂盒(AxyGEN)；引物由南京金思特生物工程技术服务有限公司合成；cDNA 末端快速扩增(RACE)试剂盒为 Clone-Tech 产品。

### 1.3 试验方法

茎叶 mRNA 的提取按照 PureLink™ Plant RNA Reagen 试剂盒说明书进行，并于 10.0 g·kg<sup>-1</sup> 琼脂葡萄糖凝胶电泳检测 RNA 提取效果(图 1)。

从 GenBank 数据库中获得纤维素合成酶基因的编码序列，通过 Clustal X2 进行多序列比对，根据序列保守性，设计 RT-PCR 引物 RR: 5'-TGYTAYGTNCARTTYCCWC-3'和 RF: 5'-GANCCRTADATCCANCC-3'(Y: C or T; N: A or T or C or G; R: A or G; W: A or T; D: A or G or T)。比对的基因编码序列：火炬松 *Pinus radiata* 登录号 AAQ63936；颤杨 *Populus tremuloides* 登录号 AA025536.1；毛果杨 *Populus trichocarpa* 登录号 XP\_002324291；马占相思 *Acacia mangium* 登录号 AAT66940.1。cDNA 的合成及聚合酶链式反应(PCR)的扩增均参照试剂盒说明书进行。凝胶回收 PCR 产物，按照 pMD-T19 载体试剂盒操作流程，将回收的 DNA 片段连接到 T 载体上，转化大肠埃希菌 DH5 $\alpha$  菌株，蓝白斑筛选，阳性克隆送南京金思特生物工程技术服务有限公司测序。测序结果经 BLAST(basic local alignment search tool)检索序列数据库，确定所克隆序列为杉木 *CsID1* 基因编码序列。根据反转录-聚合酶链式反应(RT-PCR)片段的序列测定结果，合成 *CsID1* 特异引物 GSP1: 5'-ACACATCCCGTCCGACATACACAG-3'和 GSP2: 5'-CTGGCAGATTCCATCCCATCGCAG-3'，分别用于克隆 cDNA 的 5'和 3'端序列。5'RACE 和 3'RACE 具体步骤均按 SMATR™ RACE cDNA Amplification Kit (Clontech)的说明书进行，PCR 产物凝胶回收后经 T 载体克隆测序，并与已克隆的 *CsID1* 基因编码区拼接，从而获得全长 cDNA 序列。以拼接得到的 *CsID1* 全长 cDNA 设计引物 CsID1-U: 5'-GAGACATTTGAATAACAGCGTG-3'和 CsID1-L: 5'-TTGAAGC-TAGTCAGATCCAACCA-3'，再经 PCR 扩增，凝胶电泳分析、测序，验证所获基因全长的准确性。

用 Nucleic Tools 分析推测 *CsID1* 基因编码蛋白质的氨基酸组成，用 ProtParam(<http://www.expasy.org/tools/protparam.html>)推测其分子量和等电点(PI)，用 ProtScale (<http://ca.expasy.org/tools/protscale.html>)推测其疏水性，用 TMHMM(<http://www.cds.dtu.dk/services/TMHMM-2.0/>)对其跨膜区进行预测，并用 SignalP (<http://www.cds.dtu.dk/services/SignalP/>)软件对其 N 端进行分析，并进行蛋白质亚细胞定位，同时用 ExPASy 的 HNN(<http://www.expasy.org/tools/HNN.html>)预测其二级结构进行，利用蛋白质保守结构域推测工具 InterProScan(<http://ebi.ac.uk/Tools/InterProScan/>)分析 *CsID1* 的结构域。

分析杉木 *CsID1* 与其他植物同源蛋白序列的同源性时，在美国国家生物技术信息中心(NCBI)上用

其 BLAST (<http://www.ncbi.nlm.nih.gov/blast>) 分别将杉木的 *ClCsID1* 蛋白质序列与其他植物的 *CsID* 蛋白质序列进行同源性比对。同时用 ClustalX2 软件对不同植物的 *CsID* 蛋白质序列与杉木 *CsID1* 蛋白序列进行多重序列比对, 并用 MEGA3<sup>[9]</sup> 软件构建不同植物的 *CsID* 基因进化树。

## 2 结果与分析

### 2.1 杉木 *CsID1* 基因的分离与分析

用 PureLink™ Plant RNA Reagen 试剂盒提取的 RNA (图 1), 电泳后 28SrRNA 和 18SrRNA 两条主带明显, 且 28SrRNA 条带亮度约为 18srRNA 条带亮度的 2 倍。这说明 RNA 基本上没降解, 且无弥散。同样加样孔附近没有杂带, 说明产物中没有 DNA 的存在。

取上述 RNA 利用逆转录酶获得 cDNAs (图 2), 通过 RT-PCR, 获得 1 个约 600 bp 的产物, 连接到 T 载体且转化 DH5 $\alpha$ , 挑取阳性克隆测序。测序结果用 NCBI 的 BLAST 程序检索数据库, 显示该序列与拟南芥、烟草、毛果杨的同源性都达到 80% 以上, 确定该片段为杉木的 *ClCsID1* 基因片段。用该序列设计特异引物 GSP1 和 GSP2, 分别克隆基因 cDNA 的 5' 和 3' 端序列。测序结果表明: 5' 和 3' 端序列与获得的 *ClCsID1* 基因编码序列有重叠, 为 *ClCsID1* 的 5' 和 3' 端序列。经拼接得到杉木的 *ClCsID1* 基因完整的 cDNA 序列, 共 4 150 bp (图 3)。

测序结果 BLAST 显示: 在核苷酸水平上, 杉木 *ClCsID1* 基因与毛果杨 (XM\_002325781), 颤杨 (AY-162184), 拟南芥 (AF232907), 烟草 (AF304375) 的相似性高达 73%, 73%, 72% 和 70%。其氨基酸序列的同源性与水稻 (DAA01756), 颤杨 (AAO03579), 拟南芥 (NP\_186955), 烟草 (AF304375\_1) 的同源性高达 71%, 81%, 78% 和 74%, 表明所克隆基因确实为杉木的纤维素合成酶类似蛋白 D 基因 (图 4)。

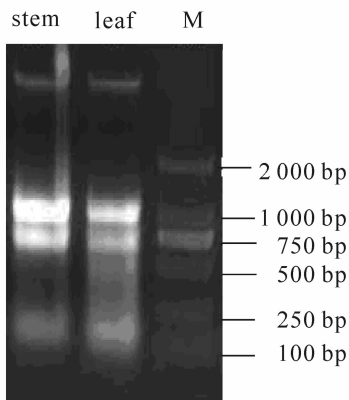


图 1 杉木总 RNA 提取结果  
Figure 1 Total RNA preparation  
*Cunninghamia lanceolata*

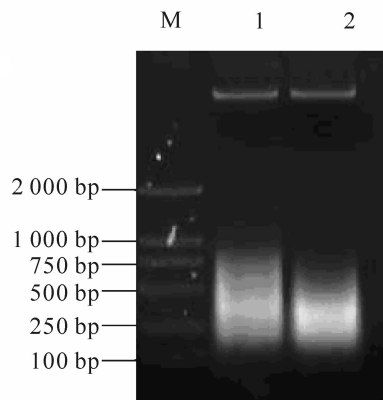


图 2 杉木 RNA 反转录产物  
Figure 2 RT product of *Cunninghamia lanceolata* RNA

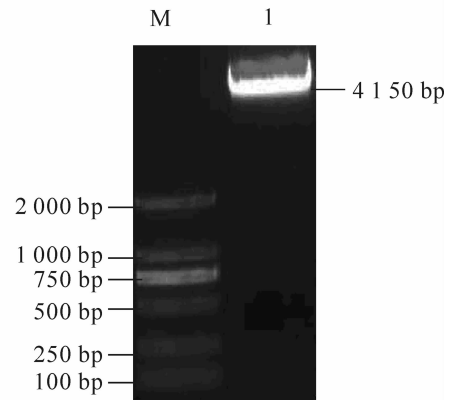
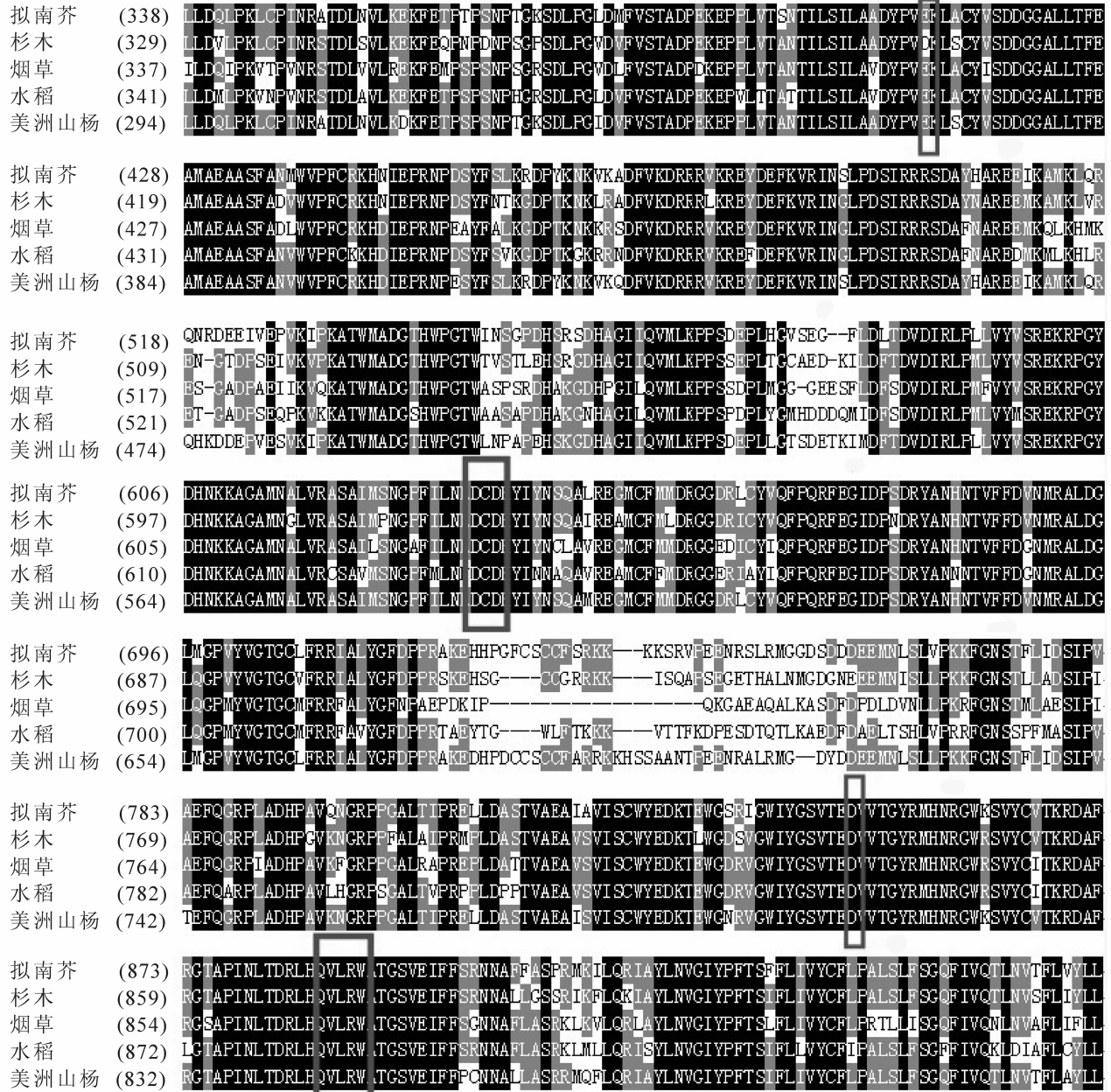


图 3 杉木 *CsID1* 全长 cDNA RT-PCR 产物的琼脂糖电泳图谱  
Figure 3 Agrose gel electropherogram of RT-PCR product of *CsID1* full-length cDNA in *Cunninghamia lanceolata*

### 2.2 杉木 *CsID1* 基因的生物信息学信息

*ClCsID1* 蛋白的结构特点和理化性质。该蛋白相对分子质量为 126 801.2 KD, 等电点为 6.68, 理论推导半衰期大于 30 h, 不稳定系数为 39.64, 属于稳定蛋白; *ClCsID1* 蛋白的二级结构中  $\alpha$ -螺旋 (helix) 占 34.84%,  $\beta$ -折叠 (sheet) 占 14.68%, 无规则卷曲 (coil) 占 50.94%; 疏水性/亲水性分析结果表明, 疏水性最大值为 3.067, 最小值为 -3.222, 总体属于亲水性蛋白; 跨膜结构预测表明, 该蛋白拥有 8 个跨膜区, 分别是 281~303, 310~332, 908~930, 943~962, 977~999, 1 029~1 051, 1 066~1 083 和 1 096~1 115, 跨膜螺旋长度最小 19 个氨基酸残基, 最大 22 个氨基酸残基, 平均长度为 21.25 个氨基酸残基。在保守区 A 中含有保守的天冬氨酸残基, 排列为 Dx...xDCD; 在保守区 B 中含有保守的天冬氨酸残基, 排列为 Dx...QVLRW。蛋白质亚细胞定位预测认定其为细胞质蛋白质 (reliability index=2; expected accuracy=74%), 而且 SignalP 软件分析该蛋白为非分泌蛋白, 不具有信号肽结构。此外, 保守功能结构域分析结果显示,

该蛋白质具有一个 cellulose\_synt 功能域, 该结构域为纤维素合成酶蛋白家族的一个特征性结构域, 具有催化尿苷二磷酸形成的功能。



拟南芥(*Arabidopsis thaliana*, NP\_186955), 烟草(*Nicotiana glauca*, AF304375\_1)水稻(*Oryza sativa*, DAA01756), 美洲山杨(*Populus tremuloides*, AA003579)。D, D, D, QVLRW 保守基序以黑色方框标出。

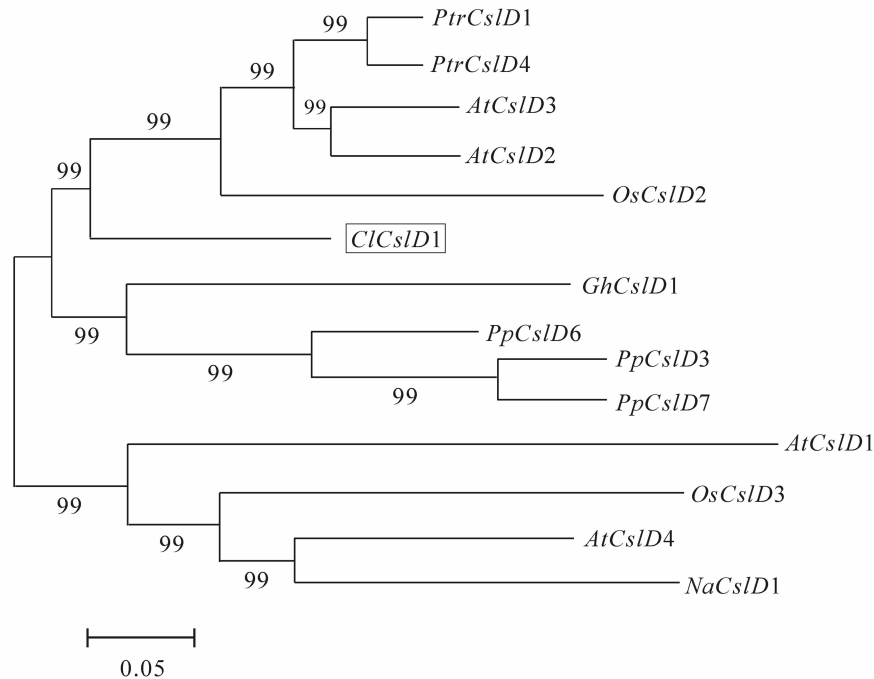
图4 杉木 C1Cs1D1 部分蛋白序列及其他植物同源蛋白序列的比较

Figure 4 Comparison of C1Cs1D1 in *Cunninghamia lanceolata* and its homologous protein from other plants

用 MEGA3 软件的邻接法(neighbor-joining, NJ)构建不同植物 CsID 蛋白的系统发育树(图 5)。从构建的发育树中可以看出: 杉木 CsID1 蛋白与 OsCsID2, AtCsID2, AtCsID3 和 PtrCsID4 距离最近。已有研究表明: AtCsID3 在拟南芥根毛的伸长区及根毛区细胞壁的抗张强度有关, 而 AtCsID2 和 OsCsID1 也在根毛区的表达丰度较强, 推测 C1Cs1D1 可能与根毛区细胞壁的合成有关, 其具体功能需要进一步验证。

### 3 讨论

近年来, 对纤维素合成酶类似蛋白基因的研究获得了初步成果。利用拟南芥突变体发现 *AtCslAT* 基



毛果杨 *Populus trichocarpa*, NP\_002325817.1; 美洲山杨 *Populus tremuloides*, AAO03579.1; 拟南芥 *Arabidopsis thaliana*, NP\_186955.1; 拟南芥 *Arabidopsis thaliana*, NP\_197193.1; 水稻 *Oryza sativa*, AZYU42.1; 杉木 *Cunninghamia lanceolata*; 陆地棉 *Gossypium hirsutum*, ADQ28096.1; 小立碗藓 *Physcomitrella patens*, XP\_001781718.1; 小立碗藓 *Physcomitrella patens*, XP\_001778678.1; 拟南芥 *Arabidopsis thaliana*, NP\_180869.1; 水稻 *Oryza sativa*, DAA01756.1; 拟南芥 *Arabidopsis thaliana*, NP\_195532.1; 烟草 *Nicotiana glauca*, AF304375.1

图 5 植物 CsID 蛋白的系统进化关系分析

Figure 5 Phylogenetic tree analysis of CsID from various plants species

因在拟南芥花粉管及胚胎发育中起重要作用, 推测此蛋白质对于细胞壁合成是必需的<sup>[2]</sup>。用 T-DNA 插入法证明: *AtCsID3* 与拟南芥根毛的伸长及根毛区细胞壁的抗张强度有关, 推测此蛋白质可能与根毛区细胞壁合成有关<sup>[4]</sup>, 而 *AtCsID2* 也在根毛区表达丰度较高。Doblin 等<sup>[8]</sup>研究发现: 烟草 *NaCsID1* 基因可能是编码花粉管纤维素合成的特异基因。Anita 等<sup>[7]</sup>利用定量 PCR 技术发现 *PtrCsID2* 基因在杨树木质部细胞的次生壁中有较高丰度的表达, 参与纤维素的合成。Butron 等<sup>[10]</sup>将仅在水稻和大麦以及其他谷类中表达的 *CsID* 基因插入到拟南芥中, 结果发现转基因拟南芥所产生的  $\beta$  葡聚糖量很低, 意味着 *CsID* 基因与纤维合成有关, 同时也表明这种细胞壁成分还需要其他的“补充因子”。植物纤维素的合成是一个复杂的过程, 而纤维素合成酶和纤维素合成酶类似蛋白基因更是一个大的基因家族。本研究的局限性在于仅以茎叶组织为材料, 很难克隆出家族中的所有成员。系统研究其分子调控过程, 需要克隆更多的纤维素合成酶及类似蛋白基因。*CiCsID1* 基因的发现为进一步在分子水平上研究杉木的纤化打下了基础。

#### 参考文献:

- [1] ENGLEHARDT J. Sources, industrial derivatives, and commercial applications of cellulose [J]. *Carbohydr Eur*, 1995, **12**: 5 - 14.
- [2] GOUBET F, MISRAHI A, PARK S K, et al. *AtCSLA7*, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis* [J]. *Plant Physiol*, 2003, **131** (2): 547 - 557.
- [3] FAVERY B, RYAN E, FOREMAN J, et al. *KOJAK* encodes a cellulose synthase-like protein required for root hair cell morphogenesis in *Arabidopsis* [J]. *Genes & Dev*, 2001, **15**: 79 - 89.
- [4] WANG X, CNOPS G, VANDERHAEGHEN R, et al. *AtCsID3*, a cellulose synthase-like gene important for root hair growth in *Arabidopsis* [J]. *Plant Physiol*, 2001, **126** (2): 575 - 586.

- [5] HAZEN S P, SCOTT-CRAIG J S, WALTON J D. Cellulose synthase-like genes of rice [J]. *Plant Physiol*, 2002, **128** (2): 336 – 340.
- [6] KIM C M, PARK S H, JE B I, *et al.* *OsCslD1*, a cellulose synthase-like *D1* gene, is required for root hair morphogenesis in rice [J]. *Plant Physiol*, 2007, **143** (3): 1220 – 1230.
- [7] SAMUGA A, JOSHI C P. Cloning and characterization of cellulose synthase-like gene, *PtrCslD2* from developing xylem of aspen trees [J]. *Physiol Plant*, 2004, **120**: 631 – 641.
- [8] DOBLIN M S, De MELIS L, NEWBIGIN E, *et al.* Pollen tubes of *Nicotiana alata* express two genes from different  $\beta$ -glucan synthase families [J]. *Plant Physiol*, 2001, **125** (4): 2040 – 2052.
- [9] KUMAR S, TAMURA K, NEI M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Brief Bioinform*, 2004, **5** (2): 150 – 163.
- [10] BURTON R A, WILSON S M, HRMOVA M, *et al.* Cellulose synthase-like *CslF* genes mediate the synthesis of cell wall(1, 3; 1, 4)- $\beta$ -D-glucans [J]. *Science*, 2006, **311** (5769): 1940 – 1942.

## 浙江省现代农业科技成果展示交流洽谈会

2011年11月11日,由浙江省科技厅和浙江农林大学共同主办的浙江省现代农业科技成果展示交流洽谈会在浙江农林大学东湖校区举行。通过举办现代农业科技成果展、“十二五”现代种业发展工程研讨会、科技创新服务平台交流座谈会等活动,搭建农业科技协作交流的综合平台,推动浙江省现代农业产业发展和农业产业机制建设。

现代农业科技成果展重点展示了浙江省14个省级农业科技创新服务平台的最新技术成果,主要包括竹产业、木材加工产业、农业种质资源与品质创新、农产品加工等相关技术成果150多项。在现代种业发展工程研讨会上,林木、畜禽、茶树、中药材、食用菌、蚕桑、水稻、旱作粮油、蔬菜、水产、果品、花卉等12位农业重点领域育种攻关协作组牵头人,分别介绍了新品种选育协作攻关方案。浙江省水稻育种攻关协作组介绍了协作组运行和管理做法及经验。参会人员对各领域育种协作攻关方案进行了讨论发言。在浙江省重大科技创新服务平台交流座谈会上,14家农业组创新服务平台总结汇报了过去一年的工作和下一阶段的工作思路。

交流洽谈会作为“2011中国浙江网上技术市场活动周暨杭州科技合作周”系列活动之一,目的是为了加快浙江省战略新兴产业的发展,加强科技对浙江省现代农业产业发展的引领和支撑作用,加快农业科技成果转化与产业化。