

陆地棉 *Dirigent-like* 基因(*GhDIR*) 的克隆与分析

赵付安^{1,2} 房卫平¹ 杨晓杰¹ 谢德意¹ 李 武¹ 唐中杰¹

(1. 河南省农业科学院 经济作物研究所, 河南 郑州 450002; 2. 河南大学 生命科学学院, 河南 开封 475100)

摘要: *Dirigent-like* 基因(*DIR*) 编码蛋白质参与木质素的形成, 与植物体的病菌防御有关。本研究以抑制差减杂交(SSH) 技术获得 *DIR* 的 EST 为信息探针, 对棉花 EST 数据库进行同源搜索和序列拼接, 获得了全长为 768 bp 的 *DIR* 基因 cDNA 全长, 将该基因命名为 *GhDIR*。该基因序列的开放阅读框(ORF) 位于 70~594 bp, 推测编码 175 个氨基酸残基的蛋白。用陆地棉基因组 DNA 和 cDNA 分别进行了 PCR 扩增验证, 克隆测序结果表明, 所得 *GhDIR* 序列与电子克隆序列完全一致, 没有内含子。序列比对和系统进化分析表明, 该基因与海岛棉的同源序列表现了最高的相似性, 与其他双子叶植物同源序列的相似性次之, 而与单子叶植物同源序列的相似性较远。实时定量 PCR 表达谱分析表明, 该基因在陆地棉幼苗根部受黄萎病菌诱导表达。

关键词: 陆地棉; 电子克隆; *DIR-like* gene; 表达分析; 大丽轮枝菌

中图分类号: Q78 文献标识码: A 文章编号: 1000-7091(2011)05-0029-05

Cloning and Analysis of Upland Cotton(*Gossypium hirsutum*) *Dirigent-like* Gene(*GhDIR*)

ZHAO Fu-an^{1,2}, FANG Wei-ping¹, YANG Xiao-jie¹, XIE De-yi¹, LI Wu¹, TANG Zhong-jie¹

(1. Institute of Cash Crops Research, Henan Academy of Agricultural Science, Zhengzhou 450002, China;
2. College of Life Science, Henan University, Kaifeng 475100, China)

Abstract: *Dirigent-like* gene encodes a protein involved in the formation of lignin and in defence to pathogen disease. Using an EST of a *dirigent-like* gene as the query probe to blast cotton EST database, a full-length 768 bp cDNA sequence of the *dirigent-like* gene was obtained and cloned by PCR, was named *GhDIR*. Sequence analysis indicated that it contained a complete open reading fragment(ORF) from 70 to 594 bp, encoding 175 amino acids, and has no intron. Sequence comparison and phylogenetic analysis showed the *GhDIR* cDNA was highly identical with its ortholog from Sea Island cotton, and the next most similar orthologs coming from other dicotyledons, but sharing low homology with monocotyledons. Quantitative PCR analysis showed that *GhDIR* was activated in roots of Upland cotton inoculated with *Verticillium dahliae*.

Key words: *Gossypium hirsutum*; In silico cloning; *DIR-like* gene; Expression analysis; *Verticillium dahliae*

棉花黄萎病(*Verticillium wilt*) 是一种土传病害, 分布广、危害重, 是目前全球范围内危害棉花的第一病害。我国是世界上的产棉大国, 棉花生产对国民经济的发展具有重要影响, 而我国棉花生产一直受枯、黄萎病的制约, 尤其是黄萎病对棉花生产的影响最为严重。因此, 如何防治棉花黄萎病成为我国棉花生产上亟待解决的一个问题。利用棉花抗病品种无疑是一种最经济、有效和环保的措施, 而培育抗病品种需要明确抗病机制和寻找高抗黄萎病的基因^[1]。目前, 由于棉花的 *R* 基因(*Resistant gene*)^[4]

尚未克隆, 用于棉花抗病转基因研究的只有 *D* 基因(*Defence gene*)^[5], 主要包括与植保素合成有关的基因、病程相关(*Pathogenesis-related*, *PR*) 蛋白基因和与结构抗性有关的基因等, 最常见是几丁质酶(*Chitinase*, *CHT*)^[6] 和 β -1, 3-葡聚糖酶(β -1, 3-*glucanase*, *GLU*) 基因^[7]。

Dirigent-like(*DIR*) 基因编码的蛋白质参与木质素的形成, 与植物体的病菌防御有关。本研究以前期构建的黄萎病菌诱导下抗病陆地棉豫棉 21 号 SSH 文库中, 上调表达 *DIR* 基因的 EST 为信息探

收稿日期: 2011-05-02

基金项目: 国家转基因生物新品种培育科技重大专项(2008ZX08005-001)

作者简介: 赵付安(1966-), 男, 河南泌阳人, 副研究员, 在读博士, 主要从事棉花分子育种研究。

通讯作者: 房卫平(1963-), 男, 河南虞城人, 研究员, 博士, 博士生导师, 主要从事棉花育种研究。

针 利用电子克隆的方法, BLAST 检索 GenBank 中的棉花 EST 数据库, 拼接同源的 EST 序列, 获得了陆地棉 *DIR* 基因的 cDNA 全长序列, 并通过 PCR 手段验证其真实性。利用生物信息学的方法对该 cDNA 序列进行了分析, 并与其他物种的同源蛋白序列进行了系统进化分析。利用实时定量 PCR 法研究了豫棉 21 号在黄萎病菌接种不同时间的表达情况。为进一步利用该基因改良棉花的黄萎病抗性研究奠定了基础。

1 材料和方法

1.1 材料

陆地棉抗黄萎病品种豫棉 21 号由河南省农业科学院经济作物所保存。棉花黄萎病 W 株系由河南农业大学植保学院汪敏博士提供。

1.2 试剂

DNA Marker, 质粒及 PCR 产物回收试剂盒购自天根生物制品公司(北京), 电泳级琼脂糖购自 Amresco 公司(美国)。Taq 酶、pMD18-T 克隆载体购自大连宝生物程公司(大连); cDNA 第一链合成采用 Invitrogen 公司的 PrimeScript™ RT-PCR Kit 试剂盒(美国); 引物合成和基因测序由北京六合华大基因科技股份有限公司完成。

1.3 总 DNA、RNA 提取及 cDNA 第一链的合成

基因组 DNA 的提取采用改良 CTAB 法^[8] 棉花根部总 RNA 提取采用天泽基因 Column Plant RNAout 试剂盒。cDNA 第一链合成采用 Invitrogen PrimeScript™ RT-PCR Kit 提供的方法。

1.4 电子克隆

以 SSH 技术得到的差异表达序列为查询探针 (Query probe), 检索棉花 EST 库, 得到同源性高的棉花 EST; 以得到的棉花 EST 为查询探针, 再次检索 dbEST, 得到由部分重叠的 EST 进行拼接, 再以拼接后的序列为查询探针, 依次进行 EST 拼接。

1.5 基因扩增

根据电子克隆获得的完整 cDNA 序列, 分别在 3' 和 5' 端非翻译区设计特异引物, RF: 5'-ATTTCTCCTCCTCACCCCAGAAATT-3' 和 RR: 5'-AGTCGACCCCGTTTTATCAAAAAA-3'。分别以豫棉 21 号基因组 DNA 和 cDNA 为模板扩增特异片段。PCR 程序为: 94℃ 预变性 2 min, 进入循环: 94℃ 变性 45 s, 57℃ 复性 45 s, 72℃ 延伸 90 s, 共进行 32 个循环。最后一轮 72℃ 延伸 10 min。

1.6 克隆、测序

用 0.8% 琼脂糖凝胶电泳纯化回收 PCR 产物,

连接到 pMD18-T simple 载体上, 转化感受态大肠杆菌 DH1 α 挑选阳性克隆进行 PCR 及酶切鉴定, 委托北京六合华大基因科技股份有限公司进行 DNA 序列测序。

1.7 生物信息学分析

用 DNAMAN 软件对 *GhDIR* 基因的开放阅读框 (ORF) 进行物种间多重比较及进化分析; 根据 ExPASy proteomics tools 系统 (<http://www.expasy.ch/tools/>) 中的 GOR4 蛋白质二级结构模型预测软件进行蛋白质二级结构预测分析。

1.8 基因表达谱分析

当棉苗两片真叶平展时, 接种黄萎病菌孢子菌液 (1.4×10^7 个/mL), 分别于接菌后 0, 6, 12, 24, 48 h 提取豫棉 21 号根的总 RNA。以各时点 cDNA 为模板, 以棉花组成型表达的 18S rRNA 基因 (JF699283) 为内部参照, 进行定量 PCR 分析。Gh-*DIR* 及内参的引物序列: DIRF: 5'-ATTTCTCCTCACCCCAGAAATT-3' 和 DIRR: 5'-AGTCGACCCCGTTTATCAAAAAA-3'; 18SF: 5'-TGAAGAACGCAGCGA AAT-3' 和 18SR: 5'-ATGCTCCAAGGAATACCAAG-3'。实时定量 PCR 反应程序采用三步法, 其中在延伸阶段收集荧光信号, 反应结束后, 添加溶解曲线, 便于分析扩增的特异性。具体反应条件如下: 94℃ 变性 10 min, 94℃ 变性 30 s, 55℃ 复性 10 s, 72℃ 延伸 30 s, 共进行 40 个循环。最后一轮 72℃ 延伸 10 min。

2 结果与分析

2.1 棉花 *DIR* 基因 cDNA 序列的电子克隆

以抑制差减杂交 (SSH) 技术获得的一条差异表达 EST 为信息探针, 对棉花 EST 数据库进行同源搜索和序列拼接, 获得的 3 条 EST 序列 (JG454946、ES844863 和 DW516607.1) 作为第一轮组装拼接的候选 EST, 其与信息探针的相似性程度均为 100%。将其组装拼接成 Contig, 以获得的 Contig 作为新的信息探针对棉花 EST 数据库进行 BLAST 检索, 无新的可利用的 EST (图 1)。



图 1 棉花 *DIR* 基因的 cDNA 序列拼接示意图

Fig. 1 Diagram of cotton *DIR* cDNA assembly
用 CAP EST Assembler 程序将上述 4 条候选

EST 序列进行 Contig 拼接,结果显示能够形成一个完整的 Contig,即获得长为 768 bp 的棉花 DIR 基因的 cDNA 序列。拼接结果见图 1。

2.2 棉花基因组 DNA 和总 RNA 提取和基因全长的获得

提取豫棉 21 号基因组 DNA,紫外分光光度计测定基因组 DNA 的 $A_{260/280}$ 值为 1.87,表明纯度较高;琼脂糖凝胶电泳结果同样表明(图 2),基因组 DNA 无降解,完整性较好。以上表明提取的基因组 DNA 可以用于扩增 *DIR* 全长基因的 PCR 模板。分别提取接种黄萎病菌以及未接种黄萎病菌的豫棉 21 号总 RNA,紫外分光光度计测定 RNA 的 260/280 值分别为 2.01 和 1.99,表明总 RNA 纯度很好,电泳结果也表明 28S/18S 亮度比约等于 2,可以用于 cDNA 第一链的合成等后续试验(图 3)。

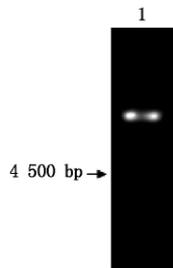


图 2 豫棉 21 号基因组 DNA 电泳检测
Fig.2 Gel electrophoresis of cotton(Yumian21) genomic DNA

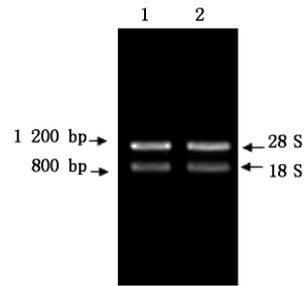


图 3 豫棉 21 号总 RNA 电泳检测

Fig.3 Gel electrophoresis of cotton(Yumian21) total RNA

利用 DIRF 和 DIRR 引物组合,分别对豫棉 21 号基因组 DNA 和 cDNA 进行扩增,所得片段长度一致,均为 684 bp 左右(图 4)。回收、克隆测序结果和推导的氨基酸序列见图 5。

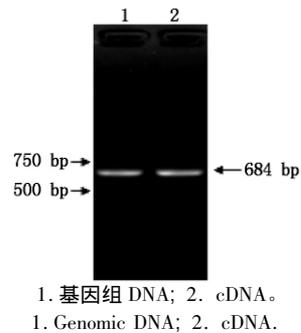


图 4 PCR 产物凝胶电泳检测

Fig.4 Gel electrophoresis of PCR products

```

1  AAGCAACAGT TCCACT AACT TTCATCT ACA TTTCTCCTC ACCCC AGAAA TT ATAGTAAA
61  GCAGAGACA ATGAGAGGAAC ATCGGTGTTG AGTTGGATTC TGATCATCTG CCTCTCCCAA
      M R G T S V L S W I L I I C L S Q
121  GTAGCAGTGC GGAGCCAATA CTACTCGGAC ACCGTACCAT ATCAACCCAG GCCAGTTGTG
      V A V R S Q Y Y S D T V P Y Q P R P V V
181  GTCACCAATC TTCACTTCTA TATGCACGAA TTTACGGGCA CTACAGCAGT CGTGCT AACC
      V T N L H F Y M H E F T G T T A V V L T
241  CAAGCTAACA TCACAAGCAA TAATACATCA GTGAC ATTTG CCACCCTAGT TGCCGTT AAT
      Q A N I T S N N T S V T F A T L V A V N
301  GATCCCCTCAGGACTGGTCC TGAGCCTGAC TCAGAGCTGA TTGGAATGT TCAGGGTCTT
      D P L R T G P E P D S E L I G N V Q G L
361  TCCCTGCTAG CCGGATCAAA TGCATCGAGC ACGCAGT ACAT AGATTTTGA ATTCAAT ACC
      S L L A G S N A S S T Q Y I D F E F N T
421  GGTAA GTTTA ACGGAAGCTC TCTAAGCGTA TTTTCAAGGG GAGAAGCTGG GCTTGCGGTG
      G K F N G S S L S V F S R G E A G L A V
481  GTCGGAGGAA GAGGCCAATT CGCGATGGCAACAGGGACTG CACTATTTAA CCCTCTCCTT
      V G G R G Q F A M A T G T A L F N P L L
541  ATAAATGCC A CCAATGTCAT T A TGGA GTTT AA TTTTACTG TAATTCATTA CTAAGATGTT
      I N A T N V I M E F N F T V I H Y
601  ATTT ACAGAG TCAATCCTCT TGGATTGTGTT GGAAGTT AT A TGA TT ATGCT CCACTTGTTA
661  GAAT AAA TGT ATGAGAAAAG TTGACACCAT TTTTITGAT AAAACGGGGTC GACTTGGGTT
      加尾信号
721  TTGAGAAATA AAAACGAATA TTGGAGTCAA AAAAAAAAAAAAAAAAAAAAAA
      加尾信号 PoLy(A)
    
```

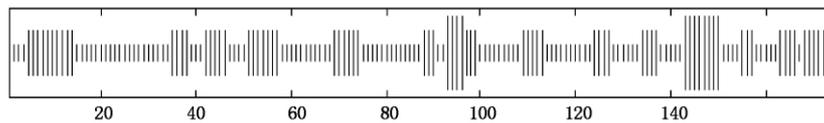
 . 起始密码子和终止密码子。 . Start and stop codon.

图 5 陆地棉 *DIR* 基因的 cDNA 序列及推导的氨基酸序列

Fig.5 The cDNA sequence and deduced amino acid sequence of *GhDIR*

2.3 开放阅读框架分析

通过 NCBI 上的 ORF finder 服务器对所获得的 cDNA 序列进行开放读码框架分析,发现其开放阅读框架位于第 70 ~ 594 位,推测编码 174 个氨基酸(图 5)。在 cDNA 5'端第 70 位存在起始密码子 ATG;3'端的 592 处存在终止密码子 TAA。3'UTR 中 662 ~ 667 位和 726 ~ 731 位典型的加尾信号



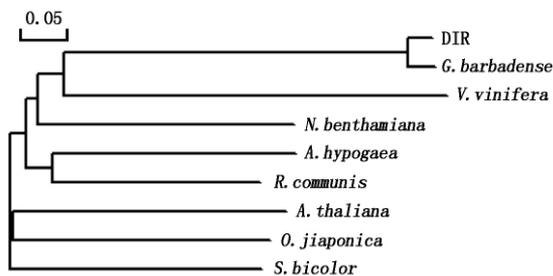
Long verticals represent Alpha helix, Medium-length verticals represent Extended strand, Short verticals represent random coil.
长竖线表示 α -螺旋; 中长竖线表示扩展链; 短竖线表示无规则卷曲。

图 6 陆地棉 DIR 基因编码蛋白二级结构预测

Fig. 6 Secondary structure prediction of the encoded protein of DIR from *Gossypium hirsutum*

2.4 不同物种间 DIR 蛋白进化分析

植物同源蛋白系统分析表明:棉属海岛棉(*Gossypium barbadense*, AAY44415.1)亲缘关系更近一些,以烟草(*Nicotiana benthamiana*, BAF02555.1)为界,与双子叶植物拟南芥(*Arabidopsis thaliana*, AEE33505.1)、蓖麻(*Ricinus communis*, XP_002537933.1)、落花生(*Arachis hypogaea*, AAZ20288.1)、葡萄(*Vitis vinifera*, CAN73537.1)亲缘关系次之,和单子叶植物水稻(*Oryza japonica*, NP_001150958.1)、高粱(*Sorghum bicolor*, XP_002441907.1)亲缘关系稍远一些(图 7)。



线段长度表示进化距离。
The length means evolution distance.

图 7 植物 DIR 蛋白系统发育树

Fig. 7 Phylogenetic tree of DIR proteins

2.5 陆地棉 DIR 基因表达谱分析

采用天泽基因公司的 Column Plant RNAout Kit 提取接菌和未接菌棉根的总 RNA,并对总 RNA 完整性和浓度进行检测,结果表明,RNA 的完整性和浓度符合实时定量 PCR 等下游试验要求(图 3)。以组成型表达内源基因 18S rRNA 为内参基因,以未接菌豫棉 21 号为对照进行定量分析,结果表明,接菌后 6 h DIR 基因表达上调,48 h 后达到 2.6 倍(图 8)。该结果表明,陆地棉抗病品种豫棉 21 号的 DIR 基因受黄萎病菌诱导上调表达,提示该基因参与陆地棉对黄萎病的响应,与抗黄萎病相关。

AATAA。

将推导的蛋白质氨基酸序列输入 GOR secondary structure prediction 蛋白质二级结构模型预测软件分析结果见图 6。该蛋白的二级结构,是由分布在整个蛋白内许多 α -螺旋结构、大量的无规则卷曲和扩展链结构组成,其中 α -螺旋结构占 6.9%,而扩展链和无规则卷曲各占 36.21% 和 56.90%。

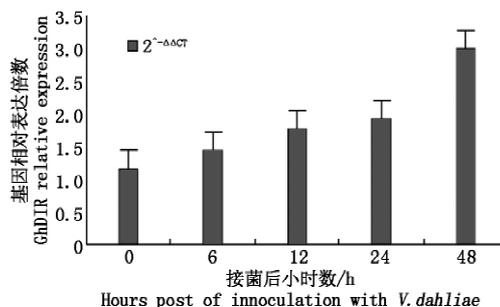


图 8 豫棉 21 号 DIR 基因的表达谱分析

Fig. 8 Expression analysis of DIR gene in different time of upland cotton(Yumian 21)

3 讨论

DIR 蛋白以拉丁文 dirigere 命名, Davin 等^[9]发现它是一个 26 kDa 的 Glycosylated 蛋白能形成一个 78 kDa 的聚合物。该聚合物没有催化中心,推测它的功能用于捕获木质素单体自由基,协助单体以立体选择方式形成二聚体,继而与氧化酶发生作用^[10]。

Wu 等^[11]通过 cDNA 微阵列筛选方法从还魂草(*Boea hygrometrica*)中克隆了 *BhDIR1*,预测它与木质素的生物合成相关。*BhDIR1* 转录本与植物对水分、高(低)温度胁迫及对外源激素的应答反应有关。在干旱胁迫下,木质素的酸可溶性下降、木质素的物理特性改变,从而影响细胞壁的抗压性和弹性。最近发现,细胞壁的抗压性和弹性在干旱胁迫状态下水的恢复中扮演重要角色。细胞壁的折叠有利于质膜的还原,减轻质膜的损伤,维持细胞结构在干旱状态下的完整性^[12,13]。

Burlat 等^[14]发现,DIR 蛋白介导木质素的合成,起防卫作用。Ralph 等^[15]发现,DIR 对于茎部昆虫的侵袭有应答反应。*Dirigent-like* 与漆酶结合,只有 (+)-pinoresinol 形成。如果 *Dirigent-like* 缺失,那么

这个反应产生外消旋的混合物^[16]。Davin 和 Lewis^[17]认为 *Dirigent-like* 通过捕获 E-松针乙醇自由基来引导立体选择结合。Zhu 等^[18]从海岛棉 (*Gossypium barbadense*) 中获得两个 dirigent-like gene 类似物的克隆 *Gbd 1* 和 *Gbd 2*, 这两个基因具有 82% 的相似性, 它们是被真菌 *Verticillium dahliae* 感染后诱导的。Ralph 等^[19]在云杉中发现 *Dirigent-like* 在茎组织的表达量最高, 在昆虫侵染茎生组织有基本的保卫机制, 还会诱导产生酚类物质杀死昆虫。云杉防御昆虫侵害, 不同的种类有不同的防御机制^[20], 其中云杉树皮为防御机制提供了一个耐久的基本的防御障碍, 其中包括软木化或木质化的周皮、石细胞、韧皮部薄壁组织细胞和富集萜类树脂的树脂输送管系统^[21, 22]。

综上所述 *Dirigent-like* 基因在多种植物中与木质素的合成相关, 而且会产生棉酚, 杀死害虫, 起到物理和化学防御作用。据推测这在棉花抗黄萎病中可能会发挥重要作用, 包括结构抗性和化学抗性。

参考文献:

- [1] 张桂寅. 棉花黄萎病抗性表现及其基因表达的研究 [D]. 保定: 河北农业大学, 2005.
- [2] 潘家驹. 棉花育种学 [M]. 北京: 中国农业出版社, 1998.
- [3] Melissa K Hill. Identification of disease response genes expressed in *Gossypium hisutum* upon infection with the wilt pathogen *Verticillium dahliae* [J]. Plant Molecular Biology, 1999, 40: 289 - 296.
- [4] 涂礼莉. 海岛棉 NBS 类抗病基因类似物的遗传多样性及进化 [J]. 遗传学报, 2003, 30(11): 1071 - 1077.
- [5] Gao Y, Guo W, Wang L, et al. Isolation and characterization of resistance and defense gene analogs in cotton (*Gossypium barbadense* L.) [J]. Sci China C Life Sci, 2006, 49(6): 530 - 542.
- [6] 程红梅, 简桂良, 倪万潮, 等. 转几丁质酶和 β -1, 3-葡聚糖酶基因提高棉花对枯萎病和黄萎病的抗性 [J]. 中国农业科学, 2005, 38(6): 1160 - 1166.
- [7] WU J H. Transgenic cotton plants of chitinase and glucanase genes and their performance of resistance to *Verticillium dahliae* [J]. Acta Genetica Sinica, 2004, 31(2): 183 - 188.
- [8] 宋国立, 崔荣霞, 王坤波. 改良 CTAB 法快速提取棉花基因组 DNA [J]. 棉花学报, 1998, 10(5): 215 - 217.
- [9] Davin L B, Wang H B, Anastasia L, et al. Stereoselective bimolecular phenoxyl radical coupling by an auxiliary (*dirigent-like*) protein without an active center [J]. Science, 1997, 17: 275 - 362.
- [10] Shang H H. A cytology study of drought tolerance mechanisms of the resurrection plant *Boea hygrometrica* (Bunge) R [D]. Br. Master Dissertation, Southwest University, 2006.
- [11] Wu R H, Wang L L, Wang Z. Cloning and expression analysis of a dirigent-like protein gene from the resurrection plant *Boea hygrometrica* [J]. Progress in Natural Science, 2009, 19: 347 - 352.
- [12] Vicié M, Olivier L, Farrant J, et al. Composition and desiccation-induced alterations of the cell wall in the resurrection plant *Craterostigma wilmsii* [J]. Physiol Plant, 2004, 120: 229 - 239.
- [13] Francesca D V, Toufik E, Roberto C, et al. Morphological and ultrastructural aspects of dehydration and rehydration in leaves of *Sporobolus stapfianus* [J]. Plant Growth Regul, 1998, 24: 219 - 228.
- [14] Vincent Burlat. Dirigent-like proteins and dirigent-like sites in lignifying tissues [J]. Phytochemistry, 2001, 57: 883 - 897.
- [15] Steven G Ralph. Dirigent-like proteins in conifer defense II: Extended genediscovery, phylogeny, and constitutive and stress-induced gene expression in spruce [J]. Phytochemistry, 2007, 68: 1975 - 1991.
- [16] Hall S C, Lewis N G. Secondary and quaternary structures of the (+) -pinosresinol-forming dirigent-like protein [J]. Biochemistry, 2002, 41: 9455 - 9461.
- [17] Davin L B, Lewis N G. Dirigent-like proteins and dirigent-like sites explain the mystery of specificity of radical precursor coupling in lignin and lignin biosynthesis [J]. Plant Physiol, 2000, 123: 453 - 462.
- [18] Zhu L, Zhang X, Tu L, et al. Isolation and characterization of two novel dirigent-like-like genes highly induced in cotton (*Gossypium barbadense* and *G. hirsutum*) after infection by *Verticillium dahliae*. J [J]. Plant Pathol, 2007, 89: 41 - 45.
- [19] Ralph S. Dirigent-like proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce [J]. Plant Molecular Biology, 2006, 60: 21 - 40.
- [20] Huber D P W, Ralph S, Bohlmann. Genomic hardwiring and phenotypic plasticity of terpenoid-based defenses in conifers [J]. J Chem Ecol, 2004, 30: 2399 - 2418.
- [21] Raffa K F, Berryman A A. Accumulation of monoterpenes associated with volatiles following inoculation of grand fir with a fungus transmitted by the fir engraver, *Scolytus ventralis* (Coleoptera: Scolytidae) [J]. Can Entom, 1982, 114: 797 - 810.
- [22] Franceschi V R, Krokene P, Krekling T, et al. Phloem parenchyma cells are involved in local and distant defense responses to fungal inoculation or bark-beetle attack in Norway spruce (Pinaceae) [J]. Am J Bot, 2000, 87: 314 - 326.