

美洲黑杨 *PdZFR* 基因的克隆和分析

王大海, 苏晓华*, 张冰玉, 黄秦军, 张香华

(中国林业科学研究院林业研究所, 国家林业局林木培育重点实验室, 北京 100091)

摘要:根据美洲黑杨 *PdZFR* 部分 cDNA 序列, 采用拟基因组文库步行法和 RACE 克隆了该基因全长 cDNA 及基因组 DNA, 并对该基因的启动子、转录起始位点、内含子分布、剪切方式、编码的氨基酸序列和该基因在染色体上的分布位置进行了分析。组织或器官特异表达分析表明: 该基因与植物发育相关。

关键词: *PdZFR*; 美洲黑杨; RACE; VGL-Walking; TATA Box; CAAT Box

中图分类号: S792.11

文献标识码: A

Cloning and Analyzing of Gene *PdZFR* from *Populus deltoides*

WANG Da-hai, SU Xiao-hua, ZHANG Bing-yu, HUANG Qin-jun, ZHANG Xiang-hua

(Research Institute of Forestry, Chinese Academy of Forestry; Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Beijing 100091, China)

Abstract: The full length cDNA and genomic DNA of *PdZFR* were cloned from *Populus deltoides* by RACE and Visual Genome Library Walking methods respectively, according to its partial cDNA sequence. Meanwhile, the gene promoter, transcription initiation, intron and exon distribution, splice pattern, the amino acid sequence encoded and the location of this gene on chromosome were analyzed. The tissue or organ-specific expression analysis showed that this gene was related to plant development.

Key word: *PdZFR*; *Populus deltoides*; RACE; VGL-Walking; TATA Box; CAAT Box

含 ZF-RING (Zinc Finger RING) 结构域的蛋白普遍存在于真核生物中, 并具有较高的表达水平。从拟南芥 (*Arabidopsis thaliana* L.) 基因组预测的 25 498 个基因, 有 387 个基因含 ZF-RING 结构域, 占 1.42%。果蝇、线虫和酵母基因组中, 据推测含 ZF-RING 结构域的基因占 0.70% ~ 0.75%^[1]。含 ZF-RING 结构域的蛋白在基因转录表达、细胞程序性凋亡调控、蛋白折叠组装等方面起重要作用。许多疾病相关基因, 如 *BRCA1*、*Cb1*、*BNI-1*、*PML*、*p53-regulator MDM2* 都含有 ZF-RING 结构域^[2-3], 这些基因主要与疾病控制和参与信号传导有关。最近研究指出, 一些含 ZF-RING 结构域的蛋白具有 E3 泛

素连接酶活性, 参与泛素代谢途径, 并在泛素代谢途径中起着关键性作用。泛素代谢途径是细胞内消耗能量的蛋白降解途径, 从蛋白水平上调控基因的表达。泛素连接酶 E3 与目标靶蛋白结合, 引导目标靶蛋白进入泛素蛋白酶复合体, 引导靶蛋白的降解^[4-5]。拟南芥中含 ZF-RING 结构域的基因 *HOS1* 参与寒胁迫诱导反应^[6]。Yang 等^[7]证实来自沙蒿 (*Artemisia desertorum* spreng) 的 ZF-RING 基因能显著提高转基因烟草的抗旱性。植物中对含 ZF-RING 结构域基因的研究处于初步阶段, 对杨树 ESTs 数据库注释进行查询表明, 存在与模式植物拟南芥含 ZF-RING 结构域基因同源的 cDNA 片断, 无具体功能

收稿日期: 2008-06-13 修回日期: 2010-06-19

基金项目: 引进国际先进林业科学技术“948”创新重大项目“杨树 RNAi 载体构建、材性相关基因克隆及功能研究”(2006-4-C01) 和国

家 973 项目“速生优质杨树的聚合育种与分子改良”(No. 2009CB119107)

作者简介: 王大海(1978—), 男, 浙江上虞市人, 博士, 研究方向: 林木基因工程。

* 通讯作者

方面的注释。到目前为止,未见有对杨树含 ZF-RING 结构域的蛋白基因进行克隆、基因结构和表达分析等方面的研究报道。

从芯片研究结果得出,含 C3HC4-Zinc finger RING 结构域的基因(*pdZFR*)表达与美洲黑杨(*Populus deltoides* Bartr.)主干不同高度处木材密度的分布相关,其相关系数为 0.89,通过测序获得了该基因的部分 cDNA 序列(结果未列出);本研究在此基础上,采用 RACE 方法(Rapid Amplification of cDNA End, cDNA 末端快速克隆)克隆其全长 cDNA 序列,采用 VGL-Walking(Visual Genome Library Walking,拟基因组染色体步移)方法从基因组上克隆该基因全长序列。对该基因序列进行比较分析,研究其基因结构、转录后的剪切方式以及在基因组上的分布;利用 RT-PCR 分析该基因在不同组织或器官中的表达。

1 试验材料与方法

1.1 试验材料

取样地点设在中国林业科学研究院内,7月中旬,此时木材未成熟木质部活动旺盛。选择树干生长饱满通直的美洲黑杨成年母树,用铲子去除母树胸径处树皮,暴露形成层和未成熟木质部组织,用裁纸刀刮取形成层以内的未成熟木质部组织,于液氮中速冻后 $-70\text{ }^{\circ}\text{C}$ 条件下保存,此材料用于该基因 RACE 全长克隆;同时摘取美洲黑杨成年母树的叶柄, $-20\text{ }^{\circ}\text{C}$ 条件下冻存,用于 VGL-Walking 试验。采集叶片、叶芽和花芽,液氮速冻后 $-70\text{ }^{\circ}\text{C}$ 条件下保存,用于 RT-PCR 半定量表达分析。

1.2 试验方法

1.2.1 未成熟木质部、花芽、叶芽和叶片总 RNA 提取 总 RNA 提取采用北京奥莱博生物技术有限责任公司生产的植物总 RNA 提取试剂盒(Cat#: ALBRE-004-100),具体操作过程严格参照其说明书。总 RNA 提取后,用 RNase-free DNase (Promega, 上海)处理 15 min,氯仿/酚抽提去除 DNase,乙醇沉淀回收 RNA。用 TAE 琼脂糖凝胶电泳检测 RNA 质量。

1.2.2 RACE 克隆 *PdZFR* 全长 cDNA 序列 RACE 具体步骤参考 Clontech 公司生产的 SMART RACE cDNA Amplification Kit 说明书,稍作改动。取 $2\text{ }\mu\text{g}$ 未成熟木质部组织 RNA 用于第一链反转录。3' RACE 模板反转录用 Oligo dT1 (5'-T25VN-3') 引物;5' RACE 模板反转录用 TSP(5'-AAGCAGTGGTCAACGCAGAGTACGCGGG-3' 和 Oligo dT2 (5'-

AAGCAGTGGTATCAACGCAGAGTAC (T) 30VN-3') 引物,引物均由上海生工合成,SuperScriptII 反转录酶购于 Invitrogen 公司。3' RACE 和 5' RACE 扩增特异引物分别为 3'-Race-ZF-RNG (5'-TTACGGCACGGAACAAGAGGTTCAAC-3'),5'-Race-ZF-RNG (5'-AGGTCCTCCAATCTGCAAATCCCAC-3');通用混合引物 UPM 为(5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3', $0.4\text{ }\mu\text{mol}\cdot\text{L}^{-1}$; 5'-CTAATACGACTCACTATAGGGC-3', $2\text{ }\mu\text{mol}\cdot\text{L}^{-1}$)均由上海生工合成。PCR 扩增体系和条件参照说明书。

PCR 反应完成后,取 $3\text{ }\mu\text{L}$ 用于电泳检测, $47\text{ }\mu\text{L}$ 用于切胶回收目的条带,切胶回收试剂盒采用北京奥莱博生物有限责任公司生产的琼脂糖凝胶及 PCR 产物纯化试剂盒(Cat#: ALBGR-003-100),回收后的特异条带连接到 T 载体 (Promega, 上海),并转化大肠杆菌 DH5 α 菌株,用 M13F 和 M13R 测序引物 PCR 筛选重组克隆,送公司进行测序。

1.2.3 *PdZFR* 基因组全长克隆

1.2.3.1 基因组 DNA 提取 取美洲黑杨叶柄 5 g ,液氮中研磨成粉末状,分 2 等份加入到含 25 mL DNA 提取缓冲液($100\text{ mmol}\cdot\text{L}^{-1}$ Tris-HCl pH 8.0, $1.5\text{ mol}\cdot\text{L}^{-1}$ NaCl, 2% CTAB, $25\text{ mmol}\cdot\text{L}^{-1}$ EDTANA2, $1\text{ mol}\cdot\text{L}^{-1}$ 尿素, $20\text{ mmol}\cdot\text{L}^{-1}$ DTT)的 50 mL 离心管内,迅速剧烈震荡混匀, $65\text{ }^{\circ}\text{C}$ 温浴 30 min ,分别用 20 mL 氯仿/酚混合物、氯仿各抽提 1 次;基因组 DNA 用 50% 乙醇沉淀,用钩针捞出后在 70% 乙醇中漂洗 3 次,真空干燥,溶于 2 mL TE。取 0.5 mL 基因组 DNA,加 $5\text{ }\mu\text{L}$ $10\text{ mg}\cdot\text{mL}^{-1}$ DNase-free RNase,混匀 $37\text{ }^{\circ}\text{C}$ 温浴 1 h ;加 $10\text{ }\mu\text{L}$ $10\text{ mg}\cdot\text{mL}^{-1}$ 蛋白酶 K,混匀, $45\text{ }^{\circ}\text{C}$ 温浴 1 h ;分别用 0.5 mL 氯仿/酚混合物、氯仿各抽提 1 次;基因组 DNA 用 50% 乙醇沉淀,用钩针捞出后在 70% 乙醇中漂洗 3 次,真空干燥,溶于 $200\text{ }\mu\text{L}$ TE,进行定量和检测完整性。

1.2.3.2 Visual Genomic Library 构建 各取 $5\text{ }\mu\text{g}$ (约 $5\text{ }\mu\text{L}$) 基因组 DNA 分别用 *EcoR* V、*Dra* I、*Pvu* II 和 *Ssp* I 酶切过夜;酶切产物分别过 Spin-Collumn 400 (Clontech, 美国)分离柱。用无水乙醇沉淀回收酶切后的基因组 DNA,经真空干燥,溶于 $10\text{ }\mu\text{L}$ 灭菌水。取 4 个 $200\text{ }\mu\text{L}$ LPCR 反应管,分别标记 *EcoR* V、*Dra* I、*Pvu* II 和 *Ssp* I,向管中加入对应基因组酶切产物 $1\text{ }\mu\text{L}$;分别加 $1\text{ }\mu\text{L}$ $10\times$ Ligase Buffer, $1\text{ }\mu\text{L}$ ATP, $2\text{ }\mu\text{L}$ Adaptor (5'-GTAATACGACTCACTATAGGGCAGCGTG-

GTCGACGGCCCGGGCTGCT-3'; 3'-H2N-CCCGACCA-PO4-5', 上海生工)和 250 U T4 DNA Ligase (TAKARA, 大连), 用 ddH₂O 补足体积至 10 μ L, 混匀, 16 $^{\circ}$ C 连接过夜; 65 $^{\circ}$ C 温浴 10 min 灭活 T4 DNA Ligase, 加 90 μ L TE 混匀, 于 -20 $^{\circ}$ C 保存。

1.2.3.3 *PdZFR* Genome Walking 取 8 个 200 μ L PCR 反应管, 分别标记 3' *EcoR* V, 3' *Dra* I, 3' *Pvu* II, 3' *Ssp* I; 5' *EcoR* V, 5' *Dra* I, 5' *Pvu* II 和 5' *Ssp* I; 从上述 *EcoR* V, *Ssp* I, *Pvu* II 和 *Dra* I 拟文库中分别取 1 μ L 加入到上述对应 200 μ L PCR 反应管中; 加 49 μ L 扩增混合液 (1 μ L 3' 或 5' -Race-ZF-RNG 10 μ M, 1 μ L AP 5' -GTA-ATACGACTCACTATAGGGC-3' 10 μ M, 1 μ L dNTP 10 mmol \cdot L⁻¹, 5 μ L 10 \times ExTaq Buffer, 1.5 U ExTaq), 混匀后立即进行如下 PCR 扩增: 96 $^{\circ}$ C 预变性 1 min; 95 $^{\circ}$ C 10 s, 68 $^{\circ}$ C 6 min (每个循环加 5 s), 32 个循环; 72 $^{\circ}$ C 10 min。扩增完毕, 取 3 μ L PCR 产物进行电泳检测。选择基因组 5' 端和 3' 端扩增片段最长的产物, 切胶回收目的条带。进行 T 载体克隆, 用 M13F 和 M13R 测序引物采用 PCR 方法筛选重组克隆, 并对阳性克隆进行测序。

1.2.4 RT-PCR 目标候选基因表达分析 各取叶片、未成熟木质部、花芽和叶芽总 RNA 2 μ g, 分别用 Oligo d (T)25 和 *MM-LV* 反转录酶进行反转录, 反转录条件和体系参照 RACE 第一链 cDNA 合成。第一链合成完毕加 90 μ L TE (pH 8.0), 混匀; 各取 5 μ L 反转录产物, 以 *UBQ* 基因 (5' -GTTGATTTTTGCTGGGAAGC -3' / 5' -GATCTTGGCCTTACGTTGT-3') 作对照, 以水作阴性对照, 以含 *pdZFR* 3' RACE 产物的质粒作阳性对照, 用 5' -Race-ZF-RNG 和 3' -Race-ZF-RNG 引物进行扩增, 反应体系 50 μ L:

不同组织 cDNA 第一链模板: 5 μ L;

5' -Race-ZF-RNG 引物 (10 μ mol \cdot L⁻¹): 1 μ L;

3' -Race-ZF-RNG 引物 (10 μ mol \cdot L⁻¹): 1 μ L;

dNTP (10 mmol \cdot L⁻¹): 1 μ L;

10 \times PCR Buffer: 5 μ L;

rTaq (TaKaKa, 大连): 0.25 μ L;

水: 至 50 μ L

扩增程序: 96 $^{\circ}$ C 预变性 1 min; 95 $^{\circ}$ C 10 s, 64 $^{\circ}$ C 10 s, 72 $^{\circ}$ C 30 s, 32 个循环; 72 $^{\circ}$ C 10 min。扩增完毕取 8 μ L 进行电泳分析。

1.2.5 数据分析 Blast 分析工具和核酸数据库 (nr/nt) 由 NCBI 网站的 FTP 下载。比对程序利用 Blast 软件包中的 Blastall 程序, 算法采用 Blastn 和 Blastx。核酸数据库 nt 和非冗余数据库 nr 经解压到硬盘指定位置, 制作成本地数据库。杨树全基因组数据库从毛果杨 (*P. trichocarpa* Torr. & Gray) 全基因组数据库下载到本地硬盘, 经解压后制作成本地杨树基因组数据库, 用于基因在基因组上定位分析。

2 结果与分析

2.1 *PdZFR* 基因全长 cDNA 克隆

从图 1 可以看出: 总 RNA 电泳图谱中 28S/18S 条带亮度比值大于 2, 表明提取的总 RNA 经 RNase-free DNase 处理后完整性较好, 无降解和杂质污染现象。图 2、3 为 *PdZFR* 基因的 5' RACE 和 3' RACE 扩增图谱。从图 2 看出: *PdZFR* 5' RACE 克隆特异性较好, 扩增条带单一, 大小为 900 bp 左右。从图 3 看出: 基因 3' RACE 在 0.50 kb 左右扩增出特异条带, 在 0.75 kb 与 1.00 kb 处扩增出非特异性条带, 这可能由于引物非特异性扩增引起。通过切胶回收 5' RACE 产物和 3' RACE 产物 500 bp 左右处特异条带, 克隆转化大肠杆菌感受态细胞, 经测序分析, 去掉重叠序列, 获得 *PdZFR* 基因 1 014 bp 的全长序列 (图 4)。利用 DNAMAN 软件对该基因进行分析。该基因全长 mRNA 序列包含 75 bp 的 5' UTR, 618 bp 完整编码序列和 293 bp 的 3' UTR; 其中, 618 bp 编码序列编码 1 条由 205 个氨基酸组成的多肽链, 其等电点为 6.94, 预测的分子量为 23.48 kD。图 5 为 *PdZFR* 基因与拟南芥 *C3HC4-type RING finger* 基因氨基酸比对情况, 二者表现出极高的同源性, 达 86%。



图 1 RACE 实验的总 RNA 电泳图谱

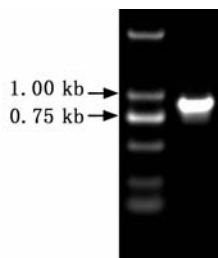


图 2 *PdZFR* 5' RACE 电泳图谱

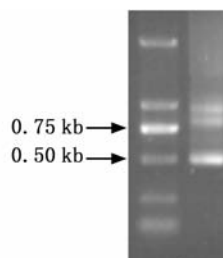


图 3 *PdZFR* 3' RACE 电泳图谱

```

1   GAAGTATTTGATTGCCGTCATTGTCATAACGAGGCGAAGAACAATATCAATGTTGATCAG
61  AAGCATAGACATGACATGCCACGCCATGAAGTCAAACAGGTGATATGCTCGCTTTGTGGC
121 ACTGAACAAGAGGTTCAACAAGTTTGTATCAACTGTGGTGTGTGCATGGGAAAGTACTTC
181 TGTGAGACTTGCAAGCTCTTTGATGATGATACATCTAAGAAACAGTATCATTGTGATGGC
241 TGTGGGATTTGCAGAATTGGAGGACCTGAGAATTTCTTCCATTGTTACAAATGTGGCTGC
301 TGCTACTCGAATCTTCTGAAGAATAGCCACCCCTGTGTAGAGGGCGCGATGCATCATGAC
361 TGCCCTGTCTGCTTTGAGTTTTTATTTGAGTCGAGATATGATGTGACTGTCTGCCATGT
421 GGACACACCATTCAACAAGAGCTGCTTAAAGGAAATGAGGGATCATTATCAATATGCTTGC
481 CCTCTTTGCTCGAAGTCAGTTTGTGATATGTCCAACGTATGGGACAAATTCGACATGGAA
541 ATTGCAGCTACACCAATGCCAGAACCTTACCTGAATAAAAATGGTTTGGATCCTTTGCAAT
601 GATTGTGGAAAGTCTCAGAAGTGCAATTCATGTAGTAGCTCAGAAATGCATGAACTGC
661 AAGTCCTATAATACTCGTCAAACAAGAAGCTGAGAAGTGGCTGCAAGGACAATGCAGCTA
721 GAAAAACAAACGGTGATGATGGTGTGATGGTGTGATGAAGACTTCTATGCTGAAAGCAATAAAA
781 TTGCATCTCGATCCATGTATATGTATATATATCCATGGATGTGGAAGGATCTGAGATCCT
841 GTGGAGGGCTTGTGGTTTCGCTTTTCTTTCCCGTTCGCTTTCTCTGATCATTTTGTCTTC
901 CTTTTATCCTCCCCCAATTGAATGACGTACTCAAGAAAACCACAGTCAAAGTGGTGGTCA
961 TGCGAATAAAGTGTCTCTCTCTTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```

图4 PdZFR 基因全长 cDNA 序列(有下划线部分为翻译编码框)

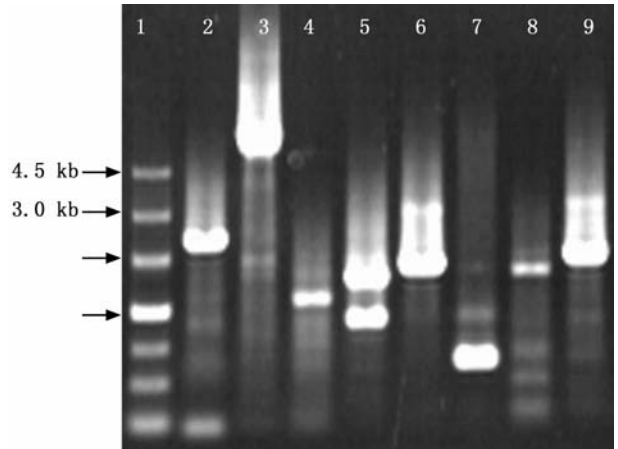
Arabidopsis thaliana C3HC4-type RING finger Pdc3HC4 ZF-RNG	YGC PHYRRRC CIRAPCCNEIFGCHHCHYEAKNNINUDQKQRHDIRHQEQUICLLCGTE -----MPRHEUKQUICSLCGTE :XXX: X:XXX XXXXX
Arabidopsis thaliana C3HC4-type RING finger Pdc3HC4 ZF-RNG	QEUQGQIC IHC GUCMGKYFCUKLYDDDTSKKQYHCDGCGICRIGGRENFFHCYKGCCTY QEUQUC INCGUCMGKYFCETCKLFDDDTSKKQYHCDGCGICRIGGRENFFHCYKGCCTY XXX X: X:XXXXXXXXXXXXX: .XXX:XXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXX
Arabidopsis thaliana C3HC4-type RING finger Pdc3HC4 ZF-RNG	SILLKNGHPCUEGAMHHDCPICFEFLFESRNDUTULPCGHTIHQKCLEEMRDHYQYACPL SNLLKNSHPCUEGAMHHDCPUFCFEFLFESRYDUTULPCGHTIHKSLKEMRDHYQYACPL X XXXX XXXXXXXXXXXXXXXX:XXXXXXXXXXXX .X:XXXXXXXXXXXX
Arabidopsis thaliana C3HC4-type RING finger Pdc3HC4 ZF-RNG	CSKSUCDMSKUWEKFDMEIAATPMPEPYQNRMUNSLQRLWEE SRGTUPCGGSEMPKLIQII CSKSUCDMSNUWDFDMEIAATPMPEPYLNKHMWILCN-----DCGKSSEUQFHUU XXXXXXXXXX: X:XXXXXXXXXXXXXXXXX X: X X . XXX X. : : : :
Arabidopsis thaliana C3HC4-type RING finger Pdc3HC4 ZF-RNG	QHSPNQRLLFLKLCYQMLRAFHLIQ AQKCMN-----CKSYNTRQTRS-- : . : . X . : X :

图5 PdZFR 基因与拟南芥 C3HC4-type RING finger 比对结果

2.2 PdZFR 基因组全长克隆

2.2.1 PdZFR 在杨树基因组中克隆 图6表明: 5'-Race-ZF-RNG primer 在 EcoR V 拟文库中克隆的片段最长, 达 5.0 kb 以上; 在 Dra I 拟文库中扩增产物长度为 2.5 kb; 在 Puv II 拟文库中克隆出 2 条特异片断, 分别为 1.5、1.0 kb 左右。3'-Race-ZF-RNG primer 在 Ssp I 拟文库中克隆的片段最大达 2.0 kb 左右; 在 Dra I 拟文库中克隆片断为 2.0 kb; 在 EcoR V 拟文库中扩增片断大小为 400 bp 左右, 而在 Puv II 拟文库中克隆片断为 1.9 kb。切胶回收 5'-Race- ZF-RNG primer 和 3'-Race- ZF-RNG primer 扩增产物中片断长度最长的条带, 连接 T 载体, 转化大肠杆菌感受态细胞, 挑取单克隆, 经测序分析, 去掉重叠序列, 最后获得长度为 5.43 kb 的片段数据。

2.2.2 PdZFR 基因转录后剪切方式 利用 Spidey 软件分析 PdZFR 基因的外显子和内含子分布。该



1. Marker; 2. 5' g-Dra I; 3. 3' g-Ecor V; 4. 5' g-Puv II; 5. 5' g-Ssp I; 6. 3' g-Dra I; 7. 3' g-Ecor V; 8. 3' g-Puv II; 9. 3' g-Ssp I

图6 PdZFR 基因在基因组拟文库中扩增的电泳图谱

基因由 10 个外显子和 9 个内含子组成(图7), 其中, 外显子 10 片断最大, 为 403 bp; 外显子 9 次之,

为 112 bp,其余片断长度都小于 100 bp,最小的仅 33 bp。9 个内含子中,有 8 个大于 100 bp,最长达 328 bp。从内含子和外显子大小分布可以看出:该基因剪切机制比较复杂。1kb 左右的基因包含 9 个内含子和 10 个外显子,而外显子的平均长度普遍较内含子小。该基因内含子 5' 端剪切位置碱基为 A(U)G(U) || GU,3' 端剪切位置碱基为 AG || A(G)U(A)。从剪切方式可以得出,内含子剪切识别位置两端趋于保守,为‘AG’和‘GU’,符合 GU-AG 规则,属于一般核 mRNA 类内含子剪切方式(图 8)。内含子与外显子片段大小分布见表 1。



图7 *PdZFR* 内含子和外显子分析结果
(黑色代表内含子,灰色代表外显子)

```

G A A G || G U G A - Intron - G C A G || A A C A
A C A G || G U A G - Intron - U C A G || G U G A
A G A G || G U G A - Intron - C U A G || G U U C
U G A U || G U A A - Intron - G C A G || A C A U
G C A G || G U A U - Intron - G C A G || A A U U
U G U G || G U A A - Intron - G A A G || G C U G
U G A G || G U U U - Intron - G C A G || U U U U
A U C A || G U G A - Intron - A C A G || A U A U
A A U G || G U U A - Intron - G C A G || G U U U
    
```

图8 *PdZFR* 基因剪切位置

2.2.3 *PdZFR* 基因在染色体上分布 采用 Blast 软件包中的 Blastall 程序,将染色体上克隆的全长序列对本地杨树基因组数据库进行比对分析,发现 *PdZFR* 基因位于杨树基因组 LG_VI 上,其中,47-4085 对应于 LG_VI 上物理位置的 16616623-16612586; 4103- 4743 对应于 LG_VI 上物理位置的 16612577-16611935;4760- 5345 对应于 LG_VI 上物理位置的 16611919- 16611332,中间的空缺序列可能与杨树基因组拼接过程中的误差有关。该基因在杨树基因组 LG_XVIII 数据簇上有 224 个碱基的同源序列。

表1 *PdZFR* 基因外显子和内含子分布情况

外显子 /个	基因组片段 对应位置	mRNA 对 应位置	外显子 长度/kb	一致性 /%	内含子 /个	内含子 长度/kb
1	1817 - 1855	1 - 39	39	100.0	1	121
2	1975 - 2034	40 - 99	60	100.0	2	112
3	2145 - 2177	100 - 132	33	100.0	3	328
4	2504 - 2581	133 - 210	78	100.0	4	197
5	2778 - 2821	211 - 254	44	100.0	5	325
6	3145 - 3185	255 - 295	41	100.0	6	123
7	3307 - 3389	296 - 378	83	100.0	7	124
8	3512 - 3603	379 - 470	92	100.0	8	98
9	3700 - 3811	471 - 582	112	98.2	9	125
10	3937 - 4339	583 - 985	403	99.5		

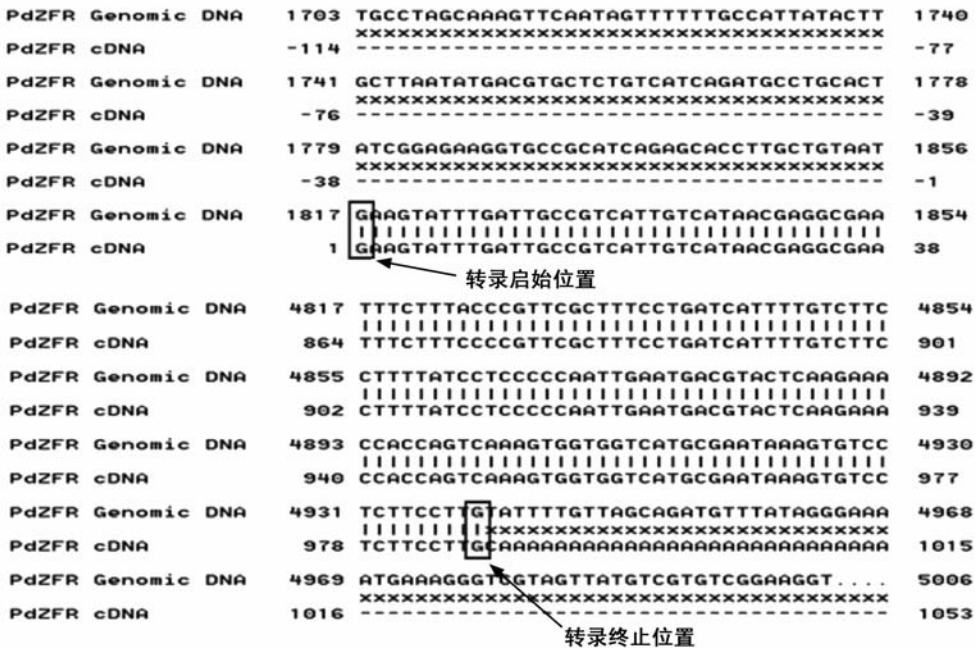


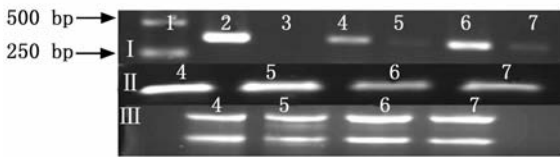
图9 *PdZFR* 基因转录起始和终止位置

2.2.4 *PdZFR* 转录起始和终止 图9为 *PdZFR* 全长 cDNA 序列与基因组上克隆的基因全长序列比对的部分结果。从图9可以看出:*PdZFR* 基因转录起

始碱基为 G(+1,1 817),从 +1 位置之前无明显的真核生物启动子核心序列 TATA 框,从 +1 (1 817) 到 +13(1 830)区域 A/T 含量明显较高,可能对基因

的表达起调节作用;而在转录终止的碱基与起始相同为 G,在转录终止位置处存在连续的 A 或 G 重复序列,有利于在转录过程中形成局部二级结构,使基因转录终止。利用 PLACE 软件对基因组上获得的 *PdZFR* 基因全长 5.43 kb 序列进行启动子分析。分析结果也表明: -100 bp 范围内无典型真核生物启动子元件 TATA 框和 CCAAT 框。

2.2.5 *PdZFR* 基因在各组织中的表达 图 10 中 I 为 *ZF-RNG* 基因在不同器官或组织中的表达情况; II 为内参 *UBQ* 基因的扩增条带电泳图谱; III 为各组织提取的总 RNA 电泳图谱。从图 10 可以看出: *PdZFR* 基因在花芽中表达量最高,未成熟木质部组织次之,叶片组织和叶芽中表达量相差无几。该基因的表达在杨树未成熟木质部和花芽中表达量很高,而在叶芽和花芽间几乎没有差异。由于实验取材为 7 月份,而此时也正是在木材形成和花芽组织发育的旺盛时期,推测该基因与发育相关。



1. Marker; 2. 阳性对照; 3. 阴性对照

4. 未成熟木质部; 5. 叶; 6. 花芽; 7. 叶芽

图 10 *PdZFR* 基因在不同组织中的表达变化

3 结论与讨论

本文利用 RACE 方法对美洲黑杨 *PdZFR* 基因进行了克隆,并且利用拟染色体步行从基因组上成功克隆了该基因包含启动子和转录终止信号在内的全长基因片断。RACE 克隆方法很多,本研究采用 MM-LV 反转录酶末端转移酶和模板转换功能在反转录过程中直接引入接头引物^[8-9]。该方法与其他方法相比省略了 mRNA 去磷酸基团、去帽子结构和寡聚核苷酸接头连接等繁琐操作步骤^[10-14],降低了操作引起的 RNA 降解问题,提高了克隆 cDNA 5' 端完整序列的概率。在已知部分 cDNA 序列基础之上,采用拟基因组文库染色体步行法从基因组上克隆了该基因的全长序列。基因组 DNA 经识别 4 个碱基的平端内切酶处理,然后两端连接人工合成的接头,制作成拟基因组文库。采用接头序列和特异引物序列扩增基因组上已知序列两翼未知区域。接头由 2 条长短不一的互补寡聚脱氧核苷酸经退火形成,其中短的寡聚脱氧核苷酸 5' 端进行磷酸化修

饰,有助与基因组 DNA 3' 端-OH 连接;3' 进行氨基化修饰,封闭在 PCR 扩增过程中 DNA 聚合酶的延伸作用。在 PCR 引物设计上,短接头序列 3' 端进行氨基化修饰,并且与长接头序列 3' 互补,保证基因特异引物延伸到接头末端后,接头引物的结合位置才形成,从而增加 PCR 扩增的特异性。该方法与文献^[14]介绍的染色体步行策略相比具有特异性高、操作简单、迅速、重复性好等特点。

在已知部分序列的基础上,对该基因在美洲黑杨基因组上的序列与杨树基因组序列进行比对分析,表明该基因位于毛果杨基因组数据簇 LG_VI 上 16616623-16612586(5192/5323,97% 同源);在 LG_XVIII 数据簇有部分同源序列(224/257,87% 同源),在其它染色体位置搜寻不到同源序列。这一结果表明,该基因表达由 LG_VI 转录而来。从比对结果来看,该基因在杨树基因组对应序列上有 3 个位置比对不理想,数据库中的数据主要表现为缺失现象。本文得到的该基因染色体上序列由美洲黑杨通过 Genome Walking 直接获得,数据比较可靠;而杨树全基因组数据最后通过大量测序电子拼接而成。由此看来,虽然杨树全基因组数据已经获得,但还需对其数据处理产生的误差进行修正,尤其是在连续几个相同碱基出现位置。

PdZFR 基因氨基酸序列与拟南芥 *C3HC4-type RING* 基因的氨基酸序列有 86% 的同源性。该基因具 40~60 个氨基酸组成的锌指结构域,包含共同序列 C-X2-C-X(9-39)-C-X(1-3)-H-X(2-3)-C-X2-C-X(4-48)-C-X2-C,该结构域可结合 2 个锌离子。RT-PCR 检测结果显示,该基因在未成熟木质部组织和花芽组织中表达量较高,尤其是在花芽中,而在成熟叶片和叶芽(枝条顶芽)组织中表达量低。由于杨树花序为柔荑花序,花着生于柔荑花序的总轴上,柔荑花序的总轴起着与树木主干相同的角色,因此在花芽发育旺盛阶段该基因表达较高。考虑到 *PdZFR* 为含锌指结构域的蛋白,而锌指蛋白在基因转录、蛋白间相互作用、外界信号传导和基因蛋白水平调控等方面起到重要作用^[15-17]。因此推测,该基因与杨树次生长相关;但鉴于目前包含 *C3HC4-type RING* 结构域的基因功能研究案例比较少,也无具体功能方面的文献报道,因此该基因的具体功能和与其相互作用的靶蛋白还有待于通过其他实验手段进一步证实。

参考文献:

- [1] Peter K, Klaus F X M, Christian S H. Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome[J]. *Genome Biology*, 2002, 3(4):12-21
- [2] Freemont P S. The RING finger: a novel protein sequence motif related to the zinc finger [J]. *Ann N Y Acad Sci*, 1993, 68: 4174-4192
- [3] Jackson P K, Eldridge A G, Freed E, *et al.* The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases [J]. *Trends Cell Biol*, 2000, 10: 429-439
- [4] Freemont P S. Ubiquitination: RING for destruction [J]. *Curr Biol*, 2000, 10: R84-R87
- [5] Lorick K L, Jensen J P, Fang S. RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination [J]. *Proc Natl Acad Sci USA*, 1999, 96: 11364-11369
- [6] Dong C H, Agarwal M, Zhang Y Y, *et al.* The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1 [J]. *Plant*, 2006, 103 (21): 8281-8286
- [7] Yang X, Sun C, Hu Y, *et al.* Molecular cloning and characterization of a gene encoding RING zinc finger ankyrin protein from drought-tolerant *Artemisia desertorum* [J]. *J Biosci*, 2008, 33: 103-112
- [8] Schramm G, Bruchhaus I, Roeder T. A simple and reliable 5'-RACE approach [J]. *Nucleic Acids Res*, 2000, 28: e96
- [9] Carl W D, Gabriela S. Dveksler PCR Primer: A Laboratory Manual [M]. New York: Cold Spring Harbor Laboratory Press, 2003, 359-389
- [10] Bertling W M, Beier F, Reichenberger E. Determination of 5' ends of specific mRNAs by DNA ligase-dependent amplification [J]. *PCR Methods Appl*, 1993, 3: 95-99
- [11] Fromont R M, Bertrand E, Pictet R. A highly sensitive method for mapping the 5' termini of mRNAs, [J]. *Nucleic Acids Res*, 1993, 21: 1683-1684
- [12] Liu X, Gorovsky M A. Mapping the 5' and 3' ends of tetrahymena-thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE) [J]. *Nucleic Acids Res*, 1993, 21: 4954-4960
- [13] Sallie R. Characterization of the extreme 5' ends of RNA molecules by RNA ligation-PCR [J]. *PCR methods Appl*, 1993, 3: 54-56
- [14] 刘博, 苏乔, 汤敏谦, 等. 应用于染色体步移的 PCR 扩增技术的研究进展 [J]. *遗传*, 2006, 28 (5): 587-595
- [15] Peng H, Begg G E, Schultz D C. Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions [J]. *J Mol Biol*, 2000, 295: 1139-1162
- [16] Kentsis A, Dwyer E C, Perez J M, *et al.* The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E [J]. *J Mol Biol*, 2001, 312: 609-623
- [17] Borden K L. RING domains: master builders of molecular scaffolds [J]. *J Mol Biol*, 2000, 295: 1103-1112