

# 杨树肉桂醇脱氢酶基因序列和表达模式分析及 CAD4 酶活性检测

任珊珊<sup>1</sup> 赵艳玲<sup>2</sup> 白华<sup>1</sup> 蒋湘宁<sup>1</sup> 盖颖<sup>1</sup>

1. 北京林业大学生物科学与技术学院, 北京 100083; 2. 华侨大学化工学院, 厦门 361021

**摘要** 应用生物信息学方法从杨树基因组数据库中筛选出 19 个杨树肉桂醇脱氢酶基因, 它们聚为 3 类。序列分析表明, 杨树 CAD 基因家族主要通过全基因组加倍和串联复制的方式进行扩张。杨树和拟南芥 CAD 家族共 6 对旁系同源基因, 在较强的净化选择压力下进化。19 个基因在杨树不同组织中均有表达, 表达量较高的是 *PoptrCAD17*、*PoptrCAD4*、*PoptrCAD10* 和 *PoptrCAD7*, 其他基因表达相对较少, 相同亚类的 CAD 基因表达变化趋势接近。从毛白杨中克隆得到 *PoptrCAD4*, 体外表达的 CAD4 酶催化活性很强, 具有明显的松柏醇底物偏好。

**关键词** 杨树; 肉桂醇脱氢酶(CAD); CAD 基因家族; 基因重复; 组织表达; 酶活检测

**中图分类号** Q 785 **文献标识码** A **文章编号** 1000-2421(2011)05-0578-07

杨树肉桂醇脱氢酶(cinnamyl alcohol dehydrogenase, CAD)是参与木质素单体合成被研究得较早的酶之一<sup>[1]</sup>, 它催化 3 种肉桂醛(芥子醛、香豆醛和松柏醛)还原生成 3 种木质素单体的前体物质芥子醇、香豆醇和松柏醇<sup>[2]</sup>。全长 CAD 基因 cDNA 于 1995 年首次从苜蓿中被克隆出来, 到目前为止, 已有 187 条完整的 mRNA 序列在 NCBI 中注册。对 CAD cDNAs 序列分析的结果表明, 它们具有高度的同源性(其中被子植物的同源性约为 80%, 裸子植物的同源性约为 70%), 这些 CAD 基因在植物基因组中以基因家族形式存在, 其中拟南芥中 9 个<sup>[3]</sup>, 水稻中 12 个, 葡萄中 18 个, 苜蓿中 17 个<sup>[4]</sup>, 棉花中 8 个<sup>[5]</sup>。拟南芥 CAD 基因家族中有 6 个可以催化 5 种肉桂醛生成肉桂醇, 另外 3 个只有在底物浓度很高时才表现出活性, 催化能力很低。拟南芥 *AtCAD4* 和 *AtCAD5* 的同源性很高, 但 *AtCAD5* 能有效催化所有 5 种肉桂醛生成肉桂醇, 而 *AtCAD4* 却不能催化芥子醛<sup>[6]</sup>。

CAD 可分为 3 类, 第 3 类存在于裸子植物和被子植物中, 而第 1 类和第 2 类只存在于被子植物中<sup>[4]</sup>。拟南芥、水稻和高粱的 CAD 蛋白中都存在

保守的结构域 ADH\_N 和 ADH\_zinc\_N<sup>[7-9]</sup>, 这说明 CAD 属于乙醇脱氢酶超家族, 即具有 NAD 结合蛋白和 Zn 结合位点。乙醇脱氢酶蛋白往往以二聚体形式存在<sup>[10]</sup>。有研究从毛果杨×美洲黑杨分离出 CAD 基因的 cDNA, 进行正义和反义转化欧洲山杨×银白杨<sup>[11-12]</sup>, 结果发现 CAD 的活性被明显抑制, 但木质素总量并未下降, 而木质素单体含量发生改变, 对拟南芥的 *AtCAD4* 和 *AtCAD5* 进行双突变后, 木质素含量却大量减少<sup>[13]</sup>, 说明植物体内 CAD 家族是协同起作用的, 单一成员表达量的改变只能改变木质素各单体的含量。

笔者对杨树基因组中 CAD 基因及其家族成员、CAD 基因在杨树不同组织和不同季节的表达活性以及对 CAD 基因克隆与高效表达体外酶活性检测进行了研究, 并推测其对木质素单体合成的影响, 以进一步了解其成员之间的亲缘关系与表达活性状况, 为进一步研究其生物学功能奠定基础。

## 1 材料与方法

### 1.1 CAD 基因家族成员的筛选与鉴定

拟南芥家族成员信息来自 Sibout 等<sup>[3]</sup>的研究。

收稿日期: 2011-05-26

基金项目: 国家自然科学基金重点项目(NSF30630053)、福建省科技计划重点项目(2010N0018)、北京林业大学新进教师科研启动基金项目(2010BLX03)和中国博士后科学基金项目(20090450015)

任珊珊, 硕士研究生, 研究方向: 生物化学与分子生物学. E-mail: rsshshley@163.com

通讯作者: 盖颖, 博士, 讲师, 研究方向: 生物化学与分子生物学. E-mail: gaiying@bjfu.edu.cn

以拟南芥 CAD 家族的核酸序列和氨基酸序列为起始序列, 对比杨树基因组数据库, 设定 E-value 值为  $10E-10$ ; 得到的序列在 GenBank 数据库<sup>[14]</sup> 中对比, 并在 Pfam 数据库<sup>[15]</sup> 中预测结构域。当该基因与其他物种的 CAD 基因同源性大于 50% 且存在 CAD 蛋白中的特定蛋白结构域 ADH\_N 或者 ADH\_zinc\_N, 则认为其对应的基因为 CAD 基因。

## 1.2 序列分析

得到的植物 CAD 基因家族成员包括基因组序列、cDNA 序列及编码序列(CDS)信息, 在 GSDS 基因结构显示系统<sup>[16]</sup> 中将基因组序列及 CDS 进行对比。应用 Clustal X 和 Bioedit 软件对植物 CAD 基因家族编码的氨基酸序列进行多序列联配。使用 MEGA 软件中的邻接法(NJ)构建 CAD 基因家族的系统进化树。应用 K-estimator 6.1 软件计算 4 对重复基因的旁系同源基因的 Ka(非同义突变率)、Ks(同义突变率)值, 确定其进化过程中所受何种选择。

## 1.3 基因表达芯片分析

以杨树 CAD 基因家族的编码序列(CDS)为搜索对象, 在 Affymetrix 公司提供的杨树基因组数据库中进行 Blastn 搜索, 设定 E-value 值为  $10E-10$ , 确定与基因家族匹配度高的序列。制作基因表达芯片的试验样品为树龄 5 a 的毛白杨 741, 分别在春、夏、秋、冬 4 个时间点采样, 每个时间点取 4 个样品, 即根、芽、主干接木质部形成层以及主干接韧皮部形成层, 每个样品取 3 个生物学重复, 芯片由博奥生物有限公司制作完成, 使用 GCOS 分析软件将图像信号转化成数字信号, 再通过 dChip 软件处理得到表达谱数据文件<sup>[17]</sup>。

## 1.4 毛白杨肉桂醇脱氢酶的高效表达、纯化和酶学特性研究

以基因工程菌株 QMCAD(由笔者所在实验室制备, 重组 CAD 来自毛白杨)为试验材料, 通过 PAGE 电泳确定 CAD 的最适 IPTG 诱导浓度及诱导时间。因为融合的 CAD 蛋白在大肠杆菌中表达时带有 6 个组氨酸标签, 可与  $Ni^{2+}$  柱结合, 所以采用亲和层析法在  $Ni^{2+}$  柱中纯化蛋白。应用分光光度法测定 400 nm 波长下, 以  $NADP^+$  为辅酶, 松柏醇和芥子醇为底物的氧化反应中 CAD 蛋白活性随时间的变化、最适反应温度、最适反应 pH 值以及金属离子对酶活性的影响。

# 2 结果与分析

## 2.1 植物 CAD 基因家族成员的鉴定

在杨树基因组数据库中共筛选出 21 个 CAD 候选基因, 其中 2 个经预测不包括 CAD 蛋白的特征序列, 所以认为杨树 CAD 基因家族共有 19 个成员(表 1), 其中 *PoptrCAD12* 和 *PoptrCAD19* 中只存在 ADH\_N 结构域, *PoptrCAD17* 和 *PoptrCAD18* 中只存在 ADH\_zinc\_N 结构域。杨树 CAD 基因家族定位于 8 条染色体上(图 1), 另外还有 2 个尚未定位在特定的染色体片段上。

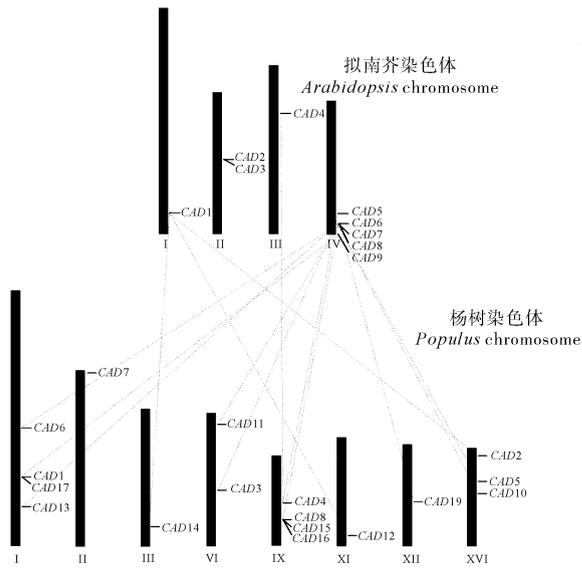
表 1 植物 CAD 基因家族信息

Table 1 The information of CAD gene family in *Populus* and *Arabidopsis*

基因 Gene	染色体位置 Loci in chromosomes	氨基酸长度/aa	分子质量/ku
		Amino acids	Molecular mass
<i>PoptrCAD1</i>	LG_I;26198485-26202457	362	39.3
<i>PoptrCAD2</i>	LG_XVI;1193899-1195990	354	38.3
<i>PoptrCAD3</i>	LG_VI;10406101-10408857	360	38.8
<i>PoptrCAD4</i>	LG_IX;4268850-4271140	357	39.0
<i>PoptrCAD5</i>	LG_XVI;4474459-4476617	362	39.1
<i>PoptrCAD6</i>	LG_I;19511432-19513448	360	38.4
<i>PoptrCAD7</i>	LG_II;1047765-1050403	360	39.4
<i>PoptrCAD8</i>	LG_IX;6478021-6480309	360	38.8
<i>PoptrCAD9</i>	scaffold_2069;1574-3584	360	39.1
<i>PoptrCAD10</i>	LG_XVI;5830946-5834884	362	38.9
<i>PoptrCAD11</i>	LG_VI;1629482-1634627	354	38.8
<i>PoptrCAD12</i>	LG_XI;14002689-14005478	308	34.4
<i>PoptrCAD13</i>	LG_I;30420496-30422191	358	38.8
<i>PoptrCAD14</i>	LG_III;17425443-17427044	355	39.1
<i>PoptrCAD15</i>	LG_IX;6486821-6488595	360	39.2
<i>PoptrCAD16</i>	LG_IX;6507390-6509409	372	40.4
<i>PoptrCAD17</i>	LG_I;26702441-26703090	139	15.7
<i>PoptrCAD18</i>	scaffold_11132;523-1428	160	17.5
<i>PoptrCAD19</i>	LG_III;8301591-8301915	71	8.1
<i>AtCAD1</i>	AG_I;27359232-27361016	355	38.7
<i>AtCAD2</i>	AG_II;9280044-9281623	376	40.9
<i>AtCAD3</i>	AG_II;9331089-9332646	376	40.9
<i>AtCAD4</i>	AG_III;6744598-6747214	365	39.1
<i>AtCAD5</i>	AG_IV;16386727-16388722	357	38.7
<i>AtCAD6</i>	AG_IV;17849648-17852246	363	39.0
<i>AtCAD7</i>	AG_IV;17852441-17854500	353	38.2
<i>AtCAD8</i>	AG_IV;17855876-17857639	359	38.9
<i>AtCAD9</i>	AG_IV;18291218-18293383	360	38.9

## 2.2 杨树 CAD 基因家族基因结构分析

CAD 基因家族中 *PoptrCAD17* ~ *PoptrCAD19*



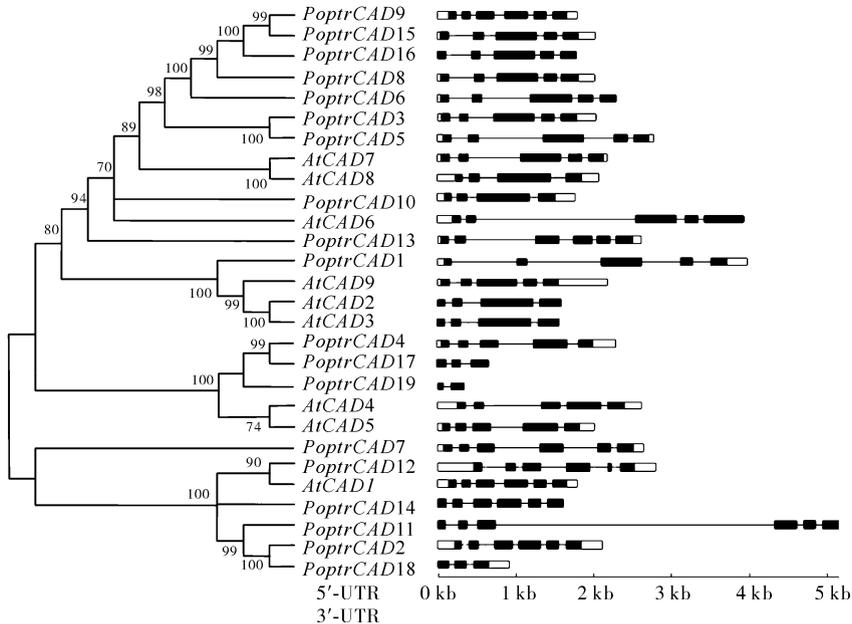
虚线代表拟南芥和杨树 CAD 基因家族的同源关系。Dash line represents homologous relationship of CAD family from *Arabidopsis* and *Populus*.

图 1 杨树和拟南芥 CAD 基因家族的染色体定位及同源关系  
Fig. 1 Loci and homologous analysis of CADs from *Populus* and *Arabidopsis*

序列长度最短,它们的外显子也最少,只有 2~3 个,其余成员的基因结构共有 3 种模式(图 2)。第 1 类包括 5 个长度保守的外显子,另外 *AtCAD6* 的第 3、4 个外显子的长度之和与第 1 类基因第 3 个外显子相同,但因为在结构中多了 1 个内含子,其外显子有 6 个。*AtCAD2*、*AtCAD3*、*AtCAD7* 和 *AtCAD8*,只有 4 个外显子,第 3 个外显子长度为第 1 类成员第 3、4 外显子长度之和,将这 5 个基因也归入第 1 类。第 2 类包括有 5 个长度不同于第 1 类的外显子。第 3 类基因的编码区有 6 个外显子,5 个内含子。*PoptrCAD12* 在 5 个外显子部位基因发生突变,产生了 1 个 GT-AT 内含子剪切点导致其编码区缩短。

2.3 拟南芥和杨树 CAD 蛋白的序列分析和聚类分析

除 *PoptrCAD17*~*PoptrCAD19* 之外,其他杨树和拟南芥 CAD 蛋白序列的相似性都达到了 75.0%~99.7%,推测序列相似性最高的几对基因 *PoptrCAD9* 和 *PoptrCAD15*、*PoptrCAD3* 和 *PoptrCAD5*、*PoptrCAD2* 和 *PoptrCAD11*、*PoptrCAD6* 和



空心部分代表 5'-UTR 及 3'-UTR,实心部分代表编码区,直线代表内含子部分。Hollow parts represent the 5'-UTRs and 3'-UTRs, solid parts represent the coding sequences (CDS), lines represent introns.

图 2 杨树和拟南芥 CAD 家族的系统进化树和基因结构

Fig. 2 Phylogenetic tree and genes structure of CADs family from *Populus* and *Arabidopsis*

*PoptrCAD8* 为旁系同源基因。*PoptrCAD17*~*PoptrCAD19* 序列长度很短,它们与其他 CAD 家族成员核苷酸序列的相似性也很低,然而,将这 3 个

CAD 基因序列在基因组序列中进行 5' 和 3' 端的适度延伸,使其与 CAD 基因家族其他成员长度相近,发现它们与其他杨树 CAD 基因相似性大大提高,

说明这 3 个 CAD 基因可能经历了快速的进化,同时发生了基因突变,导致开放性阅读框变短,但其氨基酸序列保留了基本的结构域。

### 2.4 杨树 CAD 家族的分子进化

对拟南芥和杨树中的 CAD 蛋白进行系统进化分析(图 2),拟南芥和杨树中的 CAD 共同聚成 3 类,每类中都含有拟南芥和杨树 CAD 蛋白,与 CAD 基因结构推测相同,说明 CAD 家族的分类进化是在拟南芥和杨树分化之前完成的。除了 *PoptrCAD12* 和 *AtCAD1* 有可能是垂直同源序列外,在不同的亚类中,拟南芥和杨树 CAD 大部分各自聚在一起,说明 CAD 家族在进化过程中不同的亚类相对比较保守,但在不同的物种中发生了适应性的进化。杨树基因组在 6 000~6 500 万年前曾发生过全基因组加倍事件<sup>[18]</sup>。将杨树 CAD 基因家族进行染色体定位,可以发现 *PoptrCAD3* 和 *PoptrCAD5*、*PoptrCAD2* 和 *PoptrCAD11*、*PoptrCAD6* 和 *PoptrCAD8* 位于重复的染色体片段上,说明它们很可能是在全基因组加倍过程中产生的。拟南芥 CAD 家族中 2 对旁系同源基因<sup>[3]</sup>,它们都处于 TIGR 上提供的拟南芥已有的染色体片段区域复制

的区域内。*AtCAD2*、*AtCAD3* 以及 *AtCAD6~AtCAD8* 在染色体上的位置非常接近,且序列相似性高,应该由串联重复产生,所以杨树和拟南芥中 CAD 基因家族的扩张可能是通过大片段复制和串联重复完成的。这 6 对基因的  $K_a/K_s$  值在 0.141~0.292 之间(表 2),显著小于 1,表明它们在进化过程中受到净化选择<sup>[19]</sup>,即它们应该都是功能基因,且功能保守,所以进化速率较慢。

表 2 重复基因对的  $K_a/K_s$  值

Table 2  $K_a$  and  $K_s$  of duplicated gene pairs

重复基因对 Duplicated gene pairs	$K_a$	$K_s$	$K_a/K_s$
<i>PoptrCAD9</i> 和 <i>PoptrCAD15</i>	0.001 13	0.008 02	0.141
<i>PoptrCAD2</i> 和 <i>PoptrCAD11</i>	0.083 10	0.502 11	0.166
<i>PoptrCAD3</i> 和 <i>PoptrCAD5</i>	0.0353 80	0.207 80	0.170
<i>PoptrCAD6</i> 和 <i>PoptrCAD8</i>	0.0983 10	0.417 49	0.235
<i>AtCAD4</i> 和 <i>AtCAD5</i>	0.0172 80	0.059 18	0.292
<i>AtCAD7</i> 和 <i>AtCAD8</i>	0.0372 40	0.212 80	0.175

### 2.5 杨树 CAD 家族基因芯片表达谱分析

基因表达芯片分析结果(图 3)显示,19 个杨树 CAD 基因在这 4 个部位均有表达,表达量较高的是 *PoptrCAD17*、*PoptrCAD4*、*PoptrCAD10* 和 *PoptrCAD7*,其中前 3 个表达量为其他基因的 100

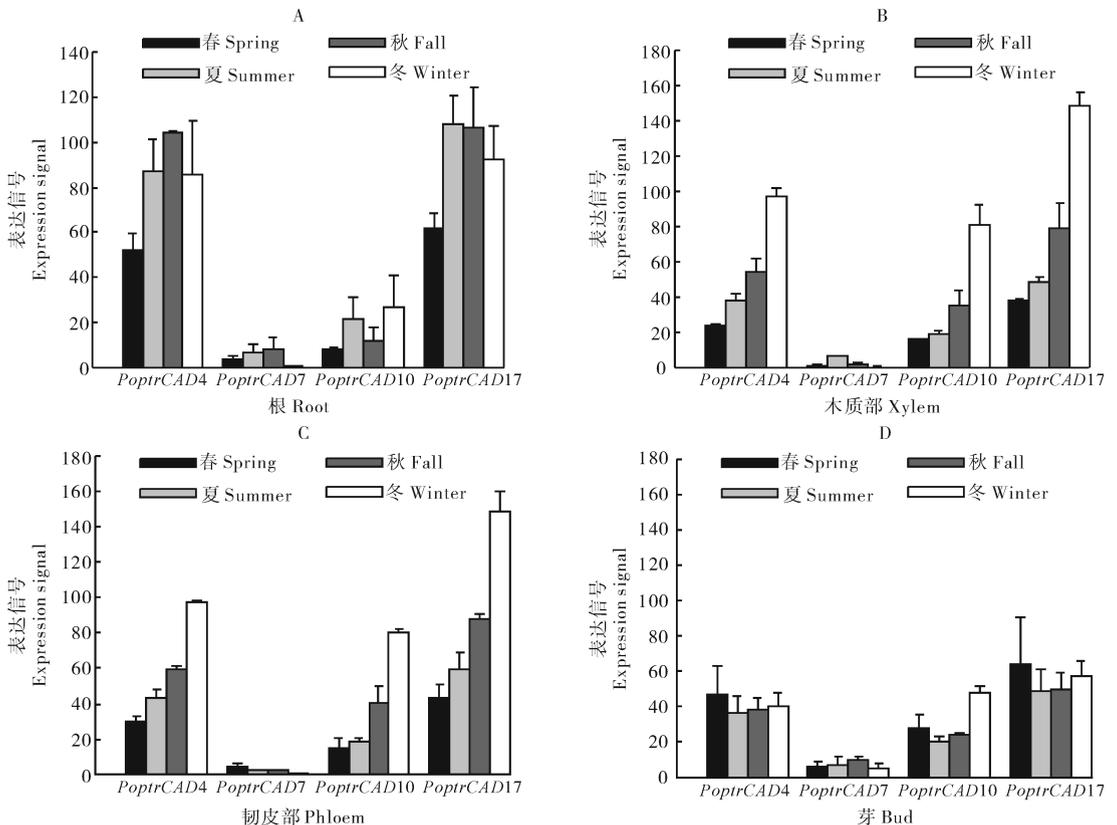


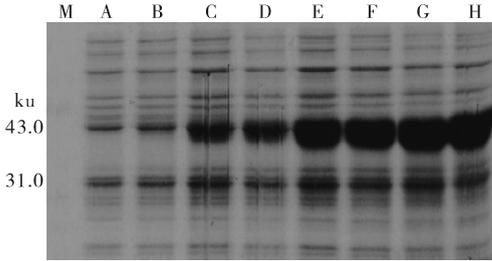
图 3 杨树 CAD 基因家族的表达信息

Fig. 3 Expression of CAD gene family of Populus

倍左右。这一结果与 Mansell 等<sup>[1]</sup>的相关研究结果一致。每个亚类 CAD 基因在不同季节、不同组织部位中表达变化情况相似,不同亚类表达情况有所不同。

### 2.6 毛白杨 CAD 的原核高效表达、纯化和酶学特性

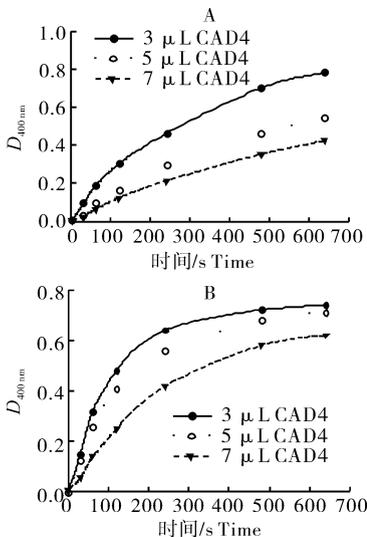
对实验室克隆出的毛白杨 CAD 基因序列进行染色体定位分析,发现其定位于杨树第 9 条染色体,



M: 蛋白 marker Protein marker; A: 未诱导 The suspension un-induced; B: 1 mol/L IPTG 诱导的菌液 The suspension induced by 1 mol/L IPTG; C: 1 mol/L IPTG 诱导超声上清 The supernatant after ultrasonicated and induced by 1 mol/L IPTG; D: 1 mol/L IPTG 诱导超声沉淀 The precipitate after ultrasonicated and induced by 1 mol/L IPTG; E: 0.4 mol/L IPTG 诱导菌液 The suspension induced by 0.4 mol/L IPTG; F: 0.4 mol/L IPTG 诱导超声上清 The supernatant after ultrasonicated and induced by 0.4 mol/L IPTG; G: 0.4 mol/L IPTG 诱导超声沉淀 The precipitate after ultrasonicated and induced by 0.4 mol/L IPTG.

图 4 不同 IPTG 浓度在 28 °C 时对蛋白可溶性的影响

Fig. 4 Expression levels of recombinant CAD on IPTG concentration

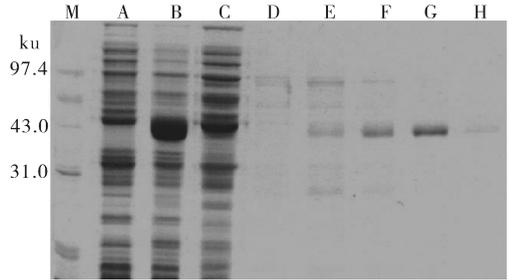


A: 反应底物为松柏醇 Coniferyl alcohol; B: 反应底物为芥子醇 Sinapinic alcohol.

图 6 不同浓度 CAD4 酶活性随时间的变化

Fig. 6 CAD4 activities on different time and different substrate concentration

序列一致性达到 98%, 即该基因为 *PoptrCAD4*。根据 PAGE 蛋白电泳图(图 4)确定 CAD4 在 28 °C, 4 mol/L IPTG 诱导 4 h 的情况下表达量最大。经纯化,洗脱液 4(图 5)中 CAD4 蛋白含量较大,且纯度很高,测其含量为 0.5 μg/μL。试验发现 CAD4 蛋白具有底物特异性,催化不同底物反应所需的时间、温度、pH 值等都不尽相同(图 6~8)。当底物为松柏醇时,CAD4 催化活性较高。CAD4 在 0~80 °C 时具有一定的活性,pH 值不在 7.0~12.0 范围内时活性丧失。另外,Zn<sup>2+</sup> 和 Cu<sup>2+</sup> 抑制 CAD4 的活性,Ca<sup>2+</sup> 和 Mg<sup>2+</sup> 对 CAD 的活性基本没有影响。



M: 蛋白 marker Protein marker; A: 未经 IPTG 诱导 Un-induced; B: 4 h 诱导菌液 The suspension induced 4 h; C: 4 h 诱导超声裂解上清液 The supernatant after ultrasonicated and induced 4 h; D: 洗脱液 1 Eluent 1; E: 洗脱液 2 Eluent 2; F: 洗脱液 3 Eluent 3; G: 洗脱液 4 Eluent 4; H: 洗脱液 5 Eluent 5.

图 5 亲和层析法一步纯化 CAD4 蛋白

Fig. 5 One-step purification of recombinant CAD4 by metal affinity chromatography

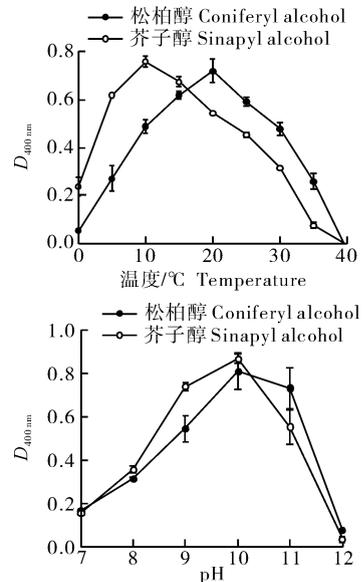


图 7 不同温度和 pH 值对 CAD4 酶活性影响

Fig. 7 CAD4 activities on different temperature and pH value

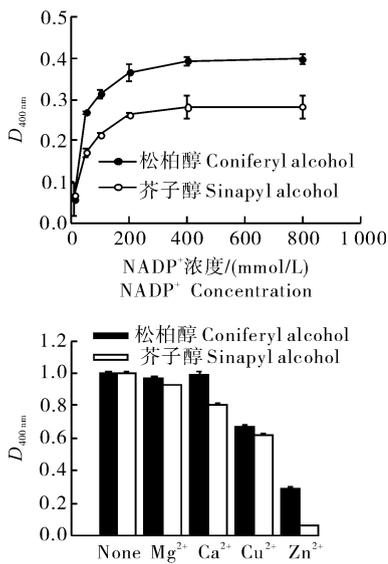


图8 NADP<sup>+</sup>浓度和金属离子对CAD4酶活性的影响  
Fig. 8 CAD activities on different concentration of NADP<sup>+</sup> and metal ions

### 3 讨论

杨树CAD基因家族共有19个成员,它们大部分还保留着CAD的特征结构域ADH\_N和ADH\_zinc\_N,也有几个基因发生了特定结构域的丢失,其中*PoptrCAD12*和*PoptrCAD19*中只存在ADH\_N结构域,*PoptrCAD17*和*PoptrCAD18*中只存在ADH\_zinc\_N结构域。杨树和拟南芥CAD基因家族各19和9个成员聚为3个亚类,每类外显子的数量和长度非常保守,这与研究结果相一致。杨树基因组在最近1次发生过全基因组加倍事件<sup>[19]</sup>,这一事件也是杨树CAD基因家族旁系同源基因*PoptrCAD3*和*PoptrCAD5*、*PoptrCAD2*和*PoptrCAD11*、*PoptrCAD6*和*PoptrCAD8*。另外,*PoptrCAD8*、*PoptrCAD15*和*PoptrCAD16*同时存在于染色体IX,形成基因簇,它们很可能是由基因串联重复形成的。拟南芥CAD家族中3对旁系同源基因,它们由拟南芥染色体片段区域复制产生,而*AtCAD6~AtCAD8*是由串联重复产生。

杨树CAD基因家族19个成员在杨树的各组织都有表达,但表达情况在不同的亚类有所不同。第1类基因在韧皮部和木质部中表达量最大,随着季节变化其表达量增加。第2类基因在根部表达量最大,随时间变化的趋势与第1类相同。第3类基因在根和顶芽中表达量最大,在夏、秋两季表达量高于另外2个季节。这说明杨树CAD基因家族3类

基因在功能上发生了显著变化,推测不同类别成员在不同部位的木质素合成中起作用,且由于CAD表达情况应与木质素单体合成量有一定的正相关性,由试验结果可推测木质素含量在不同季节不同部位也有不同。杨树CAD4在不同底物的最适反应温度也不同,以松柏醇为底物时,40℃时活性最高,以芥子醇为底物时,20℃下活性最高,其具有底物特异性,催化松柏醇生成松柏醛的反应更具活性,所以该酶主要参与木质素单体中的松柏醇单体的形成过程。

致谢 本研究在国家林业局树木花卉育种生物工程重点开放实验室、教育部树木花卉育种与基因工程实验室和林木育种国家工程实验室的帮助与支持下完成,在此向以上单位表示感谢。任珊珊、赵艳玲和白华在本研究中贡献相同。

### 参 考 文 献

- [1] MANSELL R L, GROSS G, STOCKIGT J. Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis[J]. *Plant Physiol*, 1974, 132:848-860.
- [2] LIN Z B, MA H, XU Y. Lignin biosynthesis and its molecular regulation[J]. *Progress in Natural Science*, 2003, 13(5): 321-328.
- [3] SIBOUT R, EUDES A, POLLET B, et al. Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in *Arabidopsis* isolation and characterization of the corresponding mutants [J]. *Plant Physiology*, 2003, 132(2): 848-860.
- [4] BAEKAT A, BAGNIEWSKA-ZADWORNIA A, CHOI A, et al. The cinnamyl alcohol dehydrogenase gene family in *Populus*: phylogeny, organization, and expression[J]. *BMC Plant Biology*, 2009, 26(9): 45-60.
- [5] FAN L, SHI W J, HU W R, et al. Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fibers[J]. *Journal of Integrative Plant Biology*, 2009, 51(7): 626-637.
- [6] KIM S J, KIM K W, CHO M H, et al. Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: lessons for database annotations [J]. *Phytochemistry*, 2007, 68(14): 1957-1974.
- [7] SABALLOS A, EJETA G, SANCHEZ E, et al. A genome wide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies SbCAD2 as the brown midrib6 gene[J]. *Genetics*, 2009, 181: 783-795.

- [8] GOFFNER D, VANDOORSELA E J, YAHHIAOUIN N, et al. Anovel aromatic alcohol dehydrogenase in higher plants: molecular cloning and expression[J]. *Plant Mol Biol*, 1998, 36: 755-765.
- [9] ZHANG L B, WANG G, CHANG J, et al. Effects of 1-MCP and ethylene on expression of three CAD genes and lignification in stems of harvested Tsai Tai (*Brassica chinensis*) [J]. *Food Chemistry*, 2010, 123: 32-40.
- [10] GOTTLIEB L D. Conservation and duplication of isozymes in plant [J]. *Science*, 1982, 216(4544): 373-380.
- [11] BAUCHER M, CHABBERT B, PILATE G, et al. Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar[J]. *Plant Physiology*, 1996, 112: 1479-1490.
- [12] BAUCHER M, VAN MONTAGU M, BOERJAN W, et al. Biosynthesis and genetic engineering of lignin [J]. *Plant Science*, 1998, 17(2): 125-197.
- [13] JOURDER M, CARDENAS CL, LASKAR D D, et al. Plant cell walls are enfeebled when attempting to preserve native lignin configuration with poly-p-hydroxycinnamaldehydes: evolutionary implications[J]. *Phytochemistry*, 2007, 68(14): 1932-1956.
- [14] BENSON D, KARSCH-MIZRACHI I, LIPMAN D, et al. GenBank [DB]. *Nucleic Acids Res*, 2010, 28: 15-18.
- [15] BABU M M, CAMBRIDGE M J, WESTHOF E, et al. Bateman Nucleic Acids Research [CP]. Database Issue, 2010, 38: 211-222
- [16] GUO A Y, ZHU Q H, CHEN X, et al. GSDS: a gene structure display server [CP]. *Yi Chuan*, 2007, 29(8): 1023-1026.
- [17] 赵继兴. 从杨树基因组表达谱芯片数据中挖掘树木生长发育规律的研究[J]. *安徽农学通报*, 2010, 26(6): 38-39.
- [18] TUSKAN G A, DIFAZIO S, JANSOON S, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray) [J]. *Science*, 2006, 313(5793): 1596-1604.
- [19] MORANT M, HEHN A, REICHAERT D W. Conservation and diversity of gene families explored using the CODEHOP strategy in higher plants[J]. *BMC Plant Biology*, 2002(2): 724-737.

## Sequences and expression pattern of cinnamyl alcohol dehydrogenase gene and over-expression and detection of CAD4 enzymatic activity of poplar *in vitro*

REN Shan-shan<sup>1</sup> ZHAO Yan-ling<sup>2</sup> BAI Hua<sup>1</sup> JIANG Xiang-ning<sup>1</sup> GAI Ying<sup>1</sup>

1. College of Biology and Biotechnology, Beijing Forestry University, Beijing 100083, China;

2. College of Chemical Engineering, Huaqiao University, Xiamen 361021, China

**Abstract** The 19 CAD genes of poplar screened by bioinformatics are clustered into 3 groups. Analysis of their sequences showed that the expansion of CAD gene family in *Populus* was resulted from a recent whole-genome doubling and tandem replication. There are 6 paralogs in CAD gene family of *Populus* and *Arabidopsis* evolved under the purifying selection. The 19 CAD genes are expressed in all tissues of *Populus*. *PoptCAD17*, *PoptCAD4*, *PoptCAD10* and *PoptCAD7* expressed highly while the other CAD genes expressed at a lower level. Expression of the CAD genes in the same group varied with the same trends. CAD gene cloned in this research was *PoptCAD4*. The catalytic activity of the CAD enzyme over-expressed *in vitro* were high, with the substrate preference of sinapyl alcohol.

**Key words** poplar; cinnamyl alcohol dehydrogenase (CAD); CAD gene family; gene duplication; tissue expression; enzyme activity detection

(责任编辑:陆文昌)