

植物4-香豆酸:辅酶A连接酶

赵淑娟 刘涤 胡之璧*

上海中医药大学中药研究所, 上海 201203

4-Coumarate:Coenzyme A Ligase in Plant

ZHAO Shu-Juan, LIU Di, HU Zhi-Bi*

Institute of Traditional Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

提要 就植物中4-香豆酸:辅酶A连接酶的基因表达调控和底物结合功能域及其与木质素的关系以及其基因家族的进化的研究进展做了介绍。

关键词 4-香豆酸:辅酶A连接酶, 苯丙烷类化合物生物代谢

苯丙烷类化合物生物代谢途径是与碳循环有关的最主要的次生代谢途径之一。植物细胞中20%的次生代谢产物经由这一途径产生。植物的许多天然产物如黄酮、类黄酮、木质素、色素、芪类及其它酚酸类化合物等,都是由来源于苯丙烷类化合物生物代谢途径的各种分支途径产生(Grand等1983; Wallis和Rhodes 1977)。这些物质在不同细胞、组织和器官中均起着重要作用,如木质素是木质部结构组成的主要成分,黄酮可作为叶表皮细胞防止紫外损伤的保护剂或者作为花粉发育和植物与微生物相互作用的信号分子,芪类可以保护植物细胞抗病虫侵袭等等(Douglas等1987; Dixon和Paiva 1995; Grand等1983; Hahlbrock和Scheel 1989; Lozoya等1988)。从经济学的角度来说,这一类次生代谢产物蕴藏着巨大的应用价值,如作为药物、食品添加剂、芳香剂、染料等等。正因为如此,人们对苯丙烷类化合物的生物代谢途径进行了一系列的研究,其内容涉及植物化学、酶学、遗传学和分子生物学等等。作为该途径中关键限速酶之一的4-香豆酸:辅酶A连接酶(4-coumarate:coenzyme A ligase, 4CL; EC 6.2.1.12)自然也引起了人们的关注。4CL处在总途径向分支途径的转折点,控制着苯丙烷类化合物生物代谢向不同方向进行,在植物与环境的相互作用中起必不可少的作用。

4-香豆酸:辅酶A连接酶,又称4-香豆酰:CoA合成酶、*p*-羟基肉桂酸:CoA连接酶,是苯丙烷类化合物生物代谢总途径中的最后一个酶,它通

过两步反应($4CL + \text{香豆酸} + \text{ATP} \rightarrow 4CL:\text{香豆酰} \sim pA + \text{PPi}$; $4CL:\text{香豆酰} \sim pA + \text{CoA} \rightarrow \text{香豆酰} - \text{CoA} + \text{AMP} + 4CL$)催化4-香豆酸及其羟基和甲羟基衍生物生成各自活性形式的硫酯酰CoA,而这些活性形式的硫酯酰CoA是后续各分支途径的直接前体分子(图1)(Allina等1998; An等1999; Brodelius和Xue 1997; Graham和Graham 1991)。

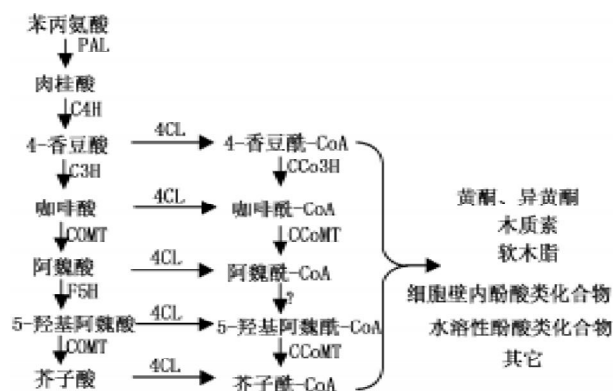


图1 苯丙烷类化合物生物代谢总途径(Allina等1998; An等1999; Brodelius和Xue 1997; Graham和Graham 1991)

PAL: 苯丙氨酸氨裂解酶; C4H: 肉桂酸4-羟基化酶; C3H: 肉桂酸3-羟基化酶; COMT: 咖啡酸*O*-甲基转移酶; F5H: 阿魏酸5-羟基化酶; CCo3H: 香豆酰CoA3-羟基化酶; CCoMT: 咖啡酰CoA3-*O*-甲基转移酶。?表示可能存在的生物转化过程。

收稿 2006-03-02 修定 2006-04-29

资助 国家自然科学基金(30300447)。

*通讯作者(E-mail: huzhibi@hotmail.com, Tel: 021-51322509)。

1 4CL同工酶的生化特性和催化活性

早在20世纪70年代初有关4CL的研究就已进行, 初期的工作主要集中在从植物中分离纯化4CL蛋白酶并对其生化特性进行研究。如Ranjeva等(1976)从矮牵牛(*Petunia hybrida* Vilm.)中分离到3个4CL同工酶Ia、Ib和Ic。三者相互间不能转化, 虽都能以*p*-香豆酸(4-香豆酸)为有效底物, 但对其它C6-C3单体化合物却表现出明显的底物特异性, Ia的最适底物为咖啡酸, 称之为咖啡酸:CoA连接酶, Ib的最适底物为芥子酸, 称之为芥子酸:CoA连接酶, Ic的最适底物为阿魏酸, 称之为阿魏酸:CoA连接酶。这3个同工酶的热稳定性和对酚酸类化合物的敏感性也不同, Ia能为*p*-香豆酰酯和咖啡酰奎宁酸酯所抑制, 对*p*-香豆酰葡萄糖和一些类黄酮不敏感; Ic能特异地为柚皮素(naringenin)所抑制, 而实验所选的化合物均不能抑制Ib即芥子酸:CoA连接酶的活性。他们还发现这些抑制反应均为非竞争性抑制, 说明底物与抑制剂在酶分子上可能有不同的结合位点。这些结果表明矮牵牛中不同4CL同工酶可能在苯丙烷类代谢不同分支代谢途径中起作用。大豆(*Glycine max* L., Knobloch和Hahlbrock 1975)、豌豆(*Pisum sativum* L., Wallis和Rhode 1977)和欧美杨(*Populus×euramericana*, Grand等1983)等的4CL同工酶研究也发现, 这些酶都是单链多肽组成的单体酶, 底物为4-香豆酸及其苯环上的羟基或甲氧基衍生物如咖啡酸、阿魏酸、芥子酸等, 最适pH一般在7.0左右, 粗略的估价分子量在55~67 kDa之间。

人们对4CL有较为明确的认识始于4CL基因的克隆。迄今已从多种植物中分离到编码4CL的基因, 如欧芹(*Petroselinum crispum*, Douglas等1987; Hauffe等1991; Lozoya等1988; Neutaedter等1999)、马铃薯(*Solanum tuberosum* L., Becker-André等1991)、火炬松(*Pinus taeda* L., Voo等1995; Zhang和Chiang 1997)、烟草(*Nicotiana tabacum* L., Lee和Douglas 1996)、大豆(*Glycine max* L., Lindermayr等2002; Uhlmann和Ebel 1993)、杂交杨(*Populus trichocarpa* Torr. & Gray×*P. deltoides* Marsh, Allina等1998)、拟南芥(*Arabidopsis thaliana*, Ehltting等1999; Hamberger

和Hahlbrock 2004; Lee等1995, 1997)等。这些植物基因组中4CL基因以小的基因家族形式存在, 一般有2~4个成员(Douglas 1996)。有些植物如马铃薯(Becker-André等1991)、欧芹(Douglas等1987; Lozoya等1988)和烟草(Lee和Douglas 1996)中4CL同工酶具有相似的分子特性, 对4-香豆酸及其羟基和甲氧基衍生物(图2)表现出相同或近乎相同的底物特异性; 而已研究的大多数植物如拟南芥(Ehltting等1999; Hamberger和Hahlbrock 2004)、大豆(Knobloch和Hahlbrock 1975; Lindermayr等2002)、颤杨(*Populus tremuloides* Michx., Hu等1998)、覆盆子(*Rubus idaeus* L., Kumer和Ellis 2003)等等, 其4CL基因家族具有结构和功能完全不同的成员, 各成员对几种常用底物表现出不同的亲和催化活性(表1), 预示着不同成员可能在苯丙烷类生物代谢不同的分支途径中起作用。

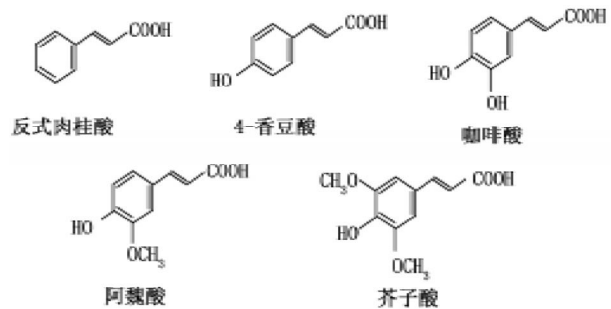


图2 4CL常用底物分子结构

大豆的重组蛋白Gm4CL1对阿魏酸盐、芥子酸盐、4-香豆酸盐、咖啡酸盐和3,4-二甲羟基肉桂酸盐表现出较强的催化活性, Gm4CL2也能催化咖啡酸盐、4-香豆酸盐、阿魏酸盐和肉桂酸, 但不能催化芥子酸盐和3,4-二甲羟基肉桂酸盐。由于这些底物相应的CoA酯是木质素和其它酚酸类化合物的前体分子, 因此推测Gm4CL1和Gm4CL2可能参与木质素和其它酚酸类化合物的生物合成(Lindermayr等2002)。Gm4CL3与Gm4CL4具有很强的相似性, 它们都能高效率地转化4-香豆酸盐和咖啡酸盐, 也能转化肉桂酸盐和阿魏酸盐, 但不能催化芥子酸盐和3,4-二甲羟基肉桂酸盐; 结合它们对来自于疫霉(*Phytophthora sojae*)及其游走孢子的 β -葡聚糖诱导子的上调特性

表1 部分4CL同工酶的底物结合特异性

植物种	同工酶	酶来源	底物结合特异性	
			最适底物	底物(亲和性由高到低)
欧芹(<i>Petroselinum crispum</i>)	Pc4CL-1	细胞培养物纯化	4-香豆酸盐	阿魏酸盐, 异阿魏酸盐, 肉桂酸盐, 咖啡酸盐, 芥子酸盐
	Pc4CL-2	细胞培养物纯化	4-香豆酸盐	Iso阿魏酸盐, 阿魏酸盐, 咖啡酸盐, 肉桂酸盐, 芥子酸盐
烟草(<i>Nicotiana tabacum</i> L. cv. Xanthi SRI)	Nt4CL-1	重组蛋白	4-香豆酸盐	阿魏酸盐, 咖啡酸盐, 芥子酸盐
	Nt4CL-2	重组蛋白	4-香豆酸盐	阿魏酸盐, 肉桂酸盐, 咖啡酸盐
颤杨(<i>Populus tremuloides</i> Michx.)	Pt4CL1	重组蛋白	4-香豆酸	阿魏酸, 咖啡酸, 5-羟基阿魏酸
	Pt4CL2	重组蛋白	4-香豆酸	咖啡酸, 阿魏酸
杂交杨(<i>Populus trichocarpa</i> Torr. & Gray× <i>P. deltoides</i> Marsh)	4CL-9	重组蛋白	4-香豆酸盐	阿魏酸盐, 咖啡酸盐, 肉桂酸盐, 5-羟基阿魏酸盐
	4CL-216	重组蛋白	4-香豆酸盐	阿魏酸盐, 咖啡酸盐, 肉桂酸盐, 5-羟基阿魏酸盐
	Ptd4CL3	重组蛋白(部分纯化)	4-香豆酸	阿魏酸, 咖啡酸, 肉桂酸
拟南芥(<i>Arabidopsis thaliana</i>)	At4CL1	重组蛋白	4-香豆酸盐	咖啡酸盐, 阿魏酸盐, 肉桂酸盐
	At4CL2	重组蛋白	咖啡酸盐	4-香豆酸盐, 肉桂酸盐
	At4CL3	重组蛋白	4-香豆酸盐	阿魏酸盐, 咖啡酸盐, 肉桂酸盐
	At4CL4	重组蛋白	芥子酸盐	咖啡酸盐, 4-香豆酸盐
大豆(<i>Glycine max</i> L. cv. Harosoy 63)	Gm4CL1	重组蛋白	阿魏酸盐	4-香豆酸盐, 芥子酸盐, 咖啡酸盐, 3,4-二甲氧基肉桂酸盐, 肉桂酸盐
	Gm4CL2	重组蛋白	咖啡酸盐	4-香豆酸盐, 阿魏酸盐, 肉桂酸盐
	Gm4CL3	重组蛋白	4-香豆酸盐	咖啡酸盐, 肉桂酸盐, 阿魏酸盐
	Gm4CL4	重组蛋白	4-香豆酸盐	咖啡酸盐, 肉桂酸盐, 阿魏酸盐

(Graham 和 Graham 1994), 推测 Gm4CL3/4 可能与防御有关的代谢如黄酮和异黄酮、细胞壁内酚酸类或其它水溶性的苯丙烷类化合物的生物合成有关(Lindermayr 等 2002)。

Ehltling 等(1999)在对拟南芥 4CL 重组蛋白的体外实验观察到, At4CL1 对 4-香豆酸盐、咖啡酸盐、阿魏酸盐的催化活性由强到弱依次递减, 但三者差别不大, 表明 At4CL1 对这几种底物都具有亲和活性; At4CL1 是 4 个成员中在成熟拟南芥木质化茎部转录表达量最多的基因, 在幼苗和子叶木质素开始积累的地方也有表达。At4CL2 对咖啡酸盐和 4-香豆酸盐也表现出强烈的催化活性, 但不能以阿魏酸盐为底物; 其 mRNA 表达量在幼苗中较高, 而在其它器官表达量却很低。At4CL4 是该基因家族中唯一能以芥子酸盐为底物而且是最适底物的成员, 它也能催化咖啡酸盐和 4-香豆酸盐(Hamberger 和 Hahlbrock 2004)。另一方面, At4CL3 重组蛋白对 4-香豆酸盐的催化活性远远大于阿魏酸盐、咖啡酸盐和肉桂酸盐, 而其在花器

官中 mRNA 表达量也远远高于其它木质化器官。由此推测: At4CL1、At4CL2 和 At4CL4 可能参与了拟南芥木质素、细胞壁内酚酸类化合物和其它苯丙烷类衍生物的生物代谢途径(Ehltling 等 1999; Hamberger 和 Hahlbrock 2004), 而 At4CL3 可能主要参与黄酮类化合物的生物合成途径(Ehltling 等 1999)。

2 4CL 基因的表达与调控

与植物次生代谢途径中的大多数酶类一样, 4CL 基因也属于调控基因(Becker-André 等 1991; Lee 等 1995), 许多环境因子诸如病虫害、伤害、紫外线等都会影响到 4CL 基因家族不同成员在植物特定组织器官中的表达情况。Ehltling 等(1999)用植物霜霉病的致病寄生真菌(*Peronospora parasitica*)的孢子感染拟南芥子叶, 发现该致病因素强烈诱导 At4CL1 和 At4CL2 mRNA 的表达, 其开始表达时间分别是感染后 12 和 24 h; 而 At4CL3 mRNA 则不被诱导。另一与抗病有关的基因 AtPOX 的 mRNA 最高表达水平则在 4 d 后出现。非

生物因素如伤害诱导 *At4CL1* 和 *At4CL2* 瞬间表达, 只是 *At4CL2* 的最高表达水平比 *At4CL1* 低, 而 *At4CL3* 不受伤害的诱导; 紫外线也能诱导的 *At4CL1* 和 *At4CL2* 瞬间表达, 紫外线照射 6 h 后其 mRNA 表达量最高; *At4CL3* 也被紫外线所诱导, 其 mRNA 表达的量相对较低, 最高表达水平也在 6 h 后出现, 一直持续到 24 h。这些结果表明 *At4CL1* 和 *At4CL2* 可能主要参与抗性有关化合物如木质素和细胞壁内酚酸类化合物等的生物代谢途径, 而 *At4CL3* 可能与光照有关的黄酮和异黄酮类化合物的生物合成有关。

4CL 是组织特异性表达基因, 同一物种中 *4CL* 基因家族中不同成员 mRNA 的表达表现出器官、组织和细胞的特异性, 这可能与不同成员参与不同的分支途径有关。Hu 等(1998)在研究颤杨 *Pt4CL* 基因中发现, *Pt4CL1* mRNA 特异性地在木质化的木质部组织中表达, 而 *Pt4CL2* 则特异地在茎和叶的表皮层中表达。这一结果也与 *Pt4CL1* 和 *Pt4CL2* 启动子驱动的 *GUS* 基因分别只在转基因烟草的木质部和表皮中表达的结果相一致(Hu 等 1998, 1999)。结合对两个基因体外功能的研究(表 1), 他们推测基因的特异性表达可能与 *Pt4CL1* 主要参与木质素的生物代谢, 而 *Pt4CL2* 主要参与黄酮类物质代谢有关。拟南芥中 *At4CL1*、*At4CL2* 和 *At4CL3* 的组织特异性表达(Ehlting 等 1999)与重组蛋白底物特异性相结合, 也反映出植物体内 *4CL* 家族各成员可能参与了不同的生物代谢途径(见上一节)。

4CL 基因还与苯丙氨酸氨裂解酶(phenylalanine ammonia lyase, PAL)和肉桂酸 4-羟基化酶(cinnamate-4-hydroxylase, C4H)基因的表达存在明显的相关性(Dixon 和 Paiva 1995; Ragg 等 1981; Weisshaar 和 Jenkin 1998)。Northern 杂交结果显示真菌诱导子和紫外线信号刺激后 *4CL* 和 *PAL* 基因 mRNA 表达量变化表现出一致性(Lois 等 1989)。拟南芥的致病性实验中也发现 *AtPAL* 和 *AtC4H* 为霜霉病致病真菌感染所诱导的 mRNA 表达特征与 *At4CL1* 和 *At4CL2* 非常相似(Ehlting 等 1999)。Logemann 等(1995)也发现含有紫外线的白光、真菌诱导子和伤害都能强烈地诱导欧芹 *PAL*、*C4H* 和 *4CL* 基因 mRNA 的瞬时积累, 并且在实验误差

允许的范围内三者 mRNA 总量的变化遵循相同的时间程序。蛋白水平的研究也证明用诱导子处理燕麦叶后 PAL、C4H 和 4CL 酶活性同时达到最高水平(Ishihara 等 1999)。PAL、C4H 和 4CL 催化相继发生的一系列生物转化, 是苯丙烷类化合物生物代谢总途径中的关键酶(图 1)。以上研究结果表明, *PAL*、*C4H* 和 *4CL* 基因的表达调控方式具有协同性, 其协同调控作用可能主要在转录水平进行(Reinold 等 1993, 1997)。

有人研究发现, *4CL*、*PAL* 和 *C4H* 基因的启动子区域有保守的“AC”富集的顺式作用元件(Hauffe 等 1991, 1993; Lois 等 1989; Neutaedter 等 1999), 称为“AC”元件, 包括 P 盒(Box P)、A 盒(Box A)、L 盒(Box L)和 H 盒(Box H)等, 它们的特征序列分别为 5'-TTC CAA A CC-3', 5'-CCGTCC-3', 5'-CCTACC ACC-3', 5'-OCTACC-3' (Logemann 等 1995; Loake 等 1992)。其中 P 盒、A 盒和 L 盒是决定这些基因为诱导子、紫外线和伤害等胁迫信号所诱导表达和组织细胞特异性表达所必需的顺式作用元件(Douglas 等 1991; Logemann 等 1995; Loake 等 1992)。研究还发现, 这些保守的区域是植物 MYB 相关的转录因子结合的特征核苷酸序列, 表明 *4CL* 基因在植物体内可能为 MYB 类转录因子所调控(Feldbrugge 等 1997; Logemann 等 1995)。研究欧芹 *4CL1* 启动子的结果表明, -210 bp (以转录起始位点为+1)的启动子片段可能是控制欧芹 *4CL1* 基因组织细胞特异性表达的最小启动子(Douglas 等 1991; Hauffe 等 1991, 1993; Neutaedter 等 1999)。活体 DNA 足迹(*in vivo* DNA footprint)实验发现该区域有两个关键结构, 位于-166~-120 之间的 FP4 (footprint 4)带有 P 盒的特征序列, 是与 MYB 类转录因子直接结合的功能域(domain); 位于-120~-78 的 FP56 (footprints 5 and 6), 以重复“C”为 10 个左右的碱基所分开为特征, 这一区域与欧芹和烟草的核蛋白表现出特异性结合, 是一个独立的顺式作用元件, 在 *4CL* 基因转录调控中起重要的作用(Douglas 等 1991; Hauffe 等 1991, 1993; Neutaedter 等 1999)。

4CL 基因还可能存在翻译后调控过程。Lee 和 Douglas (1996)在烟草 *4CL* 的研究中发现, 从烟草中提取的非纯化的 *4CL* 蛋白不具有催化肉桂酸的

活性; 与之不同的是, 其体外重组蛋白则可以肉桂酸为有效反应底物; 而将4CL重组蛋白与烟草提取物共培育后这一活性则被抑制, 这一抑制现象与提取物中具有高分子量的热不稳定性组份有关。由此他们推测, 烟草4CL酶可能存在翻译后修饰或在体内有与其它蛋白质相互作用的过程。

3 4CL氨基酸序列保守的功能域

Motif I——AMP结合功能域: 已知的4CL基因推导的氨基酸序列中都存在保守的AMP结合功能域(Branchini等2000; Challis等2000; Gocht和Marahiel 1994; Stachelhaus等1999)——“SSGTTGLPKG”。这一区域不仅在4CL中保守, 而且是所有腺苷酸形成酶如荧光素酶、乙酰辅酶A合成酶、脂酰辅酶A合成酶等的共有特征, 是这类酶催化底物形成腺苷酸中间产物所必需的氨基酸序列, 也是这一超基因家族(superfamily)的分类依据之一(Stuible等2000; Stuible和Kombrink 2001)。

Motif II——“GEICIRG”: Motif II (又称Box II) “GEICIRG”是4CL中另一绝对保守的区域(Stuible和Kombrink 2001; Zhang和Chiang 1997)。已知的4CL氨基酸序列中只有大豆Gm4CL1在该区域有轻微的不同, 为“GEICIIIG”。曾推测Motif II直接参与4CL的催化反应, 其中间的半胱氨酸残基(Cys)侧链可能是参与催化反应的基团。这一假说是在抑制实验和Cys在腺苷酸形成酶中的高度保守性基础提出的。最近的研究发现, 该位Cys突变为Ala的酶仍表现出相当的酶活性(Stuible等2000; Stuible和Kombrink 2001); 而Ohmiya和Tsuji (1997)的研究也表明该位Cys并非萤火虫的荧光素酶生物发光所必需, 但Motif II缺失的4CL酶活性会完全丧失。因此推测这一序列可能也与稳定蛋白质的活性空间构象有关(Stuible等2000)。

4 底物结合功能域sbd I和sbd II

研究的植物大多存在2~4个4CL同工酶, 多数同工酶具有底物结合特异性。Stuible和Kombrink (2001)利用At4CL1和At4CL2的底物特异性结合特性, 结合短杆菌肽S合成酶1苯丙氨酸激活结构域(phenylalanine-activating domain, PheA)的研究结果[许多腺苷酸形成酶的蛋白质立体结构

与底物识别功能域已研究得较为清晰(Broun等1998; Chang等1997; Conti等1996; Gocht和Marahiel 1994; Wang和Pichersky 1999), 结合短杆菌肽S合成酶即为其一], 已确定了与底物特异性结合密切相关的两个区域: sbd (substrate binding domain) I和sbd II。并推测, 4CL中与PheA相类似的功能域sbd I是决定4CL底物特异性的关键区域, 这一区域可能是底物与酶直接结合的一个“口袋”。sbd II则可能是通过与sbd I区域氨基酸残基的相互作用改变了sbd I的结构, 从而间接地影响了酶与底物的特异性结合。sbd II区域内Motif II羧基端的氨基酸残基可能是限制底物进入sbd I口袋的因素。

根据拟南芥4CL突变体的研究, 人们设想sbd I区域几个位置上氨基酸残基的空间位阻可能在4CL底物特异性选择中起关键作用。Stuible和Kombrink (2001)报道, At4CL1、At4CL2和At4CL3中唯有At4CL2不能以分子量相对较大的阿魏酸(图2, 底物结构式)。三者氨基酸的序列比较表明, sbd I区域存在两个关键位点的氨基酸残基差异: At4CL2序列中Met²⁸³和Lys³²⁰在At4CL1和At4CL3中对应位置分别为空间位阻相对较小的Pro和Ala。进一步实验发现, At4CL2 Met²⁸³/Pro和Met²⁸³/Ala的突变体对阿魏酸的催化活性分别表现出与At4CL1和At4CL2具有相似的特性; 而Lys³²⁰/Leu和Lys³²⁰/Ala的突变体对咖啡酸的酶活性也表现出一定程度的提高。推测sbd I区Met²⁸³和Lys³²⁰可能是位于“口袋”开口处的两个氨基酸残基, 这两个氨基酸残基的空间位阻可能是影响At4CL2与底物特异性结合的重要因素(Stuible等2000; Stuible和Kombrink 2001)。

5 4CL基因在基因工程中的应用

如引言中所述, 苯丙烷类化合物不仅对植物本身在环境中生存具有重要意义, 而且与人类日常生活密切相关。与此代谢途径有关的基因工程研究也随之展开(Hu等1999; Kajita等1997; Li等2003)。4CL处在苯丙烷类化合物生物代谢中总途径向分支途径的转折点, 是这一生物合成途径中的限速酶之一。由于在代谢途径中的特殊位置和关键作用, 所以4CL已成为苯丙烷类化合物生物合成基因工程工作中的一个有力候选者。

木质素是植物细胞壁的组成部分, 对植物的正常生长起作用, 但对造纸纸浆工业来说, 木质素却是影响造纸工业中浆和纸的品质与产率的重要因素。从植物中除去木质素的常规方法既消耗大量化学药品和能源, 又严重污染环境。而基因工程手段调控木质素的合成可能是解决这一问题的根本途径之一。Lee等(1997)采用反义基因抑制拟南芥中4CL酶的活性, 拟南芥中木质素的组成成分即发生变化, 愈疮木基丙烷(guaiacyl, G)单位含量显著减少, 而紫丁香基丙烷(syringyl, S)单位则没有变化, 转基因拟南芥中G/S比大大降低。G和S是构成木质素的不同结构单元, G/S比是反应木质素降解难易的一个指标。类似的结果还在转反义4CL基因的烟草和颤杨中得到: 转基因烟草木质素分子组成成分改变, 褐化组织中木质素含量由原先的23.4%±0.2%下降到20.3%±0.1% (Kajita等1997); 4CL酶活性受到抑制的转基因颤杨中木质素含量下降45%, 纤维素含量增加15% (Hu等1999)。Li等(2003)将反义4CL基因和正义*CA1d5H* (松柏醛5-羟基化酶)基因协同转入颤杨后, 木质素的含量减低52%, S/G比率提高64%, 纤维素成分增加30%, 这为人们改造植物中木质素生物合成提供了科学依据。

黄酮是苯丙烷类化合物生物代谢中另一类重要产物, 由于其在医药和食品添加剂中的经济价值, 近些年来人们试图在微生物中构建黄酮的生物合成途径以达到采用发酵技术生产此类化合物的目的。4CL是构建这一途径的关键酶之一(Jiang等2005; Kaneko等2003)。将不同来源的PAL、4CL和查耳酮合成酶(chalcone synthase, CHS)转入原核生物大肠杆菌(*Escherichia coli*)后表达这3个基因的大肠杆菌即能产生两种黄烷酮(flavanone): 松属素(pinocembrin)和柚皮素(naringenin), 产率分别是450和750 $\mu\text{g}\cdot\text{L}^{-1}$ (Kaneko等2003)。4CL基因与黄酮生物合成途径其它关键酶基因在大肠杆菌中联合表达也使得大肠杆菌产生黄酮醇山奈酚(flavonol kaempferol)和3', 4' B-环羟基黄酮醇槲皮素(3', 4' B-ring hydroxylated flavonol quercetin, Leonard等2005)、芹菜素(apigenin)和荜花素(genkwanin)等化合物(Leonard等2006)。真核生物酵母菌(*Saccharomyces cerevisiae*)中转入植物来源的

PAL、4CL和CAH基因后也能产生大约0.8 $\text{mg}\cdot\text{L}^{-1}$ 的松属素和7 $\text{mg}\cdot\text{L}^{-1}$ 的柚皮素; 该菌株还能生成2', 4', 6'-三羟基二氢查耳酮(2', 4', 6'-trihydroxydihydro-chalcone)和根皮素(phloretin)等副产物(Jiang等2005)。

这些研究成果不仅为阐明4CL在苯丙烷类化合物生物代谢中的功能提供了依据, 也为应用4CL基因工程中的作用提供了借鉴。多个功能基因的共转化拓宽了基因工程的思路, 为改变多酶控制的植物次生代谢途径提供了可能。

6 4CL基因家族的进化

从GenBank中提取不同植物共37种4CL基因和蓝色链霉菌(*Streptomyces coelicolor*) ScCCL、拟南芥At4CL-like的氨基酸序列, 加上我们从丹参(*Salvia miltiorrhiza* Bunge)中克隆的两个4CL基因(*Sm4CL1*和*Sm4CL2*)进行同源分析, 以Clustal方法建立的基因进化树(图3)中可以看出, 基因进化树中4CL聚集成两大类群——Class I和Class II, 不同4CL基因并不完全按照物种聚集, 即同一植物的不同4CL基因在进化树中关系并非最近。如杂交杨(*Populus trichocarpa* Torr. & Gray×*P. deltoides* Marsh) Pbd4CL1、Pbd4CL2和Pbd4CL3同属于Class I, 而Pbd4CL4属于Class II; 即使同处Class I, Pbd4CL3与Pbd4CL1、Pbd4CL2的进化关系也较远。由此推测, 4CL基因可能是一个较为古老的基因, 早在被子植物与裸子植物分化之前4CL基因就已存在, 并已在复制中发生了变异, 遂有Class I和Class II的分类(Cukovic等2001; Ehlting等1999; Kumar和Ellis 2003)。链霉菌中肉桂酸:辅酶A连接酶(ScCCL)的存在似乎也证实了这一推论(Kaneko等2003)。在相对较近的年代, Class I和Class II中4CL基因仍然发生着变异, 从而导致现代植物中4CL基因家族多成员的出现。4CL基因家族多成员的出现是植物长期适应环境的结果, 对物种的繁衍生息起着不可少的作用。

4CL基因进化树中Class I和Class II的形成, 可能更多地反映出4CL功能上的分类(Cukovic等2001; Kaneko等2003; Kumar和Ellis 2003)。实验发现, Class I中拟南芥At4CL1、At4CL2和大豆Gm4CL1主要与黄酮的生物合成相关; Class II中

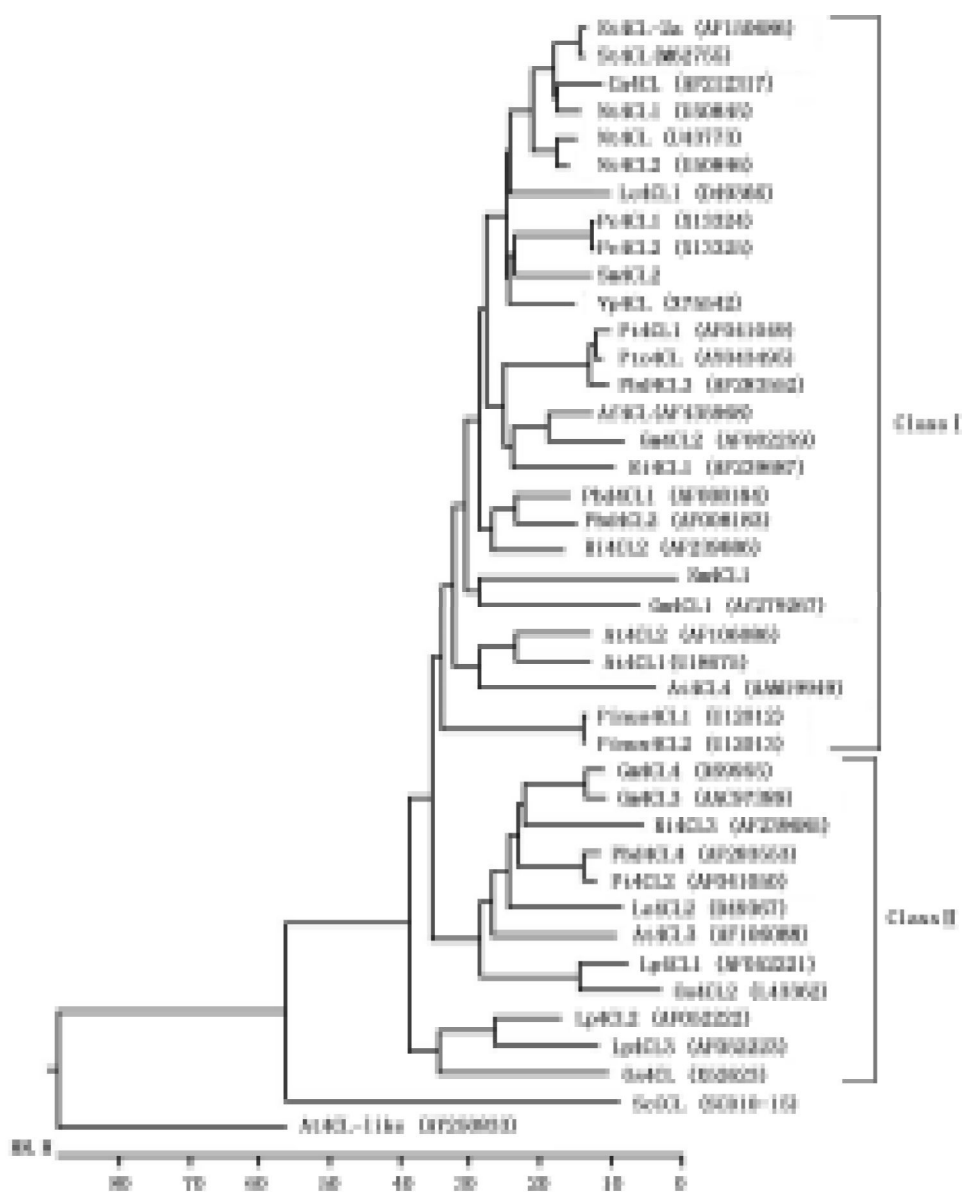


图3 CLUSTAL方法建立的4CL基因进化树

Af: 紫穗槐(*Amorpha fruticosa*); At: 拟南芥(*Arabidopsis thaliana*); Ca: 辣椒(*Capsicum annuum*); CCL: 肉桂酸:辅酶A连接酶(cinnamate:coenzyme A ligase); Gm: 大豆(*Glycine max*); Le: 紫草(*Lithospermum erythrorhizon*); Lp: 黑麦草(*Lolium perenne*); Nt: 烟草(*Nicotiana tabacum*); Os: 水稻(*Oryza sativa*); Pbδ 杂交杨(*Populus trichocarpa*×*Populus deltoides*); Pc: 欧芹(*Petroselinum crispum*); Pinus: 火炬松(*Pinus taeda*); Ptα 毛白杨(*Populus tomentosa*); Pt: 颤杨(*Populus tremuloides*); Ri: 覆盆子(*Rubus idaeus*); Sc: 蓝色链霉菌(*Streptomyces coelicolor*); Sm: 丹参(*Salvia miltiorrhiza*); St: 马铃薯(*Solanum tuberosum*); Vp: 香荚兰(*Vanilla planifolia*)。

At4CL3、Gm4CL2、Gm4CL3 和 Gm4CL4 主要参与木质素及其它苯丙烷类化合物的生物合成(Ehlting等1999; Lindermayr等2002)。

7 结语

植物次生代谢基因工程的出现是人类对次生

代谢途径的深入了解和分子生物学向纵深发展的结果, 同时它又促进了次生代谢分子生物学的发展。4CL 在苯丙烷类化合物生物代谢途径的关键作用及其此类化合物对植物本身和人类的重要意义, 使得 4CL 成为植物次生代谢领域研究的重要

对象之一。

具有不同功能的4CL同工酶的存在,为人们有针对性地选择不同的4CL基因,通过基因工程技术控制苯丙烷类化合物的物质流向提供了可能。如Jiang等(2005)将拟南芥具有广泛底物结合特性的*At4CL1*基因与*PAL*和*CHS*基因转入酵母菌中获得可以产生黄酮的转基因酵母菌株。Hu等(1998, 1999)将颤杨中主要参与木质素生物合成的*Pt4CL1*反义基因与正义*CA1d5H*基因协同转入颤杨后也使得木质素的含量明显降低(Li等2003)。

基因的表达调控是基因发挥其功能的根本,基因转录水平的调控是植物次生代谢调控的主要方式,而转录因子在此过程中起着关键作用(Vom Endt等2002)。P盒、A盒和L盒等顺式元件的存在表明4CL基因受MYB类转录因子的调控,这种调控与苯丙烷类代谢其它关键酶具有协同性。虽然目前尚无转录因子调控4CL表达的报道,但4CL表达调控方面的研究成果以及Grotewold等(1998)和Bruce等(2000)利用转录因子调控玉米细胞中花青素(anthocyanin)积累的研究为利用转录因子调控此类化合物生物合成的基因工程提供了可能。

4CL功能域氨基酸位点的突变体结合酶的体外功能的研究和蛋白质晶体结构模型,将为最终阐明蛋白质特定位点氨基酸残基的功能提供依据。这也是蛋白质组学的一个重要的研究方向——结构与功能蛋白质学。关键位点影响4CL底物特异性结合的研究成果,为人们通过点突变在体外构建具有新功能的4CL基因提供了思路。在转基因植物中表达这种具有新功能的基因,则可能会在原植物中产生新的物质代谢流向。

可以预见,在前人工作基础上进一步深入展开4CL基因的研究,不仅能揭示4CL的生物学功能,还将为4CL参与的生物代谢途径的改造和用转录因子及多基因协同转化的植物次生代谢基因工程提供依据。

参考文献

Allina SM, Pri-Hadash A, Theilmann DA, Ellis BE, Douglas CJ (1998). 4-Coumarate:coenzyme A ligase in hybrid poplar: properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes. *Plant Physiol*, 116: 743~754
An JH, Lee GY, Jung JW, Kim YS (1999). Identification of residues

essential for a two-step reaction by malonyl-CoA synthetase from *Rhizobium trifolii*. *Biochem J*, 344: 159~166
Becker-André M, Schulze-Lefert P, Hahlbrock K (1991). Structural comparison, modes of expression, and putative *cis*-acting elements of the two 4-coumarate: CoA ligase genes in potato. *J Biol Chem*, 266: 8551~8559
Branchini BR, Murtiashaw MH, Magyar RA, Anderson SM (2000). The role of lysine 529, a conserved residue of acyladenylate-forming enzyme superfamily, in firefly luciferase. *Biochemistry*, 39: 5433~5440
Brodelius PE, Xue ZT (1997). Isolation and characterization of a cDNA from cell suspension cultures of *Vanilla planifolia* encoding 4-coumarate:coenzyme A ligase. *Plant Physiol Biochem*, 35: 497~506
Broun P, Shanklin J, Whittle E, Somerville C (1998). Catalytic plasticity of fatty acid modification enzymes underlying chemical diversity of plant lipids. *Science*, 282: 1315~1317
Bruce W, Folkerts O, Garnaat C, Crasta O, Roth B, Bowen B (2000). Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P. *Plant Cell*, 12: 65~79
Challis GL, Ravel J, Townsend CA (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol*, 7: 211~224
Chang KH, Xiang H, Dunaway-Mariano D (1997). Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with *Pseudomonas* sp. strain CBS3 4-chlorobenzoate:coenzyme A ligase. *Biochemistry*, 36: 15650~15659
Conti E, Franks NP, Brick P (1996). Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*, 4: 287~298
Cukovic D, Ehltling J, VanZiffle J, Douglas CJ (2001). Structure and evolution of 4-coumarate:coenzyme A ligase (4CL) gene families. *Biol Chem*, 382: 645~654
Dixon RA, Paiva NL (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell*, 7: 1085~1097
Douglas CJ (1996). Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends Plant Sci*, 1: 171~178
Douglas CJ, Hauffe KD, Ites-Morales ME, Ellard M, Paszkowski U, Hahlbrock K, Dangl JL (1991). Exonic sequences are required for elicitor and light activation of a plant defense gene, but promoter sequences are sufficient for tissue specific expression. *EMBO J*, 10 (7): 1767~1775
Douglas CJ, Homann H, Schulz W, Hahlbrock K (1987). Structure and elicitor or U.V.-light-stimulated expression of two 4-coumarate:CoA ligase genes in parsley. *EMBO J*, 6: 1189~1195
Ehltling J, Buettner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E (1999). Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionary classes in angiosperms. *Plant J*, 19: 9~20
Feldbrugge M, Sprenger M, Hahlbrock K, Weisshaar B (1997). PcMYB1, a novel plant protein containing a DNA-binding

- domain with one MYB repeat, interacts *in vivo* with a light-regulatory promoter unit. *Plant J*, 11 (5):1079~1093
- Gocht M, Marahiel MA (1994). Analysis of core sequences in the D-Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis. *J Bacteriol*, 176: 2654~2662
- Graham MY, Graham TL (1994). Wound-associated competency factors are required for the proximal cell responses of soybean to the *Phytophthora sojae* wall glucan elicitor. *Plant Physiol*, 105: 571~578
- Graham TL, Graham MY (1991). Cellular coordination of molecular responses in plant defense. *Mol Plant-Microbe Interact*, 4: 415~422
- Grand C, Boudet A, Boudet AM (1983). Isoenzymes of hydroxycinnamate:CoA ligase from poplar stems, properties and tissue distribution. *Planta*, 158: 225~229
- Grotewold E, Chamberlin M, Snook M, Siame B, Butler L, Swenson J, Maddock S, St Clair G, Bowen B (1998). Engineering secondary metabolism in maize cells by ecotopic expression of transcription factors. *Plant Cell*, 10: 721~740
- Hahlbrock K, Scheel D (1989). Physiology and molecular biology of phenylpropanoid metabolism. *Annu Rev Plant Physiol Plant Mol Biol*, 40: 347~369
- Hamberger B, Hahlbrock K (2004). The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc Natl Acad Sci USA*, 101 (7): 2209~2214
- Hauffe KD, Lee SP, Subramaniam R, Douglas CJ (1993). Combinatorial interactions between positive and negative *cis*-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J*, 4: 235~253
- Hauffe KD, Paszkowski U, Schulze-Lefert P, Hahlbrock K, Dangl JL, Douglas CJ (1991). A parsley 4CL-1 promoter fragment specifies complex expression patterns in transgenic tobacco. *Plant Cell*, 3: 435~443
- Hu WJ, Harding SA, Lung J, Popko JL, Ralph J, Stokke DD, Tsai CJ, Chiang VL (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat Technol*, 17: 808~812
- Hu WJ, Kawaoka A, Tsai CJ, Lung J, Osakabe K, Ebinuma H, Chiang VL (1998). Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (*Populus tremuloides*). *Proc Natl Acad Sci USA*, 95: 5407~5412
- Ishihara A, Ohtsu Y, Iwamura H (1999). Induction of biosynthetic enzymes for avenanthramides in elicitor-treated oat leaves. *Planta*, 208: 512~518
- Jiang H, Wood KV, Morgan JA (2005). Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*, 71 (6): 2962~2969
- Kajita S, Hishiyama S, Tomimura Y, Katayama Y, Omori S (1997). Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate:coenzyme A ligase is depressed. *Plant Physiol*, 114: 871~879
- Kaneko M, Hwang E, Ohnishi Y, Horinouchi S (2003). Heterologous production of flavanones in *Escherichia coli*: potential for combinatorial biosynthesis of flavonoids in bacteria. *J Ind Microbiol Biotechnol* 30: 456~461
- Kaneko M, Ohnishi Y, Horinouchi S (2003). Cinnamate:coenzyme A ligase from the filamentous bacterium *Streptomyces coelicolor* A3(2). *J Bacteriol*, 185: 20~27
- Knobloch KH, Hahlbrock K (1975). Isoenzymes of *p*-coumarate:CoA ligase from cell suspension cultures of *Glycine max*. *Eur J Biochem*, 52: 311~320
- Kumar A, Ellis BE (2003). 4-Coumarate:CoA ligase gene family in *Rubus idaeus*: cDNA structures, evolution, and expression. *Plant Mol Biol*, 31: 327~340
- Lee D, Douglas CJ (1996). Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family. *Plant Physiol*, 112: 193~205
- Lee D, Ellard M, Wanner LA, Davis KR, Douglas CJ (1995). The *Arabidopsis* 4-coumarate:CoA ligase (4CL) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant Mol Biol*, 28: 871~884
- Lee D, Meyer K, Chapple C, Douglas CJ (1997). Antisense suppression of 4-coumarate:coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. *Plant Cell*, 9: 1985~1998
- Leonard E, Chemler J, Lim KH, Koffas MA (2006). Expression of a soluble flavone synthase allows the biosynthesis of phytoestrogen derivatives in *Escherichia coli*. *Appl Microbiol Biotechnol*, 70: 85~91
- Leonard E, Yan Y, Koffas MA (2005). Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. *Metabolic Engineering*, Dec 26; [Epub ahead of print]
- Li L, Zhou Y, Cheng X, Sun J, Marita JM, Ralph J, Chiang VL (2003). Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc Natl Acad Sci USA*, 100: 4939~4944
- Lindermayr C, Möllers B, Fliegmann J, Uhlmann A, Lottspeich F, Meimberg H, Ebel J (2002). Divergent members of a soybean (*Glycine max* L.) 4-coumarate:coenzyme A ligase gene family: Primary structures, catalytic properties, and differential expression. *Eur J Biochem*, 269: 1304~1315
- Loake GJ, Faktor O, Lamb CJ, Dixon RA (1992). Combination of H-box [CCTACC(N)7CT] and G-box (CACGTG) *cis* elements is necessary for feed-forward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate *p*-coumaric acid. *Proc Natl Acad Sci USA*, 89(19): 9230~9234
- Logemann E, Parniske M, Hahlbrock K (1995). Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. *Proc Natl Acad Sci USA*, 92(13): 5905~5909
- Lois R, Dietrich A, Hahlbrock K, Schulz W (1989). A phenylalanine ammonia-lyase gene from parsley: structure, regulation

- and identification of elicitor and light responsive *cis*-acting elements. *EMBO J*, 8: 1641~1648
- Lozoya E, Homann H, Douglas CJ, Schulz W, Scheel D, Hahlbrock K (1988). Primary structures and catalytic properties of isoenzymes encoded by two 4-coumarate:CoA ligase genes in parsley. *Eur J Biochem*, 176: 661~667
- Neutaedter DA, Lee SP, Douglas CJ (1999). A novel parsley 4CL1 *cis*-element is required for developmentally regulated expression and protein-DNA complex formation. *Plant J*, 18: 77~88
- Ohmiya Y, Tsuji FI (1997). Mutagenesis of firefly luciferase shows that cysteine residues are not required for bioluminescence activity. *FEBS Lett*, 404: 115~117
- Ragg H, Kuhn DN, Hahlbrock K (1981). Coordinated regulation of 4-coumarate:CoA ligase and phenylalanine ammonia-lyase mRNAs in cultured plant cells. *J Biol Chem*, 256: 10061~10065
- Ranjeva R, Boudet A, Faggiox R (1976). Phenolic metabolism in petunia tissues. IV. Properties of *p*-coumarate:coenzyme A ligase isoenzymes. *Biochimie*, 58: 1255~1262
- Reinold S, Hauffe KD, Douglas CJ (1993). Tobacco and parsley 4-coumarate:coenzyme A ligase genes are temporally and spatially regulated in a cell type-specific manner during tobacco flower development. *Plant Physiol*, 101: 373~383
- Reinold S, Hauffe KD, Douglas CJ (1997). *In situ* localization of phenylpropanoid biosynthetic mRNAs and proteins in parsley (*Petroselinum crispum*). *Bot Acta*, 110: 431~443
- Stachelhaus T, Mootz HD, Marahiel MA (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol*, 6: 493~505
- Stuible HP, Buettner D, Ehltng J, Hahlbrock K, Kombrink E (2000). Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylation-forming enzymes. *FEBS Lett*, 467: 117~122
- Stuible HP, Kombrink E (2001). Identification of the substrate specificity-conferring amino acid residues of 4-coumarate:coenzyme A ligase allows the rational design of mutant enzymes with new catalytic properties. *J Biol Chem*, 276: 26893~26897
- Uhlmann A, Ebel J (1993). Molecular cloning and expression of 4-coumarate:coenzyme A ligase, an enzyme involved in the resistance response of soybean (*Glycine max* L.) against pathogen attack. *Plant Physiol*, 102: 1147~1156
- Vom Endt D, Kijine JW, Memlink J (2002). Transcription factors controlling plant secondary metabolism: what regulates the regulators? *Phytochemistry*, 61: 107~114
- Voo KS, Whetten RW, O'Malley DM, Sederoff RR (1995). 4-coumarate:coenzyme A ligase from loblolly pine xylem. Isolation, characterization, and complementary DNA cloning. *Plant Physiol*, 108: 85~97
- Wallis PJ, Rhodes MJC (1977). Multiple forms of hydroxycinnamate:CoA ligase in etiolated pea seedling. *Phytochemistry*, 16: 1891~1894
- Wang J, Pichersky E (1999). Identification of specific residues involved in substrate discrimination in two plant *O*-methyltransferases. *Arch Biochem Biophys*, 368: 172~180
- Weisshaar B, Jenkin GI (1998). Phenylpropanoid biosynthesis and its regulation. *Curr Opin Plant Biol*, 1: 251~257
- Zhang XH, Chiang VL (1997). Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. *Plant Physiol*, 113: 65~74