

DOI: 10.11931/guihaia.gxzw201306025

凌瑶,高飞,王安虎,等. 苦荞 4-香豆酸辅酶 A 连接酶基因(*Ft4CL*)的克隆及序列分析[J]. 广西植物,2015,35(5):728—732Ling Y,Gao F,Wang AH,*et al.* Cloning and sequence analysis of 4-coumarate: CoA ligase gene from *Fagopyrum tatarium*[J]. *Guihaia*,2015,35(5):728—732

苦荞 4-香豆酸辅酶 A 连接酶基因 (*Ft4CL*) 的克隆及序列分析

凌瑶¹, 高飞², 王安虎³, 李成磊², 陈惠², 吴琦^{2*}

(1. 四川农业大学 新农村发展研究院, 四川 雅安 625014; 2. 四川农业大学 生命科学学院, 四川 雅安 625014; 3. 西昌学院, 四川 西昌 615000)

摘要: 以苦荞栽培种‘西荞 2 号’为材料,利用同源克隆和 RT-PCR 技术获得 *Ft4CL* 保守片段,采用 RACE 技术获得 *Ft4CL* 基因的 3'末端及 5'末端序列,并进一步采用生物信息学方法进行序列分析。结果表明:从苦荞花蕾总 RNA 中获得一条苦荞麦(*Fagopyrum tatarium*)4-香豆酸辅酶 A 连接酶基因(4-coumarate: Coa ligase, *Ft4CL*)的 cDNA 全长序列。生物信息学分析结果显示,*Ft4CL* 基因 ORF 全长 1 602 bp,可编码 553 个氨基酸,理论标准分子质量为 58.02 kDa,等电点(pI)为 5.23。该研究首次从苦荞中获得 *Ft4CL* 基因的 cDNA 全长序列,该基因具有植物 4CL 同源基因的典型特征,推导的氨基酸序列具有 4CL 的所有活性位点并归属于黄酮代谢支路。该研究结果可为深入研究苦荞黄酮代谢途径奠定基础,为采用代谢工程技术提高苦荞黄酮含量提供候选靶基因。

关键词: 苦荞; 4-香豆酸辅酶 A 连接酶; 基因克隆; 序列分析

中图分类号: Q943.2 文献标识码: A 文章编号: 1000-3142(2015)05-0728-05

Cloning and sequence analysis of 4-coumarate: CoA ligase gene from *Fagopyrum tatarium*

LING Yao¹, GAO Fei², WANG An-Hu³, LI Cheng-Lei², CHEN Hui², WU Qi^{2*}

(1. Institute for New Socialist Countryside Development, Sichuan Agricultural University, Ya'an 625014, China; 2. College of Life Sciences, Sichuan Agricultural University, Ya'an 625014, China; 3. Xichang College, Xichang 615000, China)

Abstract: This study focused on cloning and characterizing the 4-coumarate: CoA ligase gene (*Ft4CL*) from *Fagopyrum tatarium*. Using buckwheat species 'Xi Qiao No.2', according to the conserved sequences of 4CL from GenBank, a pair of degenerate primer was designed and synthesized. Through RT-PCR (reverse transcription PCR) technique, the conserved fragment of *Ft4CL* was amplified from the total RNA of *F. tataricum* flower buds. Then, the RACE technique (rapid-amplification of cDNA ends) was performed, and the 5' end and 3' end of *Ft4CL* were successfully amplified, respectively. The complete cDNA of *Ft4CL* was obtained by splicing the above sequences, and a pair of gene-specific primers was synthesized to amplify the ORF (open reading frame) region of *Ft4CL*. Using DNAMAN software to deduce the ORF sequence of *Ft4CL* to the amino acid sequence, its homologous with other 4CLs were analyzed by NCBI Blast tool. The secondary structure of *Ft4CL* was predicted by SOPMA (<http://pbil.ibcp.fr>), multiple sequence alignment was performed by DNAMAN software, and phylogenetic tree was built with neighbor-joining method by MEGA 5.0. The results were as follows: the similarity of *Ft4CL* with *F. esculentum*

收稿日期: 2014-10-18 修回日期: 2015-03-13

基金项目: 国家科技支撑计划项目(2013BAD20B07); 四川省科技厅育种攻关项目(2011NZ0098-17)。

作者简介: 凌瑶(1978-),女,四川雅安人,博士,副研究员,研究方向为草种质资源创新及育种,(E-mail)lingyao23@163.com。

*通讯作者: 吴琦,博士,教授,研究方向为植物分子生物学,(E-mail)wuqiwwq@gmail.com。

(HM149785) showed the highest level (up to 94%) and it ranging from 66%–75% with other plant 4CLs. Multiple sequence alignment results showed that the *Ft4CL* had conserved motifs of BOXI and BOXII near by the C-terminus, but it had relatively low similarity with other 4CLs at the C-terminus. According to the phylogenetic tree analysis results, the selected 4CLs were grouped into 2 cluster, *Ft4CL*, *Arabidopsis thaliana* 4CL1, and *A. thaliana* At4CL2 belonged to Cluster I. In conclusion, the results could provide basic data for in-depth study of *Fagopyrum tataricum* flavonoid pathway. Furthermore, this study indicated that *Ft4CL* could be a new candidate target gene for developing high flavonoid *F. tataricum* by metabolic engineering technology in future.

Key words: *Fagopyrum tataricum*; 4-coumarate; CoA ligase; gene cloning; sequence analysis

苦荞(*Fagopyrum tataricum*)是蓼科荞麦属一年生草本植物,主要分布在海拔 2 500~3 500 m 的高寒山区,原产我国及印度等地,在我国主要分布在西南山区及陕西、山西等地。苦荞作为一种药食兼用的传统作物,富含大量以芦丁为代表的黄酮类化合物。已有研究表明,黄酮类化合物具有多种生物活性,在医药、食品、保健等领域具有广阔的前景,被认为是 21 世纪具有前景的绿色食品。

植物黄酮类化合物的合成源于苯丙烷代谢途径,而 4-香豆酸辅酶 A 连接酶(4CL)是植物苯丙烷类生物合成途径流向下分支途径的最后一个酶。它通过两步反应催化 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(香豆酸 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。此后,香豆酸 CoA 在查尔酮合酶(CHS)的催化作用下生成查尔酮,开启黄酮类化合物代谢支路,生成异黄酮、黄酮醇和花青素等产物。这些次生代谢产物在植物的生长发育及抗病抗逆过程中均起重要作用。目前,4CL 基因已在烟草、杨树、拟南芥等植物中被克隆研究,其在转录水平上的丰度和蛋白质水平上的活性能有效影响植物木质素和黄酮类化合物的生物合成量。本研究以苦荞‘西荞 2 号’为材料,采用 RACE 技术分离得到其 4-香豆酸辅酶 A 连接酶基因(*Ft4CL*),并对其进行生物信息学分析,为深入研究 *Ft4CL* 基因在苦荞黄酮类化合物代谢途径中的作用奠定基础。

1 材料与方 法

1.1 材料与试剂

苦荞‘西荞 2 号’的种子由西昌学院王安虎教授提供,种植于四川农业大学试验基地,供试材料于苦荞花期取样保存。大肠杆菌 DH5 α 由本实验室保存。

Takara 公司:质粒 DNA 小量提取试剂盒、胶回

收试剂盒、Taq DNA 聚合酶、克隆载体 pMD19-T、3'-Full RACE Core Set Ver. 2.0、5'-Full RACE Kit;天泽基因工程有限公司;植物 RNA_{out}(Trizol 法)试剂盒;Fermentas 公司:逆转录试剂盒 RevertAid™ First Strand cDNA Synthesis Kit;引物由上海英骏生物公司合成;其他试剂均为进口或国产分析纯。

1.2 主要仪器与设备

MyCycler™ PCR 仪;凝胶成像系统,美国 Bio-Rad 公司;DNP-9272 型电热恒温培养箱,上海精宏试验设备有限公司;高速冷冻离心机 Thermo Electron Corporation;DYY-III-68 型稳压稳流电泳仪,北京市六一仪器厂。

1.3 方法

1.3.1 苦荞总 RNA 提取和 cDNA 第一链制备 按照植物 RNA_{out} 试剂盒方法提取苦荞花蕾总 RNA,琼脂糖凝胶电泳检验其质量。采用反转录试剂盒(RevertAid™ First Strand cDNA Synthesis Kit),以 Oligo(dT)引物反转录获得 cDNA 第一链,并保存于-20 °C 冰箱备用。

1.3.2 苦荞 *Ft4CL* cDNA 的克隆 根据 GenBank 中植物 4CL 氨基酸保守序列设计一对兼并引物,以 cDNA 第一链为模板扩增 *Ft4CL* 保守片段。将保守片段克隆于 pMD19-T 载体上,经蓝白斑筛选,挑选阳性克隆,送上海英骏公司测序,根据测序结果设计 2 条 3'RACE 引物和 2 条 5'RACE 引物(表 1)。

参照 3'RACE 试剂盒说明,以苦荞花蕾 RNA 为模板合成 3'RACE 第一链 cDNA,以试剂盒引物(3-gsp)和特异引物(3-ngsp1、3-ngsp2)进行 2 轮巢式 PCR。参照 5'RACE 试剂盒说明,合成 5'RACE 第一链 cDNA,以试剂盒引物(5-gsp)和特异引物(5-ngsp1、5-ngsp2)进行 2 轮巢式 PCR。PCR 产物经连接转化后,各挑选 3 个阳性克隆送英骏生物公司测序。

利用 Dnaman 软件拼接保守片段、3'RACE 和

表 1 引物序列及用途
Table 1 Primers used in the study

引物名称 Primer name	引物序列 Primer sequence	引物用途 Role of primer
4CLdf	5'-TTAAGGGCATTTCCTA(C/T)GA(G/T)(C/T)AT(G/T)T(G/C)GA-3'	保守片段克隆
4CLdr	5'-TAGCCGAATCACTCG(C/T)CAT(C/T)(A/C)T(C/T)GA(A/G)TA-3'	Cloning of gene conserved fragment
3-gsp	5'-GACCCAGAGGGCGACGAAAAATACAA-3'	第一轮 3'RACE The 1st round 3'RACE
3-ngsp1	5'-CGGCGACATTGGATTAGTAGATGAT-3'	第二轮 3'RACE The 2nd round 3'RACE
3-ngsp2	5'-GGTGTCTACAAGAGAATCGGTCGA-3'	第三轮 3'RACE The 3rd round 3'RACE
5-gsp	5'-GCACCGGACATGATAAGGCGAATAGAAG-3'	第一轮 5'RACE The 1st round 5'RACE
5-ngsp1	5'-TTAGGCATGATGAGGATCGCGGAACCAAC-3'	第二轮 5'RACE The 2nd round 5'RACE
5-ngsp2	5'-GCGCAAAGCAAAAACCGAGTTGAGAGAGTAG-3'	第三轮 5'RACE The 3rd round 5'RACE
4CLf	5'- CCCACCAAAGTAATATCAATTA ACT -3;	<i>Ft4CL</i> cDNA 全长扩增
4CLr	5'- TTAGCTCCTTCTTAAGGATT -3'	<i>Ft4CL</i> cDNA full-length amplification

5'RACE 测序结果,设计一对特异引物(表 1),以 cDNA 为模板进行苦荞 4-香豆酸辅酶 A 连接酶 ORF 序列的扩增。

1.3.3 苦荞 *Ft4CL* 基因的生物信息学分析 使用 NCBI Blast 进行 *Ft4CL* 核苷酸序列及其编码的蛋白质氨基酸序列同源性分析;利用 DNAMAN 对所得 4CL 基因编码的氨基酸序列进行多序列比对;用 SOPMA(<http://pbil.ibcp.fr>)对 *Ft4CL* 编码的氨基酸二级结构进行预测;利用 MEGA5.0 软件采用邻接法(neighbor-joining, NJ)构建系统进化树。

2 结果与分析

2.1 苦荞 *Ft4CL* cDNA 的克隆

以 cDNA 为模板,使用试剂盒引物 3-gsp 和 2 条特异引物 3-ngsp1、3-ngsp2,经 3'RACE PCR 扩增后,电泳检测得到 1 条约 600 bp 的特异条带(图 2)。测序结果表明,该片段长度为 534 bp,包含终止密码 TAA、3'-UTR 和 Poly A 结构。使用试剂盒引物 5-gsp 2 条特异引物 5-ngsp1、5-ngsp2,经 5'RACE PCR 扩增后,电泳检测得到 1 条约 800 bp 的特异条带。测序结果表明,该片段长度为 852 bp,包含起始密码 ATG 和 5'-UTR(图 3)。

使用 1 对特异引物 4CLf 和 4CLr,以苦荞花期 cDNA 为模板,扩增后得到约 1 500 bp 的特异条带(图 4)。该片段长度为 1 602 bp,包含 1 个由 ATG 至 TAA 的完整 ORF,与甜荞 4CL 基因片段(HM149785)同源性为 94%,与蓖麻 4CL 基因片段(*Ricinus communis*, XM_002533140)同源性为 75%,表明已成功克隆苦荞 *Ft4CL* 基因 ORF 序列。

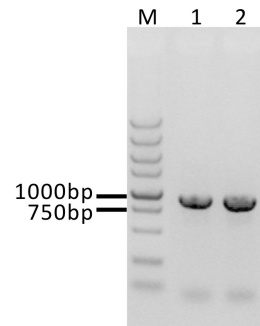


图 1 苦荞 *Ft4CL* 基因保守片段扩增

1,2. 保守片段; M. DNA 标记 D2000。

Fig. 1 Amplification of *Ft4CL* conservative fragment from tartary buckwheat

1,2. Conservative fragment;

M. DNA Marker D2000.

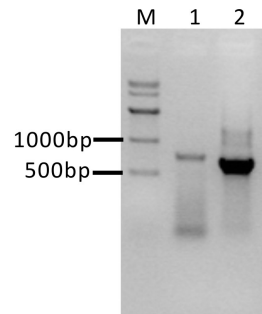


图 2 苦荞 *Ft4CL* 的 3'RACE M. DNA 标记 IV;

1. 第一轮扩增结果; 2. 第二轮扩增结果。下同。

Fig. 2 3'RACE of *Ft4CL* gene M. DNA Marker IV;

1. The first round of PCR amplification; 2. The second round of PCR amplification. The same below.

2.2 苦荞 *Ft4CL* 基因的生物信息学分析

2.2.1 苦荞 *Ft4CL* 基因核苷酸序列分析 核苷酸序列分析表明,该基因的 cDNA 序列编码区的长度为 1 602 bp。通过 GENSCAN 分析 *Ft4CL* 的 cDNA

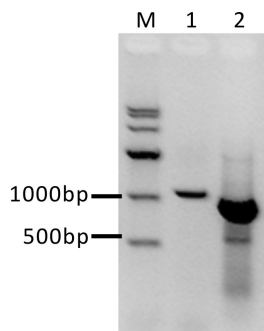


图 3 苦荞 *Ft4CL* 的 5'RACE
Fig. 3 5'RACE of *Ft4CL* gene

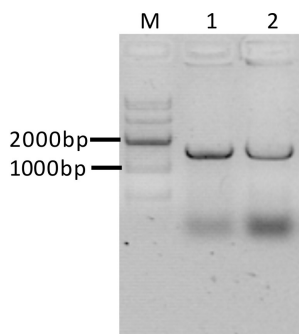


图 4 苦荞 *Ft4CL* 基因全长扩增 1, 2. *Ft4CL* cDNA 全长。
Fig. 4 *Ft4CL* ORF amplification of *F. tataricum*
1, 2. *Ft4CL* cDNA full-length

序列编码区可知,该基因编码 533 个氨基酸。Blast 结果显示,苦荞 4CL 基因全长 cDNA 与甜荞 4CL 基因 (*Fagopyrum esculentum*, HM149785) 的相似性最高达到 94%,与毛果杨 (*Populus trichocarpa*, XM_002329613) 和拟南芥 (*Ipomoea batatas*, AB469557) 的相似性仅为 72% 和 70%。

2.2.2 苦荞 *Ft4CL* 二级结构的分析及氨基酸序列的比对 DNAMAN 分析表明,该基因编码的蛋白质理论等电点 pI 为 5.23,理论相对分子质量为 58.02 kDa。SOPMA 预测表明,*Ft4CL* 二级结构富含无规则卷曲和 α -螺旋为,其中无规则卷曲 42.03%, α -螺旋 31.14%, β -折叠 20.26%, β -转角 6.57。采用 SignalP 4.0 对 *Ft4CL* 蛋白进行亚细胞定位预测,没有发现分泌途径信号肽 (SP,0.303)、叶绿体转运肽 (cTP,0.056) 以及线粒体靶肽 (mTP,0.058),表明 *Ft4CL* 蛋白可能定位于细胞质。选择模式植物拟南芥、烟草和葡萄的 4CL 氨基酸序列与 *Ft4CL* 进行多重序列比对 (图 5)。结果显示,4CL 氨基酸序列 N 端保守性较低,而 C 端保守性较高;N 端及 C 端各具有一个保守的肽基序 (BOX I 和 BOX II)。

2.2.3 4CL 系统进化树分析 经 NCBI Protein Blast 同源序列比对,苦荞 4CL 氨基酸序列与其它植物 4CL 氨基酸序列相似性较低,一般在 66%~75% 之间。其中与咖啡树 (*Coffea Arabica*, AFP49810.1) 和葡萄 (*Vitis vinifera*, CAN69130.1) 相似性为 74%,而与拟南芥 (*Arabidopsis thaliana*, AAD47193.1) 相似性仅为 66%。从 GenBank 数据库中下载拟南芥、烟草等模式植物的 4CL 氨基酸序列,利用 MEGA 5.0 软件,采用邻接法构建 4CL 的系统发育树,结果见图 6。由图 6 可知,系统发育树可明显的分为 2 个类群 (Cluster I 和 Cluster II)。苦荞 *Ft4CL* 与拟南芥 *At4CL1*、*At4CL2* 等同聚于 Cluster I,而 *At4CL3*、*At4CL4* 和黑麦草 (*Lolium perenne*) *Lp4CL* 等同聚于 Cluster II。

3 讨论与结论

苯丙烷类代谢途径是植物重要的次生代谢途径之一,4CL 的催化活性的高低关系到其下游终产物如黄酮和木质素等的生物合成量。此外,尽管 4CL 与腺苷酸化酶的氨基酸同源性不好,但具有类似的保守基序和反应机制,从而将其归类为腺苷酸化酶超家族的一员。本研究克隆得到的苦荞 *Ft4CL* 基因编码的蛋白质具有植物 4CL 的典型特征,氨基酸序列具有两个保守肽基序,其中位于 N 端的肽基序 BOX I (SSGTTGLPKG V),在 4CL 蛋白质序列中几乎绝对保守,通常认为是 4CL 中与 AMP 结合的功能域,而且与荧光素酶、长链脂肪酰基-CoA 合成酶、肽合成酶中发现的保守基序相似。位于 C 端肽基序 Box II (GEICIRG),在所有 4CL 中绝对保守,其中 *Gly*、*Glu* 和 *Cys* 是最为保守的氨基酸位点,被认为直接参与催化过程。

根据植物中各种 4CL 同工酶遗传距离和代谢功能的差异,一般将植物中的 4CL 分为两大类: I 类同工酶参与调节催化黄酮的生物合成; II 类同工酶参与调节催化木质素的生物合成,并且与可溶的或者与细胞壁结合的苯基丙酸的衍生物在结构组成上有关。Angela *et al.* (2002) 研究表明,位于 Cluster I 的拟南芥 *At4CL1*、*At4CL2* 归于黄酮支路而位于 Cluster II 的 *At4CL3* 归于木质素支路,Butmer *et al.* (1999) 研究表明,位于 Cluster II 的黑麦草 *Lp4CL* 归于木质素支路。故推测本研究克隆得到的 *Ft4CL* 基因与苦荞黄酮代谢相关,属于

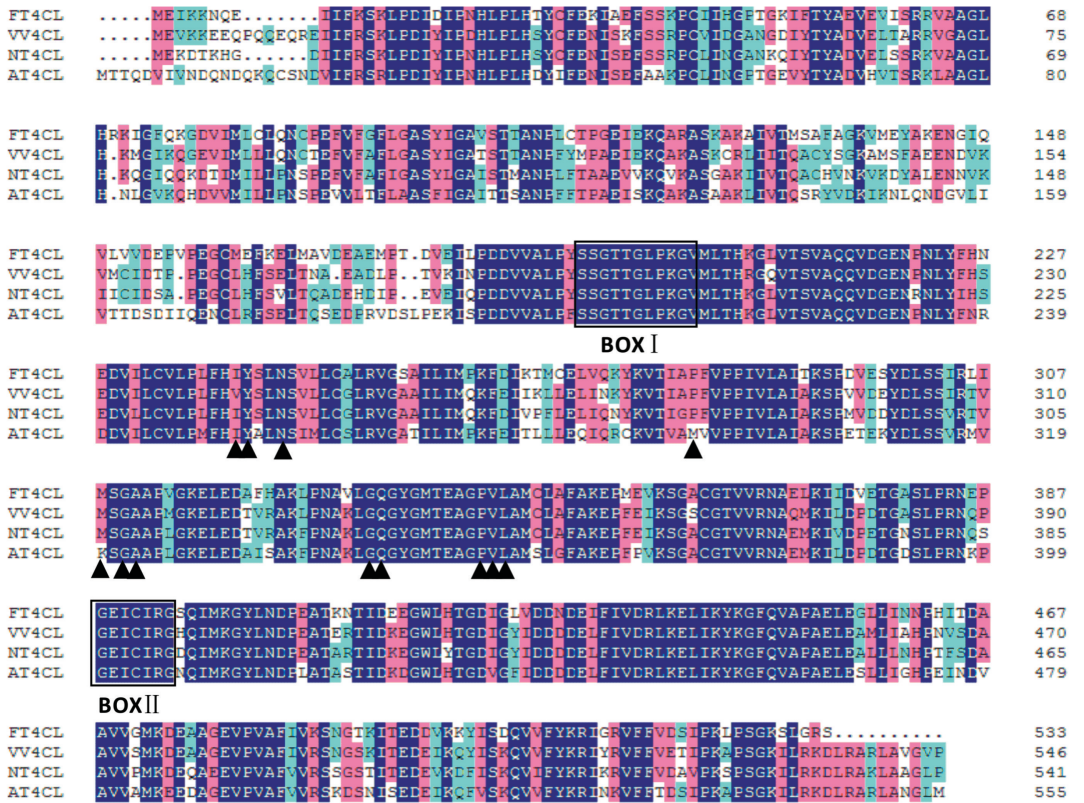


图 5 Ft4CL 编码蛋白与其他 4CL 蛋白编码的氨基酸序列的比对
 Fig. 5 Alignments of Ft4CL encoding protein with other 4CL proteins

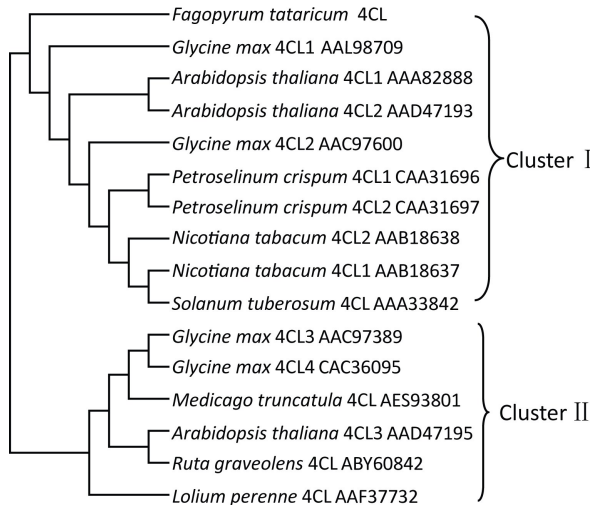


图 6 Ft4CL 与其他植物来源 4CL 蛋白序列的系统进化树构建

Fig. 6 Construction result of phylogenetic tree

4CL 的黄酮代谢支路。

黄酮类物质作为一种重要的植物次生代谢产物,在植株中的高水平积累不仅能提升植物品质,对植株本身而言,还能提高其抗逆能力,增强该植物的

环境适应能力。本研究首次成功克隆了苦荞 Ft4CL 基因 cDNA ORF 序列,为深入研究 Ft4CL 基因对苦荞类黄酮代谢的影响奠定基础,也为采用分子育种方法培育高黄酮苦荞新品种提供了候选靶基因。

参考文献:

Allina SM, Ellis BE, Theilmann DA, *et al.* 1998. 4-Coumarate: coenzyme A ligase in hybrid poplar properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes[J]. *Plant Physiol.* **116**(2): 743—754

Angela L, Robyn H, Russell M, *et al.* 2002. Isolation and characterisation of three 4-coumarate: CoA-ligase homologue cDNAs from perennial ryegrass (*Loium perenne*) [J]. *J Plant Physiol.* **159**(7): 773—779

Buttner D, Ehrling J, Wang Q, *et al.* 1999. Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms [J]. *Plant J.* **19**(1): 9—20

Buttner D, Ehrling J, Stuibler HP, *et al.* 2000. Mutational analysis of 4-coumarate: CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes [J]. *FEBS Lett.* **467**(1): 117—122

Cillard J, Cillard P, Torel J, *et al.* 1986. Antioxidant activity of flavonoids and reactivity with peroxy radical [J]. *Phytochemistry.* (下转第 767 页 Continue on page 767)