

罂粟与虞美人可待因-O-脱甲基酶基因克隆及序列分析

郝丹丹^{1,2}, 胡启跳², 袁媛^{2*}, 龚永福³, 陈芳³, 魏玉杰³, 黄璐琦²(1. 山西中医药大学中药学院, 山西 晋中 030619; 2. 中国中医科学院中药资源中心, 北京 100700;
3. 甘肃省农业工程技术研究院, 甘肃 武威 733000)

摘要: CODM 是可待因和吗啡生物合成过程中的关键酶, 本实验以 NCBI 数据库中的可待因-O-脱甲基酶基因 (codeine-O-demethylase, CODM) 序列为参考, 克隆得到了罂粟和虞美人的 CODM 核苷酸序列, 并对二者进行序列比对及生物信息学分析。结果发现, 罂粟 CODM 有 3 种基因型, 而虞美人 CODM 有 5 种基因型; 生物信息学分析显示所有 CODM 蛋白无信号肽序列, 预测这些蛋白都为非分泌性蛋白, 属于 Pcbc superfamily 超基因家族。罂粟 CODM 氨基酸序列相同, 但在不同罂粟资源中的表达量有明显差异, 推测 CODM 基因转录水平变异可能与其非编码区序列有关, 这为进一步开展罂粟生物碱类成分的合成与调控机制研究奠定基础。

关键词: 罂粟; 虞美人; 可待因-O-脱甲基酶; 生物信息学分析

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2019)07-1312-05

Cloning and sequence analysis of codeine-O-demethylase gene from *Papaver somniferum* and *Papaver rhoeas*HAO Dan-dan^{1,2}, HU Qi-tiao², YUAN Yuan^{2*}, GONG Yong-fu³, CHEN Fang³,
WEI Yu-jie³, HUANG Lu-qi²(1. School of Chinese Materia Medica, Shanxi University of Chinese Medicine, Jinzhong 030619, China;
2. National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China; 3. Gansu Institute of Agricultural Engineering Technology, Wuwei 733000, China)

Abstract: Codeine-O-demethylase (CODM) is a key enzyme in the biosynthesis of codeine and morphine. In this study, CODM gene sequences were cloned from *Papaver somniferum* and *Papaver rhoeas*, and were compared with each other by sequence alignment and bioinformatics analysis. The results showed that there were three genotypes of CODM in *Papaver somniferum* and five genotypes of CODM in *Papaver rhoeas*. Bioinformatics analysis showed that all CODM proteins had no signal peptide sequence, and these proteins were predicted to be non-secretory proteins, belonging to the Pcbc supergene family. Although the amino acid sequences of CODM in poppies are the same, the expression levels of CODM in different poppy resources are significantly different. It is speculated that the variation of transcription level of CODM may be related to its non-coding region sequence, which lays a foundation for further research on the synthesis and regulation mechanism of alkaloids in poppies.

Key words: *Papaver somniferum*; *Papaver rhoeas*; codeine-O-demethylase; bioinformatics analysis

收稿日期: 2019-03-11; 修回日期: 2019-04-13.

基金项目: 中央级公益性科研院所基本科研业务费专项资金资助 (ZZ10-008); 中央本部级重大支减项目: 名贵中药资源可持续利用能力建设项目 (2060302); 成都中医药大学西南特色中药资源重点实验室开放研究基金资助项目 (003109034001).

*通讯作者 Tel: 86-10-64087649, E-mail: y_yuan0732@163.com

DOI: 10.16438/j.0513-4870.2019-0163

罂粟 (*Papaver somniferum* L.) 与虞美人 (*Papaver rhoeas* L.) 均是罂粟科罂粟属植物, 为一年生草本植物。罂粟是一种重要的药源植物, 含有 100 多种生物碱, 已分离的生物碱有 20 多种, 是麻醉性镇痛药物吗啡、可待因和半合成类似物 (如羟考酮、氢可酮和纳曲酮) 的主要原料^[1]。其同属近源种虞美人中也含有多

种生物碱,如阿朴啡、丽春花碱、原阿片碱和海罂粟胺等,在民间被用来治疗咳嗽、腹泻等症,但研究表明虞美人中不含有吗啡和可待因生物碱^[2-4]。

*CODM*基因是苜基异喹啉类生物碱可待因、吗啡生物合成过程中的关键酶基因。研究表明,*CODM*基因可竞争性利用可待因生物合成中的上游底物蒂巴因合成吗啡,同时又可促进可待因向吗啡的转化^[5]。如Wijekoo等^[6]发现降低*CODM*水平则引起吗啡含量的下降,而可待因成分积累。Hagel等^[5]发现不同罂粟资源中*CODM*转录水平的差别可导致吗啡、可待因含量的差异。

本文以3个不同罂粟资源及不含吗啡、可待因成分的虞美人为研究对象,克隆其*CODM*基因,并对二者的序列特征进行比较,以期为进一步开展罂粟生物碱类成分的合成与调控机制研究奠定基础。

材料与方法

植物材料和试剂 罂粟和虞美人叶片均由甘肃省农业工程技术研究院提供,叶片样品自然干燥后放于密封袋中,室温保存。

Plant Genomic DNA Kit (TIANGEN公司)、pEASY-Blunt Zero Cloning Kit (Trans公司)、Gel Extraction Kit (OMEGA公司)、15k DNA Marker (TaKaRa公司)、5k DNA Marker (TaKaRa公司)、Fast pfu Fly聚合酶 (Trans公司)、2×T5 Super PCR Mix (北京擎科新业生物技术有限公司)、Trans1-T1感受态细胞 (Trans公司),引物合成及测序服务由北京睿博兴科生物技术有限公司完成。

仪器 Veriti 96孔梯度PCR扩增仪 (Applied Biosystem公司)、VORTEX-2 GENIE漩涡震荡仪 (Scientific industries公司)、SYNGENE凝胶成像系统 (GENE公司)、Nanodrop 2000微量核酸定量分析仪 (Thermo Fisher公司)、电热恒温培养箱 (上海一恒科学仪器有限公司)。

基因组DNA提取 所有样品各取20 mg,放入2 mL EP管中,液氮研磨。使用植物DNA提取试剂盒,依据操作说明提取,DNA溶解于60 μL灭菌蒸馏水中,使用核酸蛋白仪测定DNA浓度和纯度后,于-20 °C保存备用。

引物设计与PCR扩增 以NCBI公布的罂粟全基因组 (BioProject Accession: PRJNA492326) 为参考,找到*CODM*基因 (Gene ID: 113311621) 及其上下游各1 000 bp的序列,利用Primer Premier 5软件分别设计*CODM*基因的外侧全长扩增引物 (2001/2002) 及内侧测通引物 (2003/2004) 并合成 (表1)。使用引物2001/2002对不同样品DNA进行扩增,50 μL PCR反应体系

包括5×Buffer 10 μL、dNTPs 6 μL、上游及下游引物各1.5 μL、Fast pfu Fly DNA聚合酶1 μL、DNA模板2 μL,加灭菌蒸馏水补足至50 μL。反应程序为95 °C预变性2 min; 95 °C变性20 s, 54 °C退火20 s, 72 °C延伸2 min,共38个循环; 72 °C终延伸5 min。扩增产物用溴化乙锭 (EB) 染色的1.0%琼脂糖凝胶电泳检测,SYNGENE凝胶成像系统成像、观察。

Table 1 Primer sequences

Primer name	Description	Primer sequence (5'→3')
2001	<i>CODM</i> out.F1	AAAACCTGGTAGGGTAAGC
2002	<i>CODM</i> out.R1	GAATAGAAGCTGCGAGAAG
2003	<i>CODM</i> in.F2	TGAATCGGAGGACCAAAG
2004	<i>CODM</i> in.R2	GACCCAAACAGAGTCTTACA

***CODM*基因的克隆、测序** 使用琼脂糖凝胶回收试剂盒回收PCR扩增片段,切胶回收片段与克隆载体pEASY-Blunt Zero 进行连接,并转化至Trans1-T1感受态细胞中,用含有氨苄抗性的LB平板在37 °C下过夜培养,通过菌落PCR鉴定阳性转化子。PCR检测体系包括2×T5 Super PCR Mix 12.5 μL、引物M13F和M13R各1 μL、菌液1 μL,加灭菌蒸馏水补足至25 μL。反应程序为98 °C预变性10 min; 98 °C变性10 s, 54 °C退火10 s, 72 °C延伸15 s,共30个循环; 72 °C终延伸2 min。扩增产物用溴化乙锭染色的1.0%琼脂糖凝胶电泳检测,SYNGENE凝胶成像系统成像、观察。通过鉴定的阳性转化子送至北京睿博兴科生物技术有限公司进行测序,测序引物为通用引物M13F、M13R及引物2003、2004。

***CODM*蛋白的生物信息学分析** 利用ExPASy ProtParam tool (<http://web.expasy.org/protparam/>) 对*CODM*的理化性质进行预测; ProtScale (<http://web.expasy.org/protscale/>) 对其亲水性和疏水性进行分析; 使用CFSSP (<http://swissmodel.expasy.org/>) 进行蛋白质二级结构分析; Signal (<http://www.cbs.dtu.dk/services/SignalP/>) 分析*CODM*蛋白信号肽; TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) 对*CODM*蛋白进行亚细胞定位; TMHMM Server (<http://www.cbs.dtu.dk/services/TMHMM-1.0/tmhmm1.0.html>) 预测跨膜结构域; 使用SWISS-MODEL (<https://swissmodel.expasy.org/>) 基于同源建模法对该蛋白的三级结构进行预测; 通过MEGA 7.0构建Neighbor-joining系统进化树^[7]。

结果与分析

1 *CODM*基因的克隆

使用试剂盒提取的罂粟及虞美人样品基因组

DNA 的浓度均在 110 mg·L⁻¹ 以上, A_{260/280} 值均在 1.8~1.9 之间, 表明其 DNA 质量较好。用 1.0% 的琼脂糖凝胶电泳检查 DNA 质量, 所提取的 DNA 条带清晰, 符合 PCR 扩增反应的要求。以虞美人和罂粟 DNA 为模板进行 PCR 扩增, 产物经 1.0% 琼脂糖凝胶电泳检测, 在 2 500 bp 附近有明亮单一的条带, 与所设计引物对应的 CODM 基因片段大小一致。

基因测序结果表明, 3 个不同罂粟资源中 CODM 均有 3 种不同基因型 (分别标注为 PsCODM.1、PsCODM.2 和 PsCODM.3), 虞美人样品 CODM 基因有 5 种不同基因型 (分别标注为 PrCODM.1、PrCODM.2、PrCODM.3、PrCODM.4 和 PrCODM.5)。

2 CODM 基因序列分析

通过 DNAMAN 软件对罂粟和虞美人 CODM 基因进行多序列比对, 发现其核苷酸序列的一致性达到 99.89%。BLAST 分析发现, 所有 CODM 基因核苷酸序列与 MG995581.1 的同源性为 100%, 与 KY688198.1 的同源性为 87%; 与 XM_026560449.1 的同源性为 75%; 与 GQ500141.1 的同源性为 73%。

使用 DNAMAN 软件将 CODM 核苷酸序列翻译为氨基酸序列, 并进行多序列比对, 结果发现存在 3 种氨基酸序列, 即 PsCODM.1、PsCODM.2、PsCODM.3、PrCODM.1、PrCODM.2 与 PrCODM.3 的氨基酸序列一致 (共同标注为 CODM.n), 均编码 360 个氨基酸; PrCODM.4 编码 360 个氨基酸; PrCODM.5 编码 373 个氨基酸, 3 种氨基酸序列的一致性达到了 96.46%。其中, CODM.n 氨基酸序列与已公布的罂粟全基因组序列一致; PrCODM.4 的氨基酸序列由于 1283 位点碱基 C 突变为 A, 导致密码子由 Asp (天冬氨酸) 变成了 Glu (谷氨酸); PrCODM.5 氨基酸序列由于 1665 位点由碱基 G 变成了 T, 导致终止密码子变为了 Leu (亮氨酸), 增加了 12 个氨基酸 (图 1)。

3 CODM 编码蛋白结构的预测与分析

使用 ExPASy 软件对罂粟和虞美人 CODM 进行理化性质预测, 结果显示 CODM.n、CODM.4 及 CODM.5 编码蛋白的分子质量分别为 40.85 kD、40.91 kD 和 42.37 kD, 理论 PI 值分别为 5.04、5.00 和 5.05。CODM.n 和 PrCODM.5 编码的蛋白的负电荷残基总数 (Asp+Glu) 和正电荷残基总数 (Arg+Lys) 均相同, 分别为 53 和 38 个, 而 PrCODM.4 为 54 和 38 个。不稳定指数均大于 40, 属于不稳定蛋白。使用 ProtScale 软件进一步分析表明, 这些氨基酸序列中疏水最大值均为 1.900, 最小值为 -2.689。利用 SOPMA 软件预测罂粟和虞美人 CODM 二级结构, 结果显示其蛋白由无规则卷曲、α-螺旋、延伸链和 β-折叠构成, 其中比例最高的是无规则卷曲, β-折叠最少 (图 2)。

亚细胞定位预测结果表明, 虞美人和罂粟 CODM 存在于线粒体的可能性为 0.043, 细胞质基质为 0.169, 其他细胞器为 0.842, 故推定该蛋白存在于其他细胞器中。利用 Signal 和 TMHMM Server 软件分别预测该蛋白的信号肽和跨膜结构, 得出这些蛋白均不含信号肽序列, 属于非分泌蛋白。保守结构域分析结果显示, 3 种 CODM 蛋白均具有 Pcbc 结构域, 由 Pcbc superfamily 超基因家族组成。

利用 SWISS-MODEL 在线软件对虞美人和罂粟的蛋白序列进行三级结构预测, 结果显示, CODM.n 蛋白与 SWISS-MODEL 数据库目标蛋白 5o7y.1.A (T6ODM) 的序列相似度为 73.61%, QMEAN 值是 -2.23; PrCODM.4 蛋白 SWISS-MODEL 数据库目标蛋白 5o7y.1.A 的序列相似度为 73.33%, QMEAN 值是 -2.25; PrCODM.5 蛋白 SWISS-MODEL 数据库目标蛋白 5o7y.1.A 的序列相似度为 73.61%, QMEAN 值是 -2.45 (图 3)。

4 CODM 基因的系统进化分析

以 CODM.n 氨基酸序列作为参考, 利用 Blast 软件

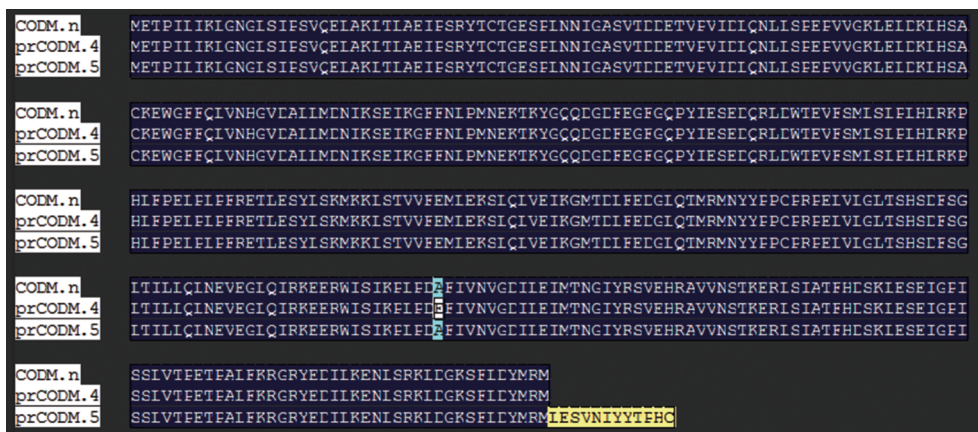


Figure 1 Clustal W analyses of CODM proteins from *Papaver somniferum* and *Papaver rhoeas*

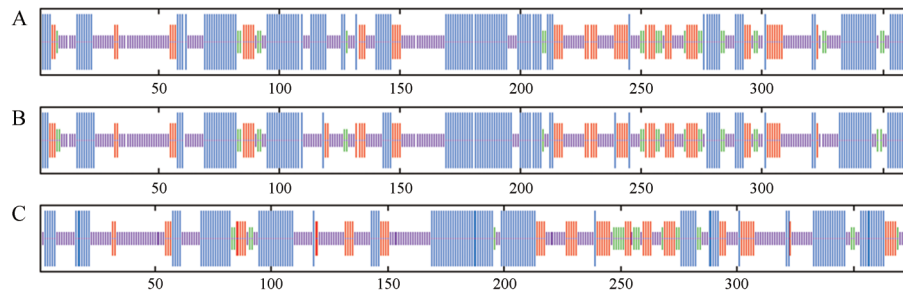


Figure 2 Predicts secondary structure of *CODM* from the amino acid sequence. A: *CODM.n*; B: *PrCODM.4*; C: *PrCODM.5*. -Helix; -Turn; -Sheet; -Coiled-coil

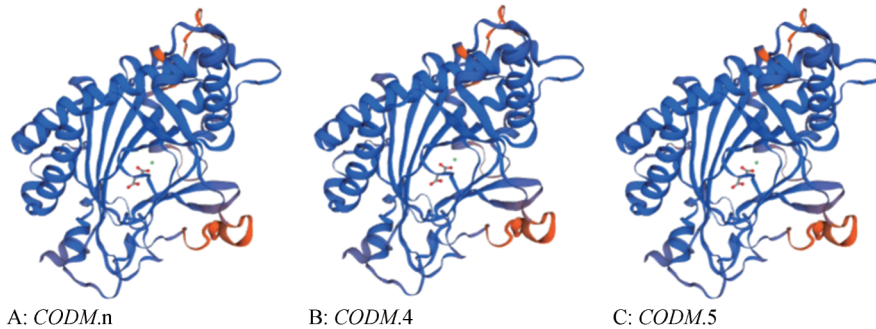


Figure 3 Predicts tertiary structure of *CODM* from the amino acid sequence

搜索与 *CODM* 蛋白同源序列, 并利用 MEGA 7.0 进行系统发育分析, 结果表明所有氨基酸序列大致分为 2 大类, 第 1 类包括罂粟 (XP_026416223.1)、博落回 (OVA18943.1) 和克隆得到的所有序列; 第 2 类是小麦 (EMS50518.1) 和拟南芥 (XP_020869163.1) (图 4)。

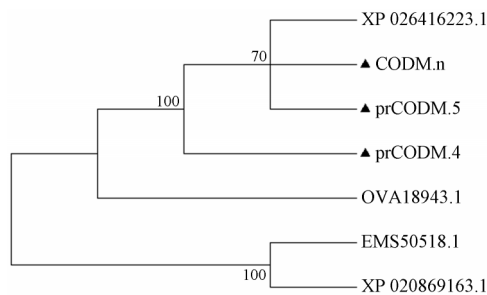


Figure 4 Phylogenetic analysis of *CODM* proteins from plants using NJ method

5 *CODM* 基因表达分析

本课题组保存了不同罂粟资源的转录组数据, 对 *CODM* 基因的转录水平进行分析, 结果发现不同罂粟资源中 *CODM* 基因的表达量存在明显差异 (表 2), 且其表达量与吗啡和可待因生物碱含量呈正相关关系, 暗示其在吗啡和可待因生物合成中发挥重要作用。

讨论

罂粟作为一种重要的药源植物, 在医疗上的应用

Table 2 Transcriptome data of *CODM* genes from different poppy resources

Sample number	Description	FPKM
1	Contains morphine, codeine	0.215 0
2	High levels of morphine and codeine	0.484 1
3	Trace amounts of morphine and codeine	0

历史悠久且有着不可替代的作用, 尤其是镇痛作用, 目前仍是麻醉性镇痛药物吗啡、可待因和半合成类似物的唯一商业来源^[8]。罂粟壳中可待因含量仅占总生物碱的 0~0.70%, 药典收录的罂粟壳中吗啡含量为 0.06%~0.40%, 这在国内远远不能满足医药界的需求。因此, 多年来研究人员在阐明苄基异喹啉类生物碱 (BIA) 代谢的吗啡分支方面做了大量的努力。*CODM* 作为苄基异喹啉类生物碱吗啡和可待因合成过程中的关键酶基因, 它可利用底物蒂巴因合成东罂粟碱再到吗啡, 也可竞争性利用可待因合成吗啡^[5], 目前对植物 *CODM* 基因的研究较少。本实验基于罂粟全基因组数据库中的 *CODM* 序列, 从罂粟和虞美人中克隆得到了 *CODM*, 结果发现罂粟中 *CODM* 有 3 个基因型, 虞美人中的 *CODM* 有 5 个基因型, Guo 等^[9]通过基因组测序表明罂粟 *CODM* 基因有 3 个拷贝, 与本实验结果相符; 而目前尚无公布虞美人的基因组序列, 通过本实验结果, 推测虞美人中 *CODM* 基因也存在多个拷贝。多序列比对发现, 这些 *CODM* 核苷酸序列的一致性高达 99.88%, 氨基酸序列的一致性达到 96.46%, 表

明该基因在罂粟科植物的不同物种间高度保守。系统发育分析表明,来自罂粟及虞美人的 *CODM.n* 与已公布的罂粟全基因组中的 *CODM* 序列聚为一类,而来自虞美人的 *PrCODM.4*、*PrCODM.5* 基因与博落回聚为一类。表达谱分析结果发现, *CODM* 基因在不同罂粟资源中的转录水平不同,且在高吗啡和可待因含量的罂粟资源中表达量最高,暗示 *CODM* 可能并不是影响罂粟中吗啡和可待因生物碱含量的唯一因素, *CODM* 基因转录水平变异可能与其非编码区序列有关,从而影响罂粟和虞美人中生物碱成分的积累。本研究结果不仅为进一步开展罂粟生物碱类成分的合成与调控机制研究奠定了基础,同时也对创新可待因专用型品种具有参考价值。

References

- [1] Hagel JM, Facchini PJ. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world [J]. *Plant Cell Physiol*, 2013, 54: 647-672.
- [2] Choe S, Kim S, Lee C, et al. Species identification of *Papaver* by metabolite profiling [J]. *Forensic Sci Int*, 2011, 211: 51-60.
- [3] Coban I, Toplan GG, Özbek B, et al. Variation of alkaloid contents and antimicrobial activities of *Papaver rhoeas* L. growing in Turkey and northern Cyprus [J]. *Pharm Biol*, 2017, 55: 1894-1898.
- [4] Oh JH, Ha IJ, Lee MY, et al. Identification and metabolite profiling of alkaloids in aerial parts of *Papaver rhoeas* by liquid chromatography coupled with quadrupole time-of-flight tandem mass spectrometry [J]. *J Sep Sci*, 2018, 41: 2517-2527.
- [5] Hagel JM, Facchini PJ. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy [J]. *Nat Chem Biol*, 2010, 6: 273-275.
- [6] Wijekoon CP, Facchini PJ. Systematic knockdown of morphine pathway enzymes in opium poppy using virus-induced gene silencing [J]. *Plant J*, 2012, 69: 1052-1063.
- [7] Chen TY, Liu J, Yuan Y, et al. Bioinformatics analysis and expressed level of Auxin response factor genes in *Scutellaria baicalensis* [J]. *Acta Pharm Sin (药学报)*, 2017, 52: 1770-1776.
- [8] Facchini PJ, Park SU. Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy [J]. *Phytochemistry*, 2003, 64: 177-186.
- [9] Guo L, Winzer T, Yang X, et al. The opium poppy genome and morphinan production [J]. *Science*, 2018, 362: 343-347.