

基于转录组信息的甘松 CONSTANS 转录因子家族分析

刘 芳, 王远航, 陈 红, 黄雄杰, 白怡婷, 崔 琪, 雷语妍, 李 莹, 李文兵, 张绍山, 陈 晨*, 刘 圆*
西南民族大学, 四川省羌彝药用资源保护与利用技术工程实验室, 青藏高原民族药用资源保护与利用国家民委重点实验室, 四川 成都 610225

摘要: **目的** 基于转录组数据筛选鉴定分析甘松 *Nardostachys jatamansi* CONSTANS (CO) 转录因子家族。**方法** 运用生物信息学技术对甘松植物的 CO 转录因子家族成员展开综合分析, 深入探究其蛋白的理化特性、基因序列结构、亚细胞位置、导肽和信号肽的预测, 以及跨膜结构域的识别。此外, 通过构建系统进化树, 进一步揭示这些基因家族成员间的进化关系。**结果** 在甘松的数据库中经过筛选、鉴定, 最终得到 10 个 CO 转录因子, 氨基酸长度为 316~440 aa, 相对分子质量在 35 561.67~48 171.96, 理论等电点范围为 5.11~8.63, 主要在细胞核中表达, 均含有 B-box 和 CCT 结构域, 均不含跨膜结构域和信号肽。构建系统进化树明确了各亚族的成员分布: 第I亚族包含 2 个成员, 第II亚族增至 3 个成员, 而第III亚族则有 5 个成员, 这一发现为进一步探讨各亚族间的进化关系及特性差异提供了重要基础。**结论** 利用了生物信息学的手段, 全面鉴定并分析了甘松 CO 转录因子家族的理化性质、蛋白结构、系统发育关系等, 对甘松 CO 转录因子家族功能的详尽研究, 为后续的深入探索提供了宝贵的参考。此外, 也为理解甘松药用成分合成的调控机制提供了基础, 并为未来培育出更高品质的甘松品种提供了科学依据。

关键词: 甘松; CONSTANS 转录因子家族; 光周期; 系统进化树; 生物信息学分析

中图分类号: R286.2 文献标志码: A 文章编号: 0253-2670(2025)04-1346-09

DOI: 10.7501/j.issn.0253-2670.2025.04.022

CONSTANS transcription factor family in *Nardostachys jatamansi* based on transcriptome information

LIU Fang, WANG Yuanhang, CHEN Hong, HUANG Xiongjie, BAI Yiting, CUI QI, LEI Yuyan, LI Ying, LI Wenbing, ZHANG Shaoshan, CHEN Chen, LIU Yuan

Tibetan Plateau Ethnic Medicinal Resources Protection and Utilization Key Laboratory of National Ethnic Affairs Commission of the People's Republic of China, Sichuan Provincial Qiang-Yi Medicinal Resources Protection and Utilization Technology Engineering Laboratory, Southwest Minzu University, Chengdu 610225, China

Abstract: Objective According to the transcriptome data, the CONSTANS (CO) transcription factor family members were identified and analyzed in *Nardostachys jatamansi*. **Methods** The bioinformatic technology was used to analyze the CO transcription factor family members of *N. jatamansi* plants, including the physicochemical properties of proteins, gene sequence architecture analysis, subcellular location, peptides and signaling peptides prediction, and identification of transmembrane domains. The phylogenetic tree was constructed to reveal the evolutionary relationship between them. **Results** A total of 10 members of the CO transcription factor family were screened and identified based on the *N. jatamansi* database. The length of amino acids was 316—440 aa, the relative molecular weight was 35 561.67—48 171.96, and the theoretical range of isoelectric point was 5.11—8.63. They are mainly expressed in the nucleus, contain B-box and CCT domains, and don't contain transmembrane domains and signal peptides. The phylogenetic tree was constructed to clarify the distribution of members of each subfamily: Subfamily I contains two members, subfamily II has three members, and subfamily III has five members. This discovery provides an important basis for further exploration of the evolutionary relationship and characteristic differences among subfamilies. **Conclusion** The physicochemical properties, protein structure and

收稿日期: 2024-08-05

基金项目: 四川省自然科学基金资助项目 (2025ZNSFSC1124); 西南民族大学青藏高原研究首批科技创新团队 (2024CXTD04); 西南民族大学人才科研项目 (RQD2021110); 中央高校基本科研业务费专项基金项目 (ZYN2024189)

作者简介: 刘 芳 (2000—), 硕士研究生, 研究方向为民族药品种和新药资源保护与利用研究。E-mail: 18582908619@163.com

*通信作者: 刘 圆 (1968—), 博士, 教授, 从事民族药品种、品质评价和新药资源保护与利用研究。E-mail: 499769896@qq.com

陈 晨 (1987—), 博士, 助理研究员, 从事药用植物资源鉴定与利用方向。E-mail: DoubleChen_SWUN@yeah.net

phylogenetic relationships of CO family transcription factors were identified and analyzed by means of bioinformatics. The detailed study on the function of CO family transcription factors provided a valuable reference for further exploration. In addition, this study also provides a basis for understanding the regulatory mechanism of medicinal ingredient synthesis of *N. jatamansi*, and provides a scientific basis for higher-quality breeding in the future.

Key words: *Nardostachys jatamansi* (D. Don) DC.; CONSTANS family of transcription factors; photoperiod; phylogenetic tree; bioinformatics analysis

甘松 *Nardostachys jatamansi* (D. Don) DC. 是败酱科甘松属的多年生草地植物, 主要分布于海拔 3 000~5 000 m 的喜马拉雅高山至亚高山地区, 在我国以四川省、云南省、西藏自治区、甘肃省和青海省为主产区^[1]。《中国药典》2020 年版明确记载, 甘松经干燥处理后的根及根茎, 具有显著的药用价值^[2]。它们不仅能开郁醒脾、理气止痛, 外用时还能有效祛湿消肿、治疗脚气肿毒并缓解牙痛。此外, 甘松在治疗食欲不振、缓解腹胀满、抗心律失常^[3]、抗抑郁^[4]等方面亦展现卓越疗效。尤为值得一提的是, 甘松在尤纳尼医及古印度阿育吠陀医学体系中占据着举足轻重的地位, 被广泛地应用于医疗领域, 具有其深厚的历史底蕴与广泛的临床价值。在我国, 它不仅是藏、维、蒙、纳西、傈僳等传统民族医学中临床常用的药材之一^[5], 还是藏香的关键成分之一, 在藏香的制作中占据重要地位^[6]。

植物完成一个完整的生命周期需要经过营养生长和生殖生长 2 个阶段, 其中, 开花是植物从营养生长转变为生殖生长的重要标志。光周期是影响植物花芽分化的重要环境因子, 对植物的生长发育、环境的适应以及产量的形成都具有重要意义^[7-9]。CONSTANS (CO) 基因在植物的光周期调控中占据核心地位, 它是感知日照长度的关键元件。这一基因将光信号与生物钟信号有效整合, 通过将光信号转变为开花信号, 正调控下游开花相关基因 suppressor of overexp-ression of cp1 (SOC1) 和 flower time locus 1 (FT) 表达, 进而调控植物开花^[10]。

Putterill 等^[11]最早从模式植物拟南芥 *Arabidopsis thaliana* L. 中分离出 CO 基因, 并通过实时荧光定量 PCR (RT-PCR) 检测到 CO 基因在根和叶中均有表达。CO 基因通常以多拷贝的形式存在于植物基因组中, 且多数包含 2~4 个外显子。在进化过程中, 这些基因展现出了显著的多样性, 在 CO 基因的 B1 亚家族中, 大部分成员具有 2 个外显子和 1 个内含子的结构; 而 B2 亚家族则呈现出一种独特的模式, 每个 CO 基因均包含 4 个外显子和 3

个内含子, 且这些结构单元的排列相位遵循“0, 0, 2”的规律。相比之下, B3 亚家族的结构更为复杂多变。多数 B3 亚家族的 CO 基因含有 2 个外显子, 但也有部分基因具备 4 个外显子。值得注意的是, 如 *CrCO₃* 基因这样的特例, 拥有高达 14 个外显子和 13 个内含子, 这一特性在系统进化树中表现为 *CrCO₃* 独立分化成一个独特的进化枝^[12]。CO 蛋白的序列中, 其特有的串联 B-box 结构域在调控植物开花过程中起着关键作用, 通过与其他蛋白的相互作用来影响开花时间 T 的表达。Dahal 等^[13]公布了拟南芥 CO 蛋白中串联 B 结构域的晶体结构解析结果。在拟南芥中, CO 基因与另外 16 个 CO-like (COL) 基因共同构成了 CO 基因家族。该家族成员显著特征在于其包含 2 个保守结构域: 靠近 N 端的 B-box 结构域和接近 C 端的 CCT (CO、COL-like、TOC1) 结构域, 这些结构域对于该家族成员的功能发挥至关重要^[13]。

甘松的药用价值主要体现在其营养器官——根茎上, 其产量和品质受到开花时期和侧根发育情况的显著影响。这些生长阶段和形态变化对于甘松药材的整体性能具有决定性作用^[14]。开花是植物由营养生长向生殖生长转变的关键阶段, 这一过程与甘松的生育周期、授粉效率、种子发育的最优时期以及甘松产量的提升均密切相关^[14-16]。CO 转录因子家族在调控开花时间、花器官发育以及侧根生长中发挥着关键作用, 然而关于甘松中 CO 转录因子家族的研究尚属空白。鉴于此, 本课题组利用课题组前期积累的转录组数据, 对甘松 CO 转录因子家族进行了深入的生物信息学功能鉴定与分析, 这一工作不仅填补了甘松 CO 转录因子研究的空白, 也为进一步揭示甘松 CO 转录因子家族在花器官形成和侧根发育中的分子调控机制提供了坚实的理论基础。

1 材料

样品由西南民族大学草地资源学院刘圆教授鉴定为甘松 *N. jatamansi* (D. Don) DC.。植株现蕾后套袋, 收集同一株甘松的种子。

2 方法

2.1 样品的处理

将甘松种子均匀播种于 6 个花盆，于光照培养箱中，设置温度 20 ℃、相对湿度 50%、8 h/16 h 光照/黑暗的条件培养，直至幼苗长出 2 片真叶。选择 3 盆喷施定量的 50 mg/mL 吲哚乙酸溶液，设置为处理组；另外 3 盆喷洒等量体积的清水，设置为对照组。处理 24 h 后采集叶片样本，每份样本 3 个生物学重复。采样后立即液氮速冻，-80 ℃ 保存。最终，将处理后的样本递交给上海欧易生物医学科技有限公司进行转录组测序。

2.2 甘松 CO 转录因子家族的获取、筛选与热图分析

甘松 CO 转录因子家族序列来源于课题组上传于 NCBI 的 SRA 数据库中的转录组数据，检索号分别为 SRR21656214、SRR21656215、SRR21656216、SRR21656217、SRR21656218、SRR21656219、SRR21656220、SRR21656221、SRR21656222、SRR21656223、SRR21656224、SRR21656225、SRR21656226、SRR21656227、SRR21656228、SRR21656229、SRR21656230、SRR21656231。通过 Swiss 注释和 NR 数据库进行搜索，设定筛选的 E 值阈值为 1×10^{-5} ，从 NCBI 数据库中检索筛选转录因子的开放阅读框（open reading frame, ORF）。利用 CDD 和 Pfam 验证分析转录因子的保守结构域，剔除掉不完整的序列后，最终获得 CO 转录因子蛋白序列。对对照组和实验组的 CO 基因表达量进行分析，通过 TBtool 软件制作表达量热图。

2.3 甘松 CO 家族蛋白的二级结构、理化特性、三级结构以及保守基序（motif）的分析

通过 Expasy 平台（<https://web.expasy.org/>

protparam/），解析 CO 蛋白序列的相对分子质量和理论等电点等关键理化参数。利用 SOPMA 工具（https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html）对蛋白序列的二级结构进行预测。利用 MEME 软件（<https://meme-suite.org/doc/meme.html>）对甘松 CO 蛋白的保守基序进行分析。甘松 CO 蛋白的三维结构通过 SWISS MODEL 平台（<https://swissmodel.expasy.org/>）进行预测。

2.4 甘松 CO 家族蛋白导肽、亚细胞定位、跨膜结构、信号肽预测分析

通过 TargetP-2.0 平台（<https://services.healthtech.dtu.dk/service.php?TargetP-2.0>）预测 CO 家族的导肽。借助 Cello（<http://cello.life.nctu.edu.tw/>）对亚细胞定位进行预测分析。利用 SignalP-5.0（<https://services.healthtech.dtu.dk/services/SignalP-5.0/>）对信号肽进行预测。使用 TM-HMMServer2.0（<http://www.cbs.dtu.dk/services/TMHMM/>）预测其跨膜结构。

2.5 甘松与拟南芥 CO 家族蛋白系统进化树构建

系统进化树构建采用 MEGA7 软件的邻位连接法（neighbor-joining, NJ）算法，利用 iTOL 平台优化。

3 结果与分析

3.1 甘松 CO 家族蛋白的筛选与热图分析

甘松相关的研究数据主要源自课题组前期完成的转录组测序。本研究从 Swiss-Prot 和 NR 数据库中筛选出 24 个 CO 转录因子家族，利用 ORFfinder 工具分析其开放阅读框，通过 CDD 和 Pfam 平台进行结构域预测分析，剔除缺乏完整保守结构域或重复的序列，最终筛选出 10 个 CO 蛋白序列。为方便后续研究，按照转录组数据的编号顺序依次命名为 NcCONSTANS01~NcCONSTANS10（表 1）。10 个

表 1 甘松 CO 转录因子家族成员理化性质分析

Table 1 Analysis of physicochemical properties of members of CO transcription factor family in *N. jatamansi*

序列	基因 ID	氨基酸长度/aa	相对分子量	PI	不稳定指数	脂肪系数	平均亲水指数
NcCONSTANS01	TRINITY_DN10805_c0_g1_i1_6	440	48 171.96	5.11	51.65	66.70	-0.401
NcCONSTANS02	TRINITY_DN14460_c0_g1_i1_2	393	44 597.24	5.21	53.11	73.97	-0.616
NcCONSTANS03	TRINITY_DN10619_c0_g1_i1_2	316	35 561.67	5.63	40.79	61.36	-0.679
NcCONSTANS04	TRINITY_DN28045_c0_g1_i2_9	321	36 277.74	8.63	47.31	56.17	-0.835
NcCONSTANS05	TRINITY_DN29110_c0_g1_i3_8	413	46 143.25	5.89	55.56	54.58	-0.732
NcCONSTANS06	TRINITY_DN21661_c0_g1_i3_7	383	43 218.49	5.34	51.91	56.79	-0.760
NcCONSTANS07	TRINITY_DN26061_c1_g1_i2_1	349	39 431.80	5.92	51.47	63.64	-0.734
NcCONSTANS08	TRINITY_DN11498_c0_g1_i1_4	318	36 051.14	5.31	50.70	64.37	-0.719
NcCONSTANS09	TRINITY_DN31058_c0_g2_i2_5	380	42 173.93	7.96	54.04	61.55	-0.481
NcCONSTANS10	TRINITY_DN32319_c0_g1_i10_5	431	47 890.56	5.77	44.73	62.71	-0.670

CO 基因在对照组和实验组中的表达水平有所不同（图 1）。例如，基因 *NcCONSTANS03*、*NcCONSTANS07*、*NcCONSTANS08* 在对照组中的表达水平较高，而在实验组中的表达水平较低；相反，基因 *NcCONSTANS02*、*NcCONSTANS04*、*NcCONSTANS05*、*NcCONSTANS06*、*NcCONSTANS09* 在对照组中的表达水平较低，而在处理组中的表达水平较高。表明 CO 基因在植物的光周期调控中起重要作用，IAA 处理通过改变植物的光周期响应机制，间接影响 CO 基因的表达。

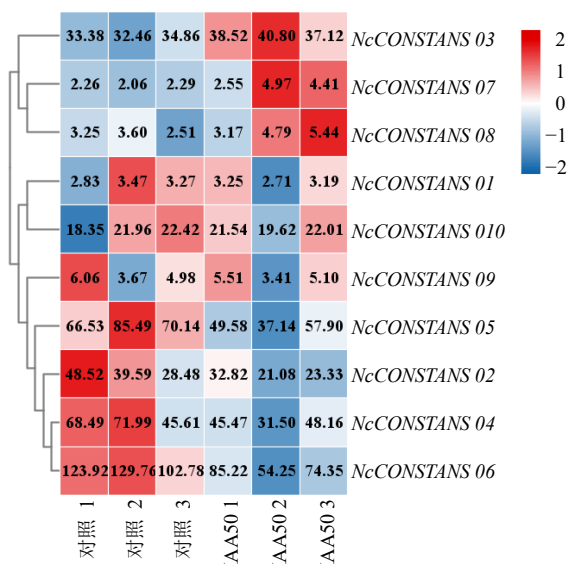


图 1 甘松 CO 基因家族热图分析

Fig. 1 Heat map analysis of CO gene family in *N. jatamansi*

3.2 甘松 CO 蛋白序列理化性质分析

利用 Protparam 分析 10 条甘松 CO 蛋白序列的理化性质，结果表明氨基酸长度介于 316~440 aa，相对分子质量在 35 561.67~48 171.96，理论等电点 (PI) 为 5.11~8.63，其中有 8 条蛋白为酸性 (PI<7)，另外 2 条为碱性 (PI>7)，脂肪系数为 40.79~55.56。甘松 CO 蛋白的平均疏水指数为 -0.835~-0.401 (均为负)，其中 80% 的蛋白 (8 条) 平均疏水指数 < -0.5，表明大多数具有亲水性质；另外 20% (2 条) 的蛋白表现出两性特性 (表 1)。

3.3 甘松 CO 蛋白亚细胞定位

利用 Cellov.2.5 在线数据库对 10 个甘松 CO 蛋白进行亚细胞定位预测。结果分析表明，10 个甘松 CO 蛋白主要定位于细胞核内，这一发现支持了它们作为转录因子在正常生理条件下能够有效调控下游基因表达的功能^[17]。然而，值得注意的是，*NcCONSTANS04* 和 *NcCONSTANS06* 2 个蛋白在细胞质中的定位占比较高，而在细胞核中的定位占比较低。这一差异表明，相比其他 8 个蛋白，*NcCONSTANS04* 和 *NcCONSTANS06* 可能在调控下游基因表达方面表现出较弱的活性。

基于以上分析，推测 CO 蛋白作为转录因子，并不依赖于蛋白转运过程，而是直接在细胞核内发挥转录调控作用。这一特性保证了它们在调控基因表达时的高效性和直接性，不同蛋白在亚细胞定位上的差异即揭示了它们在调控功能上的潜在差异 (表 2)。

表 2 甘松 CO 转录因子家族亚细胞定位

Table 2 Subcellular localization of CO transcription factor family in *N. jatamansi*

编号	预测评分											
	细胞核	胞外	液泡	细胞质	叶绿体	溶酶体	质膜	线粒体	高尔基体	细胞骨架	过氧化物酶体	内质网
NcCONSTANS01	3.650	0.658	0.216	0.154	0.084	0.078	0.077	0.042	0.015	0.011	0.008	0.006
NcCONSTANS02	2.756	0.058	0.021	0.645	0.837	0.008	0.058	0.551	0.014	0.024	0.020	0.009
NcCONSTANS03	4.753	0.040	0.007	0.068	0.033	0.009	0.014	0.056	0.002	0.002	0.015	0.002
NcCONSTANS04	2.500	0.027	0.058	1.103	0.785	0.013	0.015	0.428	0.013	0.019	0.029	0.010
NcCONSTANS05	4.330	0.274	0.192	0.059	0.022	0.034	0.051	0.018	0.007	0.008	0.002	0.003
NcCONSTANS06	2.095	0.061	0.042	1.492	0.523	0.011	0.023	0.605	0.019	0.030	0.042	0.024
NcCONSTANS07	4.656	0.137	0.025	0.064	0.022	0.007	0.013	0.054	0.014	0.003	0.001	0.004
NcCONSTANS08	4.203	0.308	0.026	0.156	0.035	0.116	0.052	0.062	0.008	0.006	0.015	0.013
NcCONSTANS09	3.437	0.989	0.033	0.090	0.106	0.079	0.154	0.075	0.009	0.006	0.005	0.017
NcCONSTANS10	3.957	0.441	0.082	0.215	0.053	0.086	0.038	0.073	0.019	0.010	0.011	0.016

3.4 甘松 CO 蛋白二级结构与三级结构分析

通过 SOPMA 平台预测甘松 CO 家族成员的二级结构（表 3）。结果表明，CO 家族的蛋白序列主要由 α -螺旋、 β -转角、无规则卷曲和延伸链等结构元素构成。其中， α -螺旋和无规则卷曲占据了二级结构的主要部分，约占总体的 85%，这一比例显著地反映了这 2 种结构元素在形成蛋白质特殊构象中的

关键作用。相对而言，无规则卷曲和 β -转角在二级结构中的占比则较小。这些结构元素的不同比例和组合方式共同决定了甘松 CO 蛋白的特定结构和功能。利用 SWISS-MODEL 在线工具对甘松中的 10 个 CO 蛋白序列进行结构预测，成功构建了三维模型，从而能够直观地观察其空间结构特征，包括 α -螺旋、 β -转角、无规则卷曲和延伸链的分布（图 2）。

表 3 甘松 CO 转录因子家族二级结构
Table 3 Secondary structure of CO transcription factor family in *N. jatamansi*

编号	α -螺旋/%	延伸链/%	β -转角/%	无规则卷曲/%
NcCONSTANS01	23.86	12.95	3.41	59.77
NcCONSTANS02	40.71	7.12	3.31	48.85
NcCONSTANS03	30.06	8.86	2.85	58.23
NcCONSTANS04	34.89	9.66	3.12	52.34
NcCONSTANS05	26.63	10.65	3.39	59.32
NcCONSTANS06	43.86	7.83	2.87	45.43
NcCONSTANS07	23.50	11.46	3.15	61.89
NcCONSTANS08	33.96	10.69	0.94	54.40
NcCONSTANS09	26.32	15.00	4.74	53.95
NcCONSTANS10	25.52	12.76	3.25	58.47

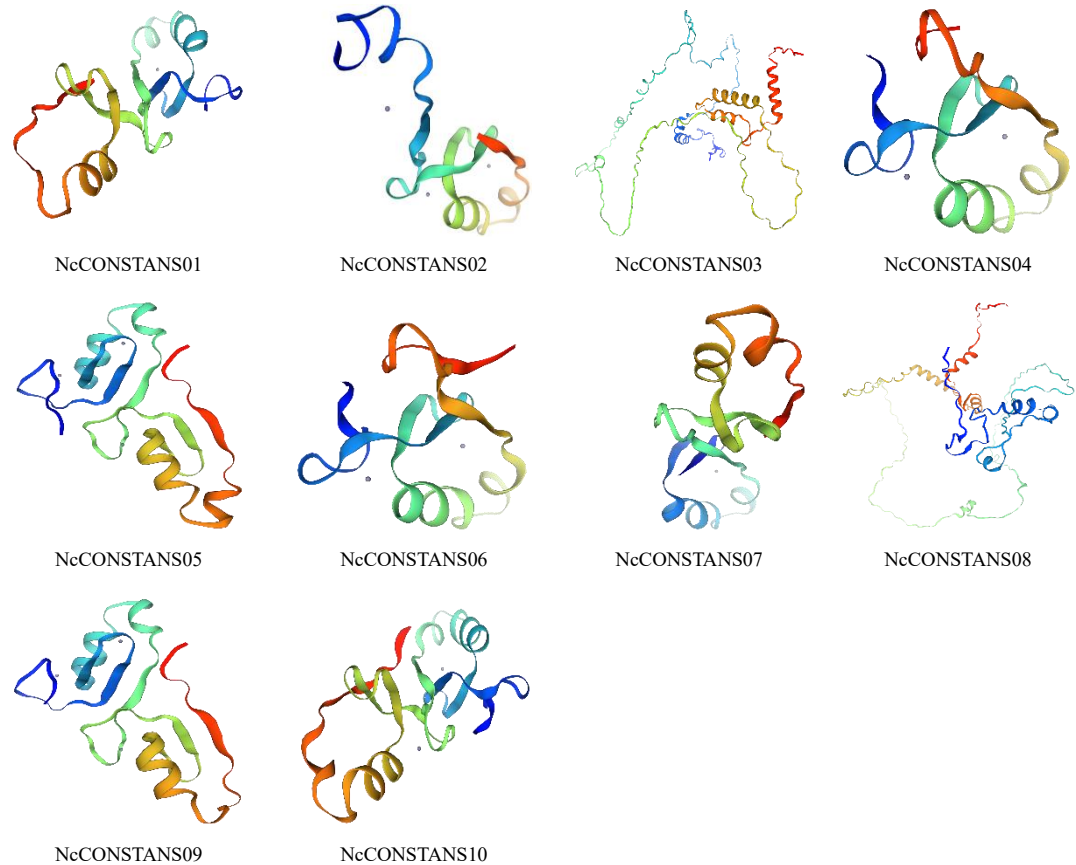


图 2 甘松 CO 转录因子家族三维结构预测
Fig. 2 Tertiary structure prediction of CO transcription factor family in *N. jatamansi*

经过深入分析，发现甘松 CO 蛋白的 4 个亚族之间，以及同一亚族内不同蛋白之间，在 α -螺旋、 β -转角、无规则卷曲和延伸链的比例上皆存在显著差异。这种结构上的多样性直接导致了它们在空间结构上的不同，进而可能赋予这些蛋白各自独特的功能特性。这些发现为深入理解甘松 CO 家族蛋白的结构-功能关系提供了重要线索。

3.5 甘松 CO 蛋白的基序特性解析

通过 MEME 在线软件，对甘松 CO 蛋白的保守

基序进行分析（图 3）。结果显示，CO 蛋白中共有 6 个显著的保守基序，将其命名为 motif1~motif6，长度为 13~48 aa。在这些保守基序中，motif2、motif1 和 motif3 被识别为 B-box 结构域，而 motif1 还同时属于 CCT 结构域。进一步分析显示，本研究中的 10 个甘松 CO 蛋白序列中，每个序列均包含 3~5 个这样的保守基序。这些保守基序和结构域的存在，为理解甘松 CO 蛋白家族的功能特性和分子机制提供了重要的线索。

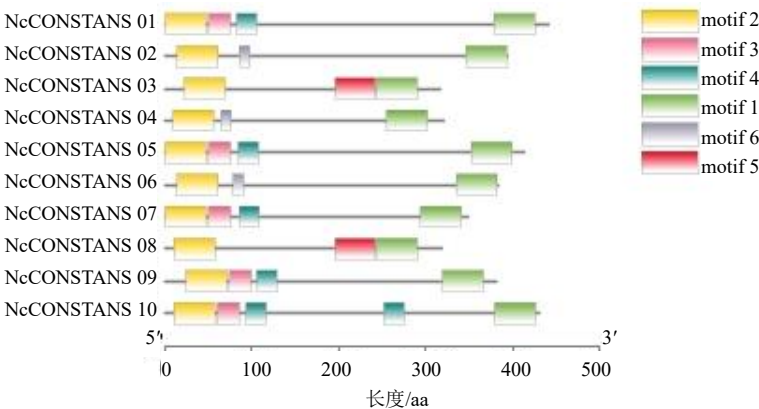


图 3 甘松 CO 转录因子家族 motif 分布

Fig. 3 Motif distribution of CO transcription factor family in *N. jatamansi*

3.6 信号肽、导肽及跨膜结构域的预测与解析

利用 SignalP-5 在线工具对 10 个 CO 蛋白的信号肽进行预测，结果表明所有 CO 蛋白均不具备信号肽特征，表明它们属于非分泌蛋白；利用 TM-HMMServer 2.0 对甘松 CO 蛋白的跨膜结构进行预

测，结果同样表明这 10 个蛋白均未展现出跨膜结构特性（表 4）。这一结果与先前关于它们亲水性的分析相吻合，进一步支持了甘松 CO 蛋白不具备跨膜运输功能，也不作为转录因子通过核孔进行转运的推论。

表 4 甘松 CO 转录因子家族导肽、信号肽、和跨膜结构域预测

Table 4 Prediction of transduction peptide, signal peptide and transmembrane domain of CO transcription factor family in *N. jatamansi*

编号	预测评分				
	信号肽	线粒体转运肽	叶绿体转运肽	类囊体转运肽	跨膜结构
NcCONSTANS01	0.002 9	0.003 4	0.000 1	0.000 0	无
NcCONSTANS02	0.005 4	0.001 8	0.011 8	0.000 6	无
NcCONSTANS03	0.004 7	0.003 0	0.000 0	0.000 0	无
NcCONSTANS04	0.003 2	0.000 1	0.000 3	0.000 3	无
NcCONSTANS05	0.000 8	0.000 0	0.000 0	0.000 0	无
NcCONSTANS06	0.005 8	0.000 8	0.000 0	0.000 1	无
NcCONSTANS07	0.001 0	0.000 3	0.000 3	0.000 1	无
NcCONSTANS08	0.001 6	0.000 1	0.000 0	0.000 0	无
NcCONSTANS09	0.030 3	0.153 8	0.018 8	0.000 0	无
NcCONSTANS10	0.002 5	0.011 0	0.000 0	0.000 0	无

3.7 甘松 CO 家族成员系统进化树构建

运用软件 MEGA7, 对从拟南芥官网数据库 (TAIR) 获取的 17 个拟南芥 CO 蛋白序列与甘松中的 10 个 CO 蛋白序列与进行同源性比较。基于同源性分析, 利用邻位法成功构建了系统进化树, 利用 iTOL 平台进行优化处理 (图 4)。结果显示, 甘松的 10 个 CO 蛋白在 3 个不同亚族中均有分布。NcCONSTANS03 和 NcCONSTANS08 蛋白被归类于

第 1 亚族 (BI), 展现出最高的保守性; NcCONSTANS02、NcCONSTANS04 和 NcCONSTANS06 则属于第 2 亚族 (BII), 保守性也相对较高; 而 NcCONSTANS01、NcCONSTANS05、NcCONSTANS07、NcCONSTANS09 和 NcCONSTANS10 则被归类于第 3 亚族 (BIII), 相较于前 2 个亚族, 保守性较低, 这一发现为理解甘松 CONSTANS 蛋白的进化关系和功能特性提供了重要线索^[18]。

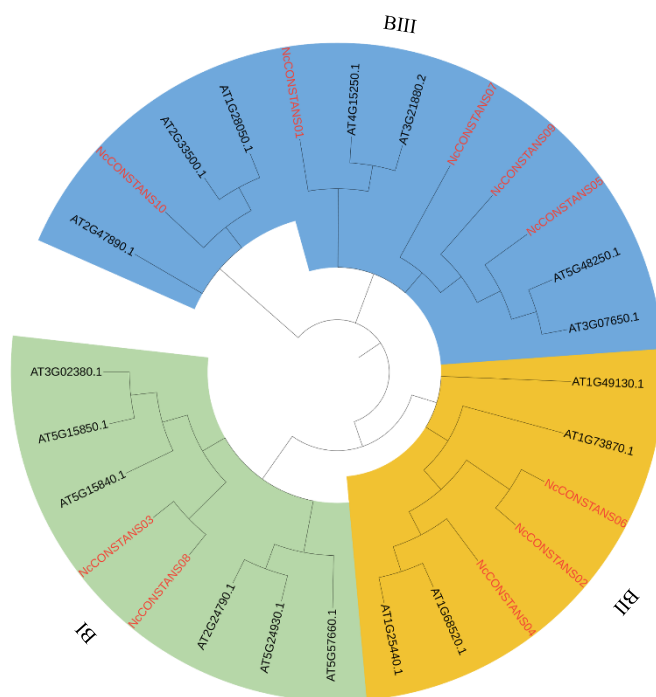


图 4 甘松与拟南芥 CO 转录因子家族比对系统进化树

Fig. 4 Phylogenetic tree of CO transcription factor family comparison between *N. jatamansi* and *Arabidopsis Thaliana* (L.) Heynh.

4 讨论

4.1 甘松 CO 转录因子家族在多种生理活动中扮演着至关重要的调控角色

CO 转录因子作为植物光周期中促进开花的关键调控因子, 在成花转变过程中起着至关重要的作用。它充当了生理钟与花分生组织决定基因之间的桥梁, 协调并调控开花相关的生理过程^[19]。在长日照条件下, 植物幼叶维管组织中的韧皮部 CO 蛋白结合至 FT 基因启动子, 激活其表达, 这是植物开花过程的核心环节。在光周期中 CO 转录因子参与调控拟南芥开花的机理目前研究的已经很透彻: ① CO 通过与 HAP3 (HEMEACTIVATORPROTEIN3) 和 HAP5 (HEMEACTIVATORPROTEIN5) 形成复合物, 共同调控 FT 基因的表达^[20]。② COP1 与 SPA1

相互结合形成复合体从而降解 CO 蛋白抑制植物开花^[21]。③ CRY2 与 COP1 和 SPA1 形成复合体, 抑制了 COP1-SPA1 复合体对 CO 的降解, 从而有利于植物开花的进程^[22]。④ CDF1 与 CO 的启动子结合, 从而抑制了 CO 基因转录的起始过程。⑤ GI 与 FKF1 形成的复合体通过降解 CDF1, 促进了 CO 基因转录的起始。⑥ FKF1 与 CO 形成复合物, 进而增强 FT 基因的表达^[23-24]。目前关于甘松的 CO 转录因子家族的研究尚属空白。本研究成功地从甘松中鉴别出 10 个功能未知的 CO 转录因子家族, 并进行了全面的分析。这一工作为后续对甘松 CO 转录因子家族的功能分析及验证奠定了坚实的基础。相较于拟南芥, 甘松中 CO 蛋白的数量较少。这种差异可能源于多个因素, 如转录组序列完整性的不

足,或是甘松中 CO 转录因子家族较高的重复率及其在拼接过程中的较高丢失率。这些发现凸显了不同物种间 CO 转录因子家族在数量和功能上的显著差异,以及它们在进化过程中可能面临的限制。

4.2 甘松 CO 转录因子家族对药材品质的影响

植物从营养生长过渡到生殖生长的关键节点是开花,而这一过程发生的时间点被称为开花时间,是植物生命周期中的一个标志性转变^[25],此阶段也称为抽薹^[26]。许多根茎类药材在抽薹开花后,其药效会明显减弱。例如,白芷 *Angelica dahurica* (Fisch. ex Hoffm.) Benth. & Hook. f. 在早期抽薹的情况下,其根部的折干率会明显低于未抽薹的植株^[27]。当归 *Angelica sinensis* (Oliv.) Diels. 的早期抽薹,这导致其在营养生长阶段无法有效积累有效成分,药用部位的肉质根发生萎缩和纤维化,最终完全丧失药用价值^[28]。在人工栽培条件下,防风 *Saposhnikovia divaricate* (Turcz.) Schischk. 植物在第二年就出现大量植株抽薹开花的现象,这导致防风植物的根部逐渐木质化,从而对药材的整体品质产生显著的不良影响^[29]。佟晓楠等^[30]研究团队利用实时荧光定量 PCR (RT-qPCR) 技术,深入探究了甜橙不同品种中 *CsCO3*、*CsCO4* 及 *CsCO7* 3 个基因的时空特异性表达模式。其研究结果显示,上述 3 个 CO 基因在甜橙进入花期时展现出更为显著的响应变化,这一发现强烈暗示了这 3 个基因在调控甜橙花期发育过程中发挥着不可或缺的关键作用。

由于甘松的药用部位为根茎,而其开花的时间影响其根部的发育及品质,通过调控光照时长,能够有效的调控甘松的开花,从而达到提高其药材的药用价值的目的。

4.3 甘松 CO 转录因子家族参与光周期调控促进药用部位的优质生产

CO 基因在植物成花过程中扮演着关键角色,它既直接又间接地促进花发育,调控植物成花的早晚^[31]。研究表明,当 *HdlmRNA* 功能缺失时,长日照条件下植物的开花会被促进,而在短日照条件下则会被抑制,这一发现为理解植物开花调控机制提供了新的视角^[32]。胡萝卜在栽培过程中易出现早期抽现象,这使得肉质根发生木质化,极大地影响了胡萝卜的产量。因此,对于寒冷地区的早期胡萝卜栽培而言,先期抽薹成为了一个亟待解决的重要问题^[33]。CO 基因在植物的光周期调控中扮演核心角色,它能直接启动并增强下游 *FTmRNA* 基因的表

达,进而促进植物的抽薹开花过程^[34]。随着抽薹进程的深入,植物的生长中心由地下转移至地上,导致根系中的营养成分被逐渐消耗,导致木质部发生明显的纤维化。此外,次生木质部和韧皮部的面积比例减少,木质化过程中薄壁细胞的数量也相应减少^[35]。在当归植株未进行抽薹时,其根系中的次生韧皮部与次生木质部的面积比例大致为 2:1。然而,一旦进入抽薹阶段,这一比例发生了显著变化,变为 1:10,显示了抽薹对植物根系结构的显著影响^[36]。研究表明,马铃薯 *StCO* 随着光周期呈现出显著的规律性变化,抑制马铃薯 *FT* 同源基因的表达水平,进而调控马铃薯块茎的形成^[37]。综上所述,光周期的调控能够影响甘松 CO 基因的表达,其被激活后,会直接促进下游基因 *FTmRNA* 的表达,而引起花器官的分化,与此同时,根部开始木质化。因此,通过合理调控光周期,可实现对甘松药用部位根茎的优质生产的控制。

甘松 CO 转录因子家族的深入探究与挖掘仍需进一步加强。本研究通过生物信息学的手段详细分析甘松 CO 转录因子家族的理化性质、二级结构、三级结构等,预测了其潜在的生物学作用及所参与的生理机制,不仅能为甘松优质高产的栽培提供理论依据,还为后续实验筛选和培育出更为优良的甘松品种提供技术支撑。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 昌维. 高原药用植物甘松的生态学研究 [D]. 重庆: 重庆大学, 2022.
- [2] 中国药典 [S]. 四部. 2020: 122-123.
- [3] 简鹏, 李庆海, 范立华. 甘松新酮对快速性心律失常大鼠心肌细胞抑制作用的实验研究 [J]. 中国临床药理学杂志, 2015, 31(22): 2240-2242.
- [4] 张小荣, 崔治家, 何海, 等. 甘松化学成分、药理作用研究进展及质量标志物预测分析 [J]. 中国中医药信息杂志, 2023, 30(11): 183-190.
- [5] 贾敏如, 张艺. 中国民族药辞典 [M]. 北京: 中国医药科技出版社, 2016: 552.
- [6] 宇妥·元丹贡布等著. 李永年译. 四部医典 [M]. 北京: 人民卫生出版社, 1983: 93-96.
- [7] Holfors A, Opseth L, Ree Rosnes A K, et al. Identification of *PaCOL1* and *PaCOL2*, two *CONSTANS*-like genes showing decreased transcript levels preceding short day induced growth cessation in Norway spruce [J]. *Plant Physiol Biochem*, 2009, 47(2): 105-115.

- [8] Almada R, Cabrera N, Casaretto J A, *et al.* *VvCO* and *VvCOL1*, two *CONSTANS* homologous genes, are regulated during flower induction and dormancy in grapevine buds [J]. *Plant Cell Rep*, 2009, 28(8): 1193-1203.
- [9] Hassidim M, Harir Y, Yakir E, *et al.* Over-expression of *CONSTANS-LIKE 5* can induce flowering in short-day grown *Arabidopsis* [J]. *Planta*, 2009, 230(3): 481-491.
- [10] 帅敏敏, 张启香, 黄有军. 光周期途径成花关键基因 *CONSTANS* 的进化机制 [J]. 浙江农林大学学报, 2019, 36(1): 7-13.
- [11] Putterill J, Robson F, Lee K, *et al.* The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors [J]. *Cell*, 1995, 80(6): 847-857.
- [12] 刘芳, 肖衡, 金玉环, 等. 短命植物小拟南芥光周期调控基因 *CONSTANS* 的克隆与表达特征分析 [J]. 石河子大学学报: 自然科学版, 2019, 37(2): 190-199.
- [13] Dahal P, Kwon E, Pathak D, *et al.* Crystal structure of a tandem B-box domain from *Arabidopsis* *CONSTANS* [J]. *Biochem Biophys Res Commun*, 2022, 599: 38-42.
- [14] 帅敏敏. 光周期途径成花关键基因 *GIGANTEA* 和 *CONSTANS* 的进化机制 [D]. 杭州: 浙江农林大学, 2018.
- [15] 崔琪, 俸明康, 丰日落, 等. 基于转录组信息的甘松 MADS-box 转录因子家族分析 [J]. 中草药, 2023, 54(3): 898-906.
- [16] 雍伟东, 谭克辉, 许智宏, 等. 高等植物开花时间决定的基因调控研究 [J]. 科学通报, 2000, 45(5): 455-466.
- [17] 黄文娟, 许鑫, 陈新松, 等. 川白芷 *NAC* 家族的生物信息及表达模式分析 [J]. 中国中药杂志, 2021, 46(7): 1769-1782.
- [18] 张清哲. 部分大豆 *CONSTANS* 家族基因的克隆和功能分析 [D]. 郑州: 河南农业大学, 2010.
- [19] 胡冬秀, 刘浩, 鲁清, 等. 花生 *CONSTANS-Like(COL)* 家族基因的克隆与表达分析 [J]. 中国油料作物学报, 2020, 42(5): 778-786.
- [20] Zuo Z C, Liu H T, Liu B, *et al.* Blue light-dependent interaction of *CRY2* with *SPA1* regulates *COP1* activity and floral initiation in *Arabidopsis* [J]. *Curr Biol*, 2011, 21(10): 841-847.
- [21] 咸焯. 月季 *CONSTANS (RoCO)* 基因的克隆与功能鉴定 [D]. 南京: 南京农业大学, 2017.
- [22] Sawa M, Nusinow D A, Kay S A, *et al.* FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis* [J]. *Science*, 2007, 318(5848): 261-265.
- [23] Ben-Naim O, Eshed R, Parnis A, *et al.* The CCAAT binding factor can mediate interactions between *CONSTANS*-like proteins and DNA [J]. *Plant J*, 2006, 46(3): 462-476.
- [24] Imaizumi T, Schultz T F, Harmon F G, *et al.* FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis* [J]. *Science*, 2005, 309(5732): 293-297.
- [25] 樊丽娜, 邓海华, 齐永文. 植物 *CO* 基因研究进展 [J]. 西北植物学报, 2008, 28(6): 1281-1287.
- [26] 郝新迪, 许梦男, 董波波, 等. 基于转录组信息的叶用莴苣 MADS-box 转录因子家族分析 [J]. 北京农学院学报, 2021, 36(4): 14-19.
- [27] 吴萍, 王晓宇, 郭俊霞, 等. 早期抽薹对白芷生长发育和品质的影响 [J]. 中国农业科技导报, 2023, 25(12): 58-66.
- [28] 高雪. 基于转录组测序挖掘当归早期抽薹关键基因及功能分析 [D]. 兰州: 甘肃农业大学, 2021.
- [29] 赵红燕. 华亭县中药材防风栽培技术 [J]. 农业科技与信息, 2015(19): 63-64.
- [30] 佟晓楠, 杨杰, 胡文娟, 等. 花期相关基因 *CO* 在甜橙基因组中的鉴定及表达分析 [J]. 分子植物育种, 2021, 19(23): 7721-7728.
- [31] 吴连成, 常丽丽, 陈晓, 等. *CO* 基因的调控表达与植物光周期反应 [J]. 中国农学通报, 2010, 26(2): 116-121.
- [32] Yano M, Katayose Y, Ashikari M, *et al.* *Hdl*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS* [J]. *Plant Cell*, 2000, 12(12): 2473-2484.
- [33] 刘思思, 孔小平. 胡萝卜抽薹相关性状遗传多样性和关联度分析 [J]. 青海大学学报, 2023, 41(4): 54-61.
- [34] 艾炎军. 光周期调控马铃薯块茎形成的全基因组甲基化分析及甲基化修饰的差异基因筛选 [D]. 武汉: 华中农业大学, 2020.
- [35] 黄珊. 当归早期抽薹分化和内源性激素变化研究 [D]. 杨凌: 西北农林科技大学, 2020.
- [36] 孟祥才, 孙晖, 孙小兰, 等. 直播、移栽和抽薹防风与野生防风药理作用比较 [J]. 现代中药研究与实践, 2012, 26(5): 31-33.
- [37] González-Schain N D, Díaz-Mendoza M, Zurczak M, *et al.* Potato *CONSTANS* is involved in photoperiodic tuberization in a graft-transmissible manner [J]. *Plant J*, 2012, 70(4): 678-690.

[责任编辑 时圣明]