

## 胡椒去甲乌药碱合酶基因编码蛋白质的生物信息学分析

邢超<sup>1,2</sup>, 姜洁<sup>1,2</sup>, 耿健强<sup>1,2</sup>, 王丹<sup>1,2</sup>, 宋丽萍<sup>1,2</sup>, 毛婷<sup>3\*</sup>

<sup>1</sup>北京市食品检验研究院(北京市食品安全监控和风险评估中心) 北京

<sup>2</sup>国家市场监督管理总局重点实验室(食品安全重大综合保障关键技术) 北京

<sup>3</sup>北京市检验检测认证中心 北京

**【摘要】**胡椒中含有一种 $\beta$ 受体激动剂类物质—去甲乌药碱,具有误食而发生兴奋剂阳性事件的风险。本研究基于胡椒基因组 blast,发现胡椒中存在三个去甲乌药碱合酶(*PnNCS*)基因。首先利用生物信息学方法对三个*PnNCS*的氨基酸序列进行了理化性质分析。随后,*PnNCS*的系统发育地位分析结果表明去甲乌药碱合酶分为八个进化分支,胡椒的三种去甲乌药碱合酶分属在不同的两个进化分支中。保守基序分析结果显示共具有三种保守基序,分别为:YGQAFVVSEDQKLDWGMLFLJTLPLVELRNMRFWPTNPPSF、FTKVDNEKMKETEVVEGGYLDGFTLYRVRFEI以及KLHLACKEWGGFFQLINHGVP。胡椒 NCS: APF31815.1 和 AS075045.1 具有一致的保守基序(FTKVDNEKMKETEVVEGGYLDGFTLYRVRFEI),胡椒 NCS: WDS62667.1 无明确保守基序。修饰位点分析结果显示胡椒 NCS 蛋白具有多个磷酸化、糖基化以及氧化还原修饰位点。顺式作用元件分析结果显示 NCS 的启动区含有应答环境胁迫与多种植物激素的顺式作用元件。本研究提供了胡椒 NCS 氨基酸的基础信息,为后续探究遗传特性提供了参考。

**【关键词】**胡椒;去甲乌药碱合酶;生物信息学预测

**【基金项目】**重大赛事食源性兴奋剂检验和防控关键技术研究及示范应用创新团队项目(市场监管总局科技创新人才计划, CXTD202307)

**【收稿日期】**2024 年 8 月 12 日

**【出刊日期】**2024 年 9 月 21 日

**【DOI】**10.12208/j.jafs.20241016

### Bioinformatics Analysis of protein Encoded by Norcochlorine synthase in *Piper nigrum*

Chao Xing<sup>1,2</sup>, Jie Jiang<sup>1</sup>, Jianqiang Geng<sup>1,2</sup>, Dan Wang<sup>1,2</sup>, Liping Song<sup>1,2</sup>, Ting Mao<sup>3\*</sup>

<sup>1</sup>Beijing Institute of Food Inspection and Research (Beijing Municipal Center for Food Safety Monitoring and Risk Assessment)

<sup>2</sup>Laboratory of Key Technologies of Major Comprehensive Guarantee of Food Safety for State Market Regulation, Beijing

<sup>3</sup>Beijing Inspection and Testing Certification Center, Beijing

**【Abstract】** There is a  $\beta$ -agonist substance in pepper - noraconine, which has the risk of positive stimulant events due to ingestion. In this study, based on the pepper genome blast, we found three normonine synthase (*PnNCS*) genes in pepper. Firstly, the amino acid sequences of three *PnNCS* were analyzed by bioinformatics method. Subsequently, phylogenetic status analysis of *PnNCS* showed that noraconitine synthases were divided into eight evolutionary clades, and the three noraconitine synthases of pepper belonged to two different evolutionary clades. The results of conservative motifs analysis showed that there were three kinds of conservative motifs, which were: YGQAFVVSEDQKLDWGMLFLJTLPLVELRNMRFWPTNPPSF, FTKVDNEKMKETEVVEGGYLDGFTLYRVRFEI and KLHLACKEWGGFFQLINHGVP. Pepper replication: APF31815.1 and AS075045.1 conservative base with a consistent sequence (FTKVDNEKMKETEVVEGGYLDGFTLYRVRFEI), pepper replication: WDS62667.1 no clearly conservative motif. The analysis of modification sites showed that the NCS protein had

第一作者简介:邢超(1996-)男,博士,研究方向为生物化学与分子生物学;

\*通讯作者:毛婷(1980-)女,北京市检验检测认证中心教授级高级工程师,博士,研究方向为食品检验、质量管理。

multiple phosphorylation, glycosylation and REDOX modification sites. The analysis of cis-acting elements showed that the initiation region of NCS contains cis-acting elements in response to environmental stress and various plant hormones. This study provided the basic information of NCS amino acids of pepper, and provided a reference for further investigation of genetic characteristics.

**【Keywords】** Pepper; Norcochlorine synthase; Bioinformatics prediction

胡椒是一种具有重要经济和生态意义的作物。由于其在世界范围内的广泛消费, 它被称为香料之王。胡椒中含有一种  $\beta$  受体激动剂类物质—去甲乌药碱, 因此具有误食而发生兴奋剂阳性事件的风险。

去甲乌药碱合酶 (Norcochlorine synthase, NCS) 是催化合成去甲乌药碱的关键酶。最初被认为是将多巴胺和 3, 4-二羟基苯乙醛转化为四羟基生物碱。自 2001 年起, 不同植物中的去甲乌药碱合酶被鉴定表征。Samanani 和 Facchini 在罂粟和 *Thalictrum flavum* sp. *glaucom* 中分离纯化出去甲乌药碱合酶。对该蛋白体外表征结果显示 PsNCS 对 4-羟基苯基乙醛表现出双曲饱和动力学而对多巴胺表现出 S 型饱和动力学, 表明 NCS 在底物结合位点之间表现出正向的协同性。这类酶通常催化代谢中的调节或限速步骤, 表明 NCS 在控制苯基异喹啉生物碱生物合成中的通量方面发挥作用<sup>[1-3]</sup>。黄连 (*Coptis japonica*) 中 NCS 蛋白功能研究显示 CjNCS 的催化反应不依赖于 2-酮戊二酸而需要亚铁离子<sup>[4]</sup>。Luk 的研究结果显示 NCS 催化多巴胺和 4-羟基苯乙醛的不对称 Pictet-Spengler 缩合反应生成去甲乌药碱<sup>[5]</sup>。随后, 研究者们优化了 NCS 体外重组蛋白纯化的方法并解析了多个 NCS 蛋白的晶体结构<sup>[6-10]</sup>。对 NCS 体内的研究表明 PsNCS 定位与罂粟韧皮部的筛网元件 (Sieve elements), 并通过病毒诱导基因沉默 (virus-induced gene silencing, VIGS) 体系实现了对 NCS 基因沉默, 通过保守结构域比较分析发现 NCS 属于 PR10/Bet v1 蛋白家族<sup>[11]</sup>, 这项研究提供了 NCS 催化生成去甲乌药碱的体内证据。Li 在构造高产苯基异喹啉生物碱的工程菌的过程中发现, 在罂粟科中编码 NCS 的基因以串联融合的形式存在<sup>[12]</sup>。Vimolmangkang 对莲花 (*Sacred lotus*) 中 NCS 基因进行了系统发育分析并测定了 SINCS 在不同组织中的表达水平和不同组织中的去甲乌药碱含量, 结果显示不同组织的 SINCS 基因转录水平与去甲乌药碱含量水平并不都具有正相关性<sup>[13]</sup>。Vimolmangkang 的结果暗示了 NCS 可能在转录后或翻译阶段被某些

元件调控。在 2017 年, 高通量测序技术被用来探究药用植物中苯基异喹啉生物碱代谢。Yang 等人利用数字基因表达技术鉴定了两个生物碱含量不同的莲藕在四个叶片发育阶段的差异表达基因<sup>[14]</sup>。Chen 等人通过比较黄连 (*Coptis chinensis*) 与祖先毛茛目的基因组发现 NCS 基因家族的扩张<sup>[15]</sup>。Hao 等人利用转录组学和代谢组学技术分析了五种中国特有的毛茛科药用植物的生物碱代谢相关的基因表达和产物含量, 发现代谢产物含量和基因表达之间具有广泛的相关性, 这支持了产生不同专门代谢产物所必须的各种基因的共同进化<sup>[16]</sup>。在最新的研究中, Feng 等人测定了药用植物延胡索不同组织中去甲乌药碱合酶的转录水平并预测了该基因启动子区域结合的顺式作用元件<sup>[17]</sup>。这些研究加深了对植物生成苯基异喹啉生物碱的遗传基础和相关基因家族进化历程的理解, 然而关于胡椒 NCS 家族的系统发育地位、编码蛋白质的结构及生物信息学预测方面还有待研究。

本研究通过对胡椒去甲乌药碱合酶基因编码蛋白质的理化特性、系统发育地位、保守基序以及调控元件分析, 以期通过遗传育种手段改善植物源性食品中去甲乌药碱含量提供理论基础。

## 1 结果

### 1.1 胡椒去甲乌药碱合酶基因的开放阅读框和编码蛋白的理化性质

由表 1 可知, 3 个 *PnNCS* 基因的起始密码子均位于第 1 个核苷酸, 终止密码子位于第 381-576 个核苷酸。*PnNCS* 蛋白共编码 127-191 个氨基酸, 分子量为 13954-21999.26, 等电点 5.52-8.94, 不稳定系数 20.06-32.55, 亲水性总平均值 -0.388 至 -0.05, 表明 3 个 *PnNCS* 蛋白均为稳定的亲水性蛋白。

### 1.2 植物去甲乌药碱合酶的系统发育

本研究在 NCBI 中收集到注释为 NCS 的 319 条氨基酸序列, 包括罂粟 (*Medicago truncatula*)、大麻 (*Cannabis sativa*)、咖啡 (*Coffea arabica*)、莲藕 (*Nelumbo nucifera*)、玉米 (*Zea mays*)、水稻

(*Oryza glaberrima*)、葡萄 (*Vitis vinifera*)、向日葵 (*helianthus annuus*)、胡椒 (*Piper nigrum*) 等植物。系统发育分析结果显示 319 条氨基酸序列被分为八个进化分支 (图 1)。其中, 在胡椒基因组中查询到存在三个 NCS, 这三个 NCS 分属于两个进化分支。

### 1.3 保守基序

在八个进化分支中随机挑选一条氨基酸序列, 共 52 条氨基酸序列以进行保守结构域分析。结果显示处于同一簇的基因具有相同的保守基序模块。分别为无保守基序、存在一种保守基序 (蓝/红)、存在两种保守基序 (绿和红) (图 2 和 3)。三种保守基序分别为: 红: YGQAFVVSEDQKLDWGMFLFJTLPVELRNMRFWPTNPPSF、蓝: FTKVDNEKMOV KETEVVEGGYLDGFTLYRVRFEI 以及绿:

KLHLACKEWGGFFQLINHGVPVE。

### 1.4 蛋白修饰与调控元件分析

3 个 *PnNCS* 的蛋白质翻译修饰位点情况如表 2 所示。磷酸化修饰位点分别有 19、24、26 个, 糖基化修饰位点分别有 16、21、16 个, 氧化还原修饰位点 (半胱氨酸残基) 分别有 1、1、2 个。

本研究选择与胡椒 NCS 进化距离近的咖啡 (*Coffea arabica*) 和柳杉 (*Cryptomeria japonica*) 的 NCS 启动子区域进行了调控元件结合位点分析。结果显示, NCS 的启动子区含有应答环境胁迫与多种植物激素的顺式作用元件。参与植物激素合成和信号转导, 调控细胞增值、分化、形态建成、细胞壁发育等生物过程的 TCP; 调控种子萌发、植物光合作用的 Dof 转录因子; 植物体内参与生长发育及植物抗逆的 C2H2、GATA 和 bZIP 转录因子等 (图 4)。

表 1 胡椒 *PnNCS* 基因的开放阅读框和编码蛋白的理化性质

登录号	分子式	氨基酸数	蛋白分子量	开放阅读框		等电点	不稳定系数	亲水性平均值
				起始	结束			
AS075045	C <sub>628</sub> H <sub>994</sub> N <sub>160</sub> O <sub>190</sub> S <sub>4</sub>	127	13954	1	381	5.49	32.55	-0.101
APF31815.1	C <sub>785</sub> H <sub>1245</sub> N <sub>201</sub> O <sub>238</sub> S <sub>5</sub>	161	17467.02	1	483	5.52	31.08	-0.05
WDS62667.1	C <sub>1009</sub> H <sub>1550</sub> N <sub>266</sub> O <sub>279</sub> S <sub>4</sub>	191	21999.26	1	576	8.94	20.06	-0.388

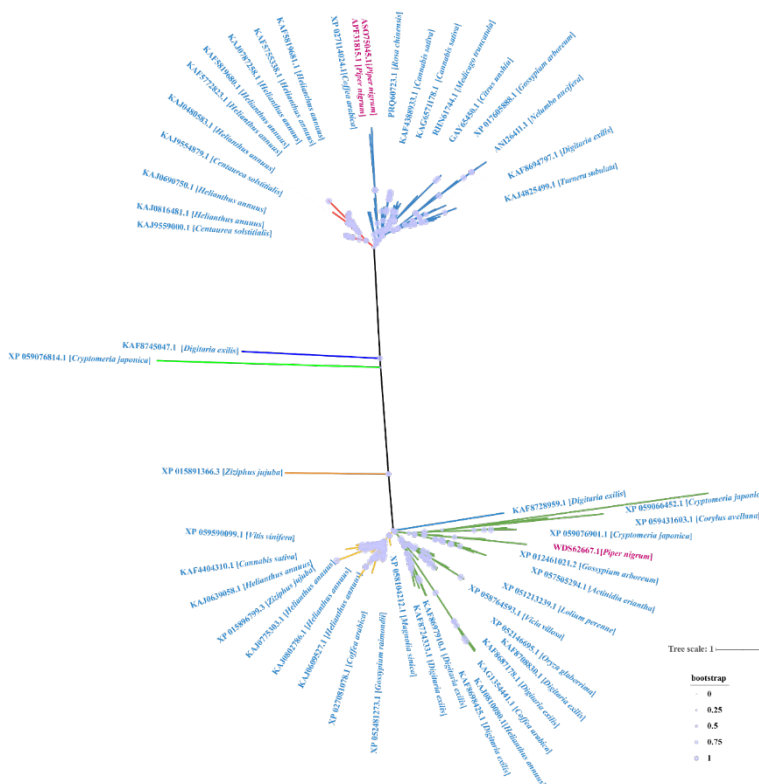


图 1 植物去甲乌药碱合成酶系统发育分析

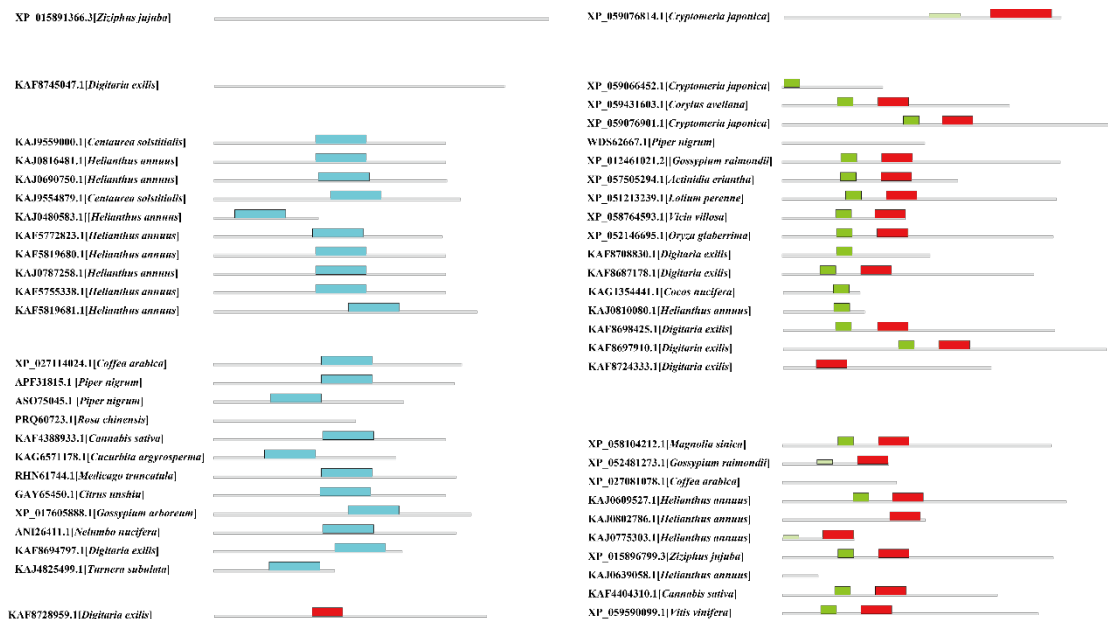


图2 代表进化分支的去甲乌药碱合成酶的保守基序分析

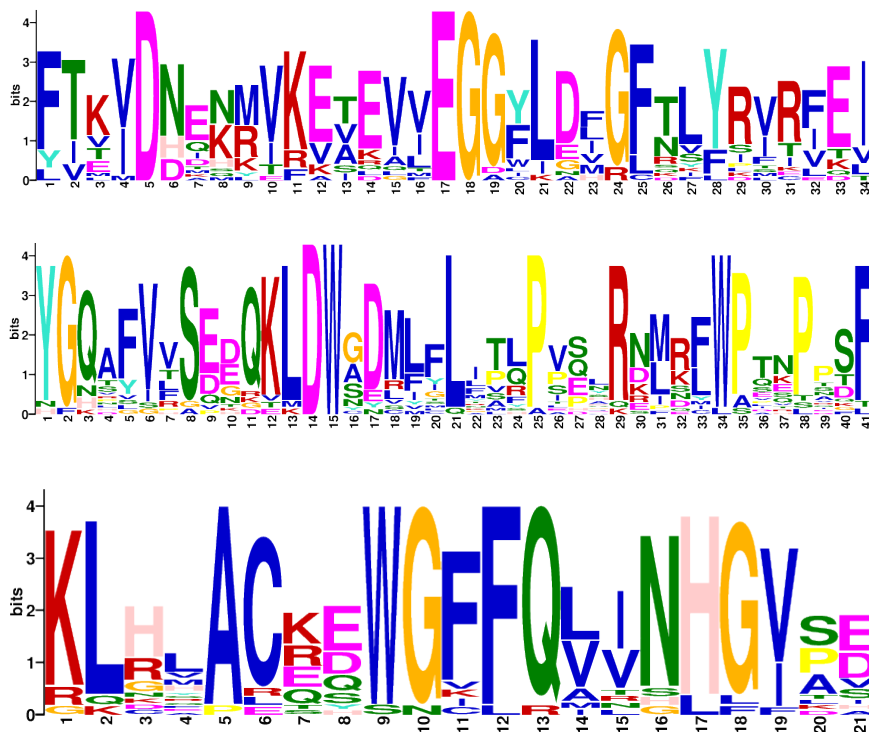


图3 代表进化分支的去甲乌药碱合成酶的保守基序

表2 三个 pNCS 蛋白质翻译后修饰位点统计

登录号	磷酸化修饰位点			糖基化修饰位点	氧化还原修饰位点
	丝氨酸	苏氨酸	酪氨酸		
AS075045	7	9	3	16	1
APF31815.1	12	9	3	21	1
WDS62667.1	10	6	10	16	2

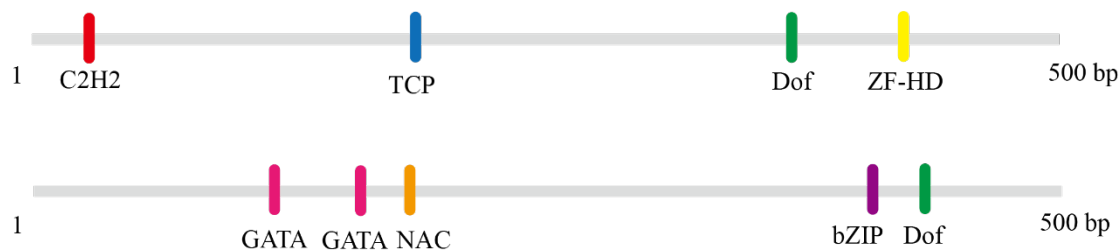


图4 启动子调控元件结合基序分析

## 2 讨论

胡椒, 是享誉世界的一种调味品。因其去甲乌药碱含量高, 运动员在参赛过程中被禁止食用含胡椒的食物。NCS 是去甲乌药碱合成的关键酶, 也是莽基异喹啉生物碱合成的第一步限速酶。之前对 NCS 的研究主要集中于药用植物, 对胡椒 NCS 的研究较少。本研究通过系统分析胡椒基因组数据, 共鉴定出 3 个胡椒 NCS 基因。对 3 个 *PnNCS* 编码蛋白的理化特性进行预测, 结果显示 *PnNCS* 蛋白均为稳定的亲水性蛋白, 表明 NCS 能够很好的发挥催化特性。

在植物中, PR10 蛋白家族<sup>[18]</sup>分为五个主要分支。典型的 PR10 (例如, 欧芹 PR1 和桦树 Bet V1) 属于在一个“双子叶植物”分支中。下一个大型分支包括三个子分支: “单子叶植物 PR10 I 型和 II 型”和一个“针叶树”分支。最后一个分支是一组植物激素结合蛋白 (PhBP), 包括几种细胞分裂素特异性结合 PR10s 和去甲乌药碱合成酶 (NCS)。NCS 是第一个被公认具有真正生物合成活性的 PR10 家族蛋白<sup>[18]</sup>。然而, 胡椒 NCS 系统发育地位还未鉴定。本研究结果表明, 3 个 *PnNCS* 分属于两个进化分支。APF31815.1 和 AS075045.1 具有一致的保守基序 (FTKVDNEKMKVETEVVEGGYLDFGFTLYRVR FEI), 与咖啡、月季花、大麻、灰籽南瓜、树棉、莲子的 NCS 同源、胡椒 NCS: WDS62667.1 无明确保守基序, 但与柳杉、猕猴桃、长柔毛野豌豆、向日葵的 NCS 同源。在保守区域的发现表明, 这些区域可能在蛋白质的结构稳定性或功能中发挥关键作用。这些发现与在罂粟等植物中<sup>[1-3]</sup>的研究结果一致, 进一步支持了其在生物学功能中的重要性。

蛋白质翻译后修饰是许多细胞信号转导事件的核心, 其深刻影响蛋白质功能。磷酸化、糖基化以及氧化还原修饰是调控和控制蛋白质活力和功能最基

本、最普遍的机制<sup>[19-21]</sup>。本研究预测了 3 个 *PnNCS* 的三种修饰位点。WDS62667.1 具有最多的磷酸化位点和氧化还原修饰位点, APF31815.1 具有最多的糖基化修饰位点, 这说明 *PnNCS* 在行使催化功能中受到多种翻译后修饰的调节。由于在 NCBI 胡椒基因组中未提供 3 个 *PnNCS* 启动子的序列, 因此选择了与胡椒 NCS 进化距离近的咖啡和柳杉 NCS 启动子区域进行了调控元件结合位点分析。结果显示, NCS 的启动子区含有应答环境胁迫与多种植物激素的顺式作用元件。本研究为进一步研究胡椒 NCS 蛋白质的翻译后修饰和调控提供了基础信息, 填补了 NCS 在胡椒中的研究。

## 3 材料与方法

### 3.1 试验材料

在 NCBI 依据胡椒基因组 blast, 查询胡椒 NCS 氨基酸序列。

### 3.2 生物信息学分析

#### 3.2.1 NCS 基因开放阅读框的确定

本研究使用了以下生物信息学工具: BLAST 用于序列比对, ORF finder 用于分析查找序列中的 ORF 区, ExPASy 用于对氨基酸序列的理化特性进行分析, ClustalW 用于进行序列比对, fasttree 用于构建系统发育树, iTOL 用于对系统发育树的美化处理, MEME 常用于基因家族分析或者展示一个序列集合的保守区域, NetPhos3.1 用于氨基酸序列磷酸化修饰位点的预测, NetOGlyc4.0 用于氨基酸序列糖基化位点的预测, Scratch Protein Predictor 用于氨基酸序列氧化还原修饰位点的预测, plantPAN4.0 用于启动子调控元件结合基序的预测。选择这些工具是基于它们在各自领域的广泛应用和高精度性能。

首先使用 BLAST 对目标序列进行同源序列比对, 查询到胡椒 NCS 氨基酸序列。接着, 利用 ORF finder 确定序列中的 ORF 区, 并使用 ExPASy 对 NCS

理化特性进行分析, 以了解三个氨基酸序列的基本信息。随后, 在 NCBI 中查询到 319 条不同物种的 NCS 氨基酸序列, 使用 ClustalW 进行多序列比对, fasttree 工具构建基于比对结果的系统发育树, 探索胡椒 NCS 的家族分布和进化关系, 使用 iTOL 对系统发育树中的序列进行标注。接着, 使用 MEME 对不同家族的代表氨基酸序列的保守基序进行了分析, 最后使用 NetPhos3.1、NetOGlyc4.0、Scratch Protein Predictor 对 3 个 PnNCS 的磷酸化、糖基化以及氧化还原修饰位点进行预测。使用 plantPAN4.0 启动子调控元件结合基序, 以探究 3 个胡椒 NCS 在蛋白质翻译后修饰以及转录调控方面的潜在作用。

### 参考文献

- [1] Samanani N, Liscombe DK, Facchini PJ. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis. *Plant J* 2004, 40(2):302-313.
- [2] Samanani N and Facchini PJ. Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzyloquinoline alkaloid biosynthesis, from opium poppy. *Planta* 2001, 213(6):898-906.
- [3] Samanani N and Facchini PJ. Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzyloquinoline alkaloid biosynthesis in plants. *J Biol Chem* 2002, 277(37):33878-33883.
- [4] Minami H, Dubouzet E, Iwasa K, et al. Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J Biol Chem* 2007, 282(9):6274-6282.
- [5] Luk LY, Bunn S, Liscombe DK, et al. Mechanistic studies on norcoclaurine synthase of benzyloquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction. *Biochemistry* 2007, 46(35):10153-10161.
- [6] Berkner H, Engelhorn J, Liscombe DK, et al. High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies. *Protein Expr Purif* 2007, 56(2):197-204.
- [7] Lee SK and Keasling JD. Heterologous protein production in *Escherichia coli* using the propionate-inducible pPro system by conventional and auto-induction methods. *Protein Expr Purif* 2008, 61(2):197-203.
- [8] Berkner H, Schweimer K, Matecko I. et al. Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *Biochem J* 2008, 413(2):281-290.
- [9] Ilari A, Franceschini S, Bonamore A. et al., Structural basis of enzymatic (S)-norcoclaurine biosynthesis. *J Biol Chem* 2009, 284(2):897-904.
- [10] Pasquo A, Bonamore A, Franceschini S. et al. Cloning, expression, crystallization and preliminary X-ray data analysis of norcoclaurine synthase from *Thalictrum flavum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2008, 64(Pt 4):281-283.
- [11] Lee EJ and Facchini P. Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. *Plant Cell* 2010, 22(10):3489-3503.
- [12] Li J, Lee EJ, Chang L, et al. Genes encoding norcoclaurine synthase occur as tandem fusions in the Papaveraceae. *Sci Rep* 2016, 6:39256.
- [13] Vimolmangkang S, Deng X, Owiti A, et al. Evolutionary origin of the NCSI gene subfamily encoding norcoclaurine synthase is associated with the biosynthesis of benzyloquinoline alkaloids in plants. *Sci Rep* 2016, 6:26323.
- [14] Yang M, Zhu L, Li L, et al. Digital Gene Expression Analysis Provides Insight into the Transcript Profile of the Genes Involved in Aporphine Alkaloid Biosynthesis in Lotus (*Nelumbo nucifera*). *Front Plant Sci* 2017, 8:80.
- [15] Chen DX, Pan Y, Wang Y, et al. The chromosome-level reference genome of *Coptis chinensis* provides insights into genomic evolution and berberine biosynthesis. *Hortic Res* 2021, 8(1):121.
- [16] Hao DC, Li P, Xiao PG, et al. Dissection of full-length transcriptome and metabolome of *Dichocarpum* (*Ranunculaceae*): implications in evolution of specialized metabolism of *Ranunculales* medicinal plants. *PeerJ* 2021, 9:e12428.

- [17] 冯飞雪,吕瑞华,李依民等.延胡索去甲乌药碱合成酶基因 CyNCS1 克隆及其表达模式分析[J/OL].西北农林科技大学学报(自然科学版),2024(03):1-9.
- [18] Dastmalchi M. Elusive partners: a review of the auxiliary proteins guiding metabolic flux in flavonoid biosynthesis. *Plant J* 2021, 108(2):314-329.
- [19] 郝晓燕, 丁璇, 赵准, 胡文冉.陆地棉 GhTCP4 基因编码蛋白质生物信息学分析[J/OL].分子植物育种. <https://link.cnki.net/urlid/46.1068.S.20240513.1736.006>.
- [20] Vu LD, Gevaert K, De Smet I. Protein Language: Post-Translational Modifications Talking to Each Other. *Trends*

*Plant Sci.* 2018 Dec;23(12):1068-1080.

- [21] Patwardhan A, Cheng N, Trejo J. Post-Translational Modifications of G Protein-Coupled Receptors Control Cellular Signaling Dynamics in Space and Time. *Pharmacol Rev.* 2021 Jan;73(1):120-151.

**版权声明:** ©2024 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



**OPEN ACCESS**