

DOI:10.14188/j.ajsh.20250317002

基于三代转录组测序的湖北黄精 *SUVH* 基因家族鉴定与分析

丁露^{1†},董伊洛^{1†},Victor Manuel Martinez Espinosa¹,田科伟^{1,2*},张德春^{1*}

(1. 三峡大学 生物技术研究中心,湖北 宜昌 443002;
2. 三峡大学 水利工程博士后科研流动站,湖北 宜昌 443002)

摘要: 组蛋白甲基化作为表观遗传调控的主要机制之一,在植物生长发育及环境响应过程中起关键作用。基于三代测序技术组装的黄精全长转录组数据,对湖北黄精组蛋白甲基化转移酶中的 *SUVH* 基因家族进行鉴定与分析,包括 *SUVH* 家族基因的序列特性、蛋白结构以及表达模式等进行了分析总结,并构建了蛋白进化树。结果表明,研究鉴定出 39 个 *SUVH* 基因,其氨基酸残基数量为 161 aa~1 268 aa,等电点在 4.77 PI~9.10 PI,分子量大小为 18 kDa~142 kDa;其中,编码酸性蛋白质的基因有 20 个,编码碱性蛋白质的基因有 19 个。该基因家族中只有 transcript/6593.p1 编码的蛋白具有一个跨膜域,该基因共编码 837 位氨基酸,第 813 至 835 位氨基酸形成跨膜螺旋。湖北黄精 *SUVH* 基因家族与拟南芥同源基因存在显著进化保守性,但其基因数目显著多于拟南芥。这些结果暗示湖北黄精可能通过基因家族扩张,从而适应复杂次生代谢调控需求。亚细胞定位结果表明,76.47% 的蛋白家族成员定位于细胞核;模体(motif)分析结果显示,3'端保守的终止元件可能参与底物识别特异性调控。此研究为阐明其表观遗传调控网络及药用次生代谢合成机制提供了关键基因资源,同时建立了药用植物功能基因组学研究的系统方法。

关键词: 湖北黄精;组蛋白甲基化;基因家族分析;生物信息学;表观遗传学

中图分类号: Q811

文献标志码: A

文章编号: 2096-3491(2025)05-0494-10

Identification and analysis of *SUVH* gene family of *Polygonatum zanlanscianense* based on third-generation transcriptome sequencing

Ding Lu^{1†}, Dong Yiluo^{1†}, Espinosa Victor Manuel Martinez¹, Tian Kewei^{1,2*}, Zhang Dechun^{1*}

(1. Biotechnology Research Center, China Three Gorges University, Yichang 443002, Hubei, China;
2. Postdoctoral Research Station in Hydraulic Engineering, China Three Gorges University, Yichang 443002, Hubei, China)

Abstract: Histone methylation, as one of the main mechanisms of epigenetic regulation, plays a key role in plant growth, development and environmental response. Based on the full-length transcriptome data of Hubei Huangjing (*Polygonatum zanlanscianense*) assembled by third-generation sequencing technology, the *SUVH* gene family in histone methyltransferase of *Polygonatum zanlanscianense* were identified and analyzed, including the sequence characteristics, protein structure, and expression pattern of *SUVH* family genes, and a protein evolution tree was constructed. The results showed that a total of 39 *SUVH* genes were identified, with the number of amino acid residues ranging from 161 aa to 1 268 aa, the isoelectric point ranging from 4.77 PI to 9.10 PI, and the molecular weight ranging from 18 kDa to 142 kDa; among them, there were 20 genes encoding acidic proteins and 19 genes encoding basic proteins. Only the protein

收稿日期: 2025-03-17 修回日期: 2025-06-30 接受日期: 2025-10-03

作者简介: 丁露(1999-),女,硕士,研究方向:植物分子生物学,E-mail: dl13972592346@163.com;董伊洛(2001-),男,硕士,研究方向:基因家族分析,E-mail: 2295619022@qq.com。†丁露和董伊洛在本文中贡献一致,为共同第一作者

*通信作者: 田科伟(1993-),男,博士,讲师,研究方向:植物发育生物学,E-mail: kw_tian@ctgu.edu.cn;张德春(1969-),男,博士,教授,研究方向:植物分子生物学,E-mail: zhangdc227@163.com

基金项目: 多组学解析玄参悬浮细胞培养合成环烯醚萜的调控机制研究(WDCM2023014)

引用格式: 丁露,董伊洛,Victor Manuel Martinez Espinosa,等. 基于三代转录组测序的湖北黄精 *SUVH* 基因家族鉴定与分析[J]. 生物资源, 2025, 47(5): 494-503.

Ding Lu, Dong Yiluo, Espinosa Victor Manuel Martinez, et al. Identification and analysis of *SUVH* gene family of *Polygonatum zanlanscianense* based on third-generation transcriptome sequencing [J]. Biotic Resources, 2025, 47(5): 494-503.

encoded by transcript/6593. p1 had a transmembrane domain in this gene family. This gene encoded a total of 837 amino acids, and amino acids from 813 to 835 formed a transmembrane helix. The *SUVH* gene family of *Polygonatum zanlanscianense* had significant evolutionary conservation with *Arabidopsis* homologous genes, but the number of its genes was significantly greater than that of *Arabidopsis*. These results suggested that *Polygonatum zanlanscianense* may adapt to the complex secondary metabolism regulation needs through gene family expansion. Subcellular localization results showed that 76.47% of the protein family members were localized in the nucleus. The results of motif analysis showed that the conserved termination element at the 3' end may be involved in the regulation of substrate recognition specificity. This study analyzes the evolution and functional characteristics of the *SUVH* gene family in *Polygonatum zanlanscianens*, provides key gene resources for clarifying its epigenetic regulatory network and the synthesis mechanism of medicinal secondary metabolism, and establishes a systematic method for the functional genomics research of medicinal plants.

Key words: *Polygonatum zanlanscianense* Pamp.; histone methylation; gene family analysis; bioinformatics; epigenetics

0 引言

黄精, 百合科 (Liliaceae) 黄精属 (*Polygonatum*), 作为中国传统药食同源珍稀植物资源, 其药理价值与开发利用日益受到学界关注^[1]。传统中医学理论认为该属植物具有补气养阴、健脾润肺、益肾填精之功效; 而现代药理学研究进一步揭示了其具有多靶点生物活性。以多花黄精 (*P. cyrtonema*) 为例, 其具有显著降血糖、免疫调节、抗疲劳、抗肿瘤及广谱抗菌抗病毒等功效^[2]。近年来, 随着公众健康意识的显著增强及大健康产业快速发展, 黄精市场需求呈现指数级增长。然而, 野生资源因过度采挖已濒临枯竭, 人工栽培已逐步成为市场供给主体 (栽培区域已覆盖中国 21 个省份, 种植面积较十年前增长 380%), 但其规范化生产体系仍存在关键科学问题亟待解决: ① 药用活性成分 (如多糖、皂苷、黄酮等次生代谢产物) 的时空积累规律尚未明确; ② 最佳药用采收期缺乏客观量化标准; ③ 食用安全性与加工适性评价体系尚未建立。上述问题严重制约了黄精产业的可持续发展, 亟待从次生代谢产物积累规律、采收期科学判定及多组学调控机制等方面开展系统研究^[3]。

表观遗传学作为重要的基因表达调控机制, 通过非 DNA 序列改变介导可遗传的表型变化。其核心在于动态甲基化等化学修饰介导的基因表达模式跨代遗传性改变, 这一调控机制能在维持基因组稳定性的前提下实现生物性状的适应性调整^[4]。组蛋白共价修饰是染色质结构调控层面的关键调控方式, 通过甲基化/去甲基化平衡、乙酰化及磷酸化等修饰方式动态调控染色质结构, 进而精确调控基因转录活性。真核生物中已鉴定的组蛋白修饰类型超过 20 种, 除甲基化与乙酰化等经典修饰外, 还包括泛素化修饰及近年发现的 ADP-核糖基化等新型修

饰, 这些修饰通过协同作用形成动态调控网络^[5]。组蛋白修饰方式多样化在多细胞真核生物基因表达的表观遗传调控中起着重要作用^[6]。组蛋白甲基化受组蛋白甲基转移酶和去甲基化酶的动态调控, 其修饰位点主要集中于组蛋白的精氨酸 (R) 和赖氨酸 (K) 残基^[7]。近年来对组蛋白修饰的研究已逐渐成为表观遗传学关注的重点。组蛋白甲基化修饰通过动态调控核小体间相互作用介导染色质结构重塑, 该修饰通过改变核小体三维拓扑结构, 精确协调 DNA 复制、修复及转录激活等分子事件^[8-10]。组蛋白甲基化修饰分子效应延伸至表观遗传调控网络, 通过定向修饰特定组蛋白残基 (如 H3K4me3/H3K27me3) 建立基因表达模式, 从而调控细胞增殖分化进程并整合胞内信号转导网络^[11]。在植物表观遗传调控中, 组蛋白 H3 和 H4 的赖氨酸 (K) 及精氨酸 (R) 残基的甲基化修饰呈现高度特异性和多样性。H3 亚基的 K4、K9、K27、K36、K79 位点以及 H4K20 位点构成赖氨酸甲基化的主要靶标, 而 H3R2、R8、R17、R26 和 H4R3 则是精氨酸甲基化的主要作用位点^[12]。甲基化程度呈现显著差异: 赖氨酸残基可实现单甲基化 (me1)、双甲基化 (me2) 和三甲基化 (me3), 而精氨酸残基仅能形成单甲基化或双甲基化修饰。值得注意的是, 精氨酸甲基化表现为动态可逆的修饰过程, 其甲基转移酶 (PRMTs) 与去甲基化酶 (如 JMJD6) 共同维持修饰平衡; 相比之下, 赖氨酸去甲基化酶 (如 JmjC 家族) 效率低, 且部分高阶修饰 (如 H3K4me3) 需依赖组蛋白替换而非直接去甲基化, 导致修饰更持久^[13]。组蛋白甲基化转移酶属于具有特定结构域的蛋白家族, 具有多种不同的识别位点^[14]。例如, 动物和植物中的组蛋白赖氨酸甲基化转移酶通常都含有 SET 结构域 (Su (var)3-9, E(z), Trithorax), 但动物中还会含有 Dot1 结构域, 组蛋白赖氨酸甲基化转移酶能够催化赖氨

酸 H3 和 H4 甲基化为单甲基化、二甲基化和三甲基化^[15]。

在组蛋白甲基转移酶(HMTs)的功能研究中,Su(var)3-9 homologs(SUVH)家族作为含有 SET 结构域蛋白的重要分支,现有研究已逐步揭示其在表观遗传调控中的关键作用。SUVH 家族广泛存在于真核生物中,植物特有的 SUVH 亚家族通过其 SRA(SET and RING-associated)结构域识别甲基化修饰的 DNA 实现 SUVH 亚家族蛋白的招募,被招募的 SUVH 蛋白利用其 SET 结构域催化组蛋白 H3K9 位点的甲基化,形成 H3K9me2 修饰,从而介导基因沉默和异染色质稳定性维持^[16-17]。研究人员发现,拟南芥 SUVH 家族成员(如 SUVH4/5/6)在 DNA 甲基化与组蛋白修饰的互作中发挥桥梁作用:SUVH4 偏好结合 CAG/CTG 甲基化位点,而 SUVH5 倾向 CCG 甲基化,这种序列特异性介导的组蛋白修饰可能通过调控靶基因所在区域的染色质影响靶基因的表达从而影响次生代谢通路。此外,SUVH6 还具有自抑制调控机制,在 SUVH6 的结构生物学研究中发现,其肽底物结合通道受到自身的自抑制调控,为理解 HMTs 的精细调控提供了结构生物学证据^[18]。另有研究表明,组蛋白甲基化也能够通过调控基因表达水平调节器官的发育、种子的萌发以及植物对逆境的耐受性等多个生物学过程^[19]。作为表观遗传调控的重要表征,组蛋白修饰对功能的影响呈现明显的位点特异性:H3K9 乙酰化与 H3K4 甲基化(H3K4me)作为转录激活标记,而 H3K9/H3K27 位点的多甲基化修饰则与基因沉默相关。其中 H3K9me3 通过稳定异染色质结构抑制转录,H3K9me2 则介导常染色质区的可逆性沉默^[20]。在作物抗逆机制研究中,通过基因组学筛选鉴定到 *TaHMT21*、*TaHMT24* 和 *TaHMT42* 等甲基转移酶基因的表达在干旱胁迫下被显著诱导,揭示了其在杂交小麦抗旱分子育种中的价值^[21]。此外,H3K27me3 在植物发育时序调控中发挥关键作用,特别是通过光周期、春化等途径调控 *FLC* 的表达,进而影响 *FLY* 和 *AP1* 等成花基因的转录^[22]。SDG26 是拟南芥中的组蛋白甲基转移酶,SDG26 功能缺失可提升拟南芥抗旱性^[23],而茶树中过表达 CsSDG36/43 通过激活氮吸收相关 NRT2 家族基因,从而间接提升低氮胁迫下的氮同化效率^[24]。

截至目前,对黄精的研究还不够深入,对于黄精组蛋白甲基化相关基因家族鲜有研究报道。尽管目

前黄精属植物中 SUVH 家族的直接研究尚未见报道,但基于模式植物的功能推测,SUVH 可能参与黄精次生代谢产物合成的调控。例如,在药用植物中,H3K9me2 修饰常与次生代谢基因簇的沉默相关,而 SUVH 介导的甲基化可能通过抑制转座子活性以维持基因组的稳定性^[25]。本研究利用生物信息学方法以湖北黄精(*Polygonatum zanlanscianense* Pamp.)二代转录组和全长三代转录组为数据支撑,对湖北黄精组蛋白甲基化酶基因家族进行鉴定和分析。期望通过对湖北黄精组蛋白甲基化酶的深入挖掘进一步加深对黄精耐逆性形成及具有药用价值次生代谢产物合成机制的理解,对品种选育和抗逆性形成的机理研究都具有重要意义,也期望能够借此研究为其他中草药的研究积累经验和方法。

1 材料与方法

1.1 数据来源及软件工具

本课题以湖北黄精二代转录组和全长三代转录组数据(实验室未发表数据)为支撑;拟南芥基因组数据下载自植物基因组数据库 Phytozome V13.0 (<https://phytozome-next.jgi.doe.gov>)、在线 motif 鉴定工具 (<https://meme-suite.org/meme/>),MEGA(v7.0.26)^[26]用于绘制基因进化树,TBtools(v1.098685)^[27]用于数据提取及格式转换, TMHMM 软件用于蛋白质跨膜结构的分析预测 (<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>),在线工具 PSORT 用于亚细胞定位分析(<http://psort1.hgc.jp/form.html>)。

1.2 实验方法

1.2.1 湖北黄精组蛋白甲基化转移酶基因家族鉴定

本研究以湖北黄精三代全长转录组数据为基础,从 Pfam 库中下载序号为 PF00856 组蛋白甲基化酶的隐马尔科夫模型,然后再从 Phytozome V13.0 库中下载拟南芥组蛋白甲基化转移酶基因家族的蛋白序列文件。使用组蛋白甲基化酶的隐马尔可夫模型 PF00856 为模板构建库,利用 HMMER(v3.3.1)程序将湖北黄精蛋白质序列信息与 Pfam 模板进行对比,筛选出湖北黄精转录组中含有组蛋白甲基化转移酶结构的候选组蛋白甲基化转移酶基因家族成员基因 ID 及序列。

将 HMMER 生成的基于隐马尔科夫模型鉴定结果与 BLAST 比对结果进行交叉验证,确保组蛋白甲基化转移酶基因识别的准确性。首先,使用 HMMER 软件中的 hmmsearch 工具以 PF00856 为模

板对三代转录组注释的蛋白序列进行鉴定得到基于隐马尔科夫模型的候选基因;使用BLAST工具以拟南芥组蛋白甲基化转移酶基因家族蛋白序列为目标对湖北黄精的三代转录组注释蛋白序列文件进行搜索得到基于BLAST的候选基因。随后,将BLAST结果与hmmsearch的结果进行交叉验证,取两种工具共同鉴定到的蛋白序列作为最终确定湖北黄精的组蛋白甲基化转移酶基因家族蛋白。这种方法有效结合了两种不同的鉴定工具,能够提高基因家族鉴定的准确性和全面性,确保所识别的基因ID更具代表性。

1.2.2 湖北黄精组蛋白甲基化基因家族蛋白序列提取

使用TBtools的Fasta Extract功能,从湖北黄精的转录组序列中提取到组蛋白甲基化转移酶基因家族的蛋白序列。具体步骤如下:将湖北黄精的蛋白质文件导入比对文件中,并以湖北黄精组蛋白甲基化转移酶的基因ID作为比对目标。利用Fasta Header Pattern Match功能进行搜索,提取出湖北黄精组蛋白甲基化转移酶家族的蛋白序列。以相同的方法,从拟南芥的蛋白质序列文件中提取出对应的蛋白序列。

1.2.3 组蛋白甲基化基因家族成员进化分析

将提取出来的湖北黄精组蛋白甲基化蛋白文件与拟南芥组蛋白甲基化转移酶蛋白文件进行合并,使用MEGA(v7.0.26)软件中的Clustal W对蛋白文件中的序列进行多序列比对,根据多序列比对结果,使用Trimal对结果进行剪枝去除保守性极低的比对序列,将比对的结果使用MEGA进行进化分析并生成进化树,调整参数,利用枚举法进行检测,节点鲁棒性设置为1000个重复,将遗传距离选择为poisson correction。

1.2.4 湖北黄精SUVH家族蛋白跨膜结构及亚细胞定位分析

使用TMHMM在线工具对湖北黄精SUVH基因家族的氨基酸序列进行分析,找到含有跨膜结构的基因ID并进行分析。使用PSORT prediction在线工具对湖北黄精SUVH基因家族蛋白进行亚细胞定位预测。

1.2.5 湖北黄精SUVH基因家族蛋白motif分析

使用TBtools提取湖北黄精SUVH基因家族蛋白序列,使用MEME在线工具对SUVH基因家族蛋白序列进行motif鉴定,设置motif值为40,得到湖北黄精SUVH基因家族蛋白motif分布模式。

2 结果与分析

2.1 家族成员鉴定

本研究通过HMMER和BLAST两种工具同时对湖北黄精组蛋白甲基化酶进行鉴定,并对HMMER生成的基于隐马尔科夫模型PF00856鉴定结果与BLAST比对结果进行交叉验证,在Pfam库中进行验证后得到湖北黄精的组蛋白甲基化转移酶家族基因。根据文献已报道的47个拟南芥组蛋白甲基化转移酶,本研究中共鉴定出95个湖北黄精组蛋白甲基化转移酶蛋白,综合催化结构域的系统发育分析、功能结构域的特征、底物特异性及进化保守性等多维度原则,95个湖北黄精组蛋白甲基化基因被分为5大类,包括I(SUVH家族)、II(Trx家族)、III(SMYD家族)、IV(Ez)以及V(SETD家族)。本研究对湖北黄精组蛋白甲基化基因家族的SUVH亚家族成员进行进一步的生物信息学分析,其基因信息见表1。其蛋白序列长度多样性丰富,最长的transcript/857.p1转录本能够编码1268个氨基酸,最短的transcript/35100.p1转录本只能编码161个氨基酸。分子量最大的有142.958 kDa,分子量最小的只有18 kDa,等电点最高的为9.10,最低的为4.77。

2.2 基因进化分析

本研究对湖北黄精和拟南芥两种植物的组蛋白甲基化基因家族的蛋白进行进化分析,并构建完整的进化树(图1)。该环形进化树清晰呈现了湖北黄精(transcript前缀序列,如transcript/626.p1、transcript7071.p1等)与拟南芥(AT前缀序列,如AT3G04380.1、AT5G43990.2等)组蛋白甲基化基因家族的蛋白进化关系,在拓扑结构与节点支持率方面,进化树以中心放射状向外延伸,分支连接点标注的100/97/90等数值代表基于1000次重复枚举法检测的节点支持率(即Bootstrap值)。高支持率分支(如100、97)集中于主干区域,反映核心进化路径具有高度可靠性;部分次级分支支持率较低(如90),表明其拓扑结构可能存在一定变异。在基因序列分布与亲缘关系方面,湖北黄精序列(如transcript/13382.p1、transcript/12601.p1)多分布于独立分支,与拟南芥序列(如AT2G24740.1、AT1G1770.1)形成显著分化,表明两者在进化历程中经历了显著的功能或结构歧化。部分拟南芥序列聚集成簇(如AT3G04380.1与AT5G43990.2位于相邻分支),提示其蛋白功能保守性较高;而湖北黄精序列分散性较强,可能反映物种特异性。

根据湖北黄精与拟南芥SUV类蛋白的进化关

表1 湖北黄精 *SUVH* 家族基因信息表

Table 1 Gene information table of the *Polygonatum zanlanscianense SUVH* family

序列 ID	编码氨基酸数目(aa)	分子量(Da)	等电点(PI)
transcript/2238.p1	1 198	130 770.00	8.54
transcript/1950.p1	1 198	130 803.09	8.54
transcript/2854.p1	1 109	122 207.13	6.38
transcript/6639.p1	919	103 205.71	9.04
transcript/7098.p1	927	104 205.96	9.07
transcript/35100.p1	161	18 000.46	4.77
transcript/14888.p1	455	50 673.65	6.33
transcript/2706.p1	927	104 290.27	9.1
transcript/15741.p1	656	72 644.88	5.81
transcript/14956.p1	656	72 644.88	5.81
transcript/15424.p1	651	72 337.39	5.37
transcript/15844.p1	454	50 585.35	5.33
transcript/18245.p3	213	23 968.46	6.81
transcript/21193.p1	571	63 453.27	8.83
transcript/18304.p1	608	67 411.77	8.83
transcript/19439.p1	568	63 037.74	8.71
transcript/19303.p1	604	66 868.34	8.54
transcript/15638.p1	684	75 439.18	8.55
transcript/13782.p1	670	74 053.62	6.71
transcript/14413.p1	684	75 356.11	8.47
transcript/14777.p1	684	75 425.15	8.55
transcript/15267.p1	679	74 765.54	8.65
transcript/13012.p1	662	73 153.99	8.14
transcript/12601.p1	662	73 095.96	8.27
transcript/7060.p1	662	73 125.94	8.14
transcript/591.p1	366	41 168.91	5.12
transcript/857.p1	1 268	142 958.64	8.04
transcript/626.p1	1 267	142 868.19	7.83
transcript/13382.p1	663	73 255.10	8.14
transcript/8845.p1	670	74 097.63	6.56
transcript/13506.p1	668	73 912.44	6.56
transcript/8969.p1	574	63 859.21	6.79
transcript/6097.p1	672	74 493.22	5.67
transcript/9643.p1	823	91 470.79	5.67
transcript/8862.p1	823	91 498.85	5.67
transcript/6593.p1	837	93 515.15	5.55
transcript/7071.p1	836	93 278.48	5.74
transcript/7522.p1	837	93 349.56	5.74
transcript/7157.p1	837	93 334.55	5.67

系,参考已有研究^[28]将蛋白划分为5个组别,Group I 包含: AT1G73100.1 (AtSUVH3)、AT5G04940.1 (AtSUVH1)、AT1G17770.1 (AtSUVH7) 等5个拟南芥 *SUVH* 蛋白,以及 transcript/591.p1、transcript/3782.p1、transcript/14413.p1 等14个湖北黄

精 *SUVH* 转录本; Group II 包含2个拟南芥 *SUVH* 蛋白: AT2G33290.1 (AtSUVH2)、AT4G13460.1 (AtSUVH9) 以及 transcript/15741.p1、transcript/15844.p1、transcript/19303.p1 等8个湖北黄精 *SUVH* 转录本。两组蛋白均参与异染色质沉默,其

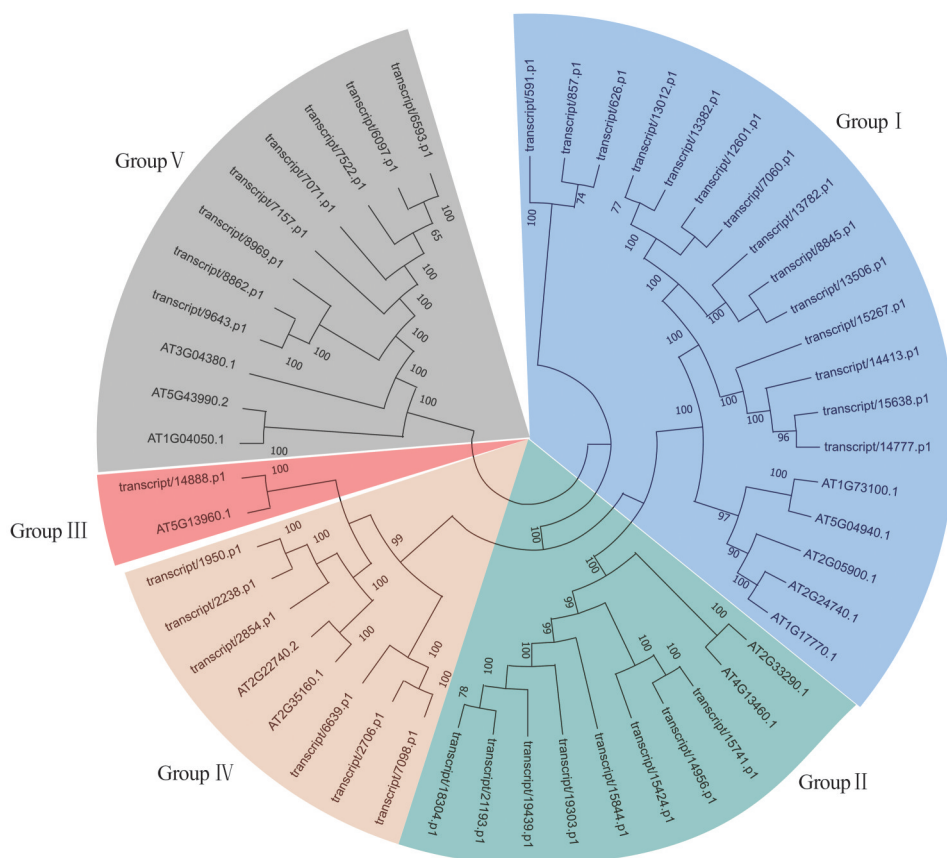


图1 湖北黄精和拟南芥 SUVH 蛋白进化树

Figure 1 Protein evolutionary analysis of the SUVH family of *Polygonatum zanlanscianense* Pamp. and *Arabidopsis*

中 Group I 占据主导作用, Group II 行使次要功能。Group III 由拟南芥 SUVH 蛋白 AT5G13960.1 (AtSUVH4) 与湖北黄精 SUVH 转录本 transcript/14888.p1 组成; Group IV 包含 2 个拟南芥 SUVH 蛋白: AT2G22740.2 (AtSUVH6)、AT2G35160.1 (AtSUVH5) 以及 transcript/2854.p1、transcript/6639.p1、transcript/2706.p1 等 6 个湖北黄精 SUVH 转录本。Group III 和 Group IV 两组蛋白主要行使二甲基转移酶功能。Group V 包含 3 个拟南芥 SUVH 蛋白: AT3G04380.1 (AtSUVH4)、AT5G43990.2 (AtSUVH2)、AT1G04050.1 (AtSUVH1) 以及 transcript/2854.p1、transcript/6639.p1、transcript/2706.p1 等 6 个湖北黄精 SUVH 转录本, 但这组蛋白功能暂不明确。

2.3 跨膜结构分析

TMHMM 在线工具对湖北黄精 39 个 SUVH 基因家族蛋白序列的分析结果显示, 仅有一个转录本 (transcript/6593.p1) 所编码的蛋白包含跨膜结构域。该转录本共编码了 837 个氨基酸, 其中第 1 至 812 位氨基酸位于膜外, 第 813 至 835 位的肽段形成跨膜螺旋, 而第 836 至 837 位氨基酸则位于膜内, 见图 2。

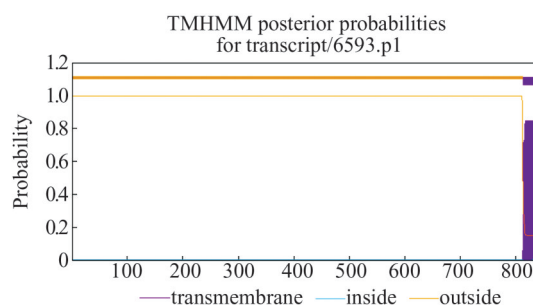


图2 湖北黄精 SUVH 基因家族蛋白跨膜结构

Figure 2 Transmembrane proteins structure of *Polygonatum zanlanscianense* Pamp. SUVH gene family

2.4 亚细胞定位分析

通过 PSORT prediction 在线工具对湖北黄精 SUVH 基因家族蛋白进行亚细胞定位预测, 该家族蛋白主要分布于细胞核、细胞质和线粒体 3 类细胞器中, 其中细胞核定位优势显著 (表 2)。在 34 个蛋白中, 26 个蛋白 (占比 76.47%) 在细胞核的预测分值超过 0.50, 表明其具有明确的核定位倾向。transcript/857.p1 和 transcript/626.p1 表现出完全的核定位特征 (分值均为 1.00), 表明这些基因可能参与核内特异性调控过程。另有 transcript/15844.p1 等基因呈现多细胞器泛定位特征,

其核定位分值为 0.26、细胞质 0.44、线粒体 0.30, 显示出潜在的功能多样性。细胞质定位的典型代表为 transcript/15741.p1(0.74)和 transcript/14956.p1(0.74), 而具有显著的线粒体定位特征的是 transcript/21193.p1(0.52)和 transcript/1834.p1(0.57)。部分基因在不同细胞器间的分值分布呈现平衡特征, 如 transcript/18245.p3 在细胞核、细胞质、线粒体的分值分别为 0.35、0.35、0.26, 未显示出明显的细胞器定位趋势。该分析为确定湖北黄精 *SUVH* 基因家族存在的亚细胞位置提供了参考依据。

2.5 motif 鉴定与分析

本研究通过 MEME 在线工具对湖北黄精 *SUVH* 基因家族进行 motif 鉴定, 得到 *SUVH* 基因家族编码氨基酸序列所具有的 motif。结果显示, 该基因家族编码的氨基酸序列中, 保守结构域的空间分布与其系统发育关系存在显著相关性。亲缘关系邻近

的转录本(如 transcript2238.p1 与 transcript1950.p1) 显示出高度一致的保守结构域组成(通过相同颜色标记的 Motif 分布模式体现), 其 motif 类型、数量及相对排列顺序均保持稳定(横轴位置 0~1 400 范围内的彩色矩形块连续分布)。具有相同保守结构域组合的转录本可能编码同一蛋白质亚基, 这种现象反映在纵轴相邻转录本间 motif 图谱的重复性(如 Motif 32 和 Motif 35 的多转录本重复出现模式)。多数转录本(如 transcript2238.p1、transcript1950.p1 等)在 3' 端(横轴右侧约 1 200~1 400)均以 Motif 1 和 Motif 18 作为保守终止元件, 形成特征性双 motif 结尾结构, 见图 3。

综上所述, 亲缘关系越近的基因其序列越保守, 所含有的结构域几乎没有差异, 他们可能编码同一个蛋白质亚基, 也暗示了其功能可能存在冗余。另外不同蛋白中的 motif 的顺序不完全一致, 但大多数以 motif 1 和 motif 18 作为 3' 端的结尾。



图 3 湖北黄精 *SUVH* 基因家族 motif 分析

Figure 3 Motif analysis of the *Polygonatum zanlanscianense* Pamp. *SUVH* gene family

3 讨论

3.1 SUVH家族的功能进化与表观遗传调控作用

SUVH家族作为组蛋白甲基转移酶的重要分支,其成员通过SET结构域催化组蛋白H3K9位点的甲基化(H3K9me2),进而介导基因沉默和染色质稳定性的维持^[25]。本研究中系统发育分析显示,湖北黄精SUVH基因与拟南芥同源基因分化显著(图1),提示其在进化中可能经历了功能特异性适应。例如,湖北黄精transcript/13382.p1等基因分布于独立进化分支,可能与次生代谢产物调控相关,而拟南芥SUVH4/5/6则偏好结合特定甲基化位点以维持基因组稳定性^[29]。值得注意的是,湖北黄精SUVH基因的分散性进化模式可能反映了物种对复杂生境的适应性策略,这一现象与ASH1L等大分子量甲基转移酶在哺乳动物中的功能歧化相似^[30]。

3.2 SUVH家族在植物逆境响应中的潜在作用

跨膜结构分析显示,仅transcript/6593.p1具有跨膜螺旋(图2),提示其可能参与膜结合信号转导。这与拟南芥SUVH6的自抑制调控机制形成对比,后者通过构象变化调节酶活性^[31]。此外,亚细胞定位预测表明76.47%的SUVH基因定位于细胞核(表2),与组蛋白修饰的核心功能一致。其中,transcript/857.p1和626.p1的完全核定位特征(分值1.00)暗示其在染色质重塑中的关键作用,类似于拟南芥SDG26通过H3K27me3调控发育时序的机制^[32]。在逆境响应方面,本研究检测到transcript/21193.p1(线粒体定位)和transcript/18245.p3(多细胞器动态分布)可能参与能量代谢与胁迫信号整合,这一发现与茶树CsSDG36/43通过氮同化增强抗逆性的机制存在功能相似性^[33]。

3.3 次生代谢调控与药用价值关联

Motif分析揭示了SUVH基因的保守结构域组合特征(图3),尤其是3'端Motif1和Motif18的终止元件模式,可能与底物识别特异性相关。例如,transcript/2238.p1与1950.p1具有相同的motif分布模式,表明二者可能参与协同调控特定次生代谢基因簇。已有研究表明,H3K9me2修饰常与药用植物(如多花黄精)次生代谢基因沉默相关^[34],而SUVH介导的甲基化可能通过抑制转座子活性维持基因组稳定性^[35]。本研究中鉴定到多个具有细胞质定位的SUVH基因(如transcript/15741.p1),其功能可能涉及非组

表2 湖北黄精SUVH基因家族亚细胞定位

Table 2 *Polygonatum zanlanscianense* Pamp. SUVH gene family subcellular localization

基因序号	可能存在的细胞器		
	细胞核	细胞质	线粒体
transcript/2238.p1	0.87		0.13
transcript/195.p1	0.87		0.13
transcript/2854.p1	0.61		0.39
transcript/6639.p1	0.83	0.04	0.13
transcript/798.p1	0.91	0.04	0.04
transcript/351.p1	0.26	0.57	
transcript/14888.p1	0.35	0.57	
transcript/276.p1	0.91	0.04	0.04
transcript/15741.p1	0.09	0.74	
transcript/14956.p1	0.09	0.74	
transcript/15424.p1	0.13	0.70	
transcript/15844.p1	0.26	0.44	0.30
transcript/18245.p3	0.35	0.35	0.26
transcript/21193.p1	0.39	0.09	0.52
transcript/1834.p1	0.35	0.09	0.57
transcript/19439.p1	0.52	0.04	0.35
transcript/1933.p1	0.30	0.48	0.13
transcript/15638.p1	0.57	0.13	0.22
transcript/13782.p1	0.74	0.22	0.04
transcript/14413.p1	0.57	0.13	0.22
transcript/14777.p1	0.57	0.13	0.22
transcript/15267.p1	0.78	0.09	0.09
transcript/1312.p1	0.57	0.22	0.13
transcript/1261.p1	0.57	0.22	0.13
transcript/76.p1	0.57	0.22	0.13
transcript/591.p1	0.22	0.70	0.04
transcript/857.p1	1.00		
transcript/626.p1	1.00		
transcript/13382.p1	0.57	0.22	0.13
transcript/8845.p1	0.74	0.22	
transcript/1356.p1	0.78	0.22	
transcript/8969.p1	0.74	0.26	
transcript/697.p1	0.70	0.09	0.09
transcript/9643.p1	0.65	0.22	0.04
transcript/8862.p1	0.65	0.22	0.04
transcript/6593.p1	0.30	0.22	0.13
transcript/771.p1	0.83	0.09	
transcript/7522.p1	0.83	0.09	
transcript/7157.p1	0.83	0.09	

蛋白底物的甲基化修饰,这一现象在ASH1L的C端结构域功能研究中也有类似报道^[30]。

4 结 论

本研究共鉴定到 95 个湖北黄精组蛋白甲基化转移酶 SET 结构域基因,进化分析显示其可分为 *SUVH*、*Trx*、*SMYD*、*Ez* 和 *SETD* 5 个大类。湖北黄精 *SUVH* 基因亚家族中共鉴定出 39 个基因,这些基因编码的氨基酸残基的数量在 161 aa~1 268 aa,等电点在 4.77 PI~9.10 PI,分子量大小在 18 kDa~142 kDa,其中编码酸性蛋白质的基因有 20 个,编码碱性蛋白质的基因有 19 个。在跨膜蛋白预测的结果中,该基因家族中只有 transcript/6593.p1 所编码的蛋白具有一个跨膜结构域,该基因共编码 837 位氨基酸,其中第 1 至 812 位氨基酸位于膜外,第 813 至 835 位氨基酸位于跨膜螺旋。

基因进化树的结果显示,湖北黄精 *SUVH* 基因家族与拟南芥 *SUVH* 基因家族具有较高的同源性,两者皆为双子叶植物,在进化上亲缘关系较近,但在湖北黄精的 *SUVH* 家族中的基因数目比拟南芥多,这可能与湖北黄精生长的过程中,除了合成自身需要的物质之外,还需要合成一些特有的次生代谢产物,因此组蛋白甲基化调控的过程可能较拟南芥更为复杂。

参考文献

- [1] 王若男,厉荣玉,郑鹏,等. 黄精多糖微生物发酵提取、表征及其抗氧化活性分析[J]. 中国食品添加剂, 2022, 33(9): 54-62.
Wang R N, Li R Y, Zheng P, et al. Microbial fermentation extraction, characterization and antioxidant activity of *Polygonatum* polysaccharide [J]. China Food Additives, 2022, 33(9): 54-62.
- [2] 赵文莉,赵晔, Tseng Yiider. 黄精药理作用研究进展[J]. 中草药, 2018, 49(18): 4439-4445.
Zhao W L, Zhao Y, Tseng Y. Research progress on pharmacological effects of *Polygonati Rhizoma* [J]. Chinese Herbal Medicine, 2018, 49(18): 4439-4445.
- [3] 崔阔澍,肖特,李慧萍,等. 我国黄精种质资源研究进展[J]. 江苏农业科学, 2021, 49(11): 35-39.
Cui K S, Xiao T, Li H P, et al. Research progress of China's *Polygonatum germplasm* resources [J]. Jiangsu Agricultural Sciences, 2021, 49(11): 35-39.
- [4] Wolffe A P, Guschin D. Chromatin structural features and targets that regulate transcription [J]. Journal of Structural Biology, 2000, 129(2/3): 102-122.
- [5] 康静婷,梁前进,梁辰,等. 表观遗传学研究进展[J]. 科技导报, 2013, 31(19): 66-74.
Kang J T, Liang Q J, Liang C, et al. Overview on rpi-genetics and its progress [J]. Science & Technology Review, 2013, 31(19): 66-74.
- [6] Millán-Zambrano G, Burton A, Bannister A J, et al. Histone post-translational modifications—cause and consequence of genome function [J]. Nature Reviews Genetics, 2022, 23(9): 563-580.
- [7] Yang L L, Zhang M, Guo J L, et al. Epigenetic regulation of histone methyltransferase SUV39H1 on the expression of recombinant protein in CHO cells [J]. Applied Microbiology and Biotechnology, 2025, 109(1): 1-10.
- [8] To T K, Kim J M. Epigenetic regulation of gene responsiveness in *Arabidopsis* [J]. Frontiers in Plant Science, 2014, 4: 548
- [9] Pikaard C S, Mittelsten Scheid O. Epigenetic regulation in plants [J]. Cold Spring Harbor Perspectives in Biology, 2014, 6(12): a019315.
- [10] Van Lijsebettens M, Grasser K D. Transcript elongation factors: shaping transcriptomes after transcript initiation [J]. Trends in Plant Science, 2014, 19(11): 717-726.
- [11] 王维,孟智启,石放雄. 组蛋白修饰及其生物学效应[J]. 遗传, 2012, 34(7): 19-27.
Wang W, Meng Z Q, Shi F X. Modification and biological role of histone [J]. Hereditas, 2012, 34(7): 19-27.
- [12] Stallcup M R. Role of protein methylation in chromatin remodeling and transcriptional regulation [J]. Oncogene, 2001, 20(24): 3014-3020.
- [13] Pham H Q H, Tao X, Yang Y. Protein arginine methylation in transcription and epigenetic regulation [J]. Frontiers in Epigenetics and Epigenomics, 2023, 1: 1245832.
- [14] 史红娟,刘杰,李群锋,等. 组蛋白甲基转移酶的生物学功能及肿瘤相关性研究[J]. 医学理论与实践, 2024, 37(19): 3280-3283.
Shi H J, Liu J, Li Q F, et al. Study on biological function of histone methyltransferase and its correlation with tumor [J]. The Journal of Medical Theory and Practice, 2024, 37(19): 3280-3283.
- [15] 陈晓晴,陈亮. 组蛋白 H3K79 甲基化转移酶 DOT1L 的功能研究进展[J]. 高师理科学刊, 2022, 42(7): 71-75.
Chen X Q, Chen L. Research progress on the function of histone H3K79 methyltransferase DOT1L [J]. Journal of Science of Teachers' College and University, 2022, 42(7): 71-75.
- [16] 潘辉,罗学刚,郭姝,等. 组蛋白甲基化及其与肿瘤间的关系[J]. 生理科学进展, 2010, 41(1): 22-26.
Pan H, Luo X G, Guo S, et al. Histone methylation and its relationship with cancer [J]. Progress in Physio-

- logical Sciences, 2010, 41(1): 22-26.
- [17] Li X, Harris C J, Zhong Z, et al. Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation [J]. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(37): E8793-E8802.
- [18] Johnson L M, Du J, Hale C J, et al. SRA-and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation [J]. Nature, 2014, 507(7490): 124-128.
- [19] 和珊, 杨贵伟, 黄莹, 等. 植物DNA甲基化及其生物学功能研究进展[J]. 农村科学实验, 2024, (2): 165-167.
- He S, Yang G W, Huang Y, et al. Research progress of plant DNA methylation and its biological functions [J]. Rural Scientific Experiment, 2024, (2): 165-167.
- [20] Parent J S, Cahn J, Herridge R P, et al. Small RNAs guide histone methylation in *Arabidopsis* embryos [J]. Genes & Development, 2021, 35(11/12): 841-846.
- [21] 陈佳敏, 刘永杰, 马锦绣, 等. 小麦组蛋白甲基化酶在杂交种中干旱胁迫表达模式分析[J]. 生物技术通报, 2022, 38(7): 51-61.
- Chen J M, Liu Y J, Ma J X, et al. Expression pattern analysis of histone methyltransferase under drought stress in hybrid wheat [J]. Biotechnology Bulletin, 2022, 38(7): 51-61.
- [22] 刘永平, 杨静, 杨明峰. 植物开花调控途径[J]. 生物工程学报, 2015, 31(11): 1553-1566.
- Liu Y P, Yang J, Yang M F. Pathways of flowering regulation in plants[J]. Chinese Journal of Biotechnology, 2015, 31(11): 1553-1566.
- [23] 马惠, 刘博宇, 阮颖, 等. 拟南芥组蛋白甲基化SDG26基因突变体抗旱生理生化研究[J]. 湖南农业大学学报(自然科学版), 2012, 38(4): 377-380.
- Ma H, Liu B Y, Ruan Y, et al. Physiological and biochemical studies on *Arabidopsis* mutant with the loss of SDG26 gene function under drought stress [J]. Journal of Hunan Agricultural University (Natural Sciences), 2012, 38(4): 377-380.
- [24] 陈清华. 茶树SDG基因家族成员鉴定及CsSDG36、CsSDG43基因功能分析[D]. 武汉: 华中农业大学, 2021.
- Chen Q H. Identification of SDG gene family members and functional analysis of CsSDG36 and CsSDG43 genes in *Camellia sinensis* [D]. Wuhan: Huazhong Agricultural University, 2021.
- [25] 谢萍, 田春艳, 张令强, 等. 组蛋白甲基转移酶的研究进展[J]. 遗传, 2007, 29(9): 1035-1041.
- Xie P, Tian C Y, Zhang L Q, et al. Progress in the study of histone methyltransferases [J]. Hereditas, 2007, 29(9): 1035-1041.
- [26] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. Molecular Biology and Evolution, 2016, 33(7): 1870-1874.
- [27] Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data [J]. Molecular Plant, 2020, 13(8): 1194-1202.
- [28] Hu M M, Li M D, Wang J B. Comprehensive analysis of the SUV gene family in allopolyploid *Brassica napus* and its diploid ancestors [J]. Genes, 2021, 12(12): 1848.
- [29] Yang Y, Yin X, Yang H, et al. Histone demethylase LSD2 acts as an E3 ubiquitin ligase and inhibits cancer cell growth through promoting proteasomal degradation of OGT [J]. Molecular Cell, 2015, 58(1): 47-59.
- [30] Vann K R, Sharma R, Hsu C C, et al. Structure-function relationship of ASH1L and histone H3K36 and H3K4 methylation [J]. Nature Communications, 2025, 16(1): 2235.
- [31] Zhang S, Zhou B, Kang Y, et al. C-terminal domains of histone demethylase JM14 interact with a pair of NAC transcription factors to mediate specific chromatin association [J]. Cell Discovery, 2015, 1(1): 1-13.
- [32] Zhu K H, Chen J C, Zhao L, et al. Dynamic control of H2A. Zub and H3K27me3 by ambient temperature during cell fate determination in *Arabidopsis* [J]. Developmental Cell, 2025: S1534-5807(25)00204-7.
- [33] Wang Q, Liu P, Jing H, et al. JM127-mediated histone H3K9 demethylation positively regulates drought-stress responses in *Arabidopsis* [J]. New Phytologist, 2021, 232(1): 221-236.
- [34] Zhao Y, Liu G, Yang F, et al. Multilayered regulation of secondary metabolism in medicinal plants [J]. Molecular Horticulture, 2023, 3(1): 11.
- [35] Liu Z W, Shao C R, Zhang C J, et al. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci [J]. PLoS Genetics, 2014, 10(1): e1003948.