

本氏烟*NbGSLs*基因家族鉴定与*NbGSL2*基因克隆

孟祺炜

浙江师范大学生命科学学院, 浙江 金华

收稿日期: 2025年3月25日; 录用日期: 2025年5月7日; 发布日期: 2025年5月19日

摘 要

胼胝质合酶(CaLS), 即葡聚糖合酶样蛋白(GSL), 是调控植物胼胝质合成的关键酶, 参与生长发育及逆境响应过程。本研究通过生物信息学方法系统鉴定了本氏烟*NbGSLs*基因家族, 并对其染色体定位、蛋白理化特性、基因结构、保守结构域、系统发育关系、共线性及顺式作用元件进行全面分析。结果显示: 1) 基因组中共鉴定12个*NbGSLs*成员, 编码蛋白(1741~2136 aa)相对分子质量为202.72~243.90 kD, 理论等电点8.31~9.23, 均为碱性蛋白; 2) 亚细胞定位预测显示, 所有成员定位于胞间连丝; 3) 基因结构分析表明, 除*NbGSL1*外均含内含子, 且普遍具有 β -1,3-葡聚糖合酶和FKS1结构域; 4) 共线性分析揭示家族内存在2个串联重复和6个片段重复事件, 并与拟南芥*AtGSLs*家族形成3个共线性对; 5) 启动子区含多种逆境响应元件, 提示*NbGSLs*可能参与植物抗逆调控。基于表达预测筛选获得*NbGSL2*基因, 成功克隆其5727 bp编码区, 该基因编码1908个氨基酸, 具有典型的跨膜结构和丰富的 α -螺旋。本研究首次系统鉴定本氏烟*NbGSLs*基因家族并成功克隆*NbGSL2*, 为解析其在韧皮部特异性表达及生物学功能奠定理论基础。

关键词

本氏烟, 基因家族, 胼胝质合酶, 生物信息学分析

Genome-Wide Identification of the *NbGSLs* Gene Family and Cloning of *NbGSL2* Gene in *Nicotiana benthamiana*

Qiwei Meng

College of Life Sciences, Zhejiang Normal University, Jinhua Zhejiang

Received: Mar. 25th, 2025; accepted: May 7th, 2025; published: May 19th, 2025

Abstract

Callose Synthases (CalS), also known as Glucan Synthase-Like proteins (GSL), are pivotal enzymes regulating callose deposition during plant development and stress responses. This study systematically identified the *NbGSLs* gene family in *Nicotiana benthamiana* through bioinformatics approaches, with comprehensive analyses of chromosomal distribution, protein physicochemical properties, gene structure, conserved domains, phylogenetic relationships, collinearity, and *cis*-acting elements. Key findings include: 1) Twelve *NbGSLs* members were identified, encoding alkaline proteins (1741~2136 aa) with molecular weights of 202.72~243.90 kD and theoretical pI values of 8.31~9.23; 2) Subcellular localization predictions indicated plasmodesmata targeting for all members; 3) Structural analysis revealed intron-containing genes (except *NbGSL1*) harboring conserved β -1,3-glucan synthase and FKS1 domains; 4) Collinearity analysis identified 2 tandem and 6 segmental duplication events within the family, with 3 orthologous pairs detected between *NbGSLs* and *Arabidopsis AtGSLs*; 5) Promoter regions contained multiple stress-responsive elements, suggesting *NbGSLs*' potential roles in abiotic stress adaptation. *NbGSL2*, predicted to exhibit phloem-specific expression, was successfully cloned with a 5727-bp coding sequence encoding a 1908-amino-acid transmembrane protein rich in α -helices. This study is the first one to systematically identify the *NbGSLs* gene family in *N. benthamiana* and successfully clone *NbGSL2*, which lays a theoretical foundation for analyzing its specific expression and biological function in the phloem.

Keywords

Nicotiana benthamiana, Gene Family, Callose Synthase, Bioinformatics Analysis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

植物细胞壁作为植物生长发育和逆境响应的动态界面，不仅为细胞提供机械支撑和物理屏障，更通过精密调控细胞与环境互作，参与多种生理过程[1]。其复杂组分中，由 β -1,3-葡萄糖苷键聚合而成的胼胝质(Callose)因其独特的时空特异性沉积特征，在植物生命周期中发挥多重生物学功能[2]。胼胝质的生物合成依赖于一类关键酶——胼胝质合酶(Callose Synthases, CalS)，该酶属于进化上高度保守的葡聚糖合酶样蛋白(Glucan Synthase-Like, GSL)家族[3]。目前已在多种植物中系统鉴定出 GSL 基因家族成员：拟南芥(*Arabidopsis thaliana*) *AtGSLs* (12 个)、水稻(*Oryza sativa*) *OsGSLs* (10 个)、甘蓝型油菜(*Brassica napus*) *BnCalSs* (32 个)、葡萄(*Vitis vinifera*) *VvCalSs* (8 个)以及棉花(*Gossypium* spp.) *Gh/Gb/Gr/GaCalSs* (15~28 个)等[3]-[7]。

分子进化分析表明，GSL 蛋白具有高度保守的 FKS1 结构域和 β -1,3-葡聚糖合酶催化域，其跨膜拓扑结构与纤维素合酶(Cellulose Synthase, CES)类似[3]。通过亲水环结构域与其他调控蛋白互作形成功能性复合体，GSLs 在细胞质膜上精确调控胼胝质的动态沉积[8]。这一过程对植物生长发育至关重要：从调控胞间连丝(Plasmodesmata)孔径的分子大小排除极限(Size Exclusion Limit, SEL) [9]，到维持气孔发育[10]、花粉母细胞减数分裂[2]以及筛管物质运输[11]等关键生理过程均需胼胝质的精确调控。例如，拟南芥 *AtGSL2* 通过控制小孢子初生细胞壁的胼胝质沉积确保雄性配子发育[2]，而水稻 *OsGSL5* 的功能缺失导致

花粉母细胞减数分裂异常,最终引发雄性不育[12]。值得注意的是,胼胝质沉积的时空失衡亦会引发病理效应,如高温胁迫下水稻叶片 *OsGSL1*、*OsGSL6*、*OsGSL7*、*OsGSL8* 和 *OsGSL9* 的异常表达导致胞间连丝阻塞,抑制光合同化物运输并最终降低产量[11]。

除发育调控外, *GSL* 基因家族在植物逆境应答中扮演双重角色:一方面,病原侵染诱导的胼胝质沉积可形成物理屏障限制病原扩散,如大豆 *GSLs* 通过调控韧皮部胼胝质沉积抑制花叶病毒的系统性传播[13];另一方面,过量沉积可能加剧生理紊乱,如柑橘黄龙病菌引发的筛板孔堵塞会破坏韧皮部运输稳态[14]。这种精细平衡凸显了系统解析 *GSL* 基因家族功能网络的重要性。

本氏烟草(*Nicotiana benthamiana*)作为植物分子生物学的模式体系,凭借其完善的遗传转化平台[15]、高质量基因组数据库[16]以及 CRISPR 编辑系统[17],在基因功能研究领域具有独特优势。尽管拟南芥 *AtGSL7* 已被证实通过调控筛管发育和损伤诱导的胼胝质沉积影响植物抗逆性[18],但本氏烟草 *NbGSLs* 基因家族的系统鉴定及其成员功能研究仍属空白。本研究通过全基因组水平鉴定 *NbGSLs* 家族成员,结合比较基因组学与进化分析筛选 *AtGSL7* 同源基因 *NbGSL2* 进行克隆与功能预测。研究结果将为解析 *GSL* 基因在茄科植物中的功能分化及作物遗传改良提供理论依据。

2. 材料与方法

2.1. *NbGSLs* 基因家族鉴定、亚细胞定位预测与理化性质分析

本氏烟基因组相关文件(*N. benthamiana* draft genome v2.6.1)下载自 Sol Genomics Net (<https://solgenomics.sgn.cornell.edu/>)数据库。Pfam 数据库(<http://pfam.xfam.org>)下载葡聚糖合酶结构域(pfam02364)的隐马尔可夫模型 hmm 文件,将获得成员的氨基酸序列提交至 NCBI-CD Search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)分析保守结构域[19],将没有葡聚糖合酶保守结构域的候选序列删除,共鉴定出 12 个独特的 *NbGSLs* 基因,根据其染色体位置命名为 *NbGSL1*~*NbGSL12*。EXPASy (<https://web.expasy.org/protparam/>)对 *NbGSLs* 家族成员氨基酸序列进行理化性质分析,预测属性涵盖蛋白氨基酸数量、分子量、等电点等[20]。WoLF PSORT (<https://www.genscript.com/wolf-psort.html>)预测亚细胞定位[21]。

2.2. 染色体定位、基因结构、保守结构域

通过 TBtools (v2.109)的 Graphics-Gene Location Visualize from GTF/GFF 模块对 *NbGSLs* 家族 12 个成员绘制染色体定位图谱,基因结构与保守结构域通过 Gene Structure View 模块实现整合[22]。

2.3. 系统发育分析、共线性分析与顺式作用元件分析

拟南芥 *AtGSLs* 蛋白序列下载自 TAIR (<https://www.arabidopsis.org/>)数据库。水稻 *OsGSLs* 家族蛋白序列下载自 NCBI (<https://www.ncbi.nlm.nih.gov/>)数据库。将 12 个 *NbGSLs*、12 个 *AtGSLs* 与 10 个 *OsGSLs* 蛋白序列整合至一起,提交 MAFFT version 7 (<https://mafft.cbrc.jp/alignment/server/index.html>)并以 Clustal format 形式导出结果。将结果在 MEGA (v7.0)中打开,使用邻接法(Bootstrap method 设置:1000)构建系统发育树[23]。

NbGSLs 物种内共线性分析由 TBtools 中的 Advanced Circos 模块完成。拟南芥的物种间共线性分析由 TBtools 中的 One Step MCScanX-Super Fast,通过提交两个物种的基因组文件与注释文件,将获得的 Ctl 文件按要求修改后,于 Dual Systemy Plot 模块提交 Ctl、Simplified GFF、Collinearity 三个文件,并对 *GSLs* 基因进行高亮,获得拟南芥与本氏烟 *GSL* 共线性结果[22]。

通过 TBtools 中的 Gtf/GFF3 Sequence Extract 模块抓取所有编码基因 ATG 上游 2000 bp 序列,再利

用 Fasta Extract (Recommended)模块从上述获得的文件中抓取 *NbGSLs* 家族 12 个成员的上游 2000 bp 序列, 再提交至 PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>), 随后获得各个基因上游顺式作用元件信息[24], 筛选后的结果通过 TBtools 的 Simple BioSequence Viewer 模块进行可视化。

2.4. 实验材料

本氏烟培养于人工气候室。

2.5. 本氏烟总 RNA 提取与 cDNA 合成

本氏烟总 RNA 提取参照 TIANGEN 总 RNA 提取试剂盒(DP419)说明书。1.5%琼脂糖凝胶电泳检测总 RNA 质量。*NbGSL2* 参考序列长达 5784 bp, 建议选用 Invitrogen™旗下 SuperScript™ IV 逆转录酶(18090200)完成长链 cDNA 合成, 方法参考说明书。

2.6. *NbGSL2* 基因克隆与生物信息学分析

根据从基因组中抓取到 *NbGSL2* (Niben261Chr01g0179010)的 CDS 序列, 设计克隆引物为 NbGSL-F: ATGGCGAGTACTAGTGGGAC; 下游引物 NbGSL-R: CAAGACTTCCACTTTGCTAC。引物由杭州有康生物科技有限公司合成。

使用诺唯赞 2 × KeyPo SE Master Mix (Dye Plus) (PK512)对 *NbGSL2* 进行基因克隆。PCR 反应体系(50 μL): 2 μL cDNA、2 μL NbGSL-F、2 μL NbGSL-R、25 μL Mix、19 μL dd H₂O; PCR 反应条件为: 95℃预变性 3 min, 95℃变性 30 s, 60℃退火 30 s, 72℃延伸 6 min, 35 个循环, 72℃最终延伸 10 min。1%琼脂糖凝胶电泳检测目的条带后, 使用 TIANGEN 琼脂糖凝胶回收试剂盒(DP214)回收 PCR 产物。将 PCR 产物与全式金 T&B Zero Cloning Kit 克隆载体(CTB501)在 25℃条件下连接 30 min。重组载体热激转化大肠杆菌 DH5α 后, 37℃摇床过夜培养。次日挑取单菌落至 500 μL 含有卡那霉素的 LB 培养基, 摇床培养 3 h 后, 将菌液送至擎科生物有限公司, 经通用引物 M13 双向测序。

TMHMM-2.0 (<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>)预测 *NbGSL2* 跨膜结构域。NCBI-CD Search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)分析 *NbGSL2* 保守结构域[19]。SOPMA (https://npsa.lyon.inserm.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)预测 *NbGSL2* 蛋白二级结构[25]; SWISS-MODE (<https://swissmodel.expasy.org>)同源建模 *NbGSL2* 蛋白三级结构[26]。

3. 结果与分析

3.1. 本氏烟 *NbGSLs* 家族鉴定与染色体定位

为鉴定本氏烟 *NbGSLs* 基因家族, 我们从 Sol Genomics Network (<https://solgenomics.net/>)下载本氏烟 *N. benthamiana* 基因组文件、基因组注释文件等, 结合葡聚糖合酶结构域的隐马尔可夫模型(pfam02364)抓取本氏烟中的 *GSLs* 基因序列与氨基酸序列。筛选掉不含葡聚糖合酶及其超家族结构域, 在剔除其结构域 N、C 端存在缺失的序列后, 从本氏烟中共鉴定到含有 12 个 *GSLs*, 分别位于本氏烟 9 条不同的染色体上(图 1), 我们按从小到大、从上到下的顺序, 分别结合染色体序号与基因在染色体上的位置, 将 12 名成员命名为 *NbGSL1*~*NbGSL12*。其中 *NbGSL1*、*NbGSL2* 和 *NbGSL3* 位于染色体 Chr01, *NbGSL4* 位于 Chr02, *NbGSL5* 位于 Chr05, *NbGSL6* 位于 Chr06, *NbGSL7* 位于 Chr09, *NbGSL8* 位于 Chr11, *NbGSL9* 位于 Chr12, *NbGSL10* 位于 Chr15, *NbGSL11* 和 *NbGSL12* 位于 Chr18。整体而言, 12 名成员不均匀分布在各条染色体上, 在 Chr1 中分布 3 个成员, Chr18 中分布 2 个, 其余都各只分布 1 个成员。

随后我们对其 CDS 长度、氨基酸长度、蛋白分子量、等电点和亚细胞定位预测等基本特征进行分析

(表 1)。12 个 *NbGSLs* 成员 CDS 最长可达 6411 bp，最低长度为 5226 bp，CDS 整体较长。蛋白的氨基酸长度范围为 1741 (*NbGSL8*)~2136 aa (*NbGSL10*)。分子量整体较大，大小浮动在 202.72 (*NbGSL8*)~243.90 kD (*NbGSL10*)。pI 范围为 8.31 (*NbGSL11*)~9.23 (*NbGSL5*)，表明 *NbGSLs* 成员为碱性蛋白质。亚细胞定位预测表明，*NbGSLs* 蛋白全部位于胞间连丝。

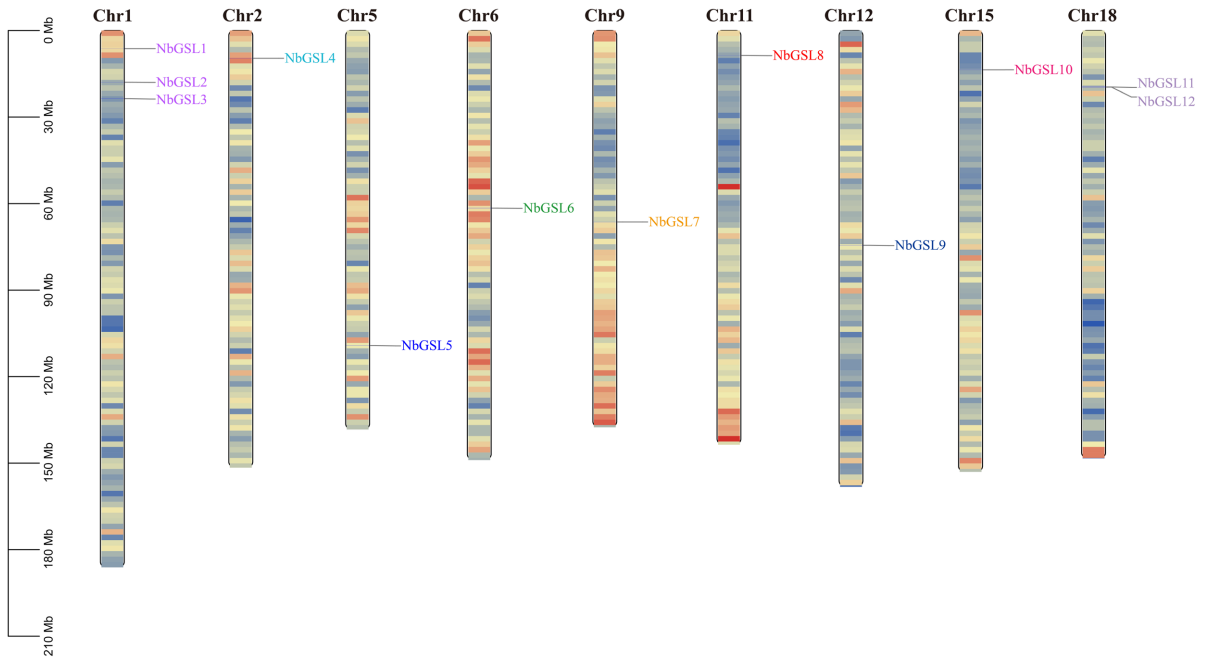


Figure 1. Chromosome mapping of the *NbGSLs* gene family in *Nicotiana benthamiana*

图 1. 本氏烟 *NbGSLs* 基因家族染色体定位

Table 1. Physical and chemical properties analysis of the *NbGSLs* gene family in *Nicotiana benthamiana*

表 1. 本氏烟 *NbGSLs* 基因家族理化性质分析

基因名称	基因号	编码区(bp)	氨基酸数(aa)	分子量(kD)	等电点	亚细胞定位
<i>NbGSL1</i>	Niben261Chr01g0064002	5307	1768	205.83	9.16	胞间连丝
<i>NbGSL2</i>	Niben261Chr01g0179010	5784	1927	223.92	8.85	胞间连丝
<i>NbGSL3</i>	Niben261Chr01g0236017	6114	2037	235.04	8.41	胞间连丝
<i>NbGSL4</i>	Niben261Chr02g0096002	5913	1970	227.79	9.15	胞间连丝
<i>NbGSL5</i>	Niben261Chr05g1092001	5871	1956	225.19	9.23	胞间连丝
<i>NbGSL6</i>	Niben261Chr06g0616013	5421	1806	206.98	8.94	胞间连丝
<i>NbGSL7</i>	Niben261Chr09g0663013	5736	1911	220.45	9.18	胞间连丝
<i>NbGSL8</i>	Niben261Chr11g0086021	5226	1741	202.72	9.17	胞间连丝
<i>NbGSL9</i>	Niben261Chr12g0745001	5694	1897	220.53	8.95	胞间连丝
<i>NbGSL10</i>	Niben261Chr15g0136001	6411	2136	243.90	9.09	胞间连丝
<i>NbGSL11</i>	Niben261Chr18g0196001	5964	1987	227.96	8.31	胞间连丝
<i>NbGSL12</i>	Niben261Chr18g0196002	5880	1959	226.59	9.13	胞间连丝

3.2. 基因结构与保守结构域分析

NbGSLs 基因家族鉴定到的 12 个成员的基因结构并非都属于断裂基因, *NbGSL1* 基因不含内含子(图 2)。成员间的基因长度也表现出大小不一的特点, 最短可达 5305 bp (*NbGSL8*), 最长可达 64068 bp (*NbGSL11*), 相差近 12 倍。除去 *NbGSL1* 与 *NbGSL8* 只含有 1、2 个外显子的特殊情况, 其他 10 位成员外显子数量浮动在 31~49 之间, 平均含有约 41 个外显子, CDS 编码区并未因基因序列长度、外显子与内含子数量上的差异而产生较大变化。蛋白保守结构域分析结果指出, 所有成员都具备葡聚糖合酶结构域与 FKS1 结构域, 除 *NbGSL1*、*NbGSL6*、*NbGSL8*、*NbGSL11* 四个蛋白外, 其余成员皆包含 Vta1 结构域, 后者占比 60%。整体上看, 该基因家族成员编码的蛋白在进化过程中保守性较强。



Figure 2. Gene structure and conserved domains of the *NbGSLs* genes in *Nicotiana benthamiana*

图 2. 本氏烟 *NbGSLs* 基因结构与保守结构域

3.3. 系统发育分析

拟南芥与水稻两种模式植物的 *GSLs* 基因家族已被鉴定, 将本氏烟 *NbGSLs* 家族成员蛋白序列与拟南芥 *AtGSLs*、水稻 *OsGSLs* 家族成员的蛋白序列构建系统发育树, 可进一步探索本氏烟 *NbGSLs* 家族成员与其他物种同源基因进化上的关系, 并可通过在拟南芥、水稻中已报道的基因功能来预测在本氏烟中同源基因的功能(图 3(a))。

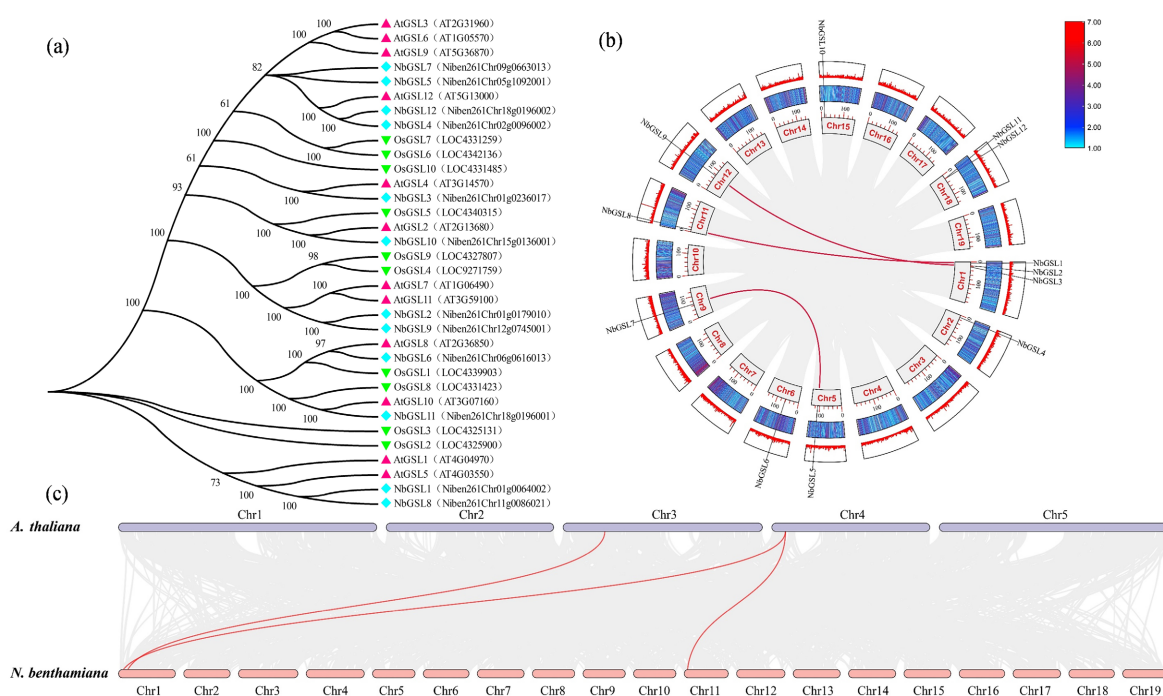
NbGSL10 是拟南芥 *AtGSL2* 与水稻 *OsGSL5* 的同源基因, 推测 *NbGSL10* 在维持减数分裂的正常运行过程中不可或缺, 其功能缺失也可能会造成花粉败育、雄性不育等。在拟南芥中已报道 *AtGSL7* 在韧皮部特异性表达, 负责调控韧皮部胼胝质的合成, 参与损伤应激等多种生理过程[10]。序列比对结果推测 *NbGSL2* 与 *NbGSL9* 可能也在韧皮部的生长发育和创伤应激等反应中发挥作用。

3.4. 共线性分析

共线性分析是揭示基因家族成员间进化关系的重要手段, 有助于理解基因复制事件及其在物种进化过程中的作用。为探索 *NbGSLs* 家族基因复制模式, 我们进行了共线性分析(图 3(b))。结果表明, *NbGSLs* 基因家族 12 个基因中有 2 个基因(*NbGSL11*、*NbGSL12*)为串联重复序列(Tandem duplication)。此外, 我们发现 6 个 *NbGSLs* 基因在 5 条染色体(Chr1、Chr5、Chr9、Chr11 和 Chr12)上存在片段重复(Segmental duplication),

存在 3 个显著的共线性对, 表明分别是 *NbGSL1* 与 *NbGSL8*、*NbGSL2* 与 *NbGSL9*、*NbGSL5* 与 *NbGSL7*。这些成对的基因可能来源于同一祖先基因, 经过物种形成过程中的基因复制事件, 保留了相似的基因序列, 这或许暗示它们在功能上的潜在冗余或分化, 成对的基因可能在植物的发育或应对环境压力具有协同作用, 但不一定都在某个组织或某个细胞内都表达, 可能因细胞分化而产生组织特异性表达的现象。

为进一步分析本氏烟 *NbGSLs* 基因家族与其他物种的进化关系, 我们选择了同为双子叶的拟南芥为对象, 分析了 *NbGSLs* 与拟南芥 *AtGSLs* 两个物种间的共线性关系(图 3(c)): *NbGSL1* 与 *AtGSL5* (AT4G03550)、*NbGSL3* 与 *AtGSL4* (AT3G14570)、*NbGSL8* 与 *AtGSL5* (AT4G03550)。有趣的是, *NbGSL1* 与 *NbGSL8* 不仅在物种内呈现共线性, 还在与拟南芥 *AtGSL5* 呈现物种间的共线性, 这或许暗示它们在进化上源自相同的祖先。



注: (a) 水稻、拟南芥、本氏烟 *GSLs* 家族系统发育树; (b) 本氏烟 *NbGSLs* 基因家族物种内共线性分析; (c) 本氏烟与拟南芥 *GSLs* 家族物种内共线性分析。

Figure 3. Phylogenetic analysis and synteny analysis

图 3. 系统发育分析与共线性分析

3.5. 顺式作用元件鉴定与分析

顺式作用元件作为基因的侧翼序列, 配合转录激活因子、抑制因子等反式作用因子, 在调控基因本底水平、抑制或高表达等方面至关重要, 调控元件并非都位于转录起始位点(Transcription Start Sites, TSS)上游, 属于核心启动子的十基序元件(Motif Ten Element, MTE)与下游启动子元件(Downstream Promoter Element, DPE)都位于 TSS 下游[27]。为了避免忽略 TSS 下游的调控元件, 我们抓取了 12 个 *NbGSLs* 家族成员 ATG 上游 2000 bp 序列, 并提交至 PlantCARE 分析元件类型。在剔除了增强子、常见的核心启动子与光响应元件后, 最终筛选得到 20 个元件, 主要分为激素应答元件、逆境胁迫应答元件与生长发育元件, 部分元件具有交叉反应(图 4)。

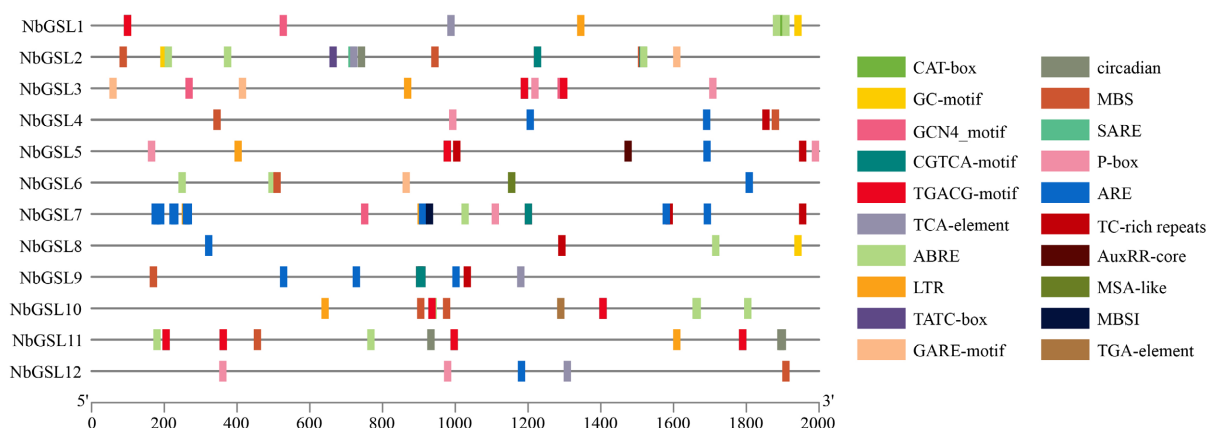


Figure 4. Analysis of cis acting elements in members of the *NbGSLs* gene family of *Nicotiana benthamiana*

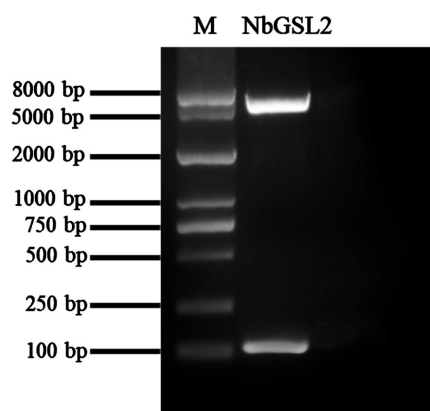
图 4. 本氏烟 *NbGSLs* 基因家族成员顺式作用元件分析

10 种激素应答元件中, CGTCA-motif、TGACG-motif 参与茉莉酸甲酯信号通路; SARE、TCA-element 参与水杨酸信号通路; ABRE 参与脱落酸信号通路; P-box、TATC-box、GARE-motif 参与赤霉素信号通路; TGA-element、AuxRR-core 参与生长素信号通路; 水杨酸、茉莉酸甲酯与脱落酸是常见的逆境胁迫激素[28]。在非生物和生物胁迫下, 植物经常引发胼胝质的积累, 这个过程在几秒钟内发生[1], 聚焦于胼胝质在胁迫中的作用, 观察到除 *NbGSL4* 外, 其余 9 位成员皆有水杨酸、茉莉酸甲酯和脱落酸应答元件, 且数量最多是 *NbGSL2* (9 个)与 *NbGSL11* (10 个)。植物在长期进化中已经形成了一套有体系的响应机制来应对各式各样的胁迫条件。TC-rich repeats 参与防御和应激反应, ARE 为厌氧胁迫下的应答元件, GC-motif 应答缺氧胁迫, 而 MBS 与 LTR 分别参与干旱与低温胁迫的响应。5 种参与生长发育的调控元件共同确保了植物生理稳态的正常运作, 它们涵盖了细胞周期的精细调控、近似昼夜节奏的精确控制, 以及对初生代谢与次生代谢过程的有效调节。CAT-box 为分生组织基因表达相关元件, GCN4-motif 参与胚乳中相关基因表达, circadian 参与昼夜节律控制, MSA-like 参与细胞周期调控, MBSI 则参与黄酮的生物合成。在开花植物中, 胼胝质对于细胞分裂不可或缺[29], 胼胝质在花药中的出现是减数分裂开始的显著组织学标志[12]。

分析基因上游调控元件能够得知该基因能够响应哪些信号的刺激, 进而推测该基因是否参与发育、胁迫或细胞周期等过程。

3.6. *NbGSL2* 基因克隆与序列分析

考虑到 *NbGSL2* 与 *NbGSL9* 存在共线性关系, 两条序列一致性高达 93.79%, 且在起始与终止密码子处都保持高相似度, 这为克隆 *NbGSL2* 带来了困难。为了解决这个问题, 我们经 TBtools 通过基因组注释文件抓取了 *NbGSL2*、*NbGSL9* 终止密码子下游 200 bp 的 UTR 序列, 通过对 UTR 的序列比对寻找靶点, 成功设计了特异性引物对 *NbGSL2*-F/*NbGSL2*-R, 并以本氏烟 cDNA 为模版克隆到 *NbGSL2* (图 5), 经单克隆菌液测序后得到其 CDS 全长 5727 bp, 编码 1908 个氨基酸。这与前文对基因组中抓取的 *NbGSL2* 参考序列有所出入, 克隆序列 CDS 较参考序列缺少 57 bp, 相应缺少 19 个氨基酸。对两者序列进行比对可以发现, 在参考序列的 852~935 bp 与 4218~4386 bp 两处序列在基因组注释文件中被标注为外显子, 而实际克隆序列中则是未被剪接上的内含子(图 6)。另有一处则与之相反, 克隆序列 1203~1298 bp 区段实为外显子, 但并不存在于参考序列上, 被误注释为内含子。*NbGSL2* 的成功克隆为本氏烟基因组注释文件的修正与完善提供了数据支撑。



注: M: DL 8000 DNA Marker; *NbGSL2*: *NbGSL2* 目的片段扩增产物。

Figure 5. Amplification of the full-length cDNA sequence of *NbGSL2* genes

图 5. *NbGSL2* 基因 cDNA 全长序列扩增

3.7. *NbGSL2* 理化性质与亲疏水性分析

使用 ExPASy 分析 *NbGSL2* 克隆序列的氨基酸理化性质, 蛋白分子式 $C_{10148}H_{15637}N_{2671}O_{2772}S_{81}$, 相对分子质量 222.01 kD。正、负电荷氨基酸数量分别为 233 个、213 个, 等电点 pI 为 8.85, 蛋白呈碱性。不稳定指数计算为 39.29 (<40), 表明 *NbGSL2* 属于稳定蛋白。脂肪族指数 92.46, 亲水性平均系数为 -0.131, 表明蛋白具有较高的疏水性与具有一定的亲水性, 可能与脂质相互作用或在细胞膜中稳定存在。跨膜结构预测表明 *NbGSL2* 具有 14 个跨膜结构(图 7(a))。亚细胞定位预测结果与参考序列预测结果一致, 仍旧定位于胞间连丝。NCBI CD-Search 表明, *NbGSL2* 克隆序列存在 β -1,3-葡聚糖合酶、FKS1 与 Vta1 三个结构域, 这与参考序列预测一致(图 7(c))。综合分析表明 *NbGSL2* 在很大程度上是定位在胞间连丝处的跨膜蛋白。

3.8. *NbGSL2* 蛋白二级、三级结构预测

水孔蛋白是一类通常形成四聚体以行使其转运水或甘油等物质的跨膜蛋白, 每个亚基上具有 6 个 α -螺旋[30]。为更好地表征 *NbGSL2* 跨膜性质, 我们通过预测 *NbGSL2* 的二级结构, 同时对其三级结构进行同源建模, 以期能从三维结构角度分析其属性, 对其属于跨膜蛋白的理论依据提供进一步的数据支撑。

通过 SOPMA 预测 *NbGSL2* 蛋白二级结构, 结果显示 *NbGSL2* 存在 α -螺旋、无规则卷曲、延伸链 3 种二级结构, 不存在 β -折叠(图 7(b))。其中有 1264 个氨基酸参与 α -螺旋(占比 66.25%), 564 个氨基酸参与无规则卷曲(占比 29.56%), 80 个氨基酸参与延伸链(占比 4.19%), 推测 *NbGSL2* 存在大量的 α -螺旋可能是其在胞间连丝处作为跨膜蛋白的结构基础。

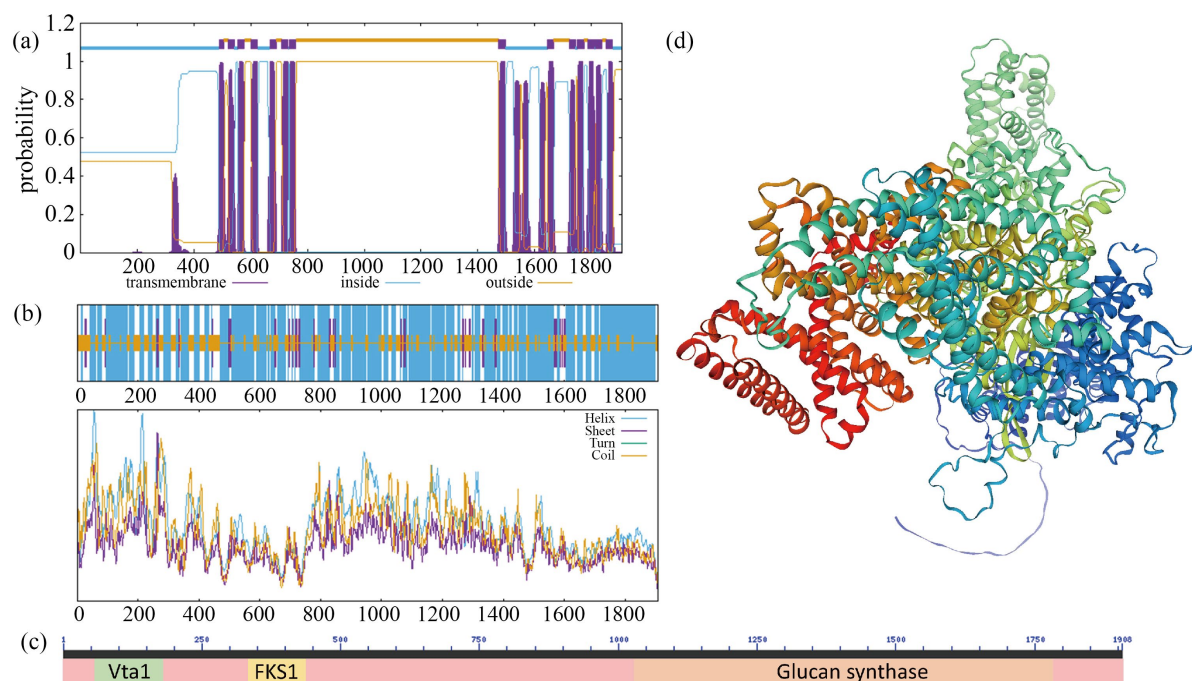
蛋白的结构与其功能相适应, 通过 SWISS-MODEL 对 *NbGSL2* 蛋白空间结构进行同源建模以期获得其高级结构(图 7(d)), 以 AlphaFold v2 算法匹配度最高的模型蛋白为拟南芥 *AtGSL7* (一致性 73.21%), GMQE 为 0.80 表明模型具有较高的预测质量。*NbGSL2* 三级结构中存在大量 α -螺旋, 图中所示其结构底部能观察到明显的无规则卷曲, 这与二级结构预测结果吻合。此外, 匹配上拟南芥 *AtGSL7* 蛋白模型也与系统发育树中 *AtGSL7*、*NbGSL2* 在进化上呈现高度亲缘性结果一致。

4. 总结与讨论

本研究首次系统鉴定了本氏烟(*Nicotiana benthamiana*)胼胝质合酶 *NbGSLs* 基因家族成员。本氏烟作

Figure 6. Comparison of *NbGSL2* reference sequence and clone sequence
图 6. *NbGSL2* 参考序列与克隆序列比对

NbGSLs 家族成员分布于 9 条染色体, 其中 Chr1 和 Chr18 呈现基因簇分布, 暗示染色体片段重组可能在家族扩张中起关键作用。基因结构分析显示, 除 *NbGSL1* 外, 其余成员均含多个内含子(31~49 个), 断裂基因结构存在一定的复杂性。所有 *NbGSLs* 蛋白均携带 β -1,3-葡聚糖酶与 FKS1 结构域, 其跨膜拓扑特征与纤维素合酶高度相似, 印证了胼胝质合酶功能的进化保守性。值得注意的是, 66.7% 的成员额外



注：(a) *NbGSL2* 蛋白疏水性分析；(b) *NbGSL2* 二级结构；(c) *NbGSL2* 保守结构域；(d) *NbGSL2* 三级结构。

Figure 7. Analysis of hydrophobicity, conserved domains, secondary and tertiary structures of *NbGSL2* protein

图 7. *NbGSL2* 蛋白疏水性、保守结构域、二级结构与三级结构分析

含有 Vta1 结构域，该结构域在其他物种中多与囊泡运输相关[29]，结合所有成员被预测定位于胞间连丝的可能性，暗示了本氏烟 *NbGSLs* 蛋白在胞间连丝调控中的独特机制。

通过系统发育树分析，我们发现 *NbGSLs* 基因家族与拟南芥 *AtGSLs* 和水稻 *OsGSLs* 基因家族存在显著的进化关系，这为预测本氏烟中 *NbGSLs* 基因的功能提供了重要线索。在共线性分析方面，我们发现了两个重要的现象。首先，*NbGSLs* 基因家族中存在 2 个串联重复序列(*NbGSL11*、*NbGSL12*)，这表明在物种进化过程中，这些基因可能经历了基因复制事件，从而保留了相似的基因序列。此外，我们还发现 6 个 *NbGSLs* 基因在 5 条染色体上存在片段重复，形成了 3 个显著的共线性对(*NbGSL1* 与 *NbGSL8*、*NbGSL2* 与 *NbGSL9*、*NbGSL5* 与 *NbGSL7*)。这些成对的基因可能来源于同一祖先基因，暗示基因复制是植物适应复杂环境的重要进化策略。这种共线性关系不仅揭示了基因家族成员间的进化关系，也为我们理解基因复制事件及其在物种进化过程中的作用提供了新的视角。基因上游调控区富集 ABRE、MBS 等 12 类逆境响应元件，进一步支持 *NbGSLs* 在各类胁迫应答中的潜在作用，这与在其他物种中报道 GSL 参与高温响应和生物胁迫等应答反应相呼应[9] [11] [18]。

值得注意的是，*NbGSL2* 的成功克隆及其与拟南芥韧皮部特异性表达基因 *AtGSL7* 的系统发育同源性，为解析本氏烟韧皮部的发育机制提供了新线索。克隆序列与基因组注释的差异直接揭示了当前本氏烟基因组数据库的局限性，为后续注释修正提供了实验证据。蛋白结构分析表明，*NbGSL2* 具有 14 个跨膜结构域和 66.25% 的 α -螺旋，与其作为膜定位胍脂质合酶的功能特性高度吻合。与拟南芥 *AtGSL7* 结构一致性达 73.21%，结合系统发育同源性结果，共同支持了 *NbGSL2* 参与韧皮部胍脂质沉积调控的可能性。

本研究首次绘制了本氏烟 GSL 基因家族的功能图谱，其成果具有三重意义：在理论层面，通过比较基因组学揭示了 *NbGSLs* 在进化中的复制与分化规律，为理解植物胍脂质合酶家族的功能演化提供了新视角；在技术层面，*NbGSL2* 的成功克隆及注释误差的发现，凸显了验证工作对基因组注释的重要性，为

茄科植物基因功能研究提供了方法论参考；在应用层面，鉴定到的逆境响应元件与共线性基因对(如 *NbGSL2/AtGSL7*)，可作为关键靶点用于作物抗逆育种或韧皮部运输改良。未来研究可进一步结合 CRISPR/Cas9 技术与时空表达谱分析，深入解析 *NbGSL2* 基因功能。这些工作将有助于完善植物胼质合成调控网络，为作物抗逆育种提供新的分子靶点。

参考文献

- [1] Wang, B., Andargie, M. and Fang, R. (2022) The Function and Biosynthesis of Callose in High Plants. *Heliyon*, **8**, e09248. <https://doi.org/10.1016/j.heliyon.2022.e09248>
- [2] Shi, X., Han, X. and Lu, T. (2015) Callose Synthesis during Reproductive Development in Monocotyledonous and Dicotyledonous Plants. *Plant Signaling & Behavior*, **11**, e1062196. <https://doi.org/10.1080/15592324.2015.1062196>
- [3] Verma, D.P.S. and Hong, Z. (2001) Plant Callose Synthase Complexes. *Plant Molecular Biology*, **47**, 693-701. <https://doi.org/10.1023/a:1013679111111>
- [4] Feng, J., Chen, Y., Xiao, X., Qu, Y., Li, P., Lu, Q., *et al.* (2021) Genome-wide Analysis of the *CalS* Gene Family in Cotton Reveals Their Potential Roles in Fiber Development and Responses to Stress. *Peer Journal*, **9**, e12557. <https://doi.org/10.7717/peerj.12557>
- [5] Liu, F., Zou, Z. and Fernando, W.G.D. (2018) Characterization of Callose Deposition and Analysis of the Callose Synthase Gene Family of Brassica Napus in Response to Leptosphaeria Maculans. *International Journal of Molecular Sciences*, **19**, Article 3769. <https://doi.org/10.3390/ijms19123769>
- [6] Yu, Y., Jiao, L., Fu, S., Yin, L., Zhang, Y. and Lu, J. (2016) Callose Synthase Family Genes Involved in the Grapevine Defense Response to Downy Mildew Disease. *Phytopathology*, **106**, 56-64. <https://doi.org/10.1094/phyto-07-15-0166-r>
- [7] Yamaguchi, T., Hayashi, T., Nakayama, K. and Koike, S. (2006) Expression Analysis of Genes for Callose Synthases and Rho-Type Small GTP-Binding Proteins That Are Related to Callose Synthesis in Rice Anther. *Bioscience, Biotechnology, and Biochemistry*, **70**, 639-645. <https://doi.org/10.1271/bbb.70.639>
- [8] 张庆雯, 王兆昊, 祁静静, 等. 植物胼质合成酶研究进展[J]. 园艺学报, 2021, 48(4): 661-675.
- [9] Barratt, D.H.P., Kölling, K., Graf, A., Pike, M., Calder, G., Findlay, K., *et al.* (2010) Callose Synthase GSL7 Is Necessary for Normal Phloem Transport and Inflorescence Growth in Arabidopsis. *Plant Physiology*, **155**, 328-341. <https://doi.org/10.1104/pp.110.166330>
- [10] Xie, B. and Hong, Z. (2011) Unplugging the Callose Plug from Sieve Pores. *Plant Signaling & Behavior*, **6**, 491-493. <https://doi.org/10.4161/psb.6.4.14653>
- [11] Luan, R., Liu, J., Tao, L., Fu, G. and Zhang, C. (2023) Comparative Transcriptome Analysis Reveals OsBGs and OsGSLs Influence Sugar Transport through Callose Metabolism under Heat Stress in Rice. *International Journal of Molecular Sciences*, **24**, Article 3175. <https://doi.org/10.3390/ijms24043175>
- [12] Somashekar, H., Mimura, M., Tsuda, K. and Nonomura, K. (2022) Rice GLUCAN SYNTHASE-LIKE5 Promotes Anther Callose Deposition to Maintain Meiosis Initiation and Progression. *Plant Physiology*, **191**, 400-413. <https://doi.org/10.1093/plphys/kiac488>
- [13] Zhang, J., Liu, N., Yan, A., Sun, T., Sun, X., Yao, G., *et al.* (2022) Callose Deposited at Soybean Sieve Element Inhibits Long-Distance Transport of Soybean Mosaic Virus. *AMB Express*, **12**, Article No. 66. <https://doi.org/10.1186/s13568-022-01402-0>
- [14] Granato, L.M., Galdeano, D.M., D'Alessandre, N.D.R., Breton, M.C. and Machado, M.A. (2019) Callose Synthase Family Genes Plays an Important Role in the Citrus Defense Response to Candidatus Liberibacter Asiaticus. *European Journal of Plant Pathology*, **155**, 25-38. <https://doi.org/10.1007/s10658-019-01747-6>
- [15] Goodin, M.M., Zaitlin, D., Naidu, R.A. and Lommel, S.A. (2008) *Nicotiana benthamiana*: Its History and Future as a Model for Plant-Pathogen Interactions. *Molecular Plant-Microbe Interactions*, **21**, 1015-1026. <https://doi.org/10.1094/mpmi-21-8-1015>
- [16] Bombarely, A., Rosli, H.G., Vrebalov, J., Moffett, P., Mueller, L.A. and Martin, G.B. (2012) A Draft Genome Sequence of *Nicotiana benthamiana* to Enhance Molecular Plant-Microbe Biology Research. *Molecular Plant-Microbe Interactions*, **25**, 1523-1530. <https://doi.org/10.1094/mpmi-06-12-0148-ta>
- [17] 杨学飞, 刘亚娟, 高原, 等. CRISPR/Cas9 系统在本氏烟草基因敲除中的应用[J]. 分子植物育种, 2017, 15(1): 30-38.
- [18] Xie, B., Wang, X., Zhu, M., Zhang, Z. and Hong, Z. (2010) *CalS7* Encodes a Callose Synthase Responsible for Callose Deposition in the Phloem. *The Plant Journal*, **65**, 1-14. <https://doi.org/10.1111/j.1365-313x.2010.04399.x>
- [19] Yang, M., Derbyshire, M.K., Yamashita, R.A. and Marchler-Bauer, A. (2019) NCBI's Conserved Domain Database and

- Tools for Protein Domain Analysis. *Current Protocols in Bioinformatics*, **69**, 1-25. <https://doi.org/10.1002/cpbi.90>
- [20] Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V. and Durinx, C. (2021) Expasy, the Swiss Bioinformatics Resource Portal, as Designed by Its Users. *Nucleic Acids Research*, **49**, W216-W227. <https://doi.org/10.1093/nar/gkab225>
- [21] Horton, P., Park, K., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., *et al.* (2007) WoLF PSORT: Protein Localization Predictor. *Nucleic Acids Research*, **35**, W585-W587. <https://doi.org/10.1093/nar/gkm259>
- [22] Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., *et al.* (2023) TBtools-II: A “One for All, All for One” Bioinformatics Platform for Biological Big-Data Mining. *Molecular Plant*, **16**, 1733-1742. <https://doi.org/10.1016/j.molp.2023.09.010>
- [23] Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, **33**, 1870-1874. <https://doi.org/10.1093/molbev/msw054>
- [24] Lescot, M. (2002) Plantcare, a Database of Plant Cis-Acting Regulatory Elements and a Portal to Tools for *in Silico* Analysis of Promoter Sequences. *Nucleic Acids Research*, **30**, 325-327. <https://doi.org/10.1093/nar/30.1.325>
- [25] Geourjon, C. and Deléage, G. (1995) SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments. *Bioinformatics*, **11**, 681-684. <https://doi.org/10.1093/bioinformatics/11.6.681>
- [26] Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., *et al.* (2018) SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Research*, **46**, W296-W303. <https://doi.org/10.1093/nar/gky427>
- [27] 谢在春, 潘秋辉, 于永春, 等. 真核蛋白编码基因核心启动子元件[J]. 医学分子生物学杂志, 2007, 4(2): 132-135.
- [28] Yang, Y., Ahammed, G., Wu, C., Fan, S. and Zhou, Y. (2015) Crosstalk among Jasmonate, Salicylate and Ethylene Signaling Pathways in Plant Disease and Immune Responses. *Current Protein & Peptide Science*, **16**, 450-461. <https://doi.org/10.2174/1389203716666150330141638>
- [29] Schneider, R., Hanak, T., Persson, S. and Voigt, C.A. (2016) Cellulose and Callose Synthesis and Organization in Focus, What's New? *Current Opinion in Plant Biology*, **34**, 9-16. <https://doi.org/10.1016/j.pbi.2016.07.007>
- [30] Verkman, A.S. and Mitra, A.K. (2000) Structure and Function of Aquaporin Water Channels. *American Journal of Physiology-Renal Physiology*, **278**, F13-F28. <https://doi.org/10.1152/ajprenal.2000.278.1.f13>