

决明 GSK 基因家族全基因组鉴定及其表达分析

冯 昭^{1,2}, 覃 洋¹, 刘世鹏¹, 吕蕊花^{1,2}, 吕瑞华¹, 胡晓晨², 毛仁俊^{3*}, 张 岗^{2*}

(1. 陕西中医药大学医学技术学院, 陕西 咸阳 712046; 2. 陕西中医药大学药学院、陕西省中医药管理局“秦药”研发重点实验室, 陕西 咸阳 712046; 3. 延安大学生命科学学院, 陕西 延安 716000)

摘要: 糖原合成激酶 3 (glycogen synthase kinase 3/SHAGGY-like kinase, GSK3) 在调节植物生长发育和胁迫响应方面发挥重要作用。为揭示药用植物决明 (*Senna tora* L.) GSK 家族成员特性, 本研究基于其全基因组数据, 结合生物信息学及基因表达研究方法, 开展决明 GSKs 基因鉴定及表达分析。共鉴定到 9 个 *StoSKs* 基因家族成员, 均具有 GSK 特征激酶结构域, 9 个成员分布在 6 条染色体上, 编码氨基酸长度为 465~943 aa, 蛋白分子质量介于 33.57~88.83 kDa, 平均等电点为 8.2。 *StoSKs* 基因家族分为 4 个进化分支, 同一进化分支中的 *StoSKs* 之间具有相同的外显子/内含子结构及保守基序。 *StoSKs* 家族成员扩张主要源于片段重复事件, 与大豆 (*Glycine max*)、蒺藜苜蓿 (*Medicago truncatula*)、拟南芥 (*Arabidopsis thaliana*)、水稻 (*Oryza sativa*) 分别存在 17、11、8 和 7 对共线基因对。 *StoSKs* 启动子区域多含有与胁迫刺激、生长发育、激素诱导相关的响应元件。转录组数据分析发现 *StoSKs* 在不同组织中均有表达, 在根中表达水平最高。qRT-PCR (quantitative real-time PCR) 分析表明, 不同进化分支的 *StoSKs* 表现出协同响应光照的表达模式, 多数 *StoSKs* 能快速响应 NaCl 胁迫处理, 表达显著上调。研究结果为下一步进行决明 GSK 基因家族的生物学功能分析提供基础。

关键词: 决明; 糖原合成激酶 3 基因家族; 生物信息学; 基因表达分析; 盐胁迫

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2023)05-1383-12

Genome-wide identification of GSK gene family in *Senna tora* L. and expression analyses

FENG Zhao^{1,2}, QIN Yang¹, LIU Shi-peng¹, LÜ Rui-hua^{1,2}, LÜ Rui-hua¹, HU Xiao-chen², MAO Ren-jun^{3*}, ZHANG Gang^{2*}

(1. College of Medical Technology, Shaanxi University of Chinese Medicine, Xianyang 712046, China; 2. Key Laboratory for Research of “Qin Medicine” of Shaanxi Administration of Chinese Medicine, College of Pharmacy, Shaanxi University of Chinese Medicine, Xianyang 712046, China; 3. College of Life Science, Yan'an University, Yan'an 716000, China)

Abstract: Glycogen synthase kinase 3/SHAGGY-like kinase (GSK3) proteins play important roles in regulating plant growth, development, and stress response. In order to reveal the characteristics of GSK family members in the medicinal plant *Senna tora* L., in this study, we conducted the identification and expression analyses of GSKs in *S. tora* based on its whole genome data, combined with bioinformatics and gene expression research methods. The results showed that a total of nine *S. tora* GSK genes were identified, all of which contained the GSK characteristic kinase domains. All members were distributed on six chromosomes, the encoding amino acid length ranged from 465 to 943 aa, the protein molecular weight was from 33.57 to 88.83 kDa, and the average

收稿日期: 2022-10-18; 修回日期: 2022-11-23.

基金项目: 陕西省自然科学基金资助项目 (2021JQ-636); 陕西中医药大学校级科研课题 (2021GP07); 陕西中医药大学学科创新团队项目 (2019-QN01).

*通讯作者 Tel: 86-29-38185165, E-mail: jay_gumling2003@aliyun.com;

Tel: 86-911-2332030, E-mail: mrjnwsuaf@126.com

DOI: 10.16438/j.0513-4870.2022-1117

isoelectric point was 8.2. The *StoSKs* were divided into four evolutionary branches, and the *StoSKs* in the same evolutionary branch shared the same exon/intron structure and conserved motifs. The expansion of the *StoSKs* gene family was mainly due to segment duplication events, and there were 17, 11, 8 and 7 pairs of collinear genes with *Glycine max*, *Medicago truncatula*, *Arabidopsis thaliana* and *Oryza sativa*, respectively. The promoter regions of *StoSKs* mostly contained responses elements related to stress stimulation, growth and development, and hormone induction. Transcriptome data analysis showed that *StoSKs* were expressed in different tissues, with the highest expression level in roots. Quantitative real-time PCR (qRT-PCR) analysis indicated that *StoSKs* in different evolutionary branches displayed a synergistic expression pattern response to light, and most of *StoSKs* could rapidly respond to NaCl stress with significantly up-regulated expression. All the results provide a basis for further analysis of the biological functions of the *GSKs* gene family in *S. tora*.

Key words: *Senna tora* L.; glycogen synthase kinase 3 gene family; bioinformatics; gene expression analysis; salt stress

糖原合成激酶 3 (glycogen synthase kinase 3/SHAGGY-like kinase, GSK3) 作为一种在真核生物中高度保守的多功能激酶, 参与真核生物许多重要的信号转导途径^[1], 其最早被证明可在兔骨骼肌肉中降低糖原合成酶的活性^[2]。近 30 年的研究表明, 动物 GSK 以多基因家族形式存在, 广泛参与肿瘤发生、胚胎发育、昼夜节律、细胞命运决定等多种生理过程^[3,4]。

植物中同样普遍存在 GSK 家族。与动物相比, 植物基因组中 GSK 的数量更多 [如拟南芥 (*Arabidopsis thaliana*) 中有 10 个, 水稻 (*Oryza sativa*) 中有 9 个]、功能更加多样^[5]。大量研究表明^[6,7], GSK 参与植物生长发育、生理代谢及抗逆胁迫等生物学过程, 如种子萌发、根系发育和气孔形成, 以及对高盐、寒冷、干旱或病原菌感染等的生理适应。在拟南芥中, 盐胁迫可强烈诱导 *AtSK13*、*AtSK31* 和 *AtSK42* 表达^[8], 而过表达 *AtSK22* 通过系统性诱导盐胁迫响应基因增强植株的耐盐性^[9]。同时, GSK 还参与植物激素信号转导, 其在油菜素内酯 (BR) 信号转导中发挥负调控功能^[10]。拟南芥 *AtSK21* 通过抑制 BR 和脱落酸 (ABA) 信号调控种子萌发和根系发育^[11]。在光照条件下, 葡萄 (*Vitis vinifera*) 的 *VvSKs* 能介导 BR 信号的反馈调节^[12]。此外, GSK 被发现参与豆科植物-根瘤菌的共生过程。日本百脉根 (*Lotus japonicus*) 的 *LSK1* 和 *LSK2* 在根瘤菌感染后得以诱导表达^[13]。大豆 (*Glycine max*) 的 *GmSK2~8* 及其同一亚家族成员在高盐胁迫下表达上调, 却抑制根瘤形成^[14]。可见, GSK 在植物响应逆境胁迫过程中起着复杂且重要的作用。许多研究都旨在筛选鉴定各种植物中的 GSK 家族成员, 而到目前为止, 大多数药用植物中 GSK 基因家族鉴定与研究仍比较少。

决明 (*Senna tora* L.) 为豆科决明属一年生草本植物, 其种子入药系决明子, 始载于《神农本草经》, 列为

上品, 具有祛风除湿、清热明目、润肠通便等功效^[15]。现代中药化学研究揭示决明中含有蒽醌类、萘并吡喃酮类、苷类及脂肪酸类等多种活性成分, 具有降压、调血脂、抗氧化、抑菌等多种药理活性^[16,17]。除了药用价值, 决明还兼具食用和观赏等经济价值。决明原产于我国, 主要分布在河南、四川、安徽、浙江、陕西等多个省区, 然而其资源规模正在锐减^[18]。决明是喜光植物, 不宜在盐碱地种植, 低温、缺乏光照、盐胁迫等非生物胁迫会严重影响其生长^[19]。随着气候环境的不断恶化及人类活动的影响, 土壤盐渍化程度在全球范围内将不断加剧, 严重制约中药生态农业的发展。培育抗逆的决明新品种并提高药材品质是应对该问题的有效手段。本草基因组学领域在药用植物基因组、系统进化或分子辅助育种等方面取得了长足进展^[20]。2020 年, 韩国科学家 Sang-Ho Kang 等^[21]率先完成了决明基因组测序, 为决明抗逆育种、次生代谢调控及药材质量控制等提供基础。鉴于 GSK 基因在药用植物生长发育和逆境胁迫响应中的重要性, 本研究基于决明基因组数据开展决明 GSK 基因家族鉴定及其表达分析, 有助于下一步研究决明 GSK 基因功能, 进而为培育高品质决明新品种提供科学依据。

材料与amp;方法

植物材料与处理方法 以采自陕西中医药大学药用植物园的决明种子为试材, 经陕西中医药大学张雨曲博士鉴定为豆科植物决明 (*Senna tora* L.)。

选取大小均匀的决明种子, 经表面消毒处理后, 平铺于 2% 的水琼脂平板, 28 °C 暗培养条件下进行催芽 48 h。随后将发芽的幼苗移至装有土壤混合物 (珍珠岩: 蛭石体积比为 2: 1) 的培养袋, 置光照培养箱中培养 [(21 ± 2) °C, 光照强度 9 000 lx, 16 h/8h 光暗交替], 每袋种 2 株, 每 5 天加入适量霍格兰营养液。约 1 月,

幼苗长至4叶期,选取长势一致的植株,用200 mmol·L⁻¹ NaCl + 霍格兰营养液进行盐胁迫处理,对照组用霍格兰营养液浇灌,处理植株0、1、3、6和12 h后收集根,每个时间点收集3袋小苗;同时,取12袋幼苗分两组,分别置于9 000 lx和黑暗条件下,12 h后收集根。所有样本均迅速用液氮冷冻,-80 °C保存备用。

决明 GSK 基因家族的鉴定 分别从TAIR (<http://www.arabidopsis.org/>)、EnsemblPlants (<http://plants.ensembl.org/index.html>) 和 NABIC (http://nabic.rda.go.kr/Species/Senna_tora2) 下载拟南芥、水稻和决明 GSK 蛋白数据。以拟南芥和水稻 GSK 蛋白序列作为参考,通过 BLAST 比对 (E-value ≤ e⁻¹⁰, identity ≥ 50%), NCBI-CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 和 SMART (<http://smart.embl-heidelberg.de/>) 验证,鉴定得到决明 GSK 基因和编码蛋白序列,按照其与拟南芥、水稻等 GSK 基因系统发育聚类顺序进行命名。

多序列比对和系统发育分析 选择拟南芥 (*Arabidopsis thaliana*)、水稻 (*Oryza sativa*)、大麦 (*Hordeum vulgare*)、日本百脉根 (*Lotus japonicus*)、大豆 (*Glycine max*)、决明 (*Senna tora*) 6 个物种的 GSK 蛋白序列,利用 ClustalX (<http://www.ebi.ac.uk/Tools/msa/muscle/>) 工具进行多序列比对分析。同时,利用 MEGA-X 分析软件,采用邻接法 (Neighbor-joining method; 1 000 bootstrap 重复),构建系统发育树。

染色体定位、基因结构、保守结构域和蛋白结构分析 利用决明基因组注释文件信息,通过 TBtools 软件^[22],分析得到已鉴定的决明 *StoSK* 基因在染色体的分布情况。使用 Gene Structure Display Server 2.0 程序 (<http://gsds.cbi.pku.edu.cn/>) 分析 *StoSK* 基因结构。同时,利用在线工具 MEME (Multiple Expectation Maximization for Motif Elicitation, <http://meme-suite.org/tools/meme>) 预测蛋白保守基序 (motif)。利用 NPS@: SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa%20_sopma.html) 和 Swiss-

Model (<http://swissmodel.expasy.org>) 分别进行 *StoSKs* 的蛋白质二级结构和三级结构预测。

共线性分析与基因重复事件分析 使用 MescanX 分析拟南芥、水稻、百脉根、大豆和决明 GSK 基因在 5 个物种间及决明物种内的共线性关系。序列覆盖率 > 90% 且相似性 > 95% 被认为属于串联重复事件。共线性结果使用 TBtools 软件绘制^[22]。

顺式作用元件分析 获取决明 *StoSKs* 基因上游 2 kb 序列作为假定的启动子区域,提交至 PlantCARE 数据库 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>),经筛选聚类分析,通过 TBtools 展示启动子区域的顺式元件。

基因表达分析 从 NCBI Sequence Read Archive (SRA) 数据库 (<https://www.ncbi.nlm.nih.gov/sra/>) 下载决明根、叶、幼嫩种子和成熟种子全转录组数据 (SAMN09657134),使用 TBtools 转录组数据分析功能,通过 FastQC 质检、Trimmomatic 去除接头,使用 Kallisto 进行基因表达分析,最终绘制表达热图^[22]。

使用 RNA 提取试剂盒 (TIANGEN 公司) 提取决明根总 RNA, SMARTScribe™ 反转录酶 (TAKARA 公司) 合成制备模板 cDNA,进行 *StoSKs* 基因的 qRT-PCR (quantitative real-time PCR) 分析。使用 20 μL 反应体系,包含 1 μL cDNA 模板、0.5 μL 上下游引物、10 μL SYBR 预混液及 8 μL DEPC 水。使用 ABI 7300 进行 qRT-PCR 反应,程序为: 95 °C 5 min, 95 °C 5 s, 60 °C 35 s, 40 个循环。使用 *StoActin* 作为内参基因,通过 2^{-ΔΔCT} 方法^[23] 计算表达量,特异性引物见表 1。

结果与分析

1 决明 GSK 基因家族成员的鉴定

以 10 个拟南芥和 9 个水稻 GSK 家族成员的氨基酸序列为参考,通过与决明基因组进行 BLASTp 比对,再经 Swiss-Port 数据库鉴定,最终得到 9 个含有 GSK 保守结构域的基因。9 个 *StoSKs* 基因分布于 6 条染色体上,8 号染色体上分布有 3 个基因,13 号染色体上分

Table 1 Primers used for quantitative real-time PCR (qRT-PCR) in this study. *StoSKs*: The glycogen synthase kinase 3/SHAGGY-like kinase (*GSK3*) genes from *S. tora* (*Sto*)

Gene	Gene ID	Primer sequence (5'-3')		Size/bp
		Forward	Reverse	
<i>StoSK1.1</i>	Sto08g234550	AGCATGCGCACACTCCTT	CGACACAACCTAGCGGGT	129
<i>StoSK1.2</i>	Sto08g235490	GCGTTGACAAGTTGCCCCG	AGTCTGCTTTGGCTGGCC	149
<i>StoSK1.3</i>	Sto08g250140	GGATGCCCCAGAGAAGCAG	AGCGTCCAAAGCTGTGCA	80
<i>StoSK1.4</i>	Sto13g440710	CAGCACCCACACCACCTT	CTGTTGCGCCATCTCCA	114
<i>StoSK2.1</i>	Sto07g210470	TGCCCGACTTCCCAATGG	TGACCACTCTGCCGCTA	123
<i>StoSK2.2</i>	Sto13g424070	GCACACCAACTCGGGAGG	CCGGGGGCATCCTTTTGT	114
<i>StoSK3.1</i>	Sto01g015100	AGCATGCGCACACTCCTT	CGACACAACCTAGCGGGT	129
<i>StoSK3.2</i>	Sto09g286690	CGGCCGCTACTGCTTTA	ACTCCTCGCATGCTCAGG	108
<i>StoSK4.1</i>	Sto02g050930	TGCACTGCACTGGAAGCA	GGCCGGCCATTAGGAAGG	80

布有 2 个基因, 1、2、7、9 号染色体上均分布有 1 个基因。基因长度从 2 608 bp (*StoSK1.2*) 到 20 101 bp (*StoSK1.1*), 编码序列 (CDS) 从 1 398 bp 到 2 832 bp, 分别对应 465 到 943 个氨基酸残基, 蛋白分子质量介于 33.57 kDa (*StoSK1.3*) 和 88.83 kDa (*StoSK1.4*)。平均等电点 (pI) 为 8.2, 介于 6.38~8.88, 所有 *StoSKs* 蛋白均具备亲水性。亚细胞定位结果预测 *StoSK1.1* 和 *StoSK1.2* 定位于细胞质, 其他蛋白定位于细胞核。

StoSKs 家族蛋白经在线软件预测包含 α -螺旋、无规则卷曲、延伸链及 β -转角 4 种典型二级结构。除 *StoSK3.1* 外, 其他成员中无规则卷曲占比最高, 其次是 α -螺旋和延伸链 (表 2)。

蛋白激酶结构域分析显示 (图 1), *StoSKs* 家族成员均含有 GSK 特征蛋白激酶保守结构域, 即包含 1 个 N 端结构域、1 个激酶结构域及 1 个 C 端结构域。激酶结构域的序列与 *AtSKs* 高度相似, 氨基酸序列相似性达到 80%~93%, 表明该家族成员在进化中高度保守。

2 决明 GSK 系统发育分析与分类

将 *StoSKs* 与拟南芥 (*A. thaliana*) *AtSKs*、水稻 (*O. sativa*) *OsSKs*、大麦 (*H. vulgare*) *HvGSKs*、马铃薯 (*S. tuberosum*) *StSKs*、日本百脉根 (*L. japonicus*) *LSKs* 及大豆 (*G. max*) *GmSKs* 共 73 个 GSK 蛋白共同构建无根系统发育树。图 2 结果显示, 9 个 *StoSKs* 成员被分为 4 个进化分支, 分支 II 是最大的分支, 共包含 27 个 GSK 蛋白, 而进化分支 IV 最小, 只有 8 个成员。分支 I 包含 4 个 *StoSKs*, 分别是 *StoSK1.1*~*1.4*, 分支 II 包含 *StoSK2.1* 和 *StoSK2.2*, 分支 III 包含 *StoSK3.1* 和 *StoSK3.2*, 分支 IV 仅有 1 个。

3 决明 GSK 基因结构与保守基序 motif 分析

同时分析拟南芥、大豆和决明 *GSK* 基因的系统发育关系和基因结构。图 3 结果显示, 4 个进化分支中, I、II、IV 中的所有基因均含有 12 个外显子, 而分支 III 中的基因则含有 13 个外显子, 表明不同物种间的 *GSK*

基因结构具有高度相似性。另外, 相较拟南芥, 同一进化分支中决明 *StoSK* 与大豆 *GmSK* 基因的结构更为相似, 说明同科植物进化中更为保守。

通过 MEME 和 TBtools 软件, 分别在拟南芥、水稻、日本百脉根、大豆和决明共 5 个物种 GSK 蛋白中鉴定到了 15 个保守 motif (图 4)。不同物种间的 GSK 蛋白在保守 motif 的组成上表现出高度相似性, 同一进化分支也表现出相似的 motif 分布, 比如, motif 12 只存在于进化分支 I 中, motif 11 和 motif 15 只存在于进化分支 III 中, motif 13 和 motif 14 只存在于进化分支 II 中的大多数成员。motif 1~9 存在于所有 *StoSK* 蛋白中。这些结果说明, 同一进化分支的 GSK 蛋白可能具有相似功能, motif 差异化分布也意味着蛋白功能和进化多样性。

4 决明 GSK 基因启动子顺式元件分析

提取 *StoSKs* 起始密码子上游约 2 kb 的序列, 使用 PlantCARE 数据库对顺式作用调控元件进行分析。从图 5 可看出, 分析得到的顺式元件可分为 4 大类: 响应胁迫、激素、生长发育和光照。其中, 与光照响应、干旱诱导、茉莉酸甲酯 (MeJA) 响应、脱落酸响应相关的顺式元件是 *StoSKs* 启动子中最丰富的元件。所有 *StoSKs* 启动子区均含有光照响应相关顺式元件。除了光照响应外, 干旱、脱落酸、赤霉素响应相关顺式元件也均存在于所有 *StoSKs* 的启动子区有所分布, 而与植物生长发育相关的顺式元件如厌氧特异性诱导、分生组织的表达、种子特异性调节、细胞周期调控等主要分布于进化分支 II 的 *StoSK2.1* 和 *StoSK2.2* 的启动子区。

5 决明 StoSK 基因共线性分析

首先分析了决明物种内 9 个 *StoSKs* 基因的重复事件, 包括串联重复和片段重复。分析发现, 在 *StoSKs* 中, 同属进化分支 I 中的 *StoSK1.1* 和 *StoSK1.2* 之间发生了基因串联重复事件, 在其他 6 个基因 (*StoSK2.1* 和 *StoSK2.2*, *StoSK2.2* 和 *StoSK4.1*, *StoSK1.2* 和 *StoSK2.2*) 之间发现了 3 次片段重复事件 (图 6)。由此推测, 片段

Table 2 Physical and chemical properties, subcellular localization and secondary structure of the deduced proteins of *StoSKs* genes. CDS: Coding sequence; pI: Isoelectric point

Protein	Gene ID	CDS length/bp	Protein length/aa	Molecular mass/ku	pI	Subcellular localization	α -Helix/%	Random coil/%	Extended strand/%	β -Turn/%
StoSK1.1	Sto08g234550	1 731	576	46 075.48	8.56	Cytoplasmic	36.63	39.11	18.07	6.19
StoSK1.2	Sto08g235490	1 935	644	47 033.6	8.82	Cytoplasmic	35.73	41.25	16.07	6.95
StoSK1.3	Sto08g250140	1 458	485	33 573.27	8.64	Nuclear	40.75	41.78	12.67	4.79
StoSK1.4	Sto13g440710	2 031	676	88 827.43	6.38	Nuclear	32.66	40.97	19.48	4.89
StoSK2.1	Sto07g210470	1 902	633	43 763.65	8.86	Nuclear	35.16	41.67	15.89	7.29
StoSK2.2	Sto13g424070	2 301	766	37 664.71	8.88	Nuclear	36.75	37.65	18.37	7.23
StoSK3.1	Sto01g015100	1 518	505	37 808.88	8.19	Nuclear	39.21	38.30	15.20	7.29
StoSK3.2	Sto09g286690	1 398	465	53 282.96	7.3	Nuclear	37.84	43.34	13.32	5.50
StoSK4.1	Sto02g050930	2 832	943	47 928.14	8.19	Nuclear	35.22	40.66	17.02	7.09

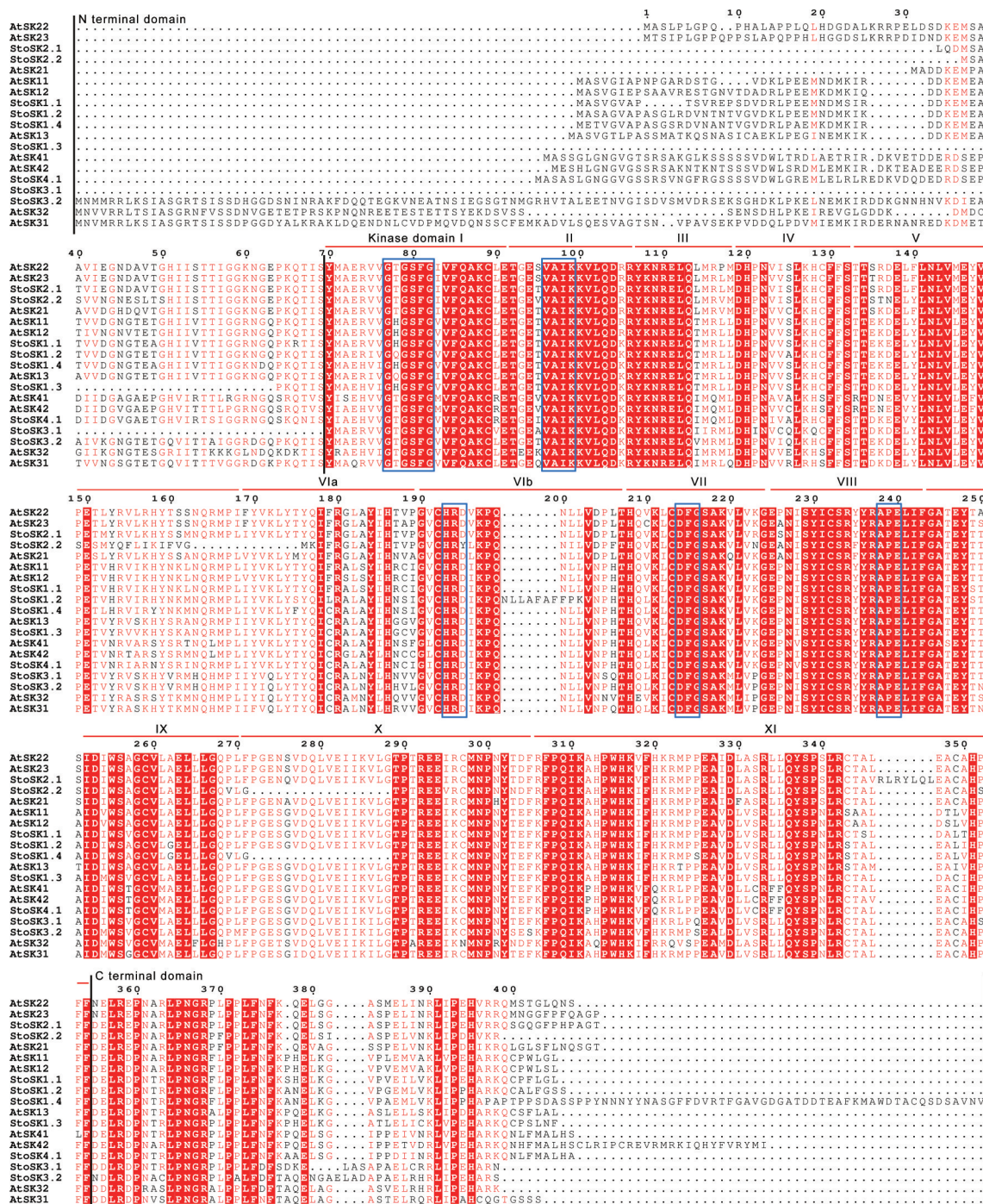


Figure 1 Sequence alignment of conserved kinase domains of GSK proteins from *S. tora* (StoSKs) and *A. thaliana* (AtSKs)

重复事件可能是 *StoSKs* 产生过程中主要的基因复制方式。

其次, 将决明 *StoSKs* 与拟南芥和水稻, 决明 *StoSKs* 与两种豆科植物蒺藜苜蓿和大豆的 *GSK* 基因分别进行种间的共线性关系分析。图 7 结果显示, 决明 *GSK* 基因与大豆和蒺藜苜蓿 *GSK* 基因之间具有更多的同源共线性关系: 与大豆之间共有 17 个同源基因对, 数量最多; 与蒺藜苜蓿存在 11 个同源基因对, 与拟南芥和

水稻分别存在 8 个和 7 个同源基因对。由此看出, *StoSKs* 与双子叶植物之间存在更多的同源基因对, 同时与豆科植物存在更近的进化关系。

以模式植物拟南芥和水稻作为参考, *StoSK1.1* 和 *StoSK1.2* 发生串联重复, 同步进化, 均可能与 *AtSK11*、*AtSK12*、*OsSK13*、*OsSK21* 具有相同起源, *StoSK1.3* 与 *AtSK13* 起源相近; 而 *StoSK1.4* 则分别与 *OsSK13*、*OsSK21* 起源相近; *StoSK2.1* 和 *StoSK2.2* 分别与 *AtSK23*

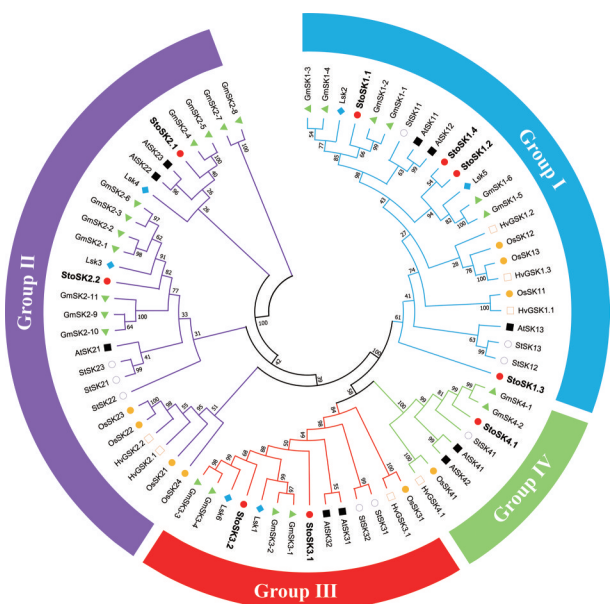


Figure 2 Phylogenetic analyses of GSK gene family from *A. thaliana* (At), *O. sativa* (Os), *L. japonicus* (Lj), *G. max* (Gm), *S. tuberosum* (St), *H. vulgare* (Hv) and *S. tora* (Sto)

和 *AtSK21* 起源相近; *StoSK4.1* 与 *AtSK42* 可能具有相同的起源。另外, 在 9 个 *StoSKs* 基因中, 有 3 个 (*StoSK1.1*、*StoSK1.2*、*StoSK2.2*) 与其他 4 个物种之间均存在同源基因对, 说明这 3 个基因在植物进化过程中

均有较高的保守性。

6 决明 *StoSK* 基因表达分析

通过对决明根、叶、幼嫩种子和成熟种子共 4 个不同组织转录组数据分析, 发现 *StoSKs* 在多个组织中均有表达 (图 8), 在根中的表达水平最高, 其次是叶片、幼嫩的种子和成熟的种子 (除了 *StoSK3.1* 和 *3.2* 在成熟的种子中的相对表达量高于幼嫩的种子)。同属进化分支 II 中的 *StoSK2.1* 和 *2.2* 表现出在所有组织中最高的表达水平, 除 *StoSK3.1* 和 *3.2* 外, 其他 7 个 *StoSKs* 基因在幼嫩种子中的表达量均显著高于成熟的种子。

进一步分析盐胁迫条件下 *StoSK* 基因在根中的表达 (图 9A), 结果显示, 经 $200 \text{ mmol} \cdot \text{L}^{-1} \text{ NaCl}$ 处理后 0~3 h, 有 6 个 *StoSKs* 的表达量显著升高, 其中 *StoSK3.2*、*StoSK4.1* 在 3 h 的表达量达到最大值, 分别较 0 h 升高约 3.8 和 2.5 倍, 而 *StoSK1.2* 在 1 h 的表达量达到最大值, *StoSK1.3*、*StoSK3.1* 的表达量在 6 h 达到峰值, *StoSK1.1*、*StoSK2.1* 的表达量在 12 h 逐渐达到最大值。而 *StoSK2.2* 表达量持续显著下调。

为了进一步验证光响应顺式元件, 分析了 *StoSKs* 在暗处理下的表达水平。除了 *StoSK1.3* 表达水平与光照条件下并无显著差异外, 其他 *StoSKs* 均表现出光响应相关表达模式。图 9B 结果表明, 经过 12 h 处理, 进

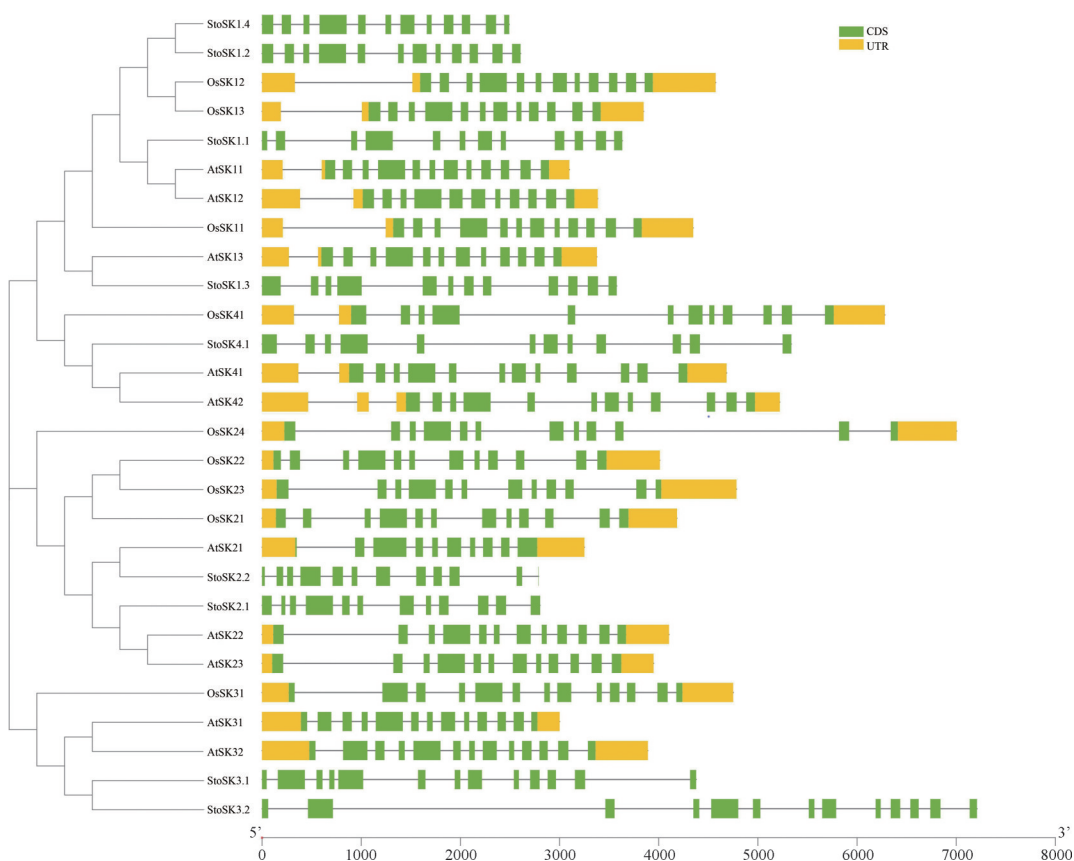


Figure 3 Gene structure analyses of GSKs from *A. thaliana* (At), *O. sativa* (Os), and *S. tora* (Sto). Scale markers represent gene length (bp)

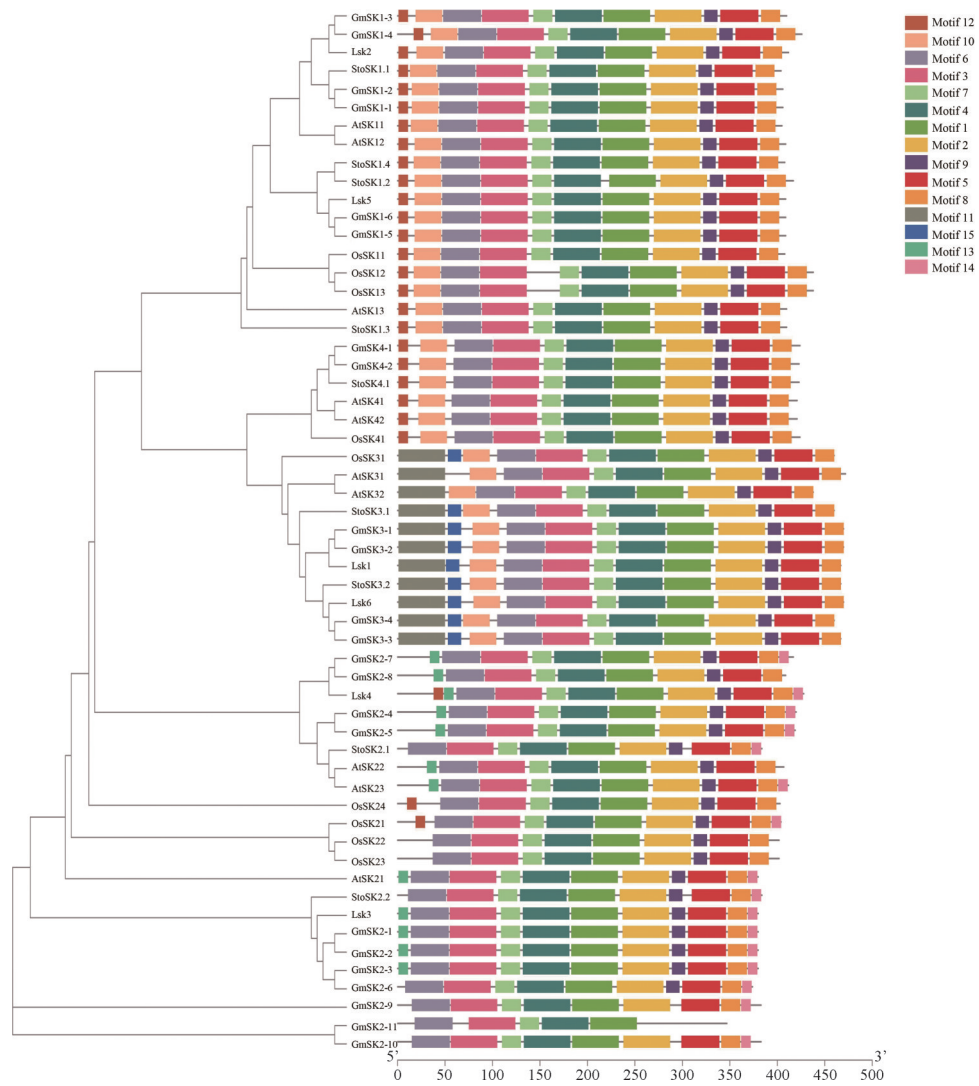


Figure 4 Conserved motif analyses of GSK proteins from *S. tora* (Sto), *A. thaliana* (At), *O. sativa* (Os), *L. japonicus* (Lj) and *G. max* (Gm). Scale markers represent protein sequence length (aa)

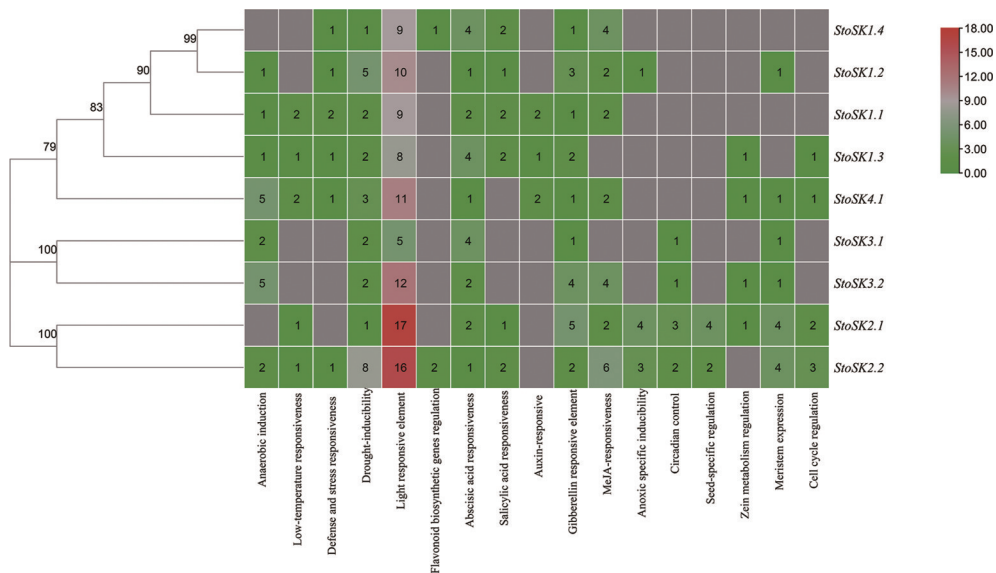


Figure 5 Analysis of putative cis-regulatory elements in promoter region of GSK genes in *S. tora*

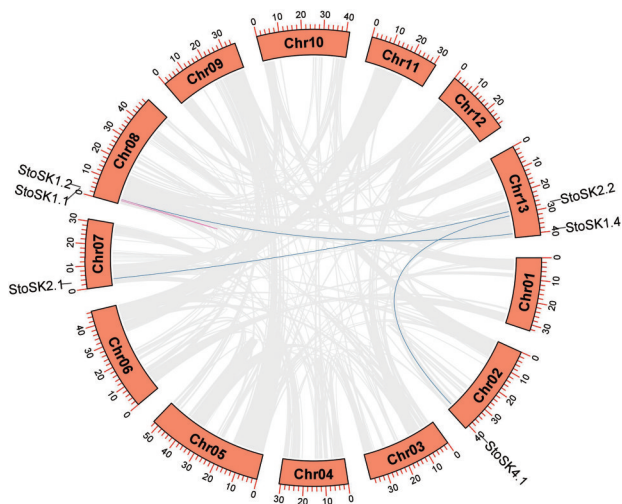


Figure 6 Segmental duplication analysis of *GSK* genes in *S. tora*. Pink link represents the *StoSK* gene pair of tandem repeats; blue curves represent *StoSK* gene pairs with segmental duplication

化分支 I 中的 *StoSK1.1*、*StoSK1.2* 和 *StoSK1.4*，以及进化分支 IV 中的 *StoSK4.1* 的表达受到显著诱导，相反，进化分支 II 和 III 中的 *StoSK2.1*、*StoSK2.2*、*StoSK3.1*、*StoSK3.2* 的转录水平均显著降低，这种相反的反应模式提示进化分支 I 和 IV 可能与进化分支 II、III 中的基因在响应光照相关的信号通路中发挥互补的作用。

讨论

GSK 蛋白激酶作为关键调节因子广泛参与细胞生长、组织发育、胁迫响应等生理生化过程^[24]。随着全基因组测序、组装和注释技术的不断发展，已在拟南芥 (*A. thaliana*)、水稻 (*O. sativa*)、日本百脉根 (*L. japonicus*)、大豆 (*G. max*)、大麦 (*H. vulgare*)^[25]、葡萄 (*Vitis vinifera*)^[12] 等物种中分别鉴定到 10、9、7、23、7、7 个 *GSK* 成员，并证实 *GSK* 在调节植物生理和发育过程中发挥重要作用^[6,26]。相比之下，关于药用植物 *GSK* 基因家族的鉴定分析研究则非常少。本研究在药用植物决明基因组中鉴定到 9 个 *GSK* 基因，成员数量与其他植物很相近。系统发育分析将 9 个 *StoSKs* 蛋白分为 4 个进化分支，与拟南芥、水稻一致^[5]。*StSK* 外显子/内含子结构与拟南芥和水稻 *GSK* 基因高度保守^[27]。大多数 *GSK* 基因都含有 12 个外显子，只有进化分支 III 的 *AtSK31*、*AtSK32*、*OsSK31* 和 *StoSK3.1*、*StoSK3.2* 含有 13 个外显子，说明进化分支 III 中这个额外的外显子有可能是在单子叶植物和双子叶植物分化之前从一个共同祖先遗传而来。9 个 *StoSKs* 成员不均等地分布在 6 条染色体上，仅有 1 对串联重复基因对和 3 对片段重复基因对，说明决明 *StoSKs* 基因家族的扩张可能与水

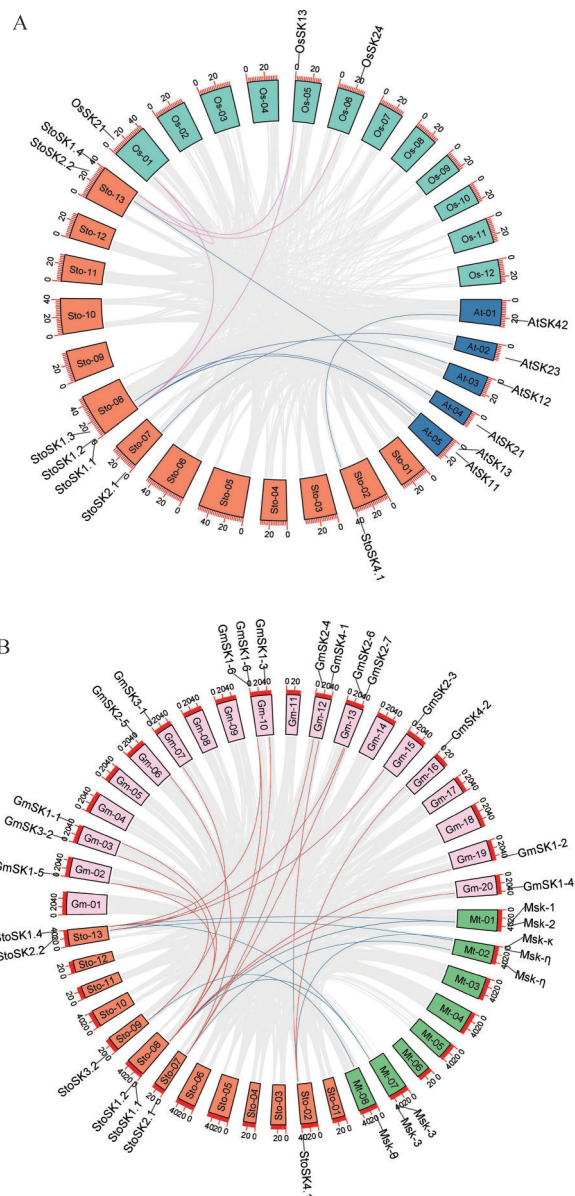


Figure 7 Synteny analyses of *GSK* gene families of *S. tora* and *A. thaliana*, *O. sativa*, *M. truncatula* and *G. max*. A: *S. tora* and *A. thaliana*, *O. sativa*; B: *S. tora* and *M. truncatula*, *G. max*. The cyan, blue, pink, green and orange rectangles represent the chromosomes of *O. sativa*, *A. thaliana*, *G. max*, *M. truncatula* and *S. tora*, respectively. The pink and blue curves represent the collinear gene pairs of *O. sativa* and *S. tora*, *A. thaliana* and *S. tora*, respectively. The red and cyan curves represent the collinear gene pairs of *G. max* and *S. tora*, *M. truncatula* and *S. tora*, respectively

稻^[28]、马铃薯^[29]一样，主要通过基因片段重复事件引起的。亚细胞定位预测 *StoSK1.1* 和 *StoSK1.2* 蛋白主要定位于细胞质，其他成员主要定位于细胞核，与拟南芥 *AtSK*^[11]、紫花苜蓿 *Msk*^[30]、大豆 *GmSK*^[14] 亚细胞定位分析结果较为一致。

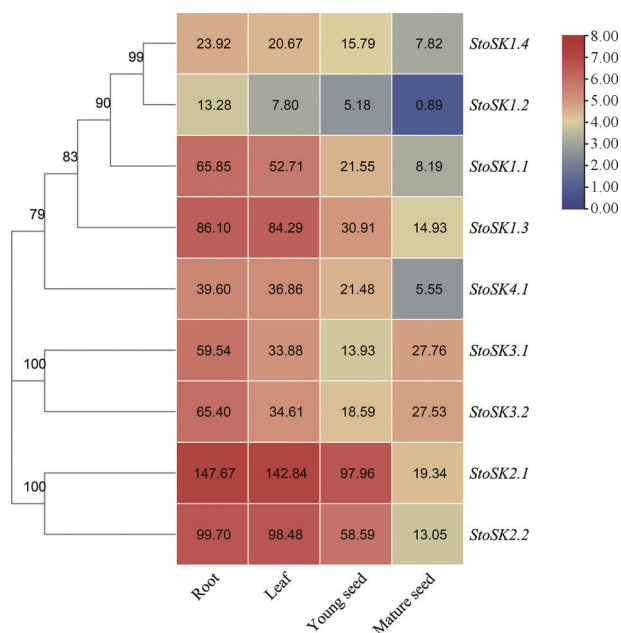


Figure 8 Heatmap of *StoSKs* expression in different tissues, growth and development stages of *S. tora*

蛋白保守结构域是其发挥功能的关键。所有 *StoSKs* 蛋白中都包含 motif 1~9, 仅 N 端和 C 端区域的 motif 存在较大差异。结合系统发育分析结果, 在同一进化分支中的 *StoSKs* 具有保守的 motif 分布和相同的外显子/内含子结构, 表明这些成员之间可能具有相似功能。可变的 N 端和 C 端为 *StoSKs* 家族功能分化提供了可能。此外, *StoSK* 蛋白包含已鉴定物种的 GSK 激酶特征序列^[31,32], 同样分为 12 个亚结构域, 并包含多个特征 motif, 如 xAxK、SYICSR 和 TREE 等, 这些 motif 对于维持酶的催化功能和活性至关重要^[33]。比如 TREE motif 仅存在于植物 GSK 蛋白激酶, 拟南芥 *AtSK21* 中 TREE 谷氨酸 263 突变为赖氨酸会引发蛋白激酶组成性激活^[34]。

StoSKs 顺式元件分析鉴定到了大量涉及激素响应、胁迫应答及生长发育等元件, 这些元件在拟南芥^[6]、棉花^[35]、马铃薯^[29] GSK 中均有发现。光响应顺式元件如 MRE、GATA motif、AE box 等在 *StoSKs* 的启动子区大量存在, qRT-PCR 分析也证实 *StoSK2.1*、*2.2*、*3.1*、*3.2* 对光敏感 (图 9B)。另外, 一些转录因子如 PIF4、MYBL2、EGL3 等^[6] 已被证实可与 GSK 启动子区相结合从而参与其转录表达, 对药用植物抗逆和生长发育发挥重要作用。

基因表达模式检测是研究功能的前提。大量证据表明植物 GSK 家族具有不同的组织表达模式。马铃薯 *StGSK* 在叶片中的表达水平较低, 在愈伤组织中的表达量相对较高^[29]; 辣椒 *CaSK23* 在叶片和果实中高

度表达, 在花和根中的表达量较低^[36]; 葡萄 *VvSK7* 在幼嫩组织中的表达水平远远高于成熟和衰老组织^[12]。本研究揭示大多数 *StoSKs* 在不同组织中的表达模式高度一致, 即根中的表达量最高, 其次是叶片。7 个 *StoSKs* 在幼嫩的种子中表达量均高于成熟的种子, 表明大多数的 *StoSKs* 在决明种子的发育过程中具有重要作用。此外, 同属进化分支 I 且互为串联重复基因的 *StoSK1.1* 和 *1.2* 在检测的组织中表达水平高低不同, 结合两者在响应盐胁迫条件下的表达趋势, 说明同一进化分支的基因之间可能存在功能分化或互补, 这种现象在拟南芥 *AtSK* 进化分支 II 中同样存在^[37]。

GSK 除了具有组织表达特性, 还参与响应干旱^[38]、低温^[39]等非生物胁迫。如水稻 *OsSK41/OsGSK5* 及其紫花苜蓿同源蛋白 *Msk4* 均已被证实可分别增强转基因水稻和拟南芥的耐盐性^[33]。与 *AtSK11* 高度同源的大豆 *GmSK1-1* 同样受到各种非生物胁迫的诱导, 可提高酵母细胞盐/渗透耐受性^[14]。在本研究中, 前期分别设置了 200 和 400 mmol·L⁻¹ 两种盐胁迫处理, 发现高浓度盐胁迫严重抑制决明根及侧根生长发育, 且 *StoSKs* 大多在 1~3 h 诱导后下降, 应答响应并不持续 (数据未展示)。而大多数 *StoSKs* 能快速响应 200 mmol·L⁻¹ 盐胁迫诱导并持续应答, 多数在 1~6 h 表达量达到最高, 其中分别与 *AtSK11* 和 *GmSK1-1* 同源的 *StoSK1.1* 和 *StoSK1.2* 能迅速且持续响应盐胁迫诱导, 表明二者可能参与决明盐胁迫应答, 而只有 *StoSK1.4* 和 *StoSK2.2* 在处理后的表达下调。结合本研究与已报道证据^[14,33], 推测植物 GSK 蛋白响应盐胁迫的作用在进化上是相对保守的。系统发育和共线性分析发现同属进化分支 II 的 *StoSK2.2* 和 *AtSK21* 是一对直系同源基因, 两者的蛋白质结构也十分相似。已有研究表明^[40], 拟南芥 *AtSK21* 的高表达可激活 *ARF7* 和 *ARF19* 从而加速植株的侧根形成。*StoSK2.2* 与 *AtSK21* 在调控胁迫反应和根的生长发育方面可能具有类似功能。

本研究共鉴定到决明 9 个 GSK 基因家族成员, 蛋白质特征序列与结构及种间的进化关系相对保守, 启动子区包含 MRE、TCA、MBS、LTR 等响应光照和非生物胁迫元件, 这些基因大多数能响应盐胁迫, 并对光暗响应方式有差异, 为后续进一步研究其生物学功能及抗逆育种奠定基础。

作者贡献: 冯昭是本研究的实验设计执行人并撰写论文初稿; 覃洋、刘世鹏、胡晓晨及毛仁俊负责实验和数据分析; 吕蕊花、吕瑞华参与实验设计和实验结果分析; 张岗和毛仁俊是项目的构思者及负责人, 指导实验设计、数据分析和论文写作与修改。全体作者都已阅读并同意最终的文本。

利益冲突: 无任何利益冲突。

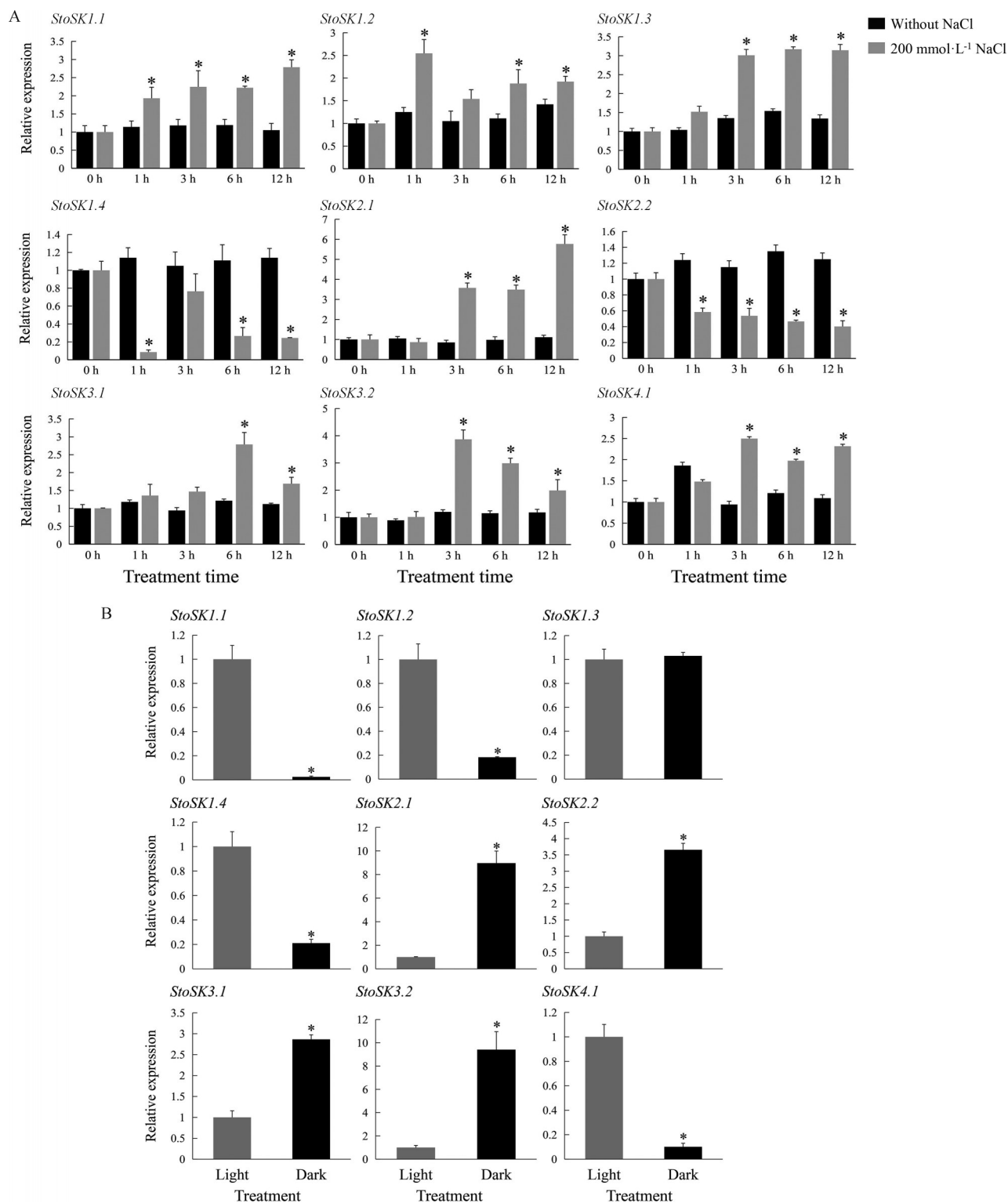


Figure 9 Expression analyses of *StoSKs* in response to 200 mmol·L⁻¹ salinity treatment (A) and in response to light and dark treatment (B). Data are presented as the mean ± standard error (SE), which were calculated from biological replicate ($n = 3$) and technical replicate ($n = 3$) mRNA. Asterisks above the bars indicate significant differences ($P < 0.05$ in A represents comparison between control and treatment at the same time point. $P < 0.05$ in B represents comparison between light and dark conditions)

References

[1] Ali A, Hoeflich KP, Woodgett JR. Glycogen synthase kinase-3: properties, functions, and regulation [J]. *Che Rev*, 2001, 101: 2527-2540.

[2] Embi N, Rylatt DB, Cohen P. Glycogen synthase kinase-3 from rabbit skeletal muscle. Separation from cyclic-AMP-dependent

- protein kinase and phosphorylase kinase [J]. *Eur J Biochem*, 1980, 107: 519-527.
- [3] Frame S, Cohen P. GSK3 takes centre stage more than 20 years after its discovery [J]. *Biochem J*, 2001, 359: 1-16.
- [4] Kaidanovich-Beilin O, Woodgett JR. GSK-3: functional insights from cell biology and animal models [J]. *Front Mol Neurosci*, 2011, 4: 40.
- [5] Yoo MJ, Albert VA, Soltis PS, et al. Phylogenetic diversification of glycogen synthase kinase 3/SHAGGY-like kinase genes in plants [J]. *BMC Plant Biol*, 2006, 6: 3.
- [6] Youn JH, Kim TW. Functional insights of plant GSK3-like kinases: multi-taskers in diverse cellular signal transduction pathways [J]. *Mol Plant*, 2015, 8: 552-565.
- [7] Yang XG, Liang WH, Li F, et al. OsGSK3 is a novel GSK3/SHAGGY-like gene from *Oryza sativa* L., involved in abiotic stress signaling [J]. *Pak J Bot*, 2012, 44: 1491-1496.
- [8] Charrier B, Champion A, Henry Y, et al. Expression profiling of the whole *Arabidopsis* Shaggy-like kinase multigene family by real-time reverse transcriptase polymerase chain reaction [J]. *Plant Physiol*, 2002, 130: 577-590.
- [9] Piao HL, Lim JH, Kim SJ, et al. Constitutive over-expression of AtGSK1 induces NaCl stress responses in the absence of NaCl stress and results in enhanced NaCl tolerance in *Arabidopsis* [J]. *Plant J*, 2001, 27: 305-314.
- [10] Li J, Nam KH. Regulation of brassinosteroid signaling by a GSK3/SHAGGY-like kinase [J]. *Science*, 2002, 295: 1299-1301.
- [11] Rozhon W, Mayerhofer J, Petutschnig E, et al. ASK θ , a group-III *Arabidopsis* GSK3, functions in the brassinosteroid signaling pathway [J]. *Plant J*, 2010, 62: 215-223.
- [12] Zeng J, Haider MS, Huang J, et al. Functional characterization of *VvSK* gene family in grapevine (*Vitis vinifera* L.) revealing their role in berry ripening [J]. *Int J Mol Sci*, 2020, 21: 4336.
- [13] Constantine G, Daniela T, Panagiota KP, et al. Lotus SHAGGY-like kinase 1 is required to suppress nodulation in *Lotus japonicus* [J]. *Plant J*, 2019, 98: 228-242.
- [14] He C, Gao H, Wang H, et al. GSK3-mediated stress signaling inhibits legume-rhizobium symbiosis by phosphorylating GmNSP1 in soybean [J]. *Mol Plant*, 2021, 14: 15.
- [15] Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China (中华人民共和国药典) [S]. 2020 Ed. Vol 1. Beijing: China Medical Science Press, 2020: 151.
- [16] Dong YJ, Jiang YQ, Liu Y, et al. Research progress on chemical composition and pharmacological effects of *Cassiae Semen* and predictive analysis on quality markers [J]. *Chin Tradit Herb Drugs (中草药)*, 2021, 52: 2719-2732.
- [17] Au TS, Yang DJ, Zhang YO, et al. Differential expression of lipid metabolism related genes in HepG2 cells treated with neotoralactone [J]. *Acta Pharm Sin (药学报)*, 2003, 38: 62-66.
- [18] Mao RJ, Qi ZH, Han RL, et al. Spaceflight-induced variation on biological traits and effective components of *Cassia obtusifolia* [J]. *Chin J Chin Mater Med (中国中药杂志)*, 2015, 40: 2571-2575.
- [19] Zhang XD, Song MY, Wang XP, et al. Response of growth and development of *Cassia Obtusifolia* to salt stress and evaluation of salt tolerance [J]. *J Hebei Agric Sci (河北农业科学)*, 2018, 22: 6.
- [20] Chen SL, Song JY. Herbgonomics [J]. *Chin J Chin Mater Med (中国中药杂志)*, 2016, 41: 9.
- [21] Kang SH, Pandey RP, Lee CM, et al. Genome-enabled discovery of anthraquinone biosynthesis in *Senna tora* [J]. *Nat Commun*, 2020, 11: 5875.
- [22] Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data [J]. *Mol Plant*, 2020, 13: 1194-1202.
- [23] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method [J]. *Methods*, 2001, 25: 402-408.
- [24] Youn JH, Kim TW, Kim EJ, et al. Structural and functional characterization of *Arabidopsis* GSK3-like kinase AtSK12 [J]. *Mol Cells*, 2013, 36: 564-570.
- [25] Kloc Y, Dmochowska-Boguta M, Zielezinski A, et al. Silencing of HvGSK1.1-A GSK3/SHAGGY-like kinase enhances barley (*Hordeum vulgare* L.) growth in normal and in salt stress conditions [J]. *Int J Mol Sci*, 2020, 21: 6616.
- [26] Saidi Y, Hearn TJ, Coates, JC. Function and evolution of 'green' GSK3/Shaggy-like kinases [J]. *Trends Plant Sci*, 2012, 17: 39-46.
- [27] Mi JY, Victor AA, Pamela SS, et al. Phylogenetic diversification of glycogen synthase kinase 3/SHAGGY-like kinase genes in plants [J]. *BMC Plant Biol*, 2006, 21: 3.
- [28] Song YL, Xiong LZ. Systematic analysis of glycogen synthase kinase 3 genes in rice reveals their differential responses to phytohormones and abiotic stresses [J]. *J Huazhong Agric Univ (华中农业大学学报)*, 2012, 31: 1-9.
- [29] Huang SH, Liu YX, Deng R, et al. Genome-wide identification and expression analysis of the GSK gene family in *Solanum tuberosum* L. under abiotic stress and phytohormone treatments and functional characterization of StSK21 involvement in salt stress [J]. *Gene*, 2021, 766: 145156.
- [30] Wrzaczek M, Rozhon W, Jonak C. A proteasome-regulated glycogen synthase kinase-3 modulates disease response in plants [J]. *J Biol Chem*, 2007, 282: 5249-5255.
- [31] Hanks SK, Quinn AM, Hunter T, et al. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains [J]. *Science*, 1988, 241: 42-52.
- [32] Bittner T, Campagne S, Neuhaus G, et al. Identification and characterization of two wheat glycogen synthase kinase 3/Shaggy-like kinases [J]. *BMC Plant Biol*, 2013, 13: 64.
- [33] Hu ZJ, Lu SJ, Wang MJ, et al. A Novel QTL qTGW3 encodes the GSK3/SHAGGY-like kinase OsGSK5/OsSK41 that interacts with OsARF4 to negatively regulate grain size and weight in rice

- [J]. Mol Plant, 2018, 11: 736-749
- [34] Cheng X, Xin M, Xu R, et al. A single amino acid substitution in STKc_GSK3 kinase conferring semispherical grains and its implications for the origin of *Triticum sphaerococcum* [J]. Plant Cell, 2020, 32: 923-934.
- [35] Wang LL, Yang ZE, Zhang B, et al. Genome-wide characterization and phylogenetic analysis of GSK gene family in three species of cotton: evidence for a role of some GSKs in fiber development and responses to stress [J]. BMC Plant Biol, 2018, 18: 330.
- [36] Qiu AL, Wu J, Lei YF, et al. CaSK23, a putative GSK3/SHAGGY-like kinase of *Capsicum annuum*, acts as a negative regulator of pepper's response to *Ralstonia solanacearum* attack [J]. Int J Mol Sci, 2018, 19: 2698.
- [37] Sun FA, Yu HQ, Qu JT, et al. Maize ZmBES1/BZR1-5 decreases ABA sensitivity and confers tolerance to osmotic stress in transgenic *Arabidopsis* [J]. Int J Mol Sci, 2020, 21: 996.
- [38] Dal SS, Stampfl H, Krasensky J, et al. Stress-induced GSK3 regulates the redox stress response by phosphorylating glucose-6-phosphate dehydrogenase in *Arabidopsis* [J]. Plant Cell, 2012, 24: 3380-3392.
- [39] Ye KY, Li H, Ding YL, et al. BRASSINOSTEROID-INSENSITIVE2 negatively regulates the stability of transcription factor ICE1 in response to cold stress in *Arabidopsis* [J]. Plant Cell, 2019, 31: 2682-2696.
- [40] Li JY, Wang ZK, Han X, et al. Cloning and expression analysis of *ArARF18* gene in *Acer rubrum* L. [J]. Genom Appl Biol (基因组学与应用生物学), 2021, 40: 10.