

DOI: 10.11931/guihaia.gxzw201601035

引文格式: 黄海霞, 苏何玲, 史云龙, 等. 甜瓜鲨烯合酶基因克隆及酶分子结构分析 [J]. 广西植物, 2017, 37(3):373-379

HUANG HX, SU HL, SHI YL, et al. Gene cloning and molecular characteristics of squalene synthase from *Cucumis melo* [J]. *Guihaia*, 2017, 37(3):373-379

甜瓜鲨烯合酶基因克隆及酶分子结构分析

黄海霞¹, 苏何玲^{1,2,*}, 史云龙¹, 肖雅伦¹, 谭燕莲¹, 梁杨浩¹, 刘永明¹, 吴耀生²
(1. 桂林医学院 生物化学与分子生物学教研室, 广西 桂林 541004; 2. 广西医科大学 生物化学与分子生物学教研室, 南宁 530000)

摘要: 葫芦素类是主要分布于葫芦科植物中具有多种医药活性的四环三萜类化合物, 目前药用葫芦素原料主要从甜瓜蒂中提取。该研究从甜瓜中克隆葫芦素类合成关键酶——鲨烯合酶(SQS)的基因, 并对其序列进行了生物信息学分析。结果表明: DNA 测序和 BLAST RefSeqGene 分析表明, 克隆的甜瓜 SQS 基因片段具有完整的该酶基因开放阅读框(ORF)序列。ORF 分析显示, 甜瓜 SQS 由 417 氨基酸残基构成, 等电点为 7.56。对推衍的甜瓜 SQS 氨基酸序列分析结果提示, 该酶二级结构以 α 螺旋为主。结构域预测结果表明, SQS 属于异戊二烯合酶家族, 具有法呢酰基二磷酸及镁离子的结合位点。三级结构预测提示, 甜瓜 SQS 为单体酶, 其活性中心主要由几个 α 螺旋围绕形成的穴状结构。磷酸化位点分析显示, S⁴⁸ 处于酶活性中心相关⁴⁷VSRSF⁵² 的模体中, 而 S¹⁹⁶ 是正选择位点, 提示这两处磷酸化位点可能是甜瓜 SQS 酶活性调节的关键部位。以甜瓜 SQS 基因 ORF 序列构建系统发生树的系统发生分类结果与形态学分类结果一致。该研究结果为葫芦素类的生物合成调控研究提供了新的线索和实验依据。

关键词: 甜瓜, 鲨烯合酶, 基因克隆, 序列分析

中图分类号: Q946.83 文献标识码: A 文章编号: 1000-3142(2017)03-0373-08

Gene cloning and molecular characteristics of squalene synthase from *Cucumis melo*

HUANG Hai-Xia¹, SU He-Ling^{1,2,*}, SHI Yun-Long¹, XIAO Ya-Lun¹,
TAN Yan-Lian¹, LIANG Yang-Hao¹, LIU Yong-Ming¹, WU Yao-Sheng²

(1. Department of Biochemistry and Molecular Biology, Guilin Medical University, Guilin 541004, Guangxi, China;

2. Department of Biochemistry and Molecular Biology, Guangxi Medical University, Nanning 530000, China)

Abstract: Cucurbitacins (Cus) are tetracyclic triterpenoid compounds which exist mainly in cucurbitaceous plants and have various pharmaceuticals-like actions. The medicinal raw materials of Cus are currently extracted mainly from the pedicellus melo. In this study, the gene of squalene synthase (SQS), the key enzyme for Cus biosynthesis, from *Cucumis melo* was cloned, and the amino acid sequence of the enzyme derived from its nucleotide sequence was analyzed with bioinformatics methods. DNA sequencing and BLAST RefSeqGene analysis indicated that the cloned fragment of the *C. melo* SQS gene contained a complete open reading frame (ORF) of the gene. ORF Finder analysis showed that the SQS of *C. melo* was constituted of 417 amino acid residues with an isoelectric point of 7.56. Analysis of the deduced amino acid sequence suggested that the main type of secondary structure of the enzyme was α helix. Domain prediction study

收稿日期: 2016-03-21 修回日期: 2016-04-23

基金项目: 国家自然科学基金(31260069); 广西高等学校优秀中青年骨干教师培养工程项目 [Supported by the National Natural Science Foundation of China (31260069); the Training Program of the Outstanding Higher Education Teachers of Guangxi from by the Education Department of Guangxi Zhuang Autonomous Region].

作者简介: 黄海霞 (1988-), 女, 山东人, 硕士, 研究方向为生物化学与分子生物学, (E-mail)893916164@qq.com。

*通信作者: 苏何玲, 教授, 研究方向为生物化学与分子生物学, (E-mail) helingsu@glmc.edu.cn。

indicated that the SQS belonged to the isoprenoid biosynthase superfamily possessing the banding sites of farnesyl diphosphate and magnesium ion. Prediction of the tertiary structure suggested that the SQS was a monomeric enzyme with a cave-like activity center formed by α helices. Protein phosphorylation analysis indicated that the phosphorylation site S⁴⁸ located in the activity center-related motif⁴⁷ VSRSF⁵² and S¹⁹⁶ was a positive selection site, suggesting that both S⁴⁸ and S¹⁹⁶ were critical sites for the regulation of SQS activity. The phylogenetic classification based on the phylogenetic tree constructed with the ORF sequence of the SQS gene showed that the result was in accordance with that of morphologic classification. Therefore, this study provides new clues and reference for research of the regulation of Cus biosynthesis.

Key words: *Cucumis melo*, squalene synthase, gene cloning, sequence analysis

葫芦素类(Cucurbitacins, Cus)为一类主要分布于葫芦科植物中的四环三萜化合物,其抗炎、抗肿瘤和免疫调节活性具有重要医学价值(王莉梅和姚铭,2015)。植物体内包括 Cus 在内的三萜化合物主要通过甲羟戊酸途径(mevalonic acid pathway, MVA)合成。该途径以乙酰 CoA 为原料,经过一系列酶促反应生成金合欢基焦磷酸或法尼基焦磷酸(farnesyl diphosphate, FPP)。2 分子 FPP 在鲨烯合酶(squalene synthase, SQS)催化下以尾尾方式连接生成鲨烯,后者经不同途径环合生成三萜化合物(Kim et al, 2011)。SQS 处于萜类代谢途径中的 FPP 到其它产物的分支点上,是生物合成三萜化合物的一个关键酶。SQS 的含量和活性直接影响到鲨烯的生物合成量,进而影响到三萜化合物的生物合成(吴耀生等, 2007)。

目前,三七、人参和红花栝楼等多种药用植物的 SQS 基因已被克隆(吴耀生等, 2007; 张明哲等, 2010; 陶晨陈等, 2015)。然而,尽管药用葫芦素原料主要从甜瓜蒂中提取,但甜瓜 SQS 基因序列及酶分子结构尚未见文献报道。本研究从甜瓜叶中克隆了 SQS 基因开放阅读框架(open reading frame, ORF)及其 3' 末端序列,并对推衍的甜瓜 SQS 氨基酸序列进行了生物信息学分析。

1 材料与方 法

1.1 材料

甜瓜种为购自驻马店市志海种业有限公司 2014 年 3 月的甜瓜一代大田用种,将种子种于温室发芽至 10 cm 后,取其叶片保存至液氮中备用。Taq DNA 聚合酶和 T4 DNA 连接酶购自大连宝生物公司,3'-Full RACE Core Set with PrimeScriptTM RTase 购自日本 TaKaRa 公司,2×Taq PCR MasterMix、核酸电泳 Marker DL2000、质粒小提试剂盒和大肠杆菌

DH5 α 购自于天根生化科技有限公司,异硫氰酸胍粉剂、X-gal、SanPrep 柱式 DNA 胶回收试剂盒及引物购自上海生物工程股份有限公司,pGEM-T Easy Vector 连接试剂盒购自美国 Promega 公司。克隆测序由上海生物工程公司完成。

1.2 方法

1.2.1 总 RNA 提取及逆转录 在陶晨陈等(2015)方法的基础上改进异硫氰酸胍法提取甜瓜叶片中的总 RNA。取 0.01~0.02 g 叶片,剪成碎片置于 2 mL 冻存管中,放入液氮罐,过夜。次日取出冻存管,立即加入异硫氰酸胍提取缓冲液(4 mol·L⁻¹异硫氰酸胍,25 mmol·L⁻¹柠檬酸钠 pH7.0,0.5% 十二烷基肌氨酸钠,用前加 β -巯基乙醇至总浓度 2%),用 S10 高速匀浆机(手提式高速分散器)匀浆。加入 100 μ L 2 mol·L⁻¹ NaAc (pH4.0), 800 μ L 饱和酚,200 μ L 氯仿,漩涡振荡混匀,冰浴 5 min,4 $^{\circ}$ C 15 000 r·min⁻¹ 离心 10 min。将上层水相转移入另一 1.5 mL 离心管中,加入 2/3 体积的 5 mol·L⁻¹的 KAc,混匀,冰浴 5 min, 4 $^{\circ}$ C 15 000 r·min⁻¹ 离心 10 min。将上层水相转移入另一 1.5 mL 离心管中,加入等体积的异丙醇,混匀。-20 $^{\circ}$ C 冰浴 5~10 min,4 $^{\circ}$ C 15 000 r·min⁻¹ 离心 10 min。弃上清,70% 的乙醇洗沉淀后加入无 RNA 酶水溶解沉淀,-80 $^{\circ}$ C 保存备用。取 1 μ g 总 RNA, Oligo(dT) 反转录后作为扩增甜瓜 SQS 的 ORF 的模板。取 1 μ g 总 RNA,按 3'-Full RACE Core Set with PrimeScriptTM RTase 说明书操作,反转录后用于扩增甜瓜 SQS 基因的 3'末端序列。

1.2.2 甜瓜 SQS 基因部分片段的扩增及 3'端序列的扩增 参照 NCBI 上 GenBank 中其它葫芦科植物 SQS 序列中的高度保守序列,设计引物 P1: 5'-ACAACAACGTTGAAGTCTTCAGGG-3', P2: 5'-CG-GTAACCGTTTGGCAGAGA-3',扩增甜瓜 SQS 基因部分片段,扩增长度为 300 bp,扩增条件:94 $^{\circ}$ C 预变性 10 min, 然后 94 $^{\circ}$ C 变性 30 s, 58 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C

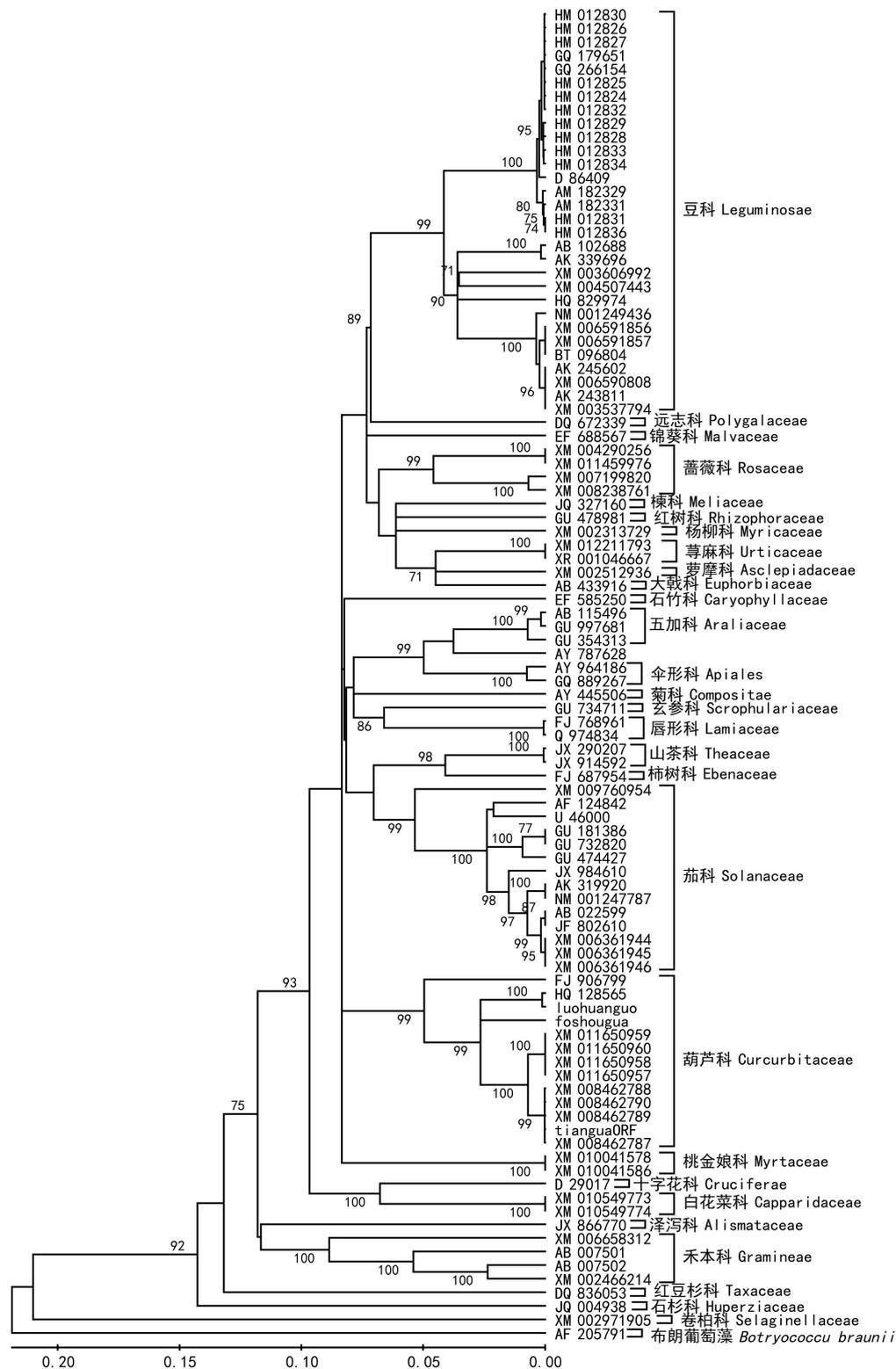


图 1 甜瓜 SQS 基因与其它植物 SQS 基因序列比对构建的系统进化树

Fig. 1 Phylogenetic tree constructed by comparing the nucleotide sequence of the SQS from *Cucumis melo* with those of the SQS genes from other plants

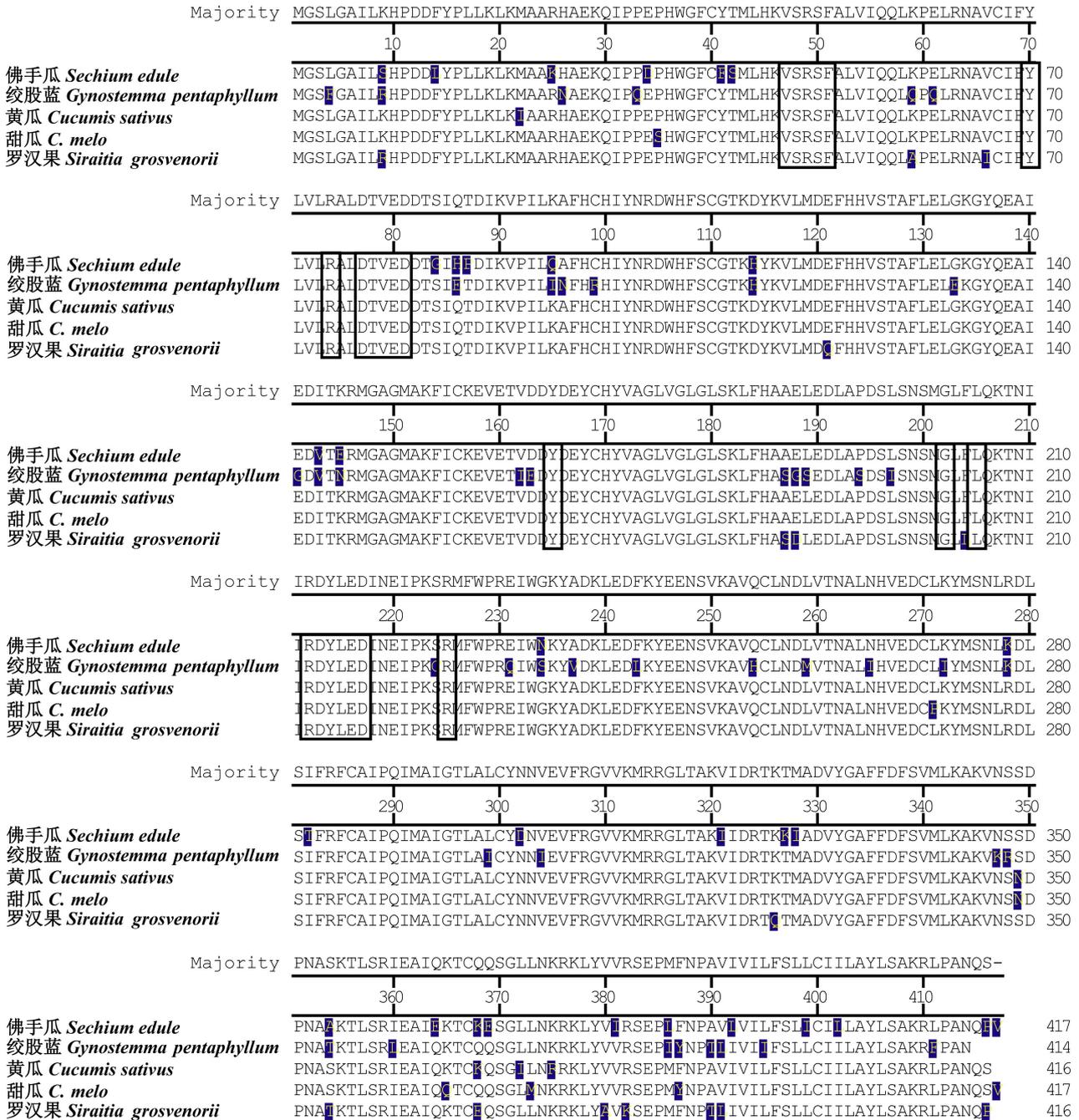


图2 甜瓜与其它葫芦科植物 SQS 氨基酸序列的比对分析

Fig. 2 Comparative analysis of the SQS amino acid sequences from *Cucumis melo* and other cucurbitaceous plants

延伸 30 s, 运行 40 个循环, 72 °C 延伸 10 min。将扩增产物直接送上海生物工程有限公司测序证实所获序列为甜瓜 SQS 基因部分序列。根据测序结果设计甜瓜 SQS 基因 3' 末端内部特异性引物 P3: 5'-GGTCAACAGTGATCCTAATGC-3' 及外部特异性引物 P4: 5'-CGATTGCGGATGTCTATGGAGC-3'。按

3'-Full RACE Core Set with PrimeScript™ RTase 说明书操作并克隆甜瓜 SQS 基因 3' 末端, 将克隆菌液送上海生物工程有限公司测序证实所获序列为甜瓜 SQS 基因 3' 末端序列。

1.2.3 甜瓜 SQS 基因 ORF 的克隆 鉴于 SQS 基因 5' 末端高度保守, 因此根据 NCBI 上 GenBank 中其

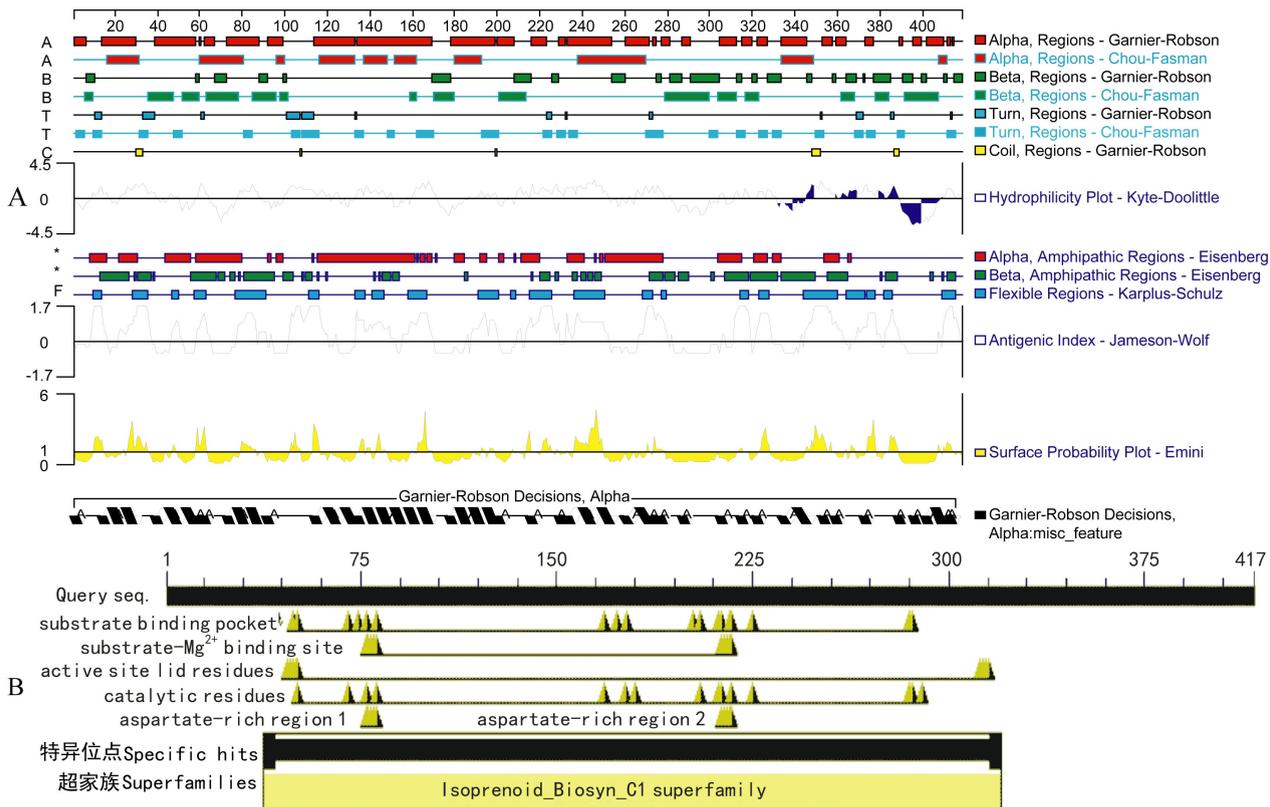


图 3 甜瓜 SQS 磷酸化位点分析

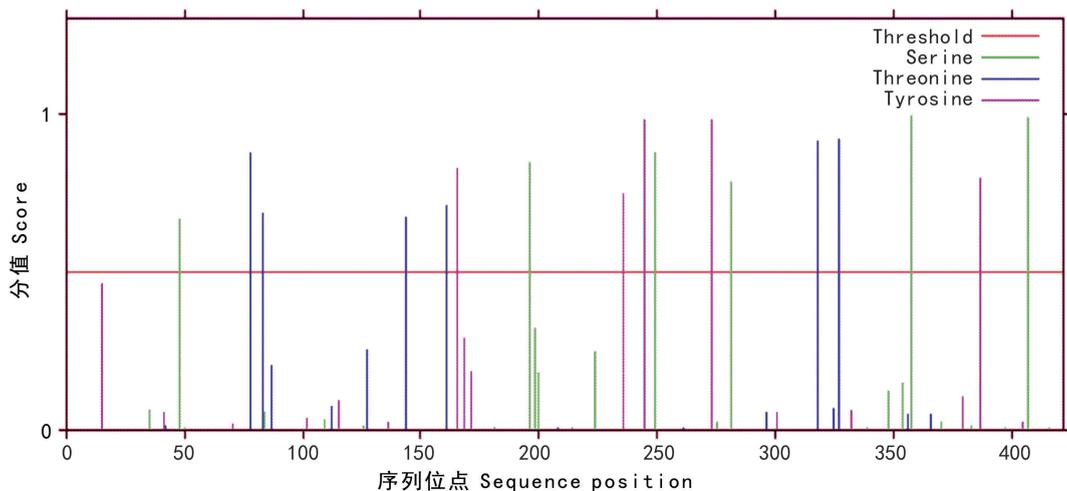
Fig. 3 Prediction analysis of the phosphorylation sites in the SQS from *Cucumis melo*

图 4 甜瓜 SQS 二级结构和结构域预测分析 A. 二级结构分析结果; B. 结构域分析结果。

Fig. 4 Prediction analysis of the secondary structure and domain of the SQS from *Cucumis melo*

A. Results of the secondary structure prediction; B. Results of the domain prediction.

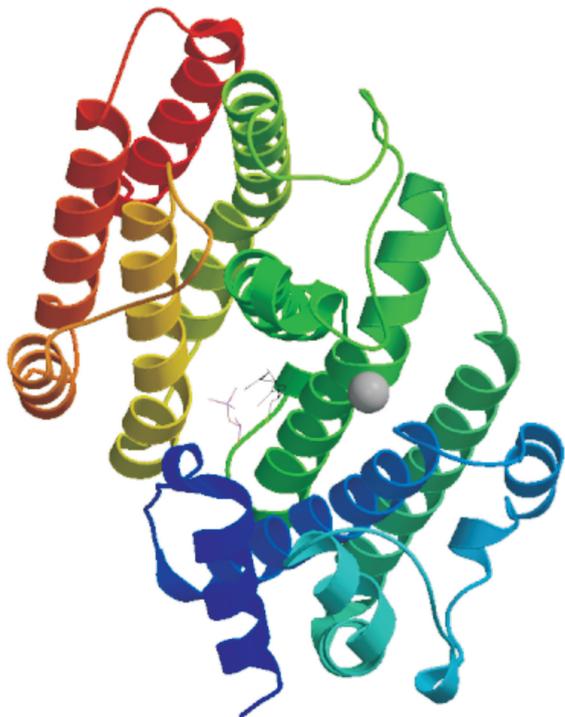


图5 甜瓜 SQS 三级结构预测分析
Fig. 5 Tertiary structure prediction of the SQS from *Cucumis melo*

它葫芦科植物 SQS 基因序列及上述操作获得的甜瓜 SQS 基因 3' 末端序列设计了扩增甜瓜 SQS 基因 ORF 的引物 P5: 5'-GATTGAGAGCGAGAAATGGG-3'; P6: 5'-TCATACAGATTGGTTGGCTG-3', 扩增条件为 94 °C 预变性 10 min; 94 °C 60 s, 58 °C 30 s, 72 °C 90 s, 35 个循环, 最后 72 °C 延伸 4 min。将扩增产物胶回收后, 按美国 Promega 公司的 pGEM-T Easy Vector 试剂盒克隆甜瓜 SQS 基因 ORF, 将获得的质粒送上海生物工程有限公司测序。

1.2.4 生物信息学分析 将获得的甜瓜 SQS 基因序列置 NCBI 上, 以 ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 寻找 ORF, 应用 DNASTAR 将获得的 ORF 序列翻译为氨基酸序列, 以 ProtScale (<http://web.expasy.org/protscale/>) 及 ProtParam (<http://web.expasy.org/cgi-bin/protparam/protparam>) 分析甜瓜 SQS 蛋白的疏水性/亲水性, NetPhos 2.0 Server (<http://www.cbs.dtu.dk/services/NetPhos/>) 对蛋白质磷酸化位点进行预测。以 DNASTAR 软件中的 protean 以及在线网站 NPS@ (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons).

htm) 分析其二级结构; 结构域预测在 NCBI CD-Search (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 网站上进行。以 TMHMM Server v. 2.0 分析其跨膜区。

通过 SWISS-MODEL 软件对甜瓜 SQS 蛋白质的三级结构进行预测, 用 RASTOP 2.0 软件编辑。

从 GenBank 中下载植物 SQS 基因, 布朗葡萄藻为 outgroup, 以 clustalx1.8 对已报道的绞股蓝、香瓜、黄瓜、人参、丹参、拟南芥、甘草、番茄、马铃薯、烟草、玉米等物种的 SQS 氨基酸序列进行比对分析, 并构建了系统进化树, 在 MEGA5.0 软件上进行 UPGMA 算法, 并进行 1 000 次的 Bootstrap 测试。

2 结果与分析

2.1 甜瓜总 RNA 提取结果

提取的甜瓜叶子总 RNA 用紫外分光光度计检测的 $A_{260/280}$ 值均在 1.8~2.0 之间, $A_{260/230}$ 值也都大于 2.0, 且琼脂糖凝胶电泳显示, 28SrRNA 与 18SrRNA 的亮度接近 2:1, 提示所提取的总 RNA 质量良好。

2.2 甜瓜 SQS 基因的扩增结果

参照 NCBI 上 GenBank 中其它葫芦科植物 SQS 基因序列中的高度保守序列, 设计引物了 P1 及 P2, 扩增了甜瓜 SQS 基因部分片段, 扩增长度与预期相符, 测序列结果在 NCBI 的 BLAST RefSeqGene 上比对提示获得 SQS 基因部分序列。按 3'-Full RACE Core Set with PrimeScript™ RTase 说明书克隆甜瓜 SQS 基因 3' 末端, 扩增结果与预期相符。

2.3 甜瓜 SQS 基因的克隆

根据获得的序列设计扩增甜瓜 SQS 基因 ORF 的引物 P5 和 P6 扩增出甜瓜 SQS 基因 ORF 片段, 结果与预期相符。将 PCR 产物克隆入 pGEM-T Easy 载体中, 测序结果在 NCBI 的 BLAST RefSeqGene 上比对提示获得甜瓜 SQS 基因 ORF 序列。

2.4 生物信息学分析

2.4.1 甜瓜 SQS 核苷酸序列分析 将获得的甜瓜基因序列上传至 NCBI (<http://www.ncbi.nlm.nih.gov/>) Analyz 中的 ORF Finder 工具对甜瓜 SQS 基因 ORF 核苷酸序列进行分析, 结果提示甜瓜 SQS 基因 ORF 编码 417 个氨基酸, 等电点为 7.56。将甜瓜 SQS 基因 ORF 与其它植物 SQS 基因进行核苷酸序列比对构建系统树, 显示系统发生分类结果与形态学分类

结果一致(图 1)。

2.4.2 甜瓜 SQS 氨基酸序列分析 结合 NCBI 中的 ORF Finder 工具以及 DNASTART 软件将甜瓜 SQS 基因 ORF 核苷酸序列推导出其氨基酸序列,甜瓜 SQS 与葫芦科其它植物氨基酸序列比对。图 2 显示,与酶活性中心相关的⁴⁷VSRSF⁵²、Y⁷⁰、R⁷⁴、R⁷⁷DTVED⁸¹、Y¹⁶⁴、G²⁰²、L²⁰⁵、²¹²RDYLED²¹⁷、R²²⁵均高度保守。NetPhos2.0 Server 分析提示甜瓜 SQS 共有 17 个磷酸化位点,其中丝氨酸磷酸化位点 6 个,分别为 S⁴⁸、S¹⁹⁶、S²⁴⁹、S²⁸¹、S³⁵⁸、S⁴⁰⁷。苏氨酸磷酸化位点 6 个,分别为 T⁷⁸、T⁸³、T¹⁴⁴、T¹⁶¹、T³¹⁸、T³²⁷。酪氨酸磷酸化位点 5 个,分别为 Y¹⁶⁵、Y²³⁶、Y²⁴⁵、Y²⁷³、Y³⁸⁷。Prot-Param 分析结果提示,甜瓜 SQS 为亲水性蛋白,多肽链中包含 50 个碱性氨基酸、49 个酸性氨基酸。极性氨基酸 100 个,疏水性氨基酸 159 个,GRAVY 值为-0.065。

2.4.3 甜瓜 SQS 蛋白质结构生物信息学分析 二级结构分析提示,该蛋白的二级结构中 α 螺旋占 64.99%,延伸主链占 7.19%,无规则卷曲占 27.82%,甜瓜 SQS 主要由 α 螺旋构成(图 4:A)。结构域预测结果提示,甜瓜 SQS 氨基酸属于异戊二烯合酶家族,具结合法呢酰基二磷酸及镁离子的结合位点,以及富含天冬氨酸模体(图 4:B)。跨膜区分析提示,甜瓜 SQS 具两个跨膜区,分别为 282 到 304 氨基酸之间及 386 到 408 氨基酸之间,1 到 281 氨基酸及 409 到 417 氨基酸位于膜外侧,305 到 385 氨基酸位于膜内侧。

三级结构预测提示,甜瓜 SQS 为单体蛋白结构(图 5),结合文献(Pandit et al, 2000; 侯嵘等, 2011),其活性中心主要由几个 α 螺旋围绕形成一穴状活性中心结构,其 17 个磷酸化位点均位于蛋白质的表面。

3 讨论

SQS 是葫芦科植物 Cus 生物合成的关键酶,其细胞定位于内质网膜(Pandit et al; 殷秀梅等, 2012)。尽管药用 Cus 原料主要来自甜瓜,但目前对甜瓜 SQS 的结构与功能仍不清楚。本研究基于克隆的甜瓜 SQS 基因 ORF 推导出甜瓜 SQS 氨基酸序列。序列分析结果提示,甜瓜 SQS 是由 417 氨基酸残基构成的单体酶,二级结构以 α 螺旋占为主。其活性中心为主要由几个 α 螺旋围绕形成的穴状结

构。该酶有两段跨膜区,分别位于 282 到 304 及 386 到 408 氨基酸残基之间,是酶蛋白锚定于内质网膜上的肽段。1 到 281 及 409 到 417 氨基酸残基的肽段位于酶蛋白的内质网膜外侧,其中包括与酶活性中心相关的⁴⁷VSRSF⁵²、Y⁷⁰、R⁷⁴、⁷⁷DTVED⁸¹、Y¹⁶⁴、G²⁰²、L²⁰⁵、²¹²RDYLED²¹⁷和 R²²⁵等保守序列。这些肽段和氨基酸残基的分布特点应与酶的催化功能相适应。

序列分析发现,一些磷酸化位点可能是甜瓜 SQS 活性调节的关键部位。该酶共有 17 个磷酸化位点,均位于蛋白质的表面。其中 S⁴⁸、S¹⁹⁶、S²⁴⁹、S²⁸¹、T⁷⁸、T⁸³、T¹⁴⁴、T¹⁶¹、Y¹⁶⁵、Y²³⁶、Y²⁴⁵和 Y²⁷³ 12 个磷酸化位点位于酶蛋白的内质网膜外侧,S³⁵⁸、T³¹⁸和 T³²⁷ 3 个磷酸化位点位于酶蛋白的内质网内侧,S⁴⁰⁷和 Y³⁸⁷ 2 个磷酸化位点位于酶蛋白的跨膜区。位于酶蛋白的内质网外侧的 12 个磷酸化位点可能参与酶的活性调节。特别是 S⁴⁸恰好位于与酶活性中心相关⁴⁷VSRSF⁵²模体中,因而提示其可能是酶活性调节的一个关键性磷酸化位点。此外,在陆生 SQS 中目前已发现 21 个正选择位点,分别为 Q³³、Q⁵⁹、E¹³³、S¹⁸⁹、S¹⁹⁴、S¹⁹⁶、I²⁶⁵、K²⁷⁸、V³⁸¹、R³⁸²、S³⁸³、E³⁸⁴、P³⁸⁵、I³⁸⁶、P³⁸⁹、T³⁹⁰、A⁴⁰⁸、R⁴¹⁰、F⁴¹¹、P⁴¹²和 N⁴¹⁴(刘墉等,2013)。其中 S¹⁹⁶分布于甜瓜 SQS 表面并位于酶蛋白的内质网膜外侧,因此,正选择位点 S¹⁹⁶可能是该酶活性调节的另一个关键性磷酸化位点。

以甜瓜 SQS 基因 ORF 序列构建系统发生树得到系统发生分类结果与形态学分类结果一致,表明甜瓜等陆生植物的 SQS 基因构建系统发生树可以较好地反映不同物种间的亲缘关系。推衍的甜瓜 SQS 氨基酸序列与其它植物比对结果显示,该酶不具备特征性的氨基酸位点。葫芦科植物 SQS 与其他植物 SQS 的氨基酸序列比对显示,葫芦科植物 SQS 氨基酸序列也不具备特征性的氨基酸位点。

总之,本研究克隆了葫芦科植物甜瓜 SQS 基因,并应用生物信息学方法分析甜瓜 SQS 分子结构特点,发现该酶磷酸化位点 S⁴⁸和 S¹⁹⁶可能是酶活性调节的关键位点。这些结果为 SQS 酶活性和 Cus 生物合成调节研究提供了新的线索和实验依据。

参考文献:

HOU R, GONG NX, LIU FT, 2011. Cloning and bioinformatics analysis of squalene synthase gene (SQS) from *Jatropha curcas* L. [J]. *J Trop Subtrop Bot*, 19(5): 391-399. [侯嵘, 龚诺希, (下转第 387 页 Continue on page 387)