

胡麻 *SUC* 基因家族的鉴定与生物信息学分析

齐燕妮, 李闻娟, 王利民, 赵 玮, 党 照, 谢亚萍, 张建平

(甘肃省农业科学院作物研究所, 甘肃 兰州 730070)

摘要: 利用多个物种的 *SUC* 基因蛋白序列在胡麻基因组内进行 BlastP 分析, 通过 Pfam 确认结构域, 获得胡麻 *SUC* 基因家族成员, 并进行基因结构分析; 对蛋白分子量、等电点、信号肽、跨膜结构域、糖基化修饰位点、亚细胞定位、Motif 及二级结构进行预测。结果鉴定得到 12 个胡麻 *SUC* 基因家族成员, 大部分成员含有 4 个以上的外显子, 10 个以上跨膜结构域, 并获 4 个特征性 Motif。进化树分析显示, 胡麻 *SUC* 分别属于 SUT1、SUT2 和 SUT4 家族。

关键词: 胡麻; *SUC* 基因家族; 基因结构分析; 系统进化分析

中图分类号: S565.9 **文献标志码:** A **文章编号:** 1001-1463(2019)08-0035-06

doi: 10.3969/j.issn.1001-1463.2019.08.008

Identification and Bioinformatics Analysis of the *SUC* Gene Family in Flax

QI Yanni, LI Wenjuan, WANG Limin, ZHAO Wei, DANG Zhao, XIE Yaping, ZHANG Jianping
(Institute of Crops, Gansu Academy of Agricultural Sciences, Lanzhou Gansu 730070, China)

Abstract: In this study, multiple *SUC* protein sequences were analyzed in flax genome using BlastP, and the structural domain was confirmed by Pfam. The members of the *SUC* gene family were obtained. The analysis of gene structure, prediction of protein molecular weight, isoelectric point, signal peptide, transmembrane domain, glycation modification site, subcellular localization, Motif and secondary structure were carried out on the family members. The results showed that 12 members of flax *SUC* gene family were identified, containing 4 characteristic motifs, most of which had more than 4 exons, more than 10 transmembrane domains. The evolutionary tree analysis showed that flax *SUC* belonged to SUT1, SUT2 and SUT4 families, respectively.

Key words: Flax; *SUC* gene family; Gene structure analysis; Phylogeny analysis

胡麻 (*Linum usitatissimum* L.) 即油用亚麻, 属于亚麻科亚麻属, 主要种植在中国、加拿大、印度^[1-2]。胡麻籽含油率在 40% 以上, 油脂中含有丰富的不饱和脂肪酸, 特别是 α - 亚麻酸含量在 55% 左右, 是人类从膳食中摄取 α - 亚麻酸的最好来源之一^[1,3]。

α - 亚麻酸具有多种营养和保健功能, 包括增强大脑健康和提高智力、预防和治疗高血压、抗癌、预防心脑血管疾病、降低血液胆固醇等^[4-5]。由于人们对健康的日益重视, 亚麻籽已经被广泛应用于奶制品、肉制品、烤制品等多种功能性食品^[6]。

收稿日期: 2019-07-19

基金项目: 国家自然科学基金(31760426)资助。

作者简介: 齐燕妮(1987—), 女, 甘肃庆阳人, 研究实习员, 硕士, 主要从事胡麻分子育种研究工作。

Email: qiyanni@gsagr.ac.cn。

通信作者: 张建平(1972—), 男, 陕西蓝田人, 研究员, 主要从事胡麻育种研究工作。Email:

zhangJPzw3@gsagr.ac.cn。

蔗糖是植物的主要光合作用产物,大部分蔗糖需运输到其他组织中代谢或储存。蔗糖可通过胞间连丝在细胞间短距离运输,然而长距离的运输则需要通过韧皮部,需要SWEET (Sugars Will Eventually be Exported Transporters)蛋白和蔗糖转运蛋白(sucrose transporters or carriers, SUT/SUC)共同来完成^[7-8]。蔗糖作为脂肪酸合成的主要底物,由蔗糖转运蛋白运输到种子细胞中^[9],因此蔗糖转运蛋白的转运效率直接影响种子含油率。蔗糖转运蛋白属于MFS (major facilitator superfamily),通常含有12个跨膜结构域^[10]。基于系统发育分析,SUT家族分为5个亚家族,其中为SUT1为双子叶植物特有,SUT3和SUT5为单子叶植物所特有,SUT2和SUT4为单子叶和双子叶植物共有^[11]。

近年来,多个植物基因组数据相继被公布,这为基因家族的研究奠定了基础。目前已经从拟南芥(*Arabidopsis thaliana*)^[12]、葡萄(*Vitis vinifera* L.)^[13]、苹果^[14]、水稻(*Oryza sativa*)^[15]等基因组中鉴定到SUT基因家族,但对胡麻SUT的研究相对较少。我们利用生物信息学方法鉴定了胡麻基因组中SUT基因家族成员,并分析其基因结构、蛋白理化性质、保守基序及系统进化,以期为克隆和深入研究胡麻SUC基因家族各成员的生物学功能奠定基础。

1 材料与方法

1.1 胡麻 SUC 基因家族鉴定

从拟南芥数据库TAIR中获得9个拟南芥SUC基因的蛋白序列,从NCBI获得油菜、大豆、水稻、烟草、玉米、葡萄等物种的SUC蛋白序列(登录号分别为AEG89530.1、NP_001335852.1、AAF90181.1、BAO47334.1、BAA83501.1、AAL32020.1)。使用BlastP,将这些序列与胡麻全基因组蛋白数据库进行比对,E值小于等于 $1E-5$,并用Pfam对比对得到的蛋白序列进行确认(SUC蛋白结构

域Pfam号为PF13347.1)。

1.2 SUC 基因家族的基因及蛋白序列分析

利用IBS在线工具绘制基因结构图,采用Expasy分析蛋白质的分子量、等电点,通过SignalP-5.0预测胡麻SUC蛋白的信号肽,采用SOSUI进行跨膜结构预测,用NetNGlyc 1.0对糖基化修饰位点在线预测。使用Plant-mPloc进行蛋白的亚细胞定位预测,利用MEME对胡麻SUC蛋白保守基序进行分析。

1.3 SUC 基因家族的系统进化分析

利用MEGA6.0软件对胡麻和拟南芥的SUC氨基酸序列进行多重比对,将比对结果用Maximum Likelihood法构建系统进化树,Bootstrap校验值为1000次。

1.4 SUC 蛋白的二级结构分析

利用SOPMA对胡麻SUC蛋白的次级结构进行分析。

2 结果与分析

2.1 胡麻 SUC 基因家族的鉴定及理化性质分析

以拟南芥9个SUC基因和油菜、大豆、水稻、烟草、玉米、葡萄SUC基因的氨基酸序列对胡麻基因组进行BLASTP分析,获得28条期望值小于 $1E-5$ 的候选序列。用Pfam对这些序列进行结构域注释,发现有12条序列含有SUC家族的保守结构域MFS_2,将这些基因分别命名为LuSUC1~LuSUC12(表1)。

比较分析12个胡麻SUC蛋白序列,结构显示不同的蛋白差异较大。蛋白长度为252(LuSUC8)~952 aa(LuSUC4),分子量为26 863.30(LuSUC8)~101 956.83 Da(LuSUC4),等电点为5.99(LuSUC10)~9.42(LuSUC3),糖基化位点为0(LuSUC7、LuSUC8)~5(LuSUC10),跨膜结构域为6(LuSUC7、LuSUC8)~13(LuSUC5)。对上述序列进行信号肽预测,发现所有序列均无信号肽。亚细

胞定位预测结果显示,所有基因都定位在细胞膜,其中 *LuSUC3* 和 *LuSUC4* 除细胞膜外还定位在叶绿体、细胞质、细胞核中。在胡麻的 12 个 SUC 蛋白二级结构中, α -螺旋所占比例最大(41.58%),其次为无规卷曲(37.48%)、 β -折叠(16.62%)、 β -转角所占比例最小(4.78%)(表2)。

2.2 胡麻 SUC 基因家族的基因结构及 motif 预测

为进一步分析胡麻 SUC 基因家族的功能,需根据其 DNA 序列和 CDS 序列分析其基

因结构(表1,图1)。除 *LuSUC6* 和 *LuSUC7* 只含有 1 个外显子,没有内含子外,其余 10 个基因均含有数量不等的外显子和内含子,外显子数量为 2(*LuSUC8*)~13(*LuSUC9*),内含子数量为 1(*LuSUC8*)~12(*LuSUC9*)。Motif 预测结果表明, Motif1、Motif2、Motif6 和 Motif7 为 SUC 家族的特征性、保守性 Motif。这 12 个基因共预测到 22 个 Motif,除 *LuSUC8* 只含有 4 个特征性 Motif 外,其余 11 个基因都含有 Motif11。另外,除 *LuSUC7* 和 *LuSUC8* 外,有 10 个基因还含有

表 1 胡麻 SUC 基因家族成员及其序列特征

基因名称	蛋白长度 /aa	分子量 /Da	等电点	外显子数量	信号肽	跨膜结构	N-糖基化位点	亚细胞定位
<i>LuSUC1</i>	520	55 023.49	8.58	5	无	11	454	细胞膜
<i>LuSUC2</i>	520	55 216.63	8.58	5	无	10	454	细胞膜
<i>LuSUC3</i>	951	101 575.91	9.42	8	无	12	468,532,576,703	细胞膜,叶绿体,细胞质,细胞核
<i>LuSUC4</i>	952	101 596.83	9.28	8	无	12	469,577,637,704	细胞膜,叶绿体,细胞质,细胞核
<i>LuSUC5</i>	543	57 572.08	7.05	4	无	13	479	细胞膜
<i>LuSUC6</i>	494	52 347.37	8.83	1	无	11	197	细胞膜
<i>LuSUC7</i>	277	29 671.46	6.94	1	无	6	无	细胞膜
<i>LuSUC8</i>	252	26 863.30	9.12	2	无	6	无	细胞膜
<i>LuSUC9</i>	612	65 292.36	6.05	13	无	11	203,333,423,599	细胞膜
<i>LuSUC10</i>	612	65 180.19	5.99	12	无	11	203,333,337,423,599	细胞膜
<i>LuSUC11</i>	507	54 703.21	8.39	4	无	11	13,156	细胞膜
<i>LuSUC12</i>	511	55 067.64	8.59	5	无	11	13,160	细胞膜

表 2 胡麻 SUC 蛋白二级结构

蛋白名称	α -螺旋	β -折叠	β -转角	无规卷曲
LuSUT1	236(45.38%)	74(14.23%)	20(3.85%)	190(36.54%)
LuSUT2	231(44.42%)	70(13.46%)	18(3.46%)	201(38.65%)
LuSUT3	366(38.49%)	184(19.35%)	62(6.52%)	339(35.65%)
LuSUT4	368(38.66%)	186(19.54%)	61(6.41%)	337(35.40%)
LuSUT5	235(43.28%)	79(14.55%)	19(3.50%)	210(38.67%)
LuSUT6	227(45.95%)	85(17.21%)	17(3.44%)	165(33.40%)
LuSUT7	116(41.88%)	43(15.52%)	14(5.05%)	104(37.55%)
LuSUT8	97(38.49%)	55(21.38%)	23(9.13%)	77(30.56%)
LuSUT9	217(35.46%)	94(15.36%)	24(3.92%)	277(45.26%)
LuSUT10	219(35.78%)	96(15.69%)	28(4.58%)	269(43.95%)
LuSUT11	222(43.79%)	86(16.96%)	21(4.14%)	178(35.11%)
LuSUT12	242(47.36%)	70(13.70%)	16(3.13%)	183(35.81%)
平均值	231.3(41.58%)	93.5(16.62%)	26.9(4.78%)	210.8(37.48%)

Motif3、Motif4、Motif5、Motif8、Motif9 和 Motif15，还有 9 个基因含有 Motif10，说明

Motif11、Motif4、Motif5、Motif8、Motif9、Motif15 和 Motif10 的功能也相当保守(图2)。

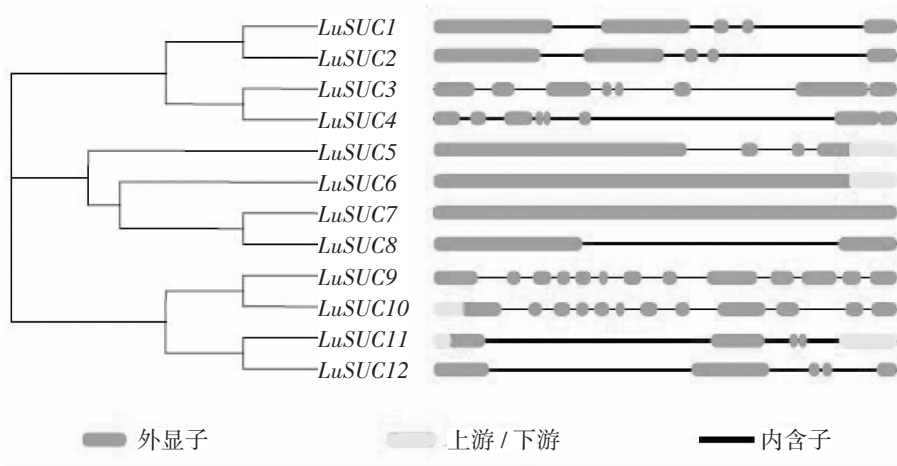


图 1 胡麻 SUC 基因结构



图 2 胡麻 SUC 基因家族 Motif 分析

在许多保守的氨基酸位点, 这些位点所在的区域应为该蛋白的保守功能区。从进化树上可以看出, 胡麻 *SUT* 基因分为 3 个家族, *LuSUC1~LuSUC8* 与拟南芥的 *AtSUC1*、*AtSUC2*、*AtSUC5*、*AtSUC6*、*AtSUC7*、*AtSUC8*、*AtSUC9* 聚在一起, 属于 SUT1 家族; *LuSUC9* 和 *LuSUC10* 与属于 SUT2 家族的拟南芥 *AtSUC3* 聚为一支; 而 *LuSUC11* 和 *LuSUC12* 与属于 SUT4 家族的拟南芥 *AtSUC4* 聚为一支。

3 结论与讨论

从胡麻基因组鉴定得到 12 个 *SUC* 基因家族成员, 并对其进行了生物信息学分析。12 个 *SUC* 基因外显子数量为 2~13, 编码蛋白长度为 252~952 aa, 含有 6~13 个跨膜结构域, 4 个特征性 Motif。系统进化树分析显示, 胡麻 *SUC* 分别属于 SUT1、SUT2 和 SUT4 家族, 与双子叶植物拟南芥分类结果相同^[12]。

总体来看, 胡麻 *SUC* 基因家族成员之间虽然具有比较保守的功能结构域, 但蛋白长度及序列、基因结构、跨膜结构域数量等也存在明显的差异, 需进行更全面的研究, 才能深入解析每个家族成员的功能。

参考文献:

- [1] OOMAH, B D. Flaxseed as a functional food source[J]. *Journal of the Science of Food and Agriculture*, 2001(81): 889-894.
- [2] 张运晖, 赵 瑛, 罗俊杰. 甘肃胡麻产业发展浅议[J]. *甘肃农业科技*, 2013(7): 54-55.
- [3] GREEN A G. Genetic control of polyunsaturated fatty acid biosynthesis in flax (*Linum usitatissimum*) seed oil[J]. *Tag. Theoretical & Applied Genetics*, 1986(72): 654-661.
- [4] SIMOPOULOS A P. Essential fatty acids in health and chronic disease[J]. *American Journal of Clinical Nutrition*, 1999(70): 560-569.
- [5] PELLIZZONMA, BILLHEIMERJT, BLOEDON AT, et al. Flaxseed reduces plasma cholesterol levels in hypercholesterolemic mouse models[J]. *J. Am. Oil. Chem. Soc.*, 2007(26): 66-75.
- [6] GOYAL A, SHARMA V, UPADHYAY N, et al. Flax and flaxseed oil: an ancient medicine & modern functional food[J]. *J. Food Sci. Technol.*, 2014(9): 1633-1653.
- [7] EOMJS, CHENLQ, SOSSOD, et al. SWEETs, transporters for intracellular and intercellular sugar translocation[J]. *Curr. Opin. Plant Biol.*, 2015(25): 53-62.
- [8] SAUER N. Molecular physiology of higher plant sucrose transporters[J]. *FEBS Lett*, 2007(581): 2309-2317.
- [9] HARITATOS E, MEDVILLE R, TURGEON R. Minor vein structure and sugar transport in *Arabidopsis thaliana*[J]. *Planta*, 2000(211): 105-111.
- [10] REINDERS A, SIVITZ A B, WARD J M. Evolution of plant sucrose uptake transporters[J]. *Front. Plant Sci.*, 2012(3): 22.
- [11] KUHN C, GROF C P L. Sucrose transporters of higher plants[J]. *Current Opinion in Plant Biology*, 2010(13): 287-298.
- [12] 高 蕾, 肖文芳, 李文燕, 等. 拟南芥蔗糖转运蛋白(SUTs)的功能研究进展[J]. *分子植物育种*, 2011(9): 251-255.
- [13] 于 菲, 卢 江, 张雅丽. 葡萄中蔗糖转运蛋白家族的序列分析及功能推测[J]. *生物信息学*, 2011(9): 35-38.
- [14] 史 鹏. 苹果蔗糖转运蛋白基因的生物信息学分析[J]. *山西农业科学*, 2019(47): 966-969; 1039.
- [15] AOKI N, HIROSE T, SCOFIELD G N, et al. The sucrose transporter gene family in rice[J]. *Plant and Cell Physiology*, 2003(44): 223-232.

(本文责编: 陈 珩)