

冯紫洲,刘春梅,衡宝山,等. 蓖麻 *RcSCP* 基因克隆与生物信息学分析[J]. 江苏农业科学,2017,45(12):27-30.

doi:10.15889/j.issn.1002-1302.2017.12.006

蓖麻 *RcSCP* 基因克隆与生物信息学分析

冯紫洲¹, 刘春梅², 衡宝山³, 张继星^{3,4}, 陈永胜^{3,4}, 王 云¹, 刘海臣³, 穆莎茉莉³, 冀照军³, 霍红雁³

(1. 内蒙古民族大学农学院, 内蒙古通辽 028000; 2. 内蒙古阿鲁科尔沁旗农业局, 内蒙古赤峰 025550;

3. 内蒙古民族大学生命科学院, 内蒙古通辽 028000; 4. 内蒙古自治区高校蓖麻产业工程技术研究中心/
内蒙古自治区蓖麻育种重点实验室/内蒙古自治区蓖麻产业协同创新培育中心, 内蒙古通辽 028000)

摘要:植物的丝氨酸羧肽酶(serine carboxypeptidases,简称SCP)是一类属于 α/β 水解酶家族的蛋白酶,在许多生化途径(包括次生代谢产物的生物合成、催化酰基转移、除草剂共轭、种子萌发相关蛋白质的降解过程)中起着重要作用。丝氨酸羧肽酶基因是植物的一种抗逆基因,在植物生长发育及抵抗逆境方面发挥着重要作用。蓖麻作为重要的油料作物,用途广泛。本研究利用cDNA末端快速克隆(RACE)方法克隆蓖麻*RcSCP*基因的3'端、5'端片段,最后设计全长引物获得蓖麻*RcSCP*全长序列,并对其进行生物信息学分析。结果表明,丝氨酸羧肽酶蛋白基因的cDNA完整开放阅读框序列为1377bp,推测它可编码458个氨基酸;蓖麻*RcSCP*蛋白含有丝氨酸羧肽酶蛋白家族保守区,预测它定位于细胞质中。

关键词:蓖麻;RACE;丝氨酸羧肽酶;克隆

中图分类号: S565.601 **文献标志码:** A **文章编号:** 1002-1302(2017)12-0027-04

丝氨酸羧肽酶(serine carboxypeptidases,简称SCP)是一类属于 α/β 水解酶家族的蛋白酶,在植物生长发育过程中参与多肽和蛋白质的加工、修饰、降解等多个重要环节,并且在许多生化途径包括次生代谢产物的生物合成、催化酰基转移、除草剂共轭、种子萌发相关蛋白质降解过程中起着重要作用^[1]。目前,已从多种植物中分离得到SCP蛋白及丝氨酸羧肽酶类蛋白(serine carboxypeptidases-like proteins,简称

SCPLs)^[2-4]。SCP基因是一种抗逆基因,在植物生长发育及抵抗逆境方面起着重要作用。叶玲飞通过对水稻耐旱方面的研究发现,水稻*OsSCP*基因缺失会导致植株耐旱性和耐高温性降低^[5],表明SCP基因在耐旱、耐高温方面发挥重要作用;刘丽等将玉米*ZmSCP*基因导入到烟草中发现,转基因烟草的抗病性、耐盐性、抗氧化性均有提高,且其自身防卫基因的表达量也有明显升高^[6],这些均表明SCP基因在植物抵抗逆境方面起着重要作用。

蓖麻(*Ricinus communis* L.)是世界十大油料作物之一,是双子叶一年生或多年生大戟科蓖麻属草本植物^[7]。蓖麻原产于非洲,属热带多年生灌木,我国的蓖麻由印度传入,至今已有1400多年的历史^[8]。蓖麻有700多种工业用途,主要用途是从蓖麻籽粒中提取蓖麻油^[9]。蓖麻根系发达,枝叶生长量大,能在沙化土壤中生长,起到防风固沙的作用。蓖麻育种的研究主要集中在诱变育种和杂交育种方面,已育成一批旨在提高产量和改善品质的新品种,在生产上取得显著的成绩^[7]。然而针对抵抗逆境育种的研究很少,关于蓖麻抗逆相关基因功能及抗逆分子育种的研究很少。研究表明,SCP在植物抵抗逆境方面起重要作用。本研究以蓖麻为材料,采用cDNA末端快速克隆(rapid-amplification of cDNA ends,简称

收稿日期:2016-09-23

基金项目:国家自然科学基金(编号:31260336);内蒙古自治区蓖麻育种重点实验室开放基金(编号:MDK2016030);内蒙古高校蓖麻产业工程技术中心项目(编号:BMJ2011007);内蒙古自治区高校蓖麻产业工程技术研究中心与军事医学科学院军事兽医研究所合作项目(编号:MV-201201-1)。

作者简介:冯紫洲(1990—),男,河北石家庄人,硕士研究生,主要从事植物基因工程研究。E-mail: fzz_1990@163.com。

通信作者:张继星,博士,教授,主要从事植物基因工程研究;Tel:(0475)8314053, E-mail: zhangjixing@imun.edu.cn。陈永胜,博士,教授,主要从事植物基因工程研究;E-mail: chenys2000@163.com。

[8] Krogh A, Larsson B, Heijne G V, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes[J]. *Journal of Molecular Biology*, 2001, 305(3): 567-580.

[9] Saeed A I, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis[J]. *Biotechniques*, 2003, 34(2): 374-378.

[10] Charfeddine M, Saïdi M N, Charfeddine S, et al. Genome-wide analysis and expression profiling of the ERF transcription factor family in potato (*Solanum tuberosum* L.) [J]. *Molecular Biotechnology*, 2015, 57(4): 348-358.

[11] Noiraud N, Mauroussat L, Lemoine R. Transport of polyols in higher plants [J]. *Plant Physiology and Biochemistry*, 2001, 39(9): 717-728.

[12] Massa A N, Childs K L, Lin H, et al. The transcriptome of the reference potato genome *Solanum tuberosum* Group Phureja clone DM1-3 516R44 [J]. *PLoS One*, 2011, 6(10): e26801.

[13] Singh A K, Sharma V, Pal A K, et al. Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.) [J]. *DNA Research*, 2013, 20(4): 403-423.

RACE)等技术,获取蓖麻丝氨酸羧基蛋白酶基因 *RcSCP*,并进行生物信息学分析,为进一步研究 *RcSCP* 基因的转化及功能验证奠定基础。

1 材料与方法

1.1 材料

蓖麻种植于内蒙古民族大学植物生物技术实训室,品种为通蓖5号,取新鲜叶片为供试材料。Ex *Taq* DNA 聚合酶、克隆载体 pMD18-T、3'RACE 和 5'RACE 试剂盒均购自宝生物工程(大连)有限公司,感受态细胞 JM109 保存于笔者所在实验室,其他试剂均为国产分析纯。

1.2 蓖麻叶片总 RNA 的提取及引物设计

使用改良的 TRIzol 提取法提取总 RNA^[10-11],几乎没有 DNA 和蛋白质的污染,其纯度较高且具有较好的效果。采用 RACE 法克隆蓖麻 *SCP* 基因 cDNA 全长。根据其他植物上的 NHX 蛋白的保守序列设计,利用 Primer Premier 5.0 软件,设计 3'RACE 简并引物 3F1、3F2;结合已经克隆得到的蓖麻 *SCP* 3'端基因片段,设计 5'RACE 特异性引物 5F1、5F2;采用 DNAMAN 软件对测定的核苷酸 5'端、3'端序列进行电子序列的拼接合并,从而获得 cDNA 全长,在该基因编码区域的两端设计全长特异性引物 F、R。各引物序列见表 1。

表 1 *RcSCP* 基因克隆引物

引物名称	引物序列(5'→3')
3F1	CRGRRGCWTGTCMTCRCSCGGMGWA
3F2	TWTGYTSCTNARCTSGCKCAWCTSA
5F1	AATCCATGAGACCAGAAGAAGCTCAG
5F2	GGATTGCCTAGAGCAACGCCITTC
F	ATGGAATAATGGCATTTGCAGTGG
R	TCAGAATGCTCCGGTAGAGGCTGC

注:W、M、B、R、Y、N、S、K 为简并碱基。

1.3 蓖麻 *SCP* 基因全长 cDNA 的克隆

蓖麻 *SCP* 基因 3'端片段克隆按照 3'-Full RACE Core Set with PrimeScript™ RTase(6106, TaKaRa) 试剂盒说明书进行。使用 3'RACE Adaptor(来自 3'-Full RACE Core)为反转录引物进行反转录,反应条件:42℃ 60 min;70℃ 15 min;4℃ 保存。第 1 轮 PCR 利用引物 3F1 和 3'RACE Outer Primer(来自 3'-Full RACE Core)进行扩增,反应条件为 94℃ 3 min;94℃ 30 s,56℃ 40 s,72℃ 60 s,20 个循环;72℃ 10 min;4℃ 保存。第 2 轮 PCR 利用引物 3F2 和 3'RACE Inner Primer(来自 3'-Full RACE Core)进行扩增,反应条件为 94℃ 3 min;94℃ 30 s,56℃ 40 s,72℃ 60 s,30 个循环;72℃ 10 min;4℃ 保存。蓖麻 *SCP* 基因 5'端片段克隆反应按照 5'-Full RACE Kit with TAP(6107, TaKaRa) 说明书进行。使用 Random 9 mers(来自 5'-Full RACE Kit)为反转录引物进行反转录获得 cDNA 第 1 条链。第 1 轮 PCR 利用引物 5F1 和 5'RACE Outer Primer(来自 5'-Full RACE Kit)进行扩增,反应条件为 94℃ 3 min;94℃ 30 s,55℃ 40 s,72℃ 60 s,20 个循环;72℃ 10 min;4℃ 保存。第 2 轮 PCR 利用引物 5F2 和 5'RACE Inner Primer(来自 5'-Full RACE Kit)进行扩增,反应条件为 94℃ 3 min;94℃ 30 s,58℃ 40 s,72℃ 60 s,30 个循环;72℃ 10 min;4℃ 保存。反转录使用 Oligo dT 为

引物,反应条件为 65℃ 5 min;42℃ 60 min;70℃ 5 min;4℃ 保存。PCR 利用引物 F、R 进行扩增,反应条件为 94℃ 3 min;94℃ 30 s,55℃ 40 s,72℃ 60 s,30 个循环;72℃ 10 min;4℃ 保存。PCR 产物用北京索莱宝科技有限公司的琼脂糖凝胶回收试剂盒回收纯化,与 pMD18-T 连接,转化 JM109 感受态细胞,菌落 PCR 和酶切验证正确的阳性克隆送至北京华大基因研究中心测序。

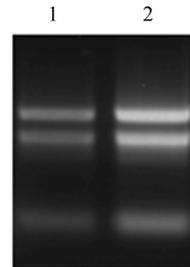
1.4 蓖麻 *SCP* 基因生物信息学分析

对测序结果进行分析,确定克隆获得蓖麻 *SCP* 基因 cDNA 序列完整开放阅读框。利用 SWISS-Model、DNAMAN、WoLF PSORT 等生物学软件进行生物信息学分析。

2 结果与分析

2.1 蓖麻叶片总 RNA 活性的检测

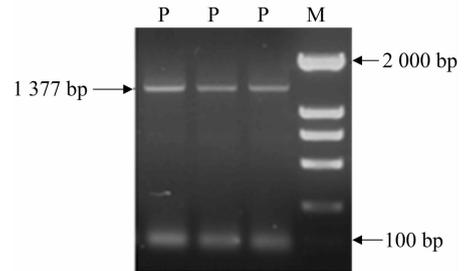
将提取好的蓖麻总 RNA 用 DEPC 水进行 10 倍稀释,然后进行琼脂糖凝胶电泳(图 1),并检测其纯度 $D_{260\text{nm}}/D_{280\text{nm}}$ 值=2.10,纯度较好,满足下一步试验要求。



1、2 为蓖麻叶片总 RNA
图 1 蓖麻总 RNA 电泳结果

2.2 蓖麻 *RcSCP* 基因克隆

采用 RACE 法克隆蓖麻 *SCP* 基因,利用设计的 3'端、5'端引物及 cDNA 全长特异性引物进行扩增,最后得到蓖麻 *SCP* 基因 cDNA 全长 1 377 bp 的片段(图 2),命名为 *RcSCP* (GenBank 登录号:KX267676)。



M—DL-2000 Marker; P—PCR 电泳条带
图 2 PCR 扩增电泳结果

2.3 蓖麻 *RcSCP* 基因生物信息学分析

利用 DNAMAN 软件对获得的蓖麻 *RcSCP* 基因全长序列进行分析。该序列包括 1 377 bp 完整的开放阅读框,可推测编码 458 个氨基酸的多肽。该多肽的分子量为 51.245 ku,等电点(pI 值)为 5.67, pH 值 7.0 时电荷为 -6.29。预测 *RcSCP* 基因的保守区,使用美国国立生物技术信息中心(NCBI)网站的保守区分析(conserved domains)(http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb/cgi)功能,结果表明,蓖麻 *RcSCP* 含有丝氨酸羧基蛋白酶蛋白家族保守区,表明 *RcSCP* 与

这些家族的蛋白具有相似的功能,是蓖麻的一种丝氨酸羧肽酶蛋白(图3)。运用 SOPMA 软件预测蓖麻 RcSCP 蛋白的二级结构,结果表明,该蛋白的 α -螺旋为 31.88%,延伸链为 27.29%, β -转角为 11.14%,无规则卷曲为 26.69%(图4)。利用在线程序 SWISS-Model 的自动模式对该蛋白的三级结构进行预测分析,得到该蛋白的三级结构(图5)。利用 DNA-

MAN 软件对蓖麻 RcSCP 与大豆(*Glycine max*, XP_003531913.1)、拟南芥(*Arabidopsis thaliana*, NP_564298.1)、麻风树(*Jatropha carcas L.*, XP_012088729.1)的丝氨酸羧肽酶蛋白的氨基酸序列进行对比,同源性分别为 77.66%、76.79%、86.39%(图6)。

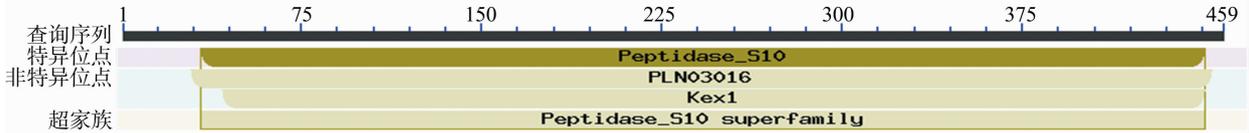


图3 SCP 蛋白保守区预测

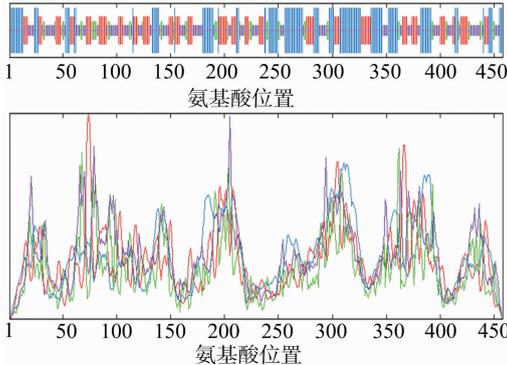


图4 RcSCP 二级结构预测

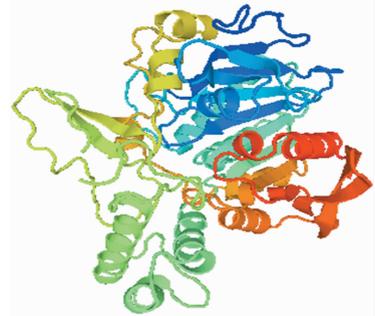


图5 RcSCP 蛋白三级结构预测

GmSCP	..MCSLLWSSIALCVAFLLELGVVHPSPSHHRITRLPGQPHVQFHQSGYVTVDDKNQR	58
AtSCP	..MSPLQWLITISFALIIIFHSLTVSSSVLHSDRVRTLPQGPRVGFQOYSGYVTVDDKKQR	58
JcSCP	MPYLTIEVMALAMVLLQLCFMKVVCHPFLFDRITQLPGQPQVGFQOYSGYVTVDNKEQR	60
RcSCPMEIMAFVAVLLQICFLMGVNSNPSLFDKIVKLPGQPOIGFHQSGYVTVDEKKQR	55
GmSCP	ALFYFAEAERKDALSKPLVWLWLNCGPGCSSLGVGAFSENGPFRFKKCKLVRNCFSWNREA	118
AtSCP	ALFYFAEAETNPSSKPLVWLWLNCGPGCSSLGVGAFSENGPFRFKCPILVKNQHSWNQEA	118
JcSCP	ALFYFAEAETDPASKPLVWLWLNCGPGCSSLGVGAFSENGPFRFSGEVLVKNQHSWNREA	120
RcSCP	ALFYFAEAETDPASKPLVWLWLNCGPGCSSLGVGAFSENGPFRFSCQVLVKNQHSWNREA	115
GmSCP	NMLYLETPIGVGFSYSTDTSSYECVNDKLTARDNLVFLQKWFVKFPEYRNRSLFITGESY	178
AtSCP	NMLYLETPIGVGFSYSTQSSHYECVNDKLTARDNLVFLQKWFVKFPHYRNRSLFITGESY	178
JcSCP	NMLYLETPIGVGFSYSTDTSSYETVNDKLTARDNLVFLQKWFVKFPOYRNRSLFITGESY	180
RcSCP	NMLYLESPIGVGFSYSVDTSFYEAVNDKLTARDNLVFSQKWFVKFPOYRNRSLFITGESY	175
GmSCP	AGHYVPQLAELMLIQFNKKEKLFNLKGIATGNPVLGFATDFNSRAEFFWSHGLISDITYKM	238
AtSCP	AGHYVPQLAELMIQYNKKHHLFNLKGIATGNPVLGFATDFNSRAEYFWSHGLISDITYKM	238
JcSCP	AGHYVPQLAELMLIEFNKHKLFNLKGIATGNPVLGFATDFNSRAEFFWSHGLISDITYKM	240
RcSCP	AGHYVPQLAELMLIEFNKQKLFNLKGVATGNPVLGFATDFNSRAEFFWSHGLISDITYKL	235
GmSCP	FTSVCNYSRYVREYYNCAVSPICSSVMVSQVTTETSRFVDKYDVTLDVCLSSVFSQTKVLN	298
AtSCP	FTSVCNYSRYVSEYYRGSMSMCSKVMSQVSTETSRFVDKYDVTLDVCLPSSLVSKVVS	298
JcSCP	FTSVCNYSRYVNEYRGSVSPICSRVMQVSRRETSRFVDKYDVTLDVCLSSVLAOSNALS	300
RcSCP	FTSVCNYSRYVSEYYRGSVSPICSRVMQVSRRETSKYVDKYDVTLDVCLSSVLSKILS	295
GmSCP	PQQVTEITDVCVEDETVNYLNRKIDVQSAMHAHLVGVORWSACSNVLDYELLDLEIPTITV	358
AtSCP	PNQGESVDVCVEDETVNYLNRKIDVQEALHARLIGVREWTVCSNVLDYQLLDVEIPTINI	358
JcSCP	PKQVGDNIDVCVEDETVNYLNRQDVQKALHARLVGVRRTVCSNILDYELLDLEIPTIAI	360
RcSCP	PHQVADNVDVCVEDETVNYLNRKIDVQMALHARLVGVHQTVCSSILDYELLDLEIPTISI	355
GmSCP	VGKLVKEGIPVLVYSGDQDSVILPTGSRTLVHKLAKELGLINTVPPYRVWFEGQQVGGWTQ	418
AtSCP	VGSLVKAGVPLVYSGDQDSVILPTGSRTLVSRLLAKQLGLRITVPPYRVWFAGQQVGGWTQ	418
JcSCP	VGRLLRAGIPVLVYSGDQDSVILPTGSRTLVVRGLAEKLGKITSVPPYRVWFEGQQVGGWTQ	420
RcSCP	VGKLTAEAGVPLVYSGDQDSVILPTGSRTLVHGLAEELGLKITVPPYRVWFEGQQVGGWTQ	415
GmSCP	VYGNILSFATIRGASHEAPFSQPERSLVLFKSFLECCPLPE	460
AtSCP	VYGNVLSFATIRGASHEAVPFSQPERSLVLFKAFLDGHPLE	460
JcSCP	VYGNILSFATIRGASHEAPFSQPERSLVLFKAFLECRPLPEA	462
RcSCP	VYGNILSFATIRGASHEAPFSQPERSLVLFKAFLECCPLPEA	457

Gm—大豆; At—拟南芥; Jc—麻疯树; Rc—蓖麻

图6 不同植物SCP 基因氨基酸序列同源性比较

利用 WoLF PSORT 软件对蓖麻 RcSCP 蛋白进行亚细胞定位预测,结果显示,蓖麻 RcSCP 在线粒体得分最高,为 4,其次是细胞质基质、内质网、细胞核,得分均为 2,细胞膜得分最低,为 1。因此,RcSCP 可能分布在细胞内的细胞质中,且主要分布在线粒体中。利用 DNAMAN 软件对不同植物来源的丝氨酸羧肽酶蛋白之间的遗传关系进行系统发育分析。结果表明,RcSCP 与麻风树、胡杨、毛果杨等双子叶植物的丝氨酸羧肽酶蛋白等的亲缘关系较近,而与玉米、水稻等单子叶植物的丝氨酸羧肽酶蛋白亲缘关系较远(图 7)。序列比较结果及进化树分析结果表明,RcSCP 应为丝氨酸羧肽酶蛋白基因。

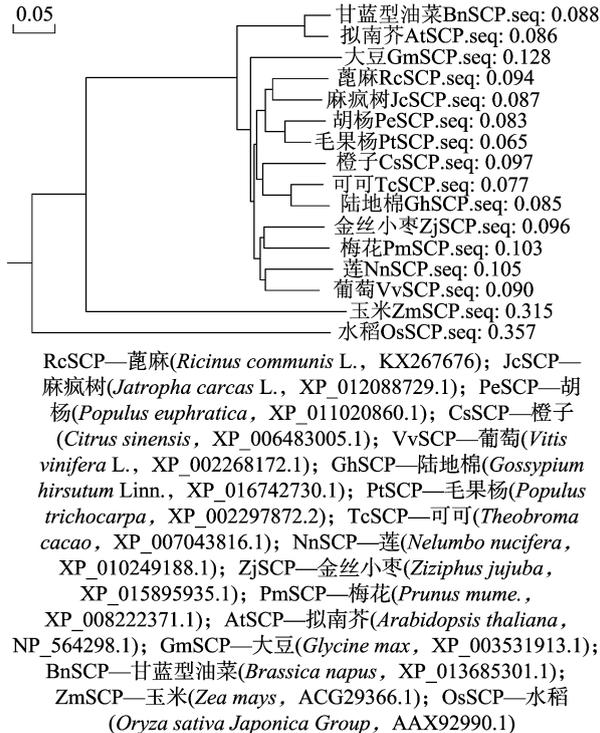


图7 SCP 的进化树分析

3 结论与讨论

经 RACE 法克隆得到 1 377 bp 的蓖麻丝氨酸羧肽酶蛋白基因 cDNA 完整的开放阅读框序列,推测它可编码 458 个氨基酸的多肽。蓖麻 RcSCP 蛋白含有丝氨酸羧肽酶蛋白家族保守区。对 RcSCP 蛋白亚细胞定位进行预测,并将其定位于细胞质中。蓖麻 RcSCP 蛋白与其他植物的丝氨酸羧肽酶蛋白具有较高同源性,而且与双子叶植物的丝氨酸羧肽酶蛋白亲缘关系较近。

丝氨酸羧肽酶是一类拥有数百个蛋白成员的水解酶类的蛋白家族,广泛存在于高等生物中^[1]。丝氨酸羧肽酶不仅在抵抗逆境方面有着极大的作用,也影响着植物的生长发育。Li 等从水稻中获得 1 个 GS5 基因,编码 1 个丝氨酸羧肽酶,发

现 GS5 基因通过调节水稻谷粒大小,进而影响水稻的产量^[12]。Bienert 等对转入 *NtSCP1*、*NtSCP2* 基因的烟草进行过表达研究发现,SCP 蛋白影响着植物细胞的伸长^[13]。随着丝氨酸羧肽酶蛋白研究的深入,丝氨酸羧肽酶已在多种植物中被克隆和表达分析,这有利于在分子水平鉴定丝氨酸羧肽酶蛋白的结构和功能特性。将来进一步对丝氨酸羧肽酶基因在植物体内作用机制的研究,有望在提高作物耐性基因工程方面取得重要突破,为生产和应用作出贡献。利用分子生物学方法研究该基因,并转化到蓖麻植株内,以期培育出抗逆效果优良的蓖麻新品种。

参考文献:

- [1] 王育华, 邹杰, 陈信波. 植物丝氨酸羧肽酶及其类蛋白的研究进展[J]. 生物学杂志, 2010, 27(6): 72-75, 102.
- [2] Liu H Z, Wang X E, Zhang H J, et al. A rice serine carboxypeptidase-like gene *OsBISCP1* is involved in regulation of defense responses against biotic and oxidative stress[J]. Gene, 2008, 420(1): 57-65.
- [3] Shirley A M, Chapple C. Biochemical characterization of sinapoylglucose: choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism[J]. The Journal of Biological Chemistry, 2003, 278(22): 19870-19877.
- [4] Cercos M, Urbez C, Carbonc J. A serine carboxypeptidase gene (*PsCP*), expressed in early steps of reproductive and vegetative development in *Pisum sativum*, is induced by gibberellins[J]. Plant Mol Biol, 2003, 51(2): 165-174.
- [5] 叶玲飞. 水稻 *OsSCP* 基因的耐逆相关功能研究[D]. 长沙: 湖南农业大学, 2013.
- [6] 刘丽, 王静, 张志明, 等. 玉米丝氨酸羧肽酶基因(*ZmSCP*)的克隆及表达分析[J]. 作物学报, 2013, 39(1): 164-171.
- [7] 郑鹭, 祁建民, 陈绍军, 等. 蓖麻遗传育种进展及其在生物能源与医药综合利用潜能[J]. 中国农学通报, 2006, 22(9): 109-113.
- [8] 程佑发. 蓖麻基因组测序初步完成[J]. 农业生物技术学报, 2010, 18(5): 866.
- [9] Anjani K. Castor genetic resources: a primary gene pool for exploitation[J]. Industrial Crops and Products, 2012, 35(1): 1-14.
- [10] 张继星, 刘鹏, 黄凤兰, 等. 蓖麻总 RNA 提取[J]. 内蒙古民族大学学报(自然科学版), 2010, 25(1): 38-39.
- [11] 狄建军, 张庆波, 孙颖飞, 等. 蓖麻种子总 RNA 提取方法比较[J]. 江苏农业科学, 2013, 41(5): 33-34.
- [12] Li Y B, Fan C C, Xing Y Z, et al. Natural variation in GS5 plays an important role in regulating grain size and yield in rice[J]. Nat Genet, 2011, 43: 1266-1269.
- [13] Bienert M D, Delannoy M, Navarre C, et al. *NtSCP1* from tobacco is an extracellular serine carboxypeptidase III that has an impact on cell elongation[J]. Plant Physiol, 2012, 158(3): 1220-1229.