

## 西伯利亚蓼几丁质酶基因 Class IV 的克隆及生物信息学分析

刘关君<sup>1,\*</sup>, 刘昌财<sup>1,2</sup>, 刘明坤<sup>1</sup>, 魏志刚<sup>1</sup>, 刘桂丰<sup>1</sup>

<sup>1</sup>林木遗传育种与生物技术教育部重点实验室, 东北林业大学, 哈尔滨 150040; <sup>2</sup>中国人民解放军 61699 部队, 湖北枝江 443200

**摘要:** 根据西伯利亚蓼抑制消减文库(SSH)中获得的几丁质酶(CHI)基因的部分序列, 采用RACE技术克隆了具完整编码区的cDNA序列, 基因全长1 017 bp, 开放阅读框编码270个氨基酸。序列分析表明, 该基因的编码蛋白(PsCHI1)以前体形式存在, N端分别有22个氨基酸的信号肽和35个氨基酸的几丁质结合域(CBD), C端199个氨基酸为催化区(CD), 连接CBD与CD的14个氨基酸为可变交联区, 成熟蛋白为不含信号肽部分, 呈碱性, 带正电荷。PsCHI1与所选其它植物class IV CHI前体序列具有高度的同源性(53%~69%), 而与class I和class II CHI的氨基酸序列同源性较低, 推测为植物class IV CHI。根据日本水稻CHI晶体结构构建了PsCHI1三维分子模型, 分析显示PsCHI1可以识别比class I和class II CHI短的几丁质片段, 并以其它植物CHI的已知结构域和功能为基础, 确定PsCHI1具有能够水解真菌细胞壁的结构, 推测其可能有抗病原微生物的功能。

**关键词:** 西伯利亚蓼; 几丁质酶(CHI); RACE; GPI锚

## Cloning and Bioinformatics Analysis of Class IV Chitinase Gene from *Polygonum sibiricum* Lxm.

LIU Guan-Jun<sup>1,\*</sup>, LIU Chang-Cai<sup>1,2</sup>, LIU Ming-Kun<sup>1</sup>, WEI Zhi-Gang<sup>1</sup>, LIU Gui-Feng<sup>1</sup>

<sup>1</sup>Key Laboratory of Forest Tree Genetic Improvement and Biotechnology, Ministry of Education, Northeast Forestry University, Harbin 150040, China; <sup>2</sup>The NO. 61699 Army of Chinese People's Liberation Army, Zhijiang, Hubei 443200, China

**Abstract:** A partial sequence of a Chitinase gene (CHI) was obtained from a random clone in *Polygonum sibiricum* SSH library and the CHI was isolated from *P. sibiricum* leaves by using rapid amplification of cDNA end (RACE) technology. The acquired gene was 1 017 bp in full length, including an open reading frame (ORF) of 810 bp which encoded 270 amino acids. Sequence analysis suggested that the protein (PsCHI1) existed as a precursor which comprised an amino-terminal (N-terminal) signal peptide of 22 amino acids, a cysteine-rich chitin-binding domain (CBD) of 35 amino acids, a variable hinge domain (VHD) of 14 amino acids, and a catalytic domain of 199 amino acids. The mature protein without signal peptide was positive charged and showed slightly highly basic. The deduced amino acid sequence showed 53%–69% homology with class IV CHI, and low homology with class I and class II CHI from selected plants. Therefore, it was identified as a typical plant class IV CHI. A homology model based on the structure of a class I chitinase from rice (*Oryza sativa* L. Japonica) was constructed, the analysis suggested that PsCHI1 enzyme recognized an even shorter segment of the substrate than class I or II enzymes. On the basis of known structure and function of plant CHI, we explained that PsCHI1 could attack against pathogen cell walls. As result, PsCHI1 was possibly involved in antimicrobial activity.

**Key words:** *Polygonum sibiricum*; CHI; RACE; GPI anchor

几丁质是昆虫外骨骼、甲壳类外壳以及真菌细胞壁的主要成分, 高等植物本身不含几丁质。植物几丁质酶(chitinase, CHI, EC3.2.1.14)是一种酸性或碱性蛋白, 分子量25~35 kDa, 通过水解病原微生物细胞壁几丁质中 $\beta$ (1-4)糖苷键产生N-乙酰氨基葡萄糖寡聚体或单体而起到抑菌作用(Collinge等1993)。大量证据显示, 纯化的CHI, 尤其是与 $\beta$ -1,3-葡聚糖酶组合在体外能够

降解真菌细胞壁且抑制真菌的生长, 在病原菌(细菌、真菌和病毒)侵染植物的过程中, 不仅CHI普遍被过量表达(Nielsen等1993), 而且其酶活性也明显增强(Grison等1996); 一些非生物胁迫如高

收稿 2008-05-04 修定 2008-05-06

资助 黑龙江省重点攻关项目(GB06B303)。

\* E-mail: liuguanjun2003@126.com; Tel: 0451-82190607-13

盐、干旱、伤害、植物激素、重金属和臭氧等也促进植物 CHI 基因的大量表达(Graham 和 Sticklen 1994)。转基因实验表明,组成性表达 CHI 增强植物对体外侵染真菌(Pasonen 等 2004)和细菌(Grison 等 1996)的抗性,从而说明 CHI 积极参与植物防御病原微生物的过程,并能够赋予植物一定的抗病性。近些年来,随着对 CHI 在植物防御生理与分子基础研究的不断深入,人们陆续从多种高等植物中克隆了 CHI,并分离纯化了不同类型的 CHI,根据其一级结构氨基酸序列的同源性和某些特征,将植物 CHI 分为 class I~class V (Meins 等 1994),且不同种类的 CHI 具有不同的催化活性、物理特征、细胞定位和作用方式。这些 CHI 对采用基因工程技术提高植物的抗病性具有一定的价值。迄今,应用 CHI 已获得了抗病力较强的杨树植株(许农等 1993)。

本文用 NaHCO<sub>3</sub> 胁迫下西伯利亚蓼地下茎消减文库中的 EST 序列(GenBank 登录号 ES599907)和 RACE 技术克隆了 *PsCHII* 基因 cDNA 序列,并作了生物信息学分析,推断了编码蛋白的结构、功能和可能的细胞定位,为进一步研究其基因功能以及与非生物胁迫关系建立了基础。

## 材料与amp;方法

西伯利亚蓼(*Polygonum sibiricum* Laxm.)耐盐种质材料 XBLYL-027 由本实验室提供,采自黑龙江省肇东县盐碱地。本文所用的大肠杆菌菌株为 *Escherichia coli* TOP10 (天根公司),克隆载体为 pGEM<sup>®</sup>-T (Promega 公司)。DNA 片段回收试剂盒、DL2000、dNTP、RNase A、Taq DNA 聚合酶、Taq Plus DNA 聚合酶(高保真)、IPTG、X-gal 购自天根生化科技有限公司;逆转录合成试剂盒 Revert Aid<sup>™</sup> First Strand cDNA Synthesis Kit 购自百泰克生物技术有限公司;BD Smart<sup>™</sup> RACE cDNA Amplification Kit 购自 BD Biosciences 公司;各种限制性内切酶、T4 DNA 连接酶购自 Promega 公司;EDTA、DEPC、MOPS 为赛百盛公司产品;SDS、Tris 为 Sigma 公司产品;其它总 RNA 提取试剂及常规药品均为进口或国产分析纯级。引物合成及测序工作由上海生工生物工程有限公司完成。

根据王玉成等(2003)文中的方法,取盐胁迫

处理后的西伯利亚蓼茎部组织提取总 RNA。非变性琼脂糖凝胶电泳参照《分子克隆实验指南》(2002)的方法进行,胶浓度为 0.8%,检测总 RNA 的完整性。用紫外分光光度计测定提取的 RNA 在波长 260 和 280 nm 处的光吸收值,确定 RNA 的纯度及浓度。

参照 SMART<sup>™</sup> RACE cDNA Amplification Kit,取 1 μg 的总 RNA 分别反转录合成 5' RACE 模板(5' RACE-Ready-cDNA, 5' rrcDNA)和 3' RACE 模板(3' RACE-Ready-cDNA, 3' rrcDNA)。

根据已知序列设计 3' RACE 特异性正向引物 CHI-S,序列为 5' CCTGTTTATTGGAACAAC-TATGG 3',反向引物为 UPM (见 RACE 说明书),以 3' rrcDNA 为模板,PCR 反应体系参照试剂盒 (Smart<sup>™</sup> RACE cDNA Amplification Kit BD Biosciences)说明书。反应条件为 94 ℃ 中变性 3 min,然后按 94 ℃ 30 s、60 ℃ 30 s、72 ℃ 1 min 进行 35 轮的循环反应,最后在 72 ℃ 下延伸 8 min。

将此 PCR 扩增产物的特异性片段进行胶回收,载体连接,转化,经 Amp<sup>+</sup> 抗性及蓝白斑筛选,随机挑取阳性克隆进行 PCR 鉴定,由上海生工生物工程技术服务有限公司完成测序。

根据上述方法将已获得的 3' RACE 扩增产物序列经 NCBI blastX 分析确认为 *PsCHII* 3' 到头序列,在其 3' 非翻译区分别设计引物 CHI-A1 (外侧引物)和 CHI-A2 (内侧引物)以获得具完整编码区的 cDNA,CHI-A1 引物序列为 5' TCAAGTAC-GATGATAATAATACATCCAC 3',CHI-A2 引物序列为 5' GAGCACAGAACACATCTTATTCAAGTACG 3',将 5' rrcDNA 作为模板,反应体系同前,CHI-A1 与 UPM 作为引物进行第 1 次 PCR 反应,反应条件为 94 ℃ 变性 3 min,然后按 94 ℃ 30 s、52 ℃ 30 s、72 ℃ 1 min 进行 35 轮的循环反应,最后于 72 ℃ 下延伸 8 min。取其 0.5 μL 产物作为模板,CHI-A2 与 NUP (见 RACE 说明书)作为引物进行第 2 次 PCR 反应,反应体系同前,反应条件为 94 ℃ 变性 3 min,然后按 94 ℃ 30 s、60 ℃ 30 s、72 ℃ 1 min 进行 35 轮的循环反应。

第 2 次 PCR 扩增产物的特异性片段进行胶回收,载体连接,转化,经 Amp<sup>+</sup> 抗性及蓝白斑筛选,随机挑取阳性克隆进行 PCR 鉴定,由上海生工生物工程技术服务有限公司完成测序。

BioEdit 软件的 CAP (contig assembly program) 进行电子拼接, ORF 程序 (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 查找序列的开放读码框 (open reading frame, ORF), 应用 BioEdit 程序选择 13 种 CHI 建立本地蛋白数据库 Local Blast 程序进行序列一致性比较, ClustalW Multiple alignment 进行多序列比较和生成系统进化树文件, Treeview 1.6.6 查看进化树文件, BioEdit 软件的 Kyte-Doolittle Scale Mean Hydrophobicity profile 程序算法进行疏水性分析, ProtParam (<http://au.expasy.org/tools/protparam.html>) 计算蛋白的分子量、净电荷和理论等电点, 选用 BlastP (<http://www.ncbi.nlm.nih.gov/BLAST/>) 比对并预测保守区, 应用 Profile search、Scan prosite 和 Motif Scan 程序 (<http://au.expasy.org/prosite>) 进行轮廓、结构域和基序分析。SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP>) 预测蛋白的信号肽及可能的切割位点, TargetP 1.1 (<http://www.cbs.dtu.dk/services/TargetP>) 进行亚细胞定位, TMHMM 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测蛋白内部跨膜螺旋, big-PI Plant Predictor ([http://mendel.imp.ac.at/sat/gpi/plant\\_server.html](http://mendel.imp.ac.at/sat/gpi/plant_server.html)) 分析 GPI 锚位点。

应用 Swiss-Model (<http://www.expasy.org/swissmod/SWISS-MODEL.html>) 和 MolMol 软件产生推断的 PsCHI1 三维结构模型。

## 实验结果

### 1 总 RNA 提取质量鉴定

总 RNA 的质量是关系反转录 cDNA 完整性的因素。提取的西伯利亚蓼茎部总 RNA 用无 RNase 的 Dnase I 处理, 除去污染的基因组 DNA, 用 1% 的非变性琼脂糖凝胶电泳检测, RNA 提取结果见图 1。可以看到 28S rRNA 与 18S rRNA 条带清晰, 亮度比例达到 2:1 以上, 说明 RNA 的完整性较好。将所提取的 RNA ( $OD_{260}/OD_{280}$  介于 1.8~2.0) 用于反转录合成 RACE 模板。

### 2 *PsCHII* 基因 3' RACE 片段和具完整编码区 cDNA 的获得

根据 SSH 文库中获得的 129 bp *PsCHII* 片段序列, 设计引物 CHI-S, 采用 3' RACE 的方法扩增 *PsCHII* 的 3' 端。经扩增得到 1 条约 500 bp 的 DNA 片段 (图 2), 测序后比对证实为 *PsCHII* 的 3'

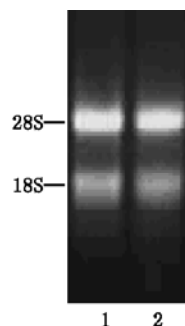


图 1 从西伯利亚蓼中提取的总 RNA  
Fig.1 Total RNA extracted from *P. sibiricum*  
1 和 2 为 2 次提取的总 RNA。

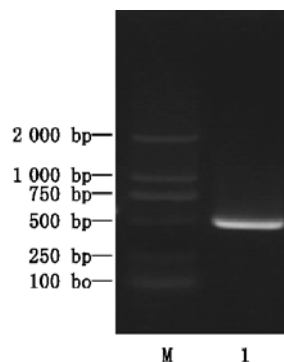


图 2 *PsCHII* 的 3' RACE 扩增产物  
Fig.2 The 3' RACE product of *PsCHII* cDNA  
M: DL2000 分子量对照; 1: 3' RACE 扩增的产物。

端到头序列。

根据 3' RACE 产物序列在其非翻译区设计引物 CHI-A1 和 CHI-A2, 采用 5' RACE 的方法直接扩增完整的 *PsCHII* 全长。经巢式 (Nested) PCR 扩增得到 1 条约 950 bp 的 DNA 片段 (图 3), 与所推测的片段大小一致, 测序后经比对, 证实为具完整 ORF 的 *PsCHII* 序列。

### 3 *PsCHII* 的序列分析

将 3' RACE 和 5' RACE 的产物序列经电子拼接获得带有 PolyA 的全长 cDNA 序列 1 017 bp, ORF 程序显示 5' 非翻译区 15 bp, 3' 非翻译区 192 bp, ORF 810 bp, 编码 270 个氨基酸 (图 4)。

经疏水性、SignalP 3.0、ProtParam 和 ScanProfile 程序分析此推断的氨基酸序列发现, *PsCHII* 所编码的蛋白 (PsCHI1) 以其前体形式存在, 由信号肽、几丁质结合域 (chitin-binding domain, CBD)、可变交联区 (variable hinge domain, VHD) 和催化区 (catalytic domain, CD) 四

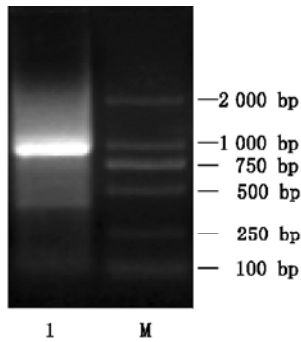


图3 *PsCH11* 的 5' RACE 扩增产物

Fig.3 The 5' RACE product of *PsCH11* cDNA

M : DL2000 分子量对照 ; 1 : 5' RACE 扩增的产物。

部分构成, 将成熟蛋白定位于液泡的 C 端延伸区, 前体等电点(IEP)为 8.5, 分子量(MW)为 28.7 kDa ; N 端 1~22 氨基酸区段具有典型的信号肽特征: 带正电荷, 疏水性很强。第 22 (Ala) 和第 23 (Gln) 氨基酸处为最有可能的信号肽切割位点, 推定为信号肽部分。

成熟蛋白缺少信号肽部分, 其第一个氨基酸为 Gln, 这与已知的 class IV 成熟蛋白第一个氨基酸一致。成熟蛋白呈碱性, 理论 IEP 为 8.3, 理论 MW 为 26.4 kDa。Motif Scan 程序分析显示成熟 *PSCH11* 的 Gln23~Ser57 氨基酸区段为 CBD 部

1	CAC AAT CAA TCA ACC <span style="border: 1px solid black; padding: 0 2px;">ATG</span> GCT CTC AAC AAA AAC CTC CTC CCT CTC GCT CTC CTC CTC GTC	60
1	<u>Met Ala Leu Asn Lys Asn Leu Leu Pro Leu Ala Leu Leu Leu Val</u>	15
61	CTA GCC ACC TCC GCC GTG GCC CAA AAC TGC GGG TGC ACC GGC AAC AAC TGC TGC AGC AAA	120
16	<u>Leu Ala Thr Ser Ala Val Ala Gln Asn Cys Gly Cys Thr Gly Asn Asn Cys Cys Ser Lys</u>	35
121	TTC GGA TAC TGT GGC ACC GGT GGC GAC TAC TGT GGC GCA GGC TGC CAA CAA GGC CCT TGC	180
36	<u>Phe Gly Tyr Cys Gly Thr Gly Gly Asp Tyr Cys Gly Ala Gly Cys Gln Gln Gly Pro Cys</u>	55
181	TAC TCT AAT GGC GGT AGT AGT AGC GAC TCC GGC GGG TCC GTG GCG TCG GTC GTG ACC GAT	240
56	<u>Tyr Ser Asn Gly Gly Ser Ser Ser Asp Ser Gly Gly Ser Val Ala Ser</u> Val Val Thr Asp	75
241	TCT TTC TTC AAT GCA ATC ATT AAC CAG GCT GCC TCT AGC TGC GCC GGA AAG AGC TTC TAC	300
76	Ser Phe Phe Asn Ala Ile Ile Asn Gln Ala Ala Ser Ser Cys Ala Gly Lys Ser Phe Tyr	95
301	ACT CGA TCG GCG TTC CTA AAC GCC GCC AGA TGG TAC CCG AGT TTT GGG ACC ACC GGC GCC	360
96	Thr Arg Ser Ala Phe Leu Asn Ala Ala Arg Trp Tyr Pro Ser Phe Gly Thr Thr Gly Ala	115
361	TCC GAT GTA CAA AAG AGG GAG ATC GCC GCC TTC TTT GCT CAT GTT ACC CAT GAG TCC GGA	420
116	Ser Asp Val Gln Lys Arg Glu Ile Ala Ala Phe Phe Ala His Val Thr His Glu Ser Gly	135
421	TAC TTT TGC TAC AAA GAG GAG ATA AAC GGA GCA TCA CAC AAC TAC TGC <span style="background-color: #cccccc;">CAA GCA ACC ACA</span>	480
136	Tyr Phe Cys Tyr Lys Glu Glu Ile Asn Gly Ala Ser His Asn Tyr Cys Gln Ala Ser Thr	155
481	<span style="background-color: #cccccc;">CAA TGG CCA TGC AAC CCA AAC AAG AAG TAC TTC GGG CGC GGT CCC CTC CAG ATC TCA TGG</span>	540
156	Gln Trp Pro Cys Asn Pro Asn Lys Lys Tyr Phe Gly Arg Gly Pro Leu Gln Ile Ser Trp	175
541	<span style="background-color: #cccccc;">AAC TAC AAC TAC GGG CCG GCA GGC AGA GAC ATC GGG TTC GAC GGG CTC AAC TCC CCC GAG</span>	600
176	Asn Tyr Asn Tyr Gly Pro Ala Gly Arg Asp Ile Gly Phe Asp Gly Leu Asn Ser Pro Glu	195
601	ACA GTA GCC AAC GAC CCC GTC GTC GCC TTC AAG ACC GCC TTC TGG TTC TGG AAG AAC AAC	660
196	Thr Val Ala Asn Asp Pro Val Val Ala Phe Lys Thr Ala Phe Trp Phe Trp Lys Asn Asn	215
661	GTC CAC TCC ATC ATC ACC GGT GGC CGG GGA TTC GGG GCC ACC ATC CGA GCC ATC AAC AGT	720
216	Val His Ser Ile Ile Thr Gly Gly Arg Gly Phe Gly Ala Thr Ile Arg Ala Ile Asn Ser	235
721	ATT GAA TGC AAT GGC GGG AAC ACT CTA GCT GTG AAT GCA AGA GTC AGT CTC TAC CAG AAG	780
236	Ile Glu Cys Asn Gly Gly Asn Thr Leu Ala Val Asn Ala Arg Val Ser Leu Tyr Gln Lys	255
781	TAT TGT CAA CAG CTT GGT GTT AAC CCT GGT GCC AAC CTC TCT TGC TAA TTT ATA TAC CAC	840
256	Tyr Cys Gln Gln Leu Gly Val Asn Pro Gly Ala Asn Leu Ser Cys End	270
841	AAT TGC AAG CTG AAA AAA AAT ATT AAA ATA AGC TGA AGA GAA AGT CCT TGA AGA TCT TCT	900
901	TTT TCA CCA CTG AAA TTA CAG TAC TAA TTA <b>GTG GAA TGT ATT ATT ATC ATC GTA CTT GAA</b>	960
961	<b>ATC TCT AGA GGA TCC CCG</b> GGT ACC GAG CTC GAA TTC GTA ATC ATG TCA TTG TTT CCC	1017

图4 *PsCH11* 具完整编码区 cDNA 序列和推定的氨基酸序列

Fig.4 Nucleotide and deduced amino acid sequences of *PsCH11* cDNA from *P. sibiricum*

下划线部分 22 个氨基酸为信号肽; 虚线部分氨基酸为 CBD; 双下划线部分氨基酸为 VHD; 其它部分为 CD。阴影部分核苷酸为从 SSH 库中获得长度为 129 bp 的 *PsCH11* 部分序列; 阴影加粗部分核苷酸(529~554 bp)为引物 CHI-S; 其后面核苷酸部分为 3' RACE 获得的片段序列。加粗核苷酸部分 950~978 bp 和 931~959 bp 分别为引物 CHI-A1 和 CHI-A2。方框表示起始密码子; End 表示终止子。

分, IEP=5.8, MW=3.5 kDa, 含有高度保守的氨基酸和较高的极性氨基酸比率, 具显著的植物 CBD (编号 PS00026) “C-x(4,5)-C-C-S-x(2)-G-x-C-G-x (3,4)-[FYW]-C” 序列特征(图 5), 8 个位置基本固定的半胱氨酸(Cys)是 PsCHI1 极为保守的 4 个二硫键的形成位点。另外, 一些较为保守的亲水性残基涉及与底物的结合, 如芳香族氨基酸 - 这里为酪氨酸(Y) (图 5)。Asn58~Ser71 区段为 VHD 部分, 负责连接 CBD 与 CD, 其氨基酸组成单一, 富含 Ser 和 Gly, 无 Pro 残基, 呈酸性, IEP=3.8, MW=1.2 kDa。C 端 Val72~Cys270 为 CD 部分, 呈碱性, 带正电荷, IEP=8.8, MW=21.8 kDa, Cys89~ Gly111 区段具 CHI 家族 19\_1 标签(chitinases family 19\_1 signature, PS00773) “C-X(4,5)-F-Y-[ST]-X(3)-[FY]-[LIVMF]-X-A-X(3)-[YF]-X(2)-F-[GSA]” 序列特征, Val203~Lys213 区段与 CHI 家族 19\_2 标签 (PS00774) “[LIVM]-[GSA]-F-X-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM]” 序列基本一致, 但在最后的一个氨基酸位置上有差异, 前者为 K, 而后者应是 L I V M 之一。Thr229~Arg231 和 Asn267~Cys270 区段分别为蛋白激酶 C 磷酸化糖基化位点(PKC phosphorylation site) 和 N-糖基化位点(ASN glycosylation site)。

TargetP 1.1、TMHMM 和 big-PI Plant Predictor 预测表明, PsCHI1 被分泌到细胞外, 无明显的跨膜区, 在 C 端序列 Arg249 位置具一个潜在的糖基磷脂酰肌醇(glycosylphosphatidylinositol, GPI) 修饰位点。根据日本水稻(*Oryza Sativa* L. Japonica) CHI (pdb2DKV) 晶体结构作为 CHI 的 CBD 和 CD 的同源模型模板构建了 PsCHI1 的晶体结构模型(图 8)。另外, 又根据 Garcia-Casado 等(1998) 对核果种子 CHI 保守位点突变的研究, 推测出与

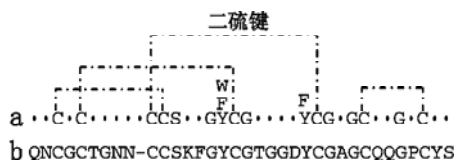


图 5 PSCHI1 的 CBD 与其它植物 class IV CBD 保守的氨基酸序列比较

Fig.5 Comparison on the conserved amino acid sequences of class IV CHI CBD between *P. sibiricum* and other plants

a : 植物 class IV CBD 保守的氨基酸序列; b : PSCHI1 的 CBD 氨基酸序列。

PsCHI1 相对应的涉及底物结合及催化活性氨基酸残基位点(图 6)。

#### 4 PsCHI1 与其它类型 CHI 的同源性比较

选择 8 个 class IV (红藜、甜菜、薯蓣、日本薯蓣、水稻、葡萄、拟南芥、胡萝卜)、2 个 class I (湖南烤烟、日本水稻)、2 个 class II (大麦、刀豆) 和 1 个 Class III (热紫链霉菌) 共 13 个 CHI 建立本地 CHI 蛋白数据库, 应用 Bioeditor 软件将 *PsCHI1* 所编码的蛋白序列与所建立的蛋白数据库进行多序列比对(图 6), 经 BLASTP 分析, 发现分子克隆到的 *PsCHI1* 所编码的蛋白与 class IV CHI 的氨基酸序列具有较高的同源性(53%~69%)。PsCHI1 与其它类型的 CHI 氨基酸序列有较低的同源性, 分别为湖南烤烟 42%、日本水稻 40%, 栽培二稜大麦 44%、刀豆 38%; 而与 class III (热紫链霉菌) 的同源性为 38%。说明 PsCHI1 与 class IV 有较近的亲缘关系, 同时发现 PsCHI1 和其它 class IV CHI 在 CD 处都有大量的缺失(图 6), 因此显示出具有明显的 class IV CHI 特征。另外, 从一种藜属植物中分离得到 4 个 CHI (Nakamura 等 1997), 其编码的 class IV CHI 与本文的 PsCHI1 序列同源性为 66%~69%。因此从氨基酸序列同源性上来看, PsCHI1 应属于 class IV CHI。

选择以上蛋白绘制出系统进化树(图 7)。从图 7 可见, PsCHI1 与其它物种 class IV CHI 在进化关系上较近, 特别是与红藜和甜菜亲缘关系最近; 而与其它类型的 CHI 亲缘关系较远。序列比对结果以及进化树分析都表明 *PsCHI1* 应为编码西伯利亚蓼几丁质酶蛋白一类的基因, 且属于 class IV CHI。

## 讨 论

本文采用 RACE 的方法从西伯利亚蓼中克隆了几丁质酶基因 *PsCHI1*, 通过同源性比对发现 *PsCHI1* 编码蛋白与植物中已经克隆的 class IV CHI 具有非常高的同源性, 而与其它类型的 CHI 同源性较低, 具有明显的 class IV CHI 特征; 进化树分析也表明 *PsCHI1* 编码蛋白与植物 class IV CHI 的同源关系较近, 处在同一分支上。推定 *PsCHI1* 编码蛋白属 class IV CHI。

class IV CHI 受真菌侵染诱导, 与植物的抗性相关(Lange 等 1996; Nielsen 等 1994; Rasmussen

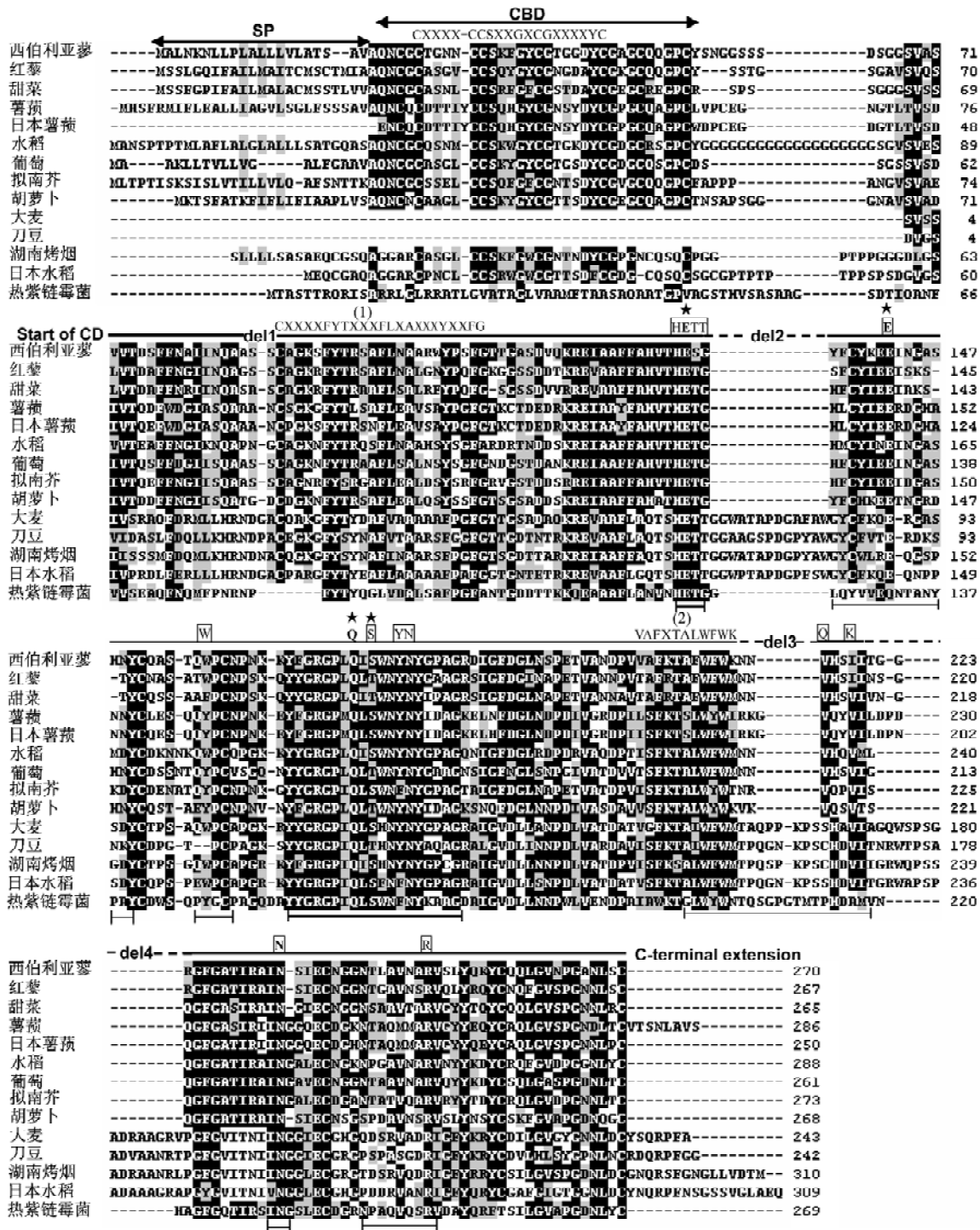


图6 PsCH11与其他13种CHI氨基酸序列的比较

Fig.6 Comparison on the amino acid sequences of PsCH11 between other plants

根据 class I/II CHI 的结构在 PsCH11 序列上标注了 4 个删除(del 1~4); SP 表示信号肽; CBD 表示几丁质结合域(PS0026); Start of CD 表示 CD 的开始; 黑色的横线标记为 CD; C-terminal extension 表示 C 端延伸区。(1)为 CHI 家族 19\_1 标签(PS00773); (2)为 CHI 家族 19\_2 标签(PS00774)。序列上方字母加粗的残基是催化活性所必需的氨基酸; 标记有星号的是对于催化活性重要的氨基酸; 带方框的氨基酸残基是推定的与底物结合位点; 序列下面有线条标记的是活性位点(Garcia-Casado 等 1998)。

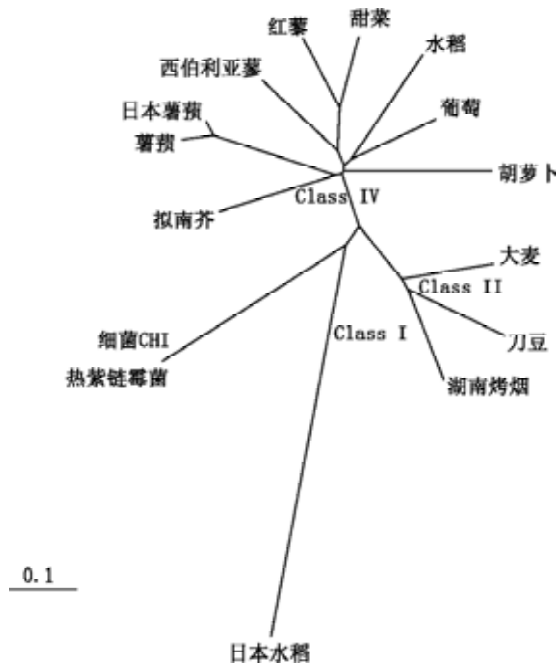


图7 PsCHI1 与其它 13 种 CHI 的进化树

Fig.7 Phylogenetic tree between of PsCHI1 and other CHI

等 1992), 在植物的防御反应中起作用(de Gerhardt 等 1997)。其赋予植物较强的抗病功能与易于亲合和水解几丁质的结构密不可分。CBD 是结合几丁质所必需的(Iseli 等 1993), 在不同 CHI 的 CBD 内部, Tyr、Trp、Phe 残基是高度保守的, 因而认为是直接与几丁质的 *N*-乙酰氨基葡萄糖的吡喃糖环结合位点(Lee 等 2007), 据此推断在 PsCHI1 的 CBD 内部 Tyr38 和 Tyr45 是结合几丁质的必需氨基酸; CD 一般由 300 个氨基酸组成, 负责催化功能, 其活性中心的质子受体一般为谷氨酸残基, 根据烟草(Iseli-Gamboni 等 1998)、大麦(Andersen 等 1997)、核果种子、芸苔(Tang 等 2004) CHI CD 内部两个极为保守的 Glu 残基(与 PsCHI1 等位的是 Glu133、Glu142)突变中的结果, 推测它是两个催化所必需氨基酸。据此我们也推测 PsCHI1 是利用 CD 中保守的 Glu133 和 Glu142 分别作为质子供体和亲核试剂直接参与几丁质的催化水解反应, 从而起到抵抗病原微生物的作用。因此, 认为 PsCHI1 在结构上具备结合和水解真菌细胞壁主要成分几丁质的活性位点。

为了构建和推测 PsCHI1 的三维结构模型, 必须寻找与目的蛋白同源性较高且已测定的 CHI 晶体结构。但目前尚无与其同源性较高的 class IV

CHI 晶体结构, 可以获得的只有来自于大麦(Hart 等 1995; Song 和 Suh 1996)和刀豆(Hahn 等 2000) class II CHI (只有 CD) 的 X 射线结构。虽然它们与 PsCHI1 的 CD 氨基酸序列有一定的同源性(分别为 45% 和 38%), 但它们没有 CBD 结构, 只能作为 CD 三维结构模型。由于水稻 class I CHI (包括 CBD 和 CD) X 射线结构刚刚得到测定, 且这两个蛋白的 CBD 和 CD 氨基酸序列具有较高的同源性(分别为 56% 和 40%), 因此, 用其晶体结构作为 CHI 的 CBD 和 CD 的同源模型模板构建 PsCHI1 晶体结构模型(图 8)表明, PsCHI1 三维结构 CD 中两个保守的 Glu 催化残基之间的距离与水稻 class I CHI 相比较短, 之间有一个更窄的裂缝, 且因为与其它 class IV CHI 一样, PsCHI1 属于内切几丁质酶, 可从几丁质的内部将其切断, 根据这一结构分析 PsCHI1 可能易识别较短的几丁质链, 也就是说在病原微生物侵袭植物时 PsCHI1 能够极易抓住暴露的几丁质的一小部分而进行水解作用, 它很可能与 class IV CHI 充分有效水解真菌细胞壁直接相关(Karasuda 等 2003)。

虽然 PsCHI1 具备水解真菌细胞壁的有力结构, 在植物的防御反应中起作用, 今后仍需考虑的是, PsCHI1 片段来自于盐碱胁迫下西伯利亚蓼地下茎的 SSH 文库, 其过量表达表明它可能参与植物抗盐碱的过程, 但其作用机制还不清楚, 有待进一步研究。

另外, 为了进一步确证 PsCHI1 在细胞内或外的定位, 在此序列中获得可能的 C 端糖基磷脂酰肌醇(GPI)锚定信息, 这对今后设计突变实验是重要的, 随后此可置换某一氨基酸对其产生的影响。GPI 作为一种普通的翻译后修饰, 由于其自身能够决定蛋白质的细胞内定位并且限定蛋白的功能范围, 所以, 了解一个蛋白的 GPI 脂质修饰是有意义的(Kinoshita 和 Inoue 2000)。本文采用 big-PI Plant Predictor 推测 PsCHI1 前体蛋白 GPI 修饰位点, 发现 C 端序列 Arg249 位置具一个潜在的 GPI 修饰位点, 因此推测 PsCHI1 的分泌途径为: PsCHI1 在粗面内质网核糖体合成, 同时进入内质网腔, 而后, GPI 分子的一部分与 PsCHI1 多肽 C 末端接合后将从前体蛋白切去 C 末端前肽, 最后通过分泌泡转移并定位到细胞外侧质膜, 因此成熟 PsCHI1 不但是其前体蛋白去除信号肽, 而且

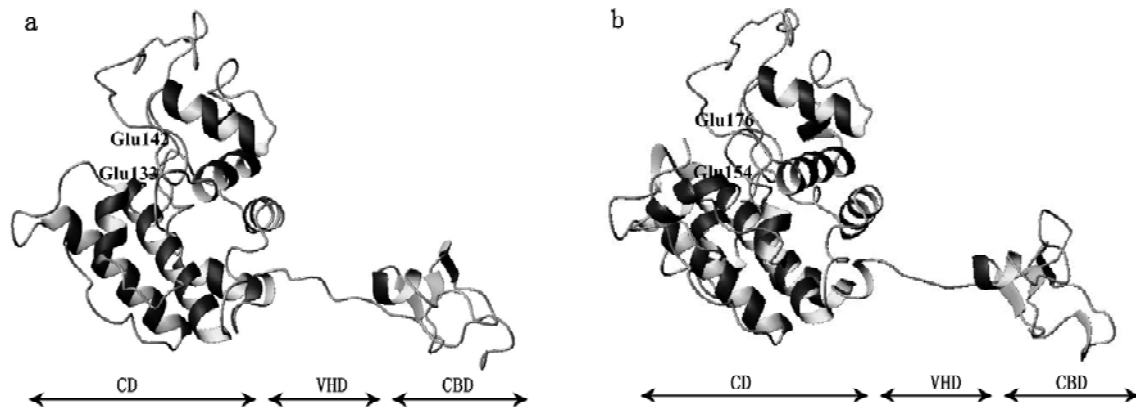


图8 CHI的同源结构模型

Fig.8 Homology modeling of CHI

CD 为催化区; VHD 为可变交联区; CBD 为几丁质结合域。a: 基于日本水稻 CHI (pdb 2DKV) 晶体结构构建的 PsCHI1 三维分子同源结构模型, 其中 Glu133 和 Glu142 为推定 PsCHI1 催化活性所必需的氨基酸; b: 日本水稻 CHI (pdb 2DKV) 晶体结构模型, 其中 Glu154 和 Glu176 为日本水稻 CHI 催化活性所必需的氨基酸。

也可能会失去一部分 C 末端前肽, 这一结果对今后进一步研究 PsCHI1 CD 特别是 C 末端保守氨基酸功能来说是有一定意义的。不久以前, Xu 等 (2007) 曾应用 Big-Pi Predictor 对水稻和拟南芥中所有已知的不同类型的 CHI 中潜在的 GPI 位点进行查找, 但都没有查到此结构。

### 参考文献

- 王玉成, 薄海侠, 杨传平 (2003). 胡杨、柽柳总 RNA 提取方法的建立. 东北林业大学学报, 31 (5): 99~100
- 许农, 黄敏仁, 陈道明 (1993). 杨树基因工程研究进展. 南京林业大学学报, 17 (4): 78~83
- Andersen MD, Jensen A, Robertus JD, Leah R, Skriver K (1997). Heterologous expression and characterization of wild-type and mutant forms of a 26 kDa endochitinase from barley (*Hordeum vulgare* L.). *Biochem J*, 322 (3): 815~822
- Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K (1993). Plant chitinase. *Plant J*, 3: 31~40
- de Gerhardt LBA, Sabetto-Martins G, Contarini MG, Sandroni M, de Ferreira RP, de Lima VM, Cordeiro MC, de Oliveira DE, Margis-Pinheiro M (1997). *Arabidopsis thaliana* class IV chitinase is early induced during the interaction with *Xanthomonas campestris*. *FEBS Lett*, 419: 69~75
- Garcia-Casado G, Collada C, Allona I, Casado R, Pacios LF, Aragoncillo C, Gomez L (1998). Site-directed mutagenesis of active site residues in a class I endochitinase from chestnut seeds. *Glycobiology*, 8: 1021~1028
- Graham LS, Sticklen MB (1994). Plant chitinases. *Can J Bot*, 72: 1057~1083
- Grison R, Grezes-Besset B, Schneider M, Lucante N, Olsen L, Leguay JJ, Toppan A (1996). Field tolerance to fungal pathogens of *Brassica napus* constitutively expressing a chimeric chitinase gene. *Nature Biotechnol*, 14: 643~646
- Hahn M, Hennig M, Schlesier B, Hohne W (2000). Structure of jack bean chitinase. *Acta Cryst*, D56: 1096~1099
- Hart PJ, Pfluger HD, Monzingo AF, Hollis T, Robertus JD (1995). The refined crystal structure of an endochitinase from *Hordeum vulgare* L. seeds at 1.8 Å resolution. *J Mol Biol*, 248: 402~413
- Iseli B, Boller T, Neuhaus TM (1993). The N-terminal cysteine-rich domain of tobacco class I chitinase is essential for chitin binding but not for catalytic or antifungal activity. *Plant Physiol*, 103: 221~226
- Iseli-Gamboni B, Boller T, Neuhaus JM (1998). Mutation of either of two essential glutamates converts the catalytic domain of tobacco class I chitinase into a chitin-binding lectin. *Plant Sci*, 134: 45~51
- Karasuda S, Tanaka S, Kajihara H, Yamamoto Y, Koga D (2003). Plant chitinase as a possible biocontrol agent for use instead of chemical fungicides. *Biosci Biotechnol Biochem*, 67: 221~224
- Kinoshita T, Inoue N (2000). Dissecting and manipulating the pathway for glycosylphosphatidylinositol-anchor biosynthesis. *Curr Opin Chem Biol*, 4: 632~638
- Lange J, Mohr U, Wiemken A, Boller T and Vogeli-Lange R (1996). Proteolytic processing of class IV chitinase in the compatible interaction of bean roots with *Fusarium solani*. *Plant Physiol*, 111: 1135~1144
- Lee YS, Park IH, Yoo JS, Chung SY, Lee YC, Cho YS, Ahn SC, Kim CM, Choi YL (2007). Cloning, purification, and characterization of chitosanase from *Bacillus* sp. DAU101. *Bioresour Tech*, 98: 2734~2741



- Meins F, Fritig B, Linthorst HJM, Mikkelsen JD, Neuhaus JM and Ryals J (1994). Plant chitinase genes. *Plant Mol Bio Rep*, 12: S22~S28
- Nakamura S, IWAI T, Honkura R, Ugaki M, Ohshima M, Ohashi Y (1997). Four chitinase cDNAs from *Chenopodium amaranticolor*. *Plant Biotech*, 14 (1): 85~86
- Nielsen KK, Bojsen K, Roepstorff P, Mikkelsen JD (1994). A hydroxyproline-containing class IV chitinase of sugar beet is glycosylated with xylose. *Plant Mol Biol*, 25: 241~257
- Nielsen KK, Mikkelsen JD, Kragh KM, Bojsen K (1993). An acidic class III chitinase in sugar beet: Induction by *Cercospora beticola*, characterization, and expression in transgenic tobacco plants. *Mol Plant Microb Interact*, 6: 495~506
- Pasonen HL, Seppanen SK, Degefu Y, Rytkonen A, von Weissenberg K, Pappinen A (2004). Field performance of chitinase transgenic silver birches (*Betula pendula*): resistance to fungal diseases. *Theor Appl Genet*, 109: 562~570
- Rasmussen U, Bojsen K, Collinge DB (1992). Cloning and characterization of a pathogen-induced chitinase in *Brassica napus*. *Plant Mol Biol*, 20: 277~287
- Song HK, Suh SW (1996). Refined structure of the chitinase from barley seeds at 2.0 and angstrom resolution. *Acta Crystallogr*, D52: 289~298
- Tang CM, Chye ML, Ramalingam S, Ouyang SW, Zhao KJ, Ubhayasekera W, Mowbray SL (2004). Functional analyses of the chitin-binding domains and the catalytic domain of *Brassica juncea* chitinase BjCHI1. *Plant Mol Biol*, 56: 285~298
- Xu FH, Fan CM, He YQ (2007). Chitinases in *Oryza sativa* ssp. *Japonica* and *Arabidopsis thaliana*. *Acta Genet Sin*, 34 (2): 138~150