

花粉发育相关基因 *BjMF6* 的分子克隆及其生物信息学分析

张豫超 叶意群 黄 鹂 曹家树*

(浙江大学蔬菜研究所细胞与分子生物学实验室, 杭州 310029)

摘要 根据白菜花粉特异的多聚半乳糖醛酸酶(PG)基因 *BcMF6* 的全长序列设计引物, 通过PCR直接扩增的方法从茎瘤芥(*Brassica juncea* Czern. et Coss var. *tumida* Tsen et Lee)‘浙桐1号’和分蘖芥(*Brassica juncea* var. *multiceps* Tsen et Lee)‘雪里蕻’中克隆到了 *BcMF6* 的同源基因 *BjMF6t* 和 *BjMF6m*。 *BjMF6t* 和 *BjMF6m* 的序列完全相同, 故合称 *BjMF6*。该基因与 *BcMF6* 在DNA水平的相似性高达99.6%。Blast分析发现它与拟南芥外切PG基因在氨基酸水平上有很高的相似性, 并且具有所有PG基因特有的功能结构域和活性位点, 推测 *BjMF6* 属于PG基因。通过序列分析和系统树构建进一步证明了该基因可能是与花粉发育相关的PG基因。

关键词 芥菜; *BjMF6*; 花粉发育; 克隆; 多聚半乳糖醛酸酶

花粉发育是一个多步骤的复杂过程, 受众多基因的表达调控。在这个过程中, 有大量基因被激活^[1]。到目前为止, 已经有近百个花粉特异基因被克隆, 尤其是近年来, 随着大规模基因组分析方法应用于花粉发育的研究, 使人们找到了更多的花粉表达基因。

本实验室先前从白菜中克隆得到一个花粉特异表达的多聚半乳糖醛酸酶(polygalacturonase, PG)基因 *BcMF6* (GenBank 登录号: DN237915)。PG是一种细胞壁结合蛋白, 可以催化果胶分子中 α -(1,4)-聚半乳糖醛酸的裂解, 在植物发育许多阶段, 尤其是那些需要细胞分离的阶段, 参与果胶降解。已有研究表明, PG基因家族的多个成员特异地存在于花粉中, 它们可能在花粉成熟和花粉管伸长中发挥重要的作用^[2-6]。PG基因是一大基因家族, 目前, 到底有多少PG家族的成员在花粉发育中起作用, 以及它们在花粉发育中的具体作用还不甚清楚。

茎瘤芥(*Brassica juncea* Czern. et Coss var. *tumida* Tsen et Lee)和分蘖芥(*B. juncea* Czern. et Coss var. *multiceps* Tsen et Lee)为芸薹属作物, 与白菜的亲缘关系较近, 为此本研究根据白菜 *BcMF6* 的序列设计引物, 通过同源序列克隆的方法获得茎瘤芥和分蘖芥中 *BcMF6* 的同源基因, 并利用生物信息学方法分析该基因的结构特征。研究不同芸薹属作物中的PG同源基因, 有助于更好地了解PG基因在芸薹属植物花粉发育中的作用, 进一步了解它们之间的相互关系,

为更好的解析花粉发育全过程打下基础。

1 材料与方法

1.1 材料

本实验所用材料为茎瘤芥‘浙桐1号’和分蘖芥‘雪里蕻’, 2005年9月25日播于苗床, 10月25日移栽于试验田, 田间管理按常规进行。在苗龄50天时, 从群体中随机选取单株, 每株取2~3片鲜嫩叶片, 在开花期时选取健康饱满的花蕾, 用75%的酒精棉擦洗干净, 做好记号, 用锡箔纸包好, 立即投入液氮中固定, -75℃保存备用。

Trizol[®] 购自Life Technology公司。cDNA合成试剂盒(SMARTTM PCR cDNA Synthesis Kit)购自Clontech公司。*Eco*RI、EDTA、DEPC、IPTG、X-gal等购自上海生工生物工程技术有限公司。Taq DNA聚合酶、T4 DNA连接酶、pGEM-T-easy-vector 购自Promega公司。*E. coli* DH 5 α 菌株为本实验室保存。

1.2 DNA、总RNA的提取与cDNA的合成

DNA提取参照本实验室^[7]的方法。各植物组织总RNA用Trizol抽提, 具体操作方法参照Trizol产品

收稿日期: 2006-04-27 接受日期: 2006-10-10

国家自然科学基金项目(No.30671426)和浙江省重大科技项目(No.2005C12019-02)资助项目

通讯作者: Tel/Fax: 0571-86971188, E-mail: jshcao@zju.edu.cn

手册。cDNA 第一链和第二链的合成参照 SMART™ PCR cDNA Synthesis Kit 使用手册进行。

1.3 引物合成

根据 *BcMF6* 的 cDNA 序列的 5'-UTR 和 3'-UTR 序列设计一对引物 PI, 扩增目的基因的 DNA 和 cDNA 全长序列。引物 PI 序列为: P1: 5'-ATTGTAAT-CCCCAGTAAC-3'; P2: 5'-GTGTATTGGTT-ATCGCATC-3'。引物委托上海生工生物工程技术服务有限公司合成。

1.4 目的基因的 PCR 扩增、克隆和测序

分别以‘浙桐1号’和‘雪里蕻’总 DNA 以及花蕾双链 cDNA 为模板, 以 PI 为引物, 扩增 DNA 全长和 cDNA 全长。PCR 反应体系: 0.5 μl DNA 或 cDNA 模板, 2.5 μl 10× PCR 反应缓冲液, 0.5 μl 10 μmol/L dNTP, 上下游引物(10 μmol/L)各 0.5 μl, 1.5 U Taq DNA 聚合酶, 加无菌水至总体积 25 μl。PCR 扩增条件: 94 °C 2 min; 94 °C 30 s; 58 °C 30 s; 72 °C 2 min; 35 个循环后, 72 °C 延伸 10 min。

PCR 产物用 Vitagene 公司的 DNA 凝胶回收试剂盒回收, 目的片段连接到 pGEM-T-easy-vector, 连接好的质粒转化 *E.coli* DH 5α 感受态细胞, 在含有 IPTG 和 X-gal 的氨苄青霉素 LB 培养基上挑选白色重组克隆。碱裂解法抽提质粒, 用 *EcoRI* 进行单酶切鉴定和 PCR 鉴定后, 送上海博亚生物公司测序。所得基因分别命名为 *BjMF6t*、*BjMF6m*。

1.5 核苷酸序列分析、同源比对和系统树构建

DNAStar 软件分析 *BjMF6t*、*BjMF6m* 最大开放阅读框, 推导氨基酸序列; 并通过蛋白质专家分析系统(<http://ca.expasy.org>)对其编码蛋白质序列进行结构特征分析; 利用 ClustalX(version 1.8)软件将 *BjMF6t*、*BjMF6m* 与数据库中的其他 PG 基因序列(表 1)进行同源序列的多重排定, MEGA3.1 构建 NJ 系统树, 节点支持率使用 1 000 次 bootstrap 检验。

2 结果

2.1 白菜 *BcMF6* 同源基因 *BjMF6t* 和 *BjMF6m* 基因的克隆与序列分析

分别以‘浙桐1号’和‘雪里蕻’总 DNA 以及花蕾双链 cDNA 为模板, 以 PI 为引物, 扩增 *BjMF6t*、*BjMF6m* DNA 全长和 cDNA 全长(图 1)。PCR 扩增结果与预期的条带大小一致。将目的基因克隆到 pGEM-T-easy-vector 上, 转化 *E.coli* DH 5α。通过蓝白斑筛选阳性克隆, 碱法提取重组质粒, *EcoRI* 酶

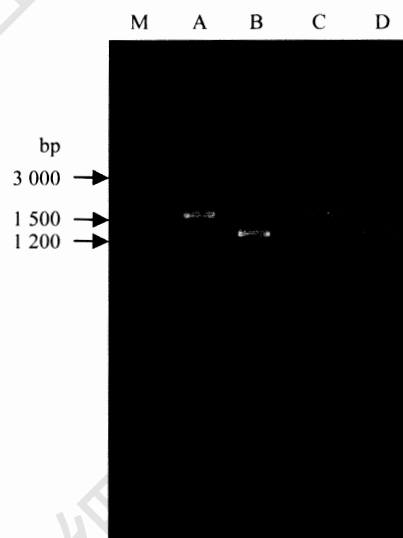


图 1 *BjMF6t*、*BjMF6m* 的 PCR 扩增结果

M, DNA 分子量; A, *BjMF6t* DNA; B, *BjMF6t* cDNA; C, *BjMF6m* DNA; D, *BjMF6m* cDNA。

切鉴定, 结果表明目的基因已被成功克隆。重组质粒插入片段全序列测定结果见图 2。测序结果表明, *BjMF6t* 与 *BjMF6m* 序列完全相同, DNA 全长 1 557 bp, cDNA 实际长度为 1 255 bp, 故以下分析均用 *BjMF6*(GenBank 登录号: DQ658218)表示。

用 ClustalX 软件比较 *BjMF6* 的 cDNA 和 DNA 序列发现, 该基因含有 4 个外显子和 3 个内含子。3 个内含子的长度分别为 81 bp、94 bp 和 127 bp, 外显子-内含子交接点处序列均遵守 GT-AG 规律。利用 DNAStar 软件对 *BjMF6* 序列进行分析, 发现其最大开放阅读框为 1 194 bp, 编码 397 个氨基酸(图 2), 其中碱性氨基酸(K, R)43 个, 酸性氨基酸(D, E)34 个, 疏水氨基酸(A, I, L, F, W, V)134 个, 极性氨基酸(N, C, Q, S, T, Y)114 个。编码的假定蛋白质的分子量为 4.28 kDa, 等电点为 8.531; 不稳定指数为 30.75, 归为稳定蛋白。应用 Hphob./Kyte & Doolittle 算法, 预测蛋白质的疏水性平均值为 -0.077, 脂溶指数(aliphatic index)为 90.83, 表明其为疏水性脂溶蛋白。通过 Prosite 数据库查询, 发现该序列具有 3 个 N-糖基化位点(172~175, 257~260, 293~296), 5 个酪蛋白激酶 II(CK2)磷酸化位点(42~45, 156~159, 174~177, 209~212, 215~218), 8 个豆蔻酰化位点(2~7, 17~22, 32~37, 83~88, 89~94, 112~117, 191~196, 235~240), 3 个蛋白激酶 C(PKC)磷酸化位点(84~86, 141~143, 317~319), 1 个 PG 活性位点(227~240), 1 个染色体浓缩调控蛋白(regulator of chromosome condensation, RCC)信号 2(RCC1_2)位点(228~238)(图 3)。通过

M G V H F G I S V •

1 ATTGTAATCC CCAGTAACATA AAAAAATCA TATAATG GGT GTACATTTTG GAATATCOGT
• L F V V C L L G L S A N A K D F T I T A •

61 ATTATTTGTT GTTGTGTTGC TAGGTTTATC AGCAATGCT AAGATTCA CAATCACTGC
• P P G S D I T S •

121 TCCTCCAGGT TCTGATATCA CCAGCGTAAG TTCACAACCA ACCTTTGACT TGACATTAGT
• L L L K T •

181 TTCTTTGTTT TCCAAGCAAG GTTAAATCT TTTGGGATCG ATGCAGTTAT TGTGAAAC
• F N E A C Q F P T R S S V L I P K G E Y •

241 ATTCAATGAG GCATGCCAGT TCCGACCAG GAGTAGCGTA TTGATCCCTA AAGGTGAATA
• K L R Q I E M M G P C K A P I R I T L Q •

301 CAAGCTTCGG CAAATAGGA TGATGGGTCC ATGCAAGCT CCATAAGGA TTACTCTTCA
• G T V K A D G N V N G N D Y V V A P R R •

361 AGGCACTGTG AAAGCTGAG GAAAGCTTAA CGGGAATGAC TACTGGGTGC CTTTCCGAG
• I N G F K L N G G G I F D G E G N A A W •

421 GATTAATGGG TTTAAGTTGA ACGGAGGTGC TATCTTTGAC GGTGAAGTA ACGTGCATG
• R A N N C H K M A L T Q C K K L P I •

481 GAGGCTAAT AATTGCCACA AAATGGCATT AACTCAGTGC AAGAACTCC CTATCGTAGC
541 TGTCTATCTA TCTCTATGTT TACATAATAA TCCGTTAAAT ATAAACAGT TTTGTTAATT
S I R P D F V T D A •

601 TGTTTATGTT TTGTGTCATT TTTAACAGA GCATACGGTT TGATTTCTG ACTGACGCTA
• I R G I T S L D S K N F H I N V L G A R •

661 AGATAAGAGG CATAACCTCG TTGATTCAA AGAACTCCA CATTAACTG CTGGAGCAA
• N M T L E E I T I I A P E E S P N T D G •

721 GGAACATGAC ATTGGAAGAA ATCAGATCA TTGCCCTGA AGAGAGTCC AACACCGAGC

• I H V G R S V G V Q I I N S N I K Y G D •

781 GAATCCAAGT GGAAGGAGT GTTGGAGTCC AGATCATCAA CTCAAACATC AAACAGGAG
• D C I S I G D G T R D L L V E R V T C G •

841 ATGACTGCAT CTCTATTGGA GACGGGACGA GAGACCTTCT TGTGAAAGA GTTACATGGC
• P G H G I S I G S L G L Y V K E E D V T •

901 GTCCGGGACA TGGAACTAGT ATTGGAAGCC TCGGATTATA CGTGAAGGAG GAAGACGTCA
• G I R V V N S T L I N T D N G V R I K T •

961 CTGGCATCAG GGTGTAAC TCCACCTCA TAAACACTGA CAATGGTGA AGGATCAAGA
• W P S A A C S T T A S G I H F E N I I L •

1021 CATGCCCGTC TGCAGCTGC TCCACCACCG CCTCGGTAT CCATTTTGG AATATTATCC
• K N V T N P I L I D Q E Y C P W N R C N •

1081 TCAAGAAAGT TACCAACCT ATCTCATOG ACCAAGATA CTGCCCTGG AACCGATGCA
• K N •

1141 ACAAGAAAGT GAGTACACTT TCAAGAAA TTTAATGAT CATAGTTTT CTATGAGATA
1201 TGCATAGATC AATCCATGGG AACTGCTTTC ATTGGATGAT GACTCCAAGA CGTTTTTTTT
K P S T I K L V D I S F K H I

1261 TTTGTTTGT TGTAGAAAC ATCAACAATC AAGCTGGTG ACATAAGCTT CAAGCATATC
R G T S G N K D A V K L L C S K G F P C

1321 AGAGGAACAT CAGGGAACAA AGACGCGT AAGCTTTTGT GCAGCAAGG ATTTCGTCG
K N V Q I G D I D I K Y T G A D G P A T

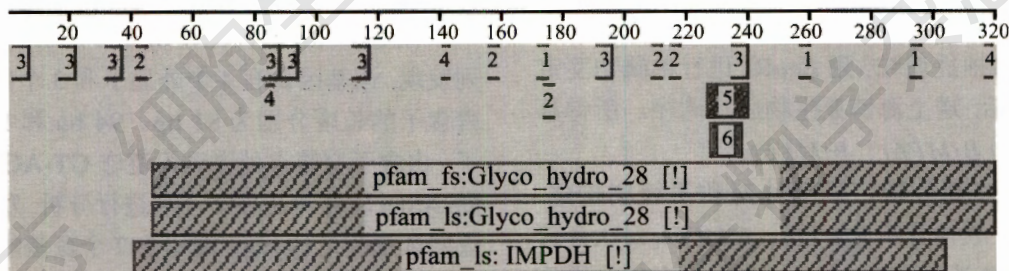
1381 AAGAAGCTGC AGATTGGTGA CATTGACATT AAGTACACG GAGCGCATG TCCAGCGACA
P Q C S N V S P T L M G T Q V P K A C S

1441 TTCCAATGCT CCAACGATC TCCACGTTG ATGGGAAC T AGTCCCTAA AGCATGTAGC
S P V T K L P G Q •

1501 AGCCCACTGA CTAAGTACC AGGCCAGTAA GAAATCTCGA TGGGATAACC AATACAA

图2 *BjMF6* 序列及推测的氨基酸序列

BjMF6 含4个外显子和3个内含子, 最大开放阅读框为1 194 bp, 编码397个氨基酸。氨基酸序列中的黑线部分为PG基因所特有4个功能结构域。代表氨基酸的字母标在第2个密码子上方。

图3 *BjMF6* 预测蛋白质结构域分析

1, N-糖基化位点; 2, CK2 磷酸化位点; 3, 豆蔻酰化位点; 4, PKC 磷酸化位点; 5, PG 活性位点; 6, RCC1_2 位点。

PredictProtein预测*BjMF6*的二级结构, 表明该蛋白质含6.30%的 α 螺旋, 48.36%的 β 折叠和45.34%的环状结构。

2.2 同源比对与系统树构建

从GenBank数据库下载其他植物的PG基因(表1), 将*BjMF6*编码的氨基酸序列与数据库中的其他PG序列进行同源序列比对, 发现该基因具有所用PG基因所特有的4个功能结构域(图2), 其中结构域I(SPNTDG)、(GDDC)中的3个天冬氨酸残基和结构域III(CGPGHGISIGSLG)中的H残基被认为与催化反应有关, 高度保守的带正电的结构域IV(RIK)和基质中的羧基相互作用^[8,9], 说明该基因是典型的PG基因。此外, 所推测的*BjMF6*编码的氨基酸序列中包

含14个Cys结合位点, 其中大多数在比对的所有PG基因中都是保守的, 而半胱氨酸可能在维持胞外蛋白的三维结构中发挥重要作用^[4]。将*BjMF6*序列与数据库中的其他植物PG基因序列进行同源比对发现, 该基因与拟南芥中PG基因*PGA3*类似基因NM_124188同源性最高, 在氨基酸水平上具有89%的相似性, 此外, 与拟南芥花粉特异表达的外切PG基因*PGA3*和花粉高表达基因*PGA5*在氨基酸水平上分别具有86%和84.7%的相似性。用MEGA3.1软件对*BjMF6*与这些PG基因的CDS序列构建系统树发现, 系统树明显分为A、B、C三个分支(图4)。*BjMF6*与NM_124188、*PGA3*和*PGA5*等同聚在分支C上。有研究和数据显示, 分支C上的其他PG基因均在花

表 1 研究中所引用的植物 PG 基因

物种	序列登录号	基因名称
<i>Arabidopsis thaliana</i> L.	NM_124188	-
<i>Arabidopsis thaliana</i> L.	X73222	<i>PGA2</i>
<i>Arabidopsis thaliana</i> L.	AJ005584	<i>PGA3</i>
<i>Arabidopsis thaliana</i> L.	NM_100158	<i>PGA4</i>
<i>Arabidopsis thaliana</i> L.	AJ003135	<i>PGA5</i>
<i>Arabidopsis thaliana</i> L.	NM_129948	<i>PGA7</i>
<i>Arabidopsis thaliana</i> L.	NM_180028	<i>PGA10</i>
<i>Brassica napus</i> L.	Z49971	<i>PGA</i>
<i>Brassica napus</i> L.	L19879	<i>Sta 44-4</i>
<i>Brassica napus</i> L.	AY280869	<i>Sta44G2</i>
<i>Brassica napus</i> L.	AJ250919	<i>PGAZ</i>
<i>Brassica oleracea</i> L.	AF518563	<i>PG1</i>
<i>Brassica rapa</i> L.	AJ428543	<i>BrPG1</i>
<i>Cucumis melo</i> L.	AF062466	<i>MPG 2</i>
<i>Cucumis melo</i> L.	AJ002532	<i>ADPG1</i>
<i>Gossypium hirsutum</i> L.	U09717	<i>G9</i>
<i>Lycopersicon esculentum</i> Mill.	U70480	<i>TAPG2</i>
<i>Medicago sativa</i> L.	U20431	<i>MePG</i>
<i>Medicago sativa</i> L.	AY219848	<i>PG11</i>
<i>Nicotiana tabacum</i> L.	X71020	<i>Npg1</i>
<i>Prunus persica</i> L.	X76735	<i>PPMEPGA</i>
<i>Pyrus communis</i> L.	AB084462	<i>PC-PG2</i>

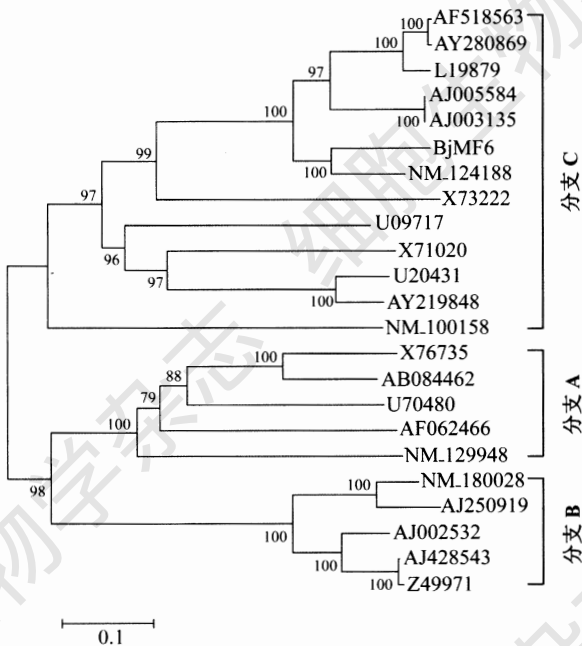


图 4 用 NJ 法构建的植物 PG 基因进化树

粉或花蕾中表达^[10-13]。分支 A 和分支 B 中的 PG 基因则主要在果实和脱落区中表达^[14]。

3 讨论

通过同源序列克隆法, 我们从茎瘤芥和分蘖芥中克隆了白菜花粉特异的 PG 基因 *BcMF6* 的同源基因

BjMF6t 和 *BjMF6m*。*BjMF6t* 和 *BjMF6m* 两者的 DNA 序列完全相同, 故可将其合称为 *BjMF6*, 说明该基因在芸薹属芥菜种中的同源性很高。序列分析发现 *BjMF6* 含有 4 个外显子和 3 个内含子, 可能编码的含有 397 个氨基酸的蛋白质为疏水脂溶性的稳定蛋白。前人研究发现, 花粉特异的 PG 氨基酸大小范围在 362 和 445 个氨基酸之间^[15], *BjMF6* 编码的氨基酸大小刚好在此范围内。序列中具有 3 个 N-糖基化位点, 5 个 CK2 磷酸化位点, 8 个豆蔻酰化位点, 3 个 PKC 磷酸化位点, 1 个 RCC1_2 位点。大量研究表明, N-糖基化位点^[16,17]、N-豆蔻酰化位点^[18,19]与细胞内信号转导、蛋白质定位以及黏附等过程有关, 而 N-糖基化位点还与转运有关^[20,21]。尤其值得注意的是, 在第 227 到 240 个氨基酸之间有一个确定的 PG 活性位点, 并且具有所有 PG 基因所特有的 4 个功能结构域。这些结果说明该基因属于 PG 基因。同源比对发现, *BjMF6* 与拟南芥中 PG 基因 *PGA3* 类似基因 NM_124188 同源性最高, 与拟南芥 PG 基因 *PGA3*、*PGA5* 在氨基酸水平上也具有很高的相似性。将 *BjMF6* 与 GenBank 中其他植物 PG 基因的 CDS 序列构建系统树发现, 系统树明显分为三个分支。Hadfield 等^[14]对众多的已克隆植物 PG 基因的氨基酸序列进行分析, 将 PG 基因分成三大分化单位, 即分化单位 A, 分化单位 B 和分化单位 C。分化单位 A 由非花粉组织中表达的基因组成, 主要存在于果实和脱落区, 编码缺乏一段预期的前序列的蛋白质; 分化单位 B 也存在于果实和脱落区, 不同的是, 其编码的 PG 基因具有一段前序列; 分化单位 C 完全由花粉表达的外切 PG 基因组成, 编码的蛋白质无前序列。在我们编制的进化树上, 分支 A 上的 PG 基因如 NM_129948 (*PGA7*) 在果实和脱落区表达^[10], 编码的蛋白质缺乏前序列, 此分支对应于分化单位 A; 分支 B 对应于分化单位 B, 其上的 PG 基因也在果实和脱落区表达, 但编码的蛋白质在 N 端具有延展序列, 在翻译后才被去除, 如 AJ0025329 (*ADPG1/PGA9*) 和 NM_180028 (*PGA10*)^[10]。*BjMF6* 聚在分支 C 上, 与其聚得最近的是 AJ005584 (*PGA3*) 和 AJ003135 (*PGA5*), *PGA3* 是花粉特异的外切 PG 基因, *PGA5* 尽管在果荚发育早期也有表达, 但主要在花粉中表达^[10]。距离次远的是甘蓝型油菜花粉表达的 PG 基因 L19879 (*Sta 44-4*)、AY280869 (*Sta44G2*)^[4,11], 以及从芥菜花粉中分离得到的 AF518563 (*PG1*) (GenBank 登录号: AF518563); 聚在这一分支上的还有拟南芥花粉特异的外切 PG 基因 X73222 (*PGA2*) 和

NM_100158 (*PGA4*)^[10], 在棉花花粉发育后期特异表达的 U09717 (*G9*)^[12], 在烟草花粉发育后期特异表达的 X71020 (*NpgI*)^[13], 以及在紫花苜蓿花粉中特异表达的 U20431 (*MePG*)和 AY219848 (*PGII*)(GenBank 登录号分别为 U20431 和 AY219848)。在花粉中表达的 PG 基因被认为可能与花粉成熟和花粉管的伸长、生长有关^[3,6]。

以上这些结果表明 *BjMF6* 可能编码芥菜花粉表达的外切 PG, 并在其花粉成熟和花粉管的伸长中起作用。尽管它们的确切功能需要进一步的确定, 但序列和生物信息学分析提示它是一个重要的花粉发育相关基因, 也为其确切功能的鉴定提供了线索, 进一步的研究正在进行中。

参考文献 (References)

- [1] Twell D. In: O'Neill SD et al. (eds.) *Plant Reproduction, Annual Plant Reviews Vol.6*, Sheffield: Sheffield Academic Press and

Boca Raton: CRC Press, 2002, 86

- [2] Niogret MF et al. *Plant Mol Biol*, 1991, **17**: 1155
 [3] Allen RL et al. *Plant J*, 1993, **3**: 261
 [4] Robert LS et al. *Plant Mol Biol*, 1993, **23**: 1273
 [5] Futamura N et al. *Plant Cell Physiol*, 2000, **41**: 16
 [6] Honys D et al. *Plant Physiol*, 2003, **132**: 640
 [7] 曹家树等. *园艺学报*, 1995, **22**: 47
 [8] 陆胜民等. *园艺学报*, 1999, **26**: 369
 [9] Bussink HL et al. *Curr Gent*, 1991, **19**: 467
 [10] Torki M et al. *Mol Gen Genet*, 1999, **261**: 948
 [11] Hong HP et al. *Plant Cell Rep*, 1997, **16**: 373
 [12] John ME et al. *Plant Mol Biol*, 1994, **26**: 1989
 [13] Tebbutt SJ et al. *Plant Mol Biol*, 1994, **25**: 283
 [14] Hadfield KA et al. *Plant Physiol*, 1998, **117**: 337
 [15] Qiu X et al. *Sex Plant Reprod*, 1996, **9**: 123
 [16] Ding DX et al. *J Biol Chem*, 1995, **270**: 24580
 [17] Buteau H et al. *Mol Endocrinol*, 1998, **12**: 544
 [18] Borgese N et al. *J Cell Biol*, 1996, **135**: 1501
 [19] Boisson B et al. *J Biol Chem*, 2003, **278**: 43418
 [20] Asano T et al. *J Biol Chem*, 1991, **266**: 24632
 [21] Li LB et al. *J Biol Chem*, 2004, **279**: 21012

Molecular Cloning and Bioinformatic Analysis of Pollen Development Related Gene *BjMF6*

Yu-Chao Zhang, Yi-Qun Ye, Li Huang, Jia-Shu Cao*

(Laboratory of Cell & Molecular Biology, Institute of Vegetable Science, Zhejiang University, Hangzhou 310029, China)

Abstract In the paper, *BjMF6t* and *BjMF6m* gene, homologous with the *BcMF6* gene encoding pollen-specific polygalacturonase (PG) of Chinese cabbage, was cloned from *Brassica juncea* Czern. et Coss var. *tumida* Tsen et Lee and *Brassica juncea* var. *multiceps* Tsen et Lee. The two genes were found to have an identical DNA sequence, so, they were given the same name *BjMF6*. The *BjMF6* gene showed 99.6% overall DNA similarity with that of *BcMF6*. Blast result showed that it also shared high similarity to PG genes in *Arabidopsis thaliana*. Sequence analysis revealed that one potential PG active position and four highly conserved domains of the plant and fungal PGs were present in *BjMF6* gene. Furthermore, phylogenetic analysis indicated that *BjMF6* might fall into the category of clade-C, which is comprised entirely of genes expressed in pollen. All these results suggested that *BjMF6* might encode a PG related to pollen development.

Key words *Brassica juncea*; *BjMF6*; pollen development; cloning; bioinformatic analysis

Received: April 27, 2006 Accepted: October 10, 2006

This work was supported by the National Natural Science Foundation of China (No.30671426) and Key Sci-technology Project of Zhejiang Province (No.2005C12019-02)

*Corresponding author. Tel/Fax: 86-571-86971188, E-mail: jshcao@zju.edu.cn