

烟草 *ovate* 同源基因的全长 cDNA 克隆与序列分析

桂蓓^{1,2}, 王瑛^{1,*}

¹ 中国科学院武汉植物园, 武汉 430074; ² 中国科学院研究生院, 北京 100049

提要: 根据番茄中控制果实形状的主效数量性状基因 *ovate* 的序列, 用生物信息学方法从茄科植物烟草中获得直系同源 *ovate* 基因(*NTovate*)的特异片段, 经鉴定, 此基因在烟草中至少有2个拷贝。在此基础上用cDNA末端快速扩增(RACE)方法, 获得其中1个拷贝的1059 bp *NTovate* 全长 cDNA 序列。序列分析表明, *NTovate* cDNA 序列编码352个氨基酸, 其蛋白序列与番茄 *ovate* 蛋白序列和拟南芥 *ovate* 蛋白家族 AtOFP7 蛋白分别为70%和36%的序列一致率, 而与此家族中其他蛋白以及水稻 *ovate* 蛋白仅在保守的 *ovate* 结构域有较低的同源性。此基因已在 GenBank 中登录(EU043369)。

关键词: 烟草; *ovate* 同源基因; 全长 cDNA; 序列分析

Cloning and Sequence Analysis of *ovate* Orthologous Gene in Tobacco (*Nicotiana tabacum* L.)

GUI Bei^{1,2}, WANG Ying^{1,*}

¹Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China; ²Graduate School of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Tomato *ovate* gene has been cloned and proved to control the fruit shape development. In this study, *ovate* orthologous gene in tobacco (*Nicotiana tabacum*), which belongs to Solanaceae and doesn't have fruit development, was cloned and compared with homologs in tomato and other model species. Two sequence fragments with 90% identity were identified using bioinformatics methods. Two copies of tobacco *ovate* ortholog (*NTovate*) were then confirmed by sequence specific PCR. A 1059 bp full length cDNA of *NTovate1* was obtained using rapid amplification of cDNA ends (RACE). Further analysis on *NTovate1* cDNA indicated that it encoded a peptide containing 352 amino acids. The amino acid sequence comparison with tomato *ovate* protein and AtOFP7 (*Arabidopsis thaliana* *ovate* family protein 7) showed that the sequence identity was 70% and 36%, respectively. Other AtOFPs and *ovate* family protein-like in rice only showed homolog in the *ovate* domain with much lower sequence identity. The full length cDNA sequence of *NTovate1* was registered in GenBank with accession number EU043369.

Key words: tobacco (*Nicotiana tabacum*); *ovate* orthologous gene; full-length cDNA; sequence analysis

茄科(Solanaceae)植物约有3000种, 它包括番茄(*Solanum lycopersicum*)、茄子(*Solanum melongena*)、马铃薯(*Solanum tuberosum*)、辣椒(*Capsicum annuum*)、粘果酸浆(*Physalis ixocarpa*)等食用蔬菜和烟草(*Nicotiana tabacum*)等经济植物以及花烟草(*Nicotiana alata*)和矮牵牛(*Petunia hybrida*)等观赏植物(Knapp 2002)。迄今, 有关烟草的基因组结构和核苷酸组成的研究仍十分不足, 它的基因组是茄科模式植物番茄的5倍, 约含45亿对碱基, 是人类基因组的1.5倍; 与人类基因组类似, 绝大部分烟草基因组的碱基对以重复非编码序列的形式存在(Opperman 2003)。鉴定一些重要的与烟草品质相关的基因(包括产量和质量

等), 有利于提高不同环境下生长的烟草质量, 以及用于传统和分子育种。

目前, 在番茄中已发现了500多个与花发育、果实发育、果实大小及产量以及次生代谢产物等有关的数量性状(Eshed 和 Zamir 1995; Bernacchi 和 Tanksley 1997; Grandillo 等 1999; Frary 等 2004; www.sgn.cornell.edu)。多个控制果实发育的质量性状和数量性状的相关基因已得到

收稿 2007-08-28 修定 2007-11-16

资助 国家自然科学基金(30570172)和武汉市晨光项目(20045006071-34)。

* 通讯作者(E-mail: yingwang@wbgcas.cn; Tel: 027-87510675)。

克隆和深入的研究,例如,显性突变基因 *nr* (*never-ripe*)通过切断乙烯受体途径而导致果实不能成熟(Wilkinson等1995;Yen等1995);主效数量性状基因 *fw2.2*是一个果实发育前期的负调节基因,它的表达能显著增加果实大小(Cong等2002;Nesbitt和 Tanksley 2002)。另一个主效数量性状基因 *ovate* 则为果实形状的调节基因,此基因的一个点突变会导致功能缺失,番茄形状即由圆形变为梨形,过量表达的 *ovate* 还影响叶片和花的生长和发育。据推测, *ovate* 蛋白质本身并不控制果实的形状差异,可能是一种植物生长抑制调节蛋白(Liu等2002)。基因组研究表明, *ovate* 同源基因在水稻和拟南芥基因组中也存在(Ku等2000,2001),拟南芥中 *ovate* 基因的过量表达,也会抑制叶片、花和角果的生长发育,显示出发育迟缓和不对称宽纹理叶片、卵形的花瓣萼片以及短角果等表型变化(Hackbusch等2005)。近缘和远缘植物的同源基因序列、结构和功能的比较研究,可为模式植物信息应用到经济作物中的基因改良和新品种培育提供参考。基于此,本文根据茄科模式植物番茄的 *ovate* 序列,克隆了同科近缘的烟草中 *ovate* 同源基因(*Nicotiana tabacum ovate* ortholog, *NTovate*),并分析比较浆果发育和非浆果发育的物种中 *ovate* 同源基因的序列特征,以期能为研究近缘物种中控制果实发育基因的功能进化提供参考。

材料与方法

烟草(*Nicotiana tabacum* L.'Petite Havana')种子发芽后于温室培养2周左右,取其幼嫩子叶作为实验材料。大肠杆菌菌株 DH5 α 为本实验室保存,克隆载体质粒 pMD18-T 购自 TaKaRa 公司, DNA 纯化回收试剂盒购自 OMEGA 公司, Trizol Reagent 购自 Invitrogen 公司, BD SMARTTM RACE cDNA 合成试剂盒和 BD PowerScript 反转录酶购自 TaKaRa 公司。引物采用 Prime 5.0 设计,由上海生物工程公司合成。

用番茄 *ovate* 蛋白序列(AAN17752)tblastn 烟草基因组数据库(Tobacco Genome Initiative, TGI, <http://tgi.ncsu.edu>, 由美国北卡州立大学(NCSU)农业与生命科学学院植物病理学系所构建)。从中

选取序列一致率高于 60% 和 E 值低于 1×10^{-10} 以上的序列,用序列拼接软件 Cap3 进行组装,获得烟草 *ovate* 同源基因片段,参数设置为重叠片段长度大于 20,序列相似度大于 95%。

用 CTAB 法(Rogers 和 Bendich 1985)提取烟草嫩叶总 DNA,根据上述所得序列设计基因特异引物对相应片段进行 PCR 扩增,验证 *NTovate* 序列预测结果。PCR 反应条件:94 预变性 5 min;94 变性 20 s,55 退火 20 s,72 延伸 35 s,扩增 35 个循环;72 延伸 7 min。PCR 产物连接到 pMD-18T 载体上,转化大肠杆菌 DH5 α 感受态细胞,挑选阳性克隆经 PCR 验证后测序。

根据已经获得的 *Ntovate* 一个片段序列,在序列保守区设计基因特异引物。根据 RNA 提取试剂盒 Trizol Reagent 的说明提取烟草叶片总 RNA,用分光光度法和琼脂糖凝胶电泳检测 RNA 质量。采用 BD SMARTTM RACE cDNA 合成试剂盒(TaKaRa 公司),在使用说明书的基础上,根据改良后的方法进行两轮 5' 与 3' cDNA 末端快速扩增(rapid amplification of cDNA ends, RACE),先以特异引物 GSP 与 Kit 末端通用引物 UPM 为引物进行第一轮扩增,再以稀释 10 倍的产物为模板,以巢式特异引物 NGSP 与 Kit 末端巢式通用引物 NUP 为引物进行第二轮扩增。PCR 反应条件:94 预变性 5 min;94 变性 30 s,68 退火 30 s,72 延伸 1 min,每循环一次退火温度降低 0.5,扩增 31 个循环降至 53;再 94 变性 30 s,53 退火 30 s,72 延伸 1 min,扩增 10 个循环;最后 72 延伸 10 min。两轮反应条件相同,扩增产物的克隆和测序同上。将 5' 与 3'-RACE 末端测序结果用 DNAMAN 软件进行拼接,得到全长序列。

所得序列再用 Blast 软件与 NCBI 的核酸与蛋白 Nr 数据库进行比对,并用 ClustalW 软件对同源蛋白序列进行比较和分析;使用 NCBI 提供的 Open Reading Frame Finder (ORF-finder)与 The Conserved Domain Database (CDD) (Marchler-Bauer 等 2005)分别进行基因开放阅读框预测与保守结构域预测;蛋白质疏水性分析使用软件 Kyte-doolittle (Kyte 和 Doolittle 1982);蛋白质信号肽、理论分子量、等电点和疏水性预测分别采用 SignalP、

Compute pI/Mw 以及 Protscal 程序(<http://au.expasy.org/tools>) ; 蛋白质跨膜预测分析采用 TMPRED 软件(<http://www.ch.embnet.org/cgi-bin/TMPRED>) ; 蛋白质蛋白信号定位采用 GenomeNet 的在线预测程序 Psort Prediction (<http://psort.hgc.jp/>)。所有软件使用的参数为默认设置。

实验结果

1 *NTovate* 特异片段的获得和鉴定

通过番茄 *ovate* 蛋白序列与 TGI 数据库搜索比对, 得到 10 个序列符合序列一致率高于 60% 和 E 值低于 1×10^{-10} 的条件, 拼接后获得大小分别为

1 000 与 1 086 bp 的 2 条特异片段 *contig1* 和 *contig2*, 中间 862 bp 的区域具有 90% 的序列一致率, 而两端的序列没有同源性。根据 2 条片段的同源和非同源序列设计引物: P1 (5' GAATAGGTAGAAG-GTCGTC 3') 为两片段同源区共有的上游非特异引物, P2 (5' CTCGTCGTTTAAGGTGTA 3') 和 P3 (5' CGTACGTACTTCGTTTCA 3') 分别为 *contig1* 和 *contig2* 非同源区的下游特异引物(图 1)。用 P1 和 P2、P1 和 P3 分别扩增出 600 bp 的 *contig1* 与 520 bp 的 *contig2* 相应目的片段(图 2), 测序结果与基因预测结果完全一致, 验证了 *ovate* 同源基因在烟草中的 2 个拷贝。

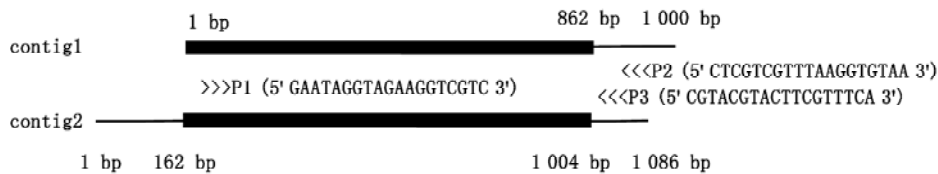


图1 *NTovate* 的序列拼接

Fig.1 The sequence assemble results of *NTovate*

粗线条表示同源区域, 细线条表示非同源区域。P1 为上游非特异引物, P2 和 P3 为下游特异引物。

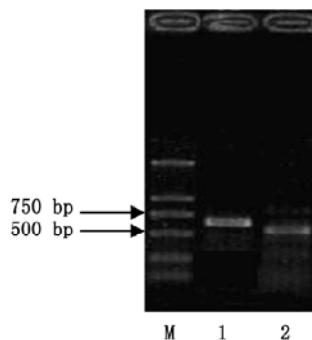


图2 *NTovate* 特异片段的 PCR 鉴定

Fig.2 PCR identification of *NTovate* special fragments

M: DL2000 标记; 1: *contig1* 片段; 2: *contig2* 片段。

2 *NTovate1* 全长 cDNA 克隆与序列分析

根据 *NTovate* *contig2* 片段序列设计了 RACE 的第一轮基因特异引物 GSP5 (5' GAGAAAGCCTCAACAATCAACCCA 3')、GSP3 (5' ACCAAGAGGAGAGTGAAACATTCCAGACGA 3') 和第二轮巢式基因特异引物 NGSP5 (5' CTATCGCAAAGCTCTCCTTCACTTTCCTTCC 3')、NGSP3 (5' AGAAGGTCGTCAGTTTCTACGTCATCGGACAG 3')。经

5' 与 3'-RACE 扩增, 5' 端获得 2 条带, 经测序确定目的条带为带 A, 带 B 比带 A 的 5' 上游缺失了 226 bp 序列, 其他序列相同, 可能是 mRNA 逆转录时 5' 端有部分降解, 3' 端获得单一特异条带(图 3)。

测序结果得到长度为 924 bp 的 5' 端片段与 460 bp 的 3' 端片段, 基于两片段间 115 bp 的重叠区序列, 并与 *contig2* 序列进行比对, 拼接得到全长为 1219 bp 的基因片段。DNAStar 软件分析表明, 该基因在 101~1 159 bp 之间有一开放阅读框, 共编码 352 个氨基酸, 5' 端非编码区有 100 bp, 3' 非编码序列有 60 bp, 其中包含 30 bp 的 polyA 尾。该序列与 *contig2* 中相应序列完全一致, 与番茄 *ovate* 基因 cDNA (AY140893) 核酸序列一致率达 83%。保守区域分析结果表明, 该基因推测出的蛋白序列包含一个功能未知的 DUF623 结构域, 含有 60 个氨基酸, 在番茄 *ovate* 蛋白和拟南芥 *ovate* 蛋白家族中也有发现, 又称之为 *ovate* 结构域 (Wang 等 2007)。由此证明获得了 *NTovate* 一个拷贝的全长 cDNA 序列, 命名为 *NTovate1* 基因(图 4)。

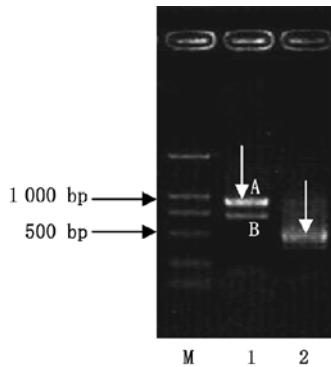


图3 *NTovate1* 的 RACE-PCR

Fig.3 RACE-PCR products of *NTovate1*

M: DL 2000 标记; 1: 5'-RACE 扩增; 2: 3'-RACE 扩增。箭头表示目的片段。

3 *NTovate1* 编码的氨基酸序列比较

将 *NTovate1* 编码的氨基酸序列 Blastp 搜索 NCBI 的 Nr 蛋白质数据库的结果显示, 其与番茄 *ovate* 蛋白具有 70% 的序列一致率, 与拟南芥 *ovate* 家族蛋白 AtOFP7 蛋白的同源性仅次于番茄, 其序列一致率为 36%。在 60 个氨基酸的 *ovate* 保守结构域中, *NTovate1* 与番茄 *ovate* 和拟南芥 AtOFP7 分别具有 4 个和 17 个氨基酸的差异(图 5)。与拟南芥 *ovate* 家族中其他 14 个蛋白相比, 除 *ovate* 结构域有一定的同源性外, 其他区段没有同源性, 在 *ovate* 结构域中氨基酸序列一致率从 40%~63% 不等。Blastp 结果发现在水稻中具有 62 个含有 *ovate* 结构域的未知蛋白, 其中预测的水稻 *ovate* 相似蛋白(BAD87078)保守域与 *NTovate1* 具有 33% 的序列一致率。在苜蓿中也曾发现 3 个含此结构域的未知蛋白, 保守域序列一致率最高达到 62%。这说明烟草与同科近缘的番茄相比, 除了在 *ovate* 保守域具有最高的同源性外, 在其他序列上具有一定的同源性, 可能体现了茄科中无浆果发育与有浆果发育作物的进化差异; 而与远缘的拟南芥、苜蓿、水稻相比, 仅在保守域中有较低同源性, 序列的一致率按亲缘关系的远近在拟南芥、苜蓿、水稻中依次降低。

4 *NTovate1* 蛋白的理化性质与结构

经 Compute pI/Mw 程序预测, *NTovate1* 编码的蛋白质分子量为 40.1 kDa, 等电点为 9.66, 经 SignalP 程序预测, 无信号肽序列, 为非分泌蛋白。Kyte-doolittle 疏水性分析结果显示, *NTovate1*

在 50~75 氨基酸间有个疏水区域。跨膜分析预测结果显示 *NTovate1* 不含跨膜区, 蛋白信号位点在线预测程序发现了一个核定位信号, 表明 *NTovate1* 可能是一个核蛋白。

讨 论

随着结构基因组学和功能基因组学的发展, 基因序列和结构的研究最终还是要归结到基因功能上来。用现有模式植物中同源基因的序列分析与比较, 寻找它们与产量和品质等重要经济性状之间的关系, 可以研究其功能的进化和演变, 并最终探索如何用它们来改良非模式物种的产量和品质(薛勇彪 2001)。

本文根据已知的番茄 *ovate* 蛋白序列和 TGI 数据库中的序列信息, 利用同源搜索拼接得到 *NTovate* 基因的序列片段, 并鉴定为 2 个拷贝, 说明烟草基因组中至少有 2 个 *ovate* 同源基因。经 RACE 扩增, 得到了其中 1 个拷贝 *NTovate1* 的全长 cDNA 序列。其蛋白序列分析结果表明, 与番茄 *ovate* 蛋白和拟南芥 AtOFP7 分别有 70% 和 36% 的相似性, 而且均具有 *ovate* 蛋白典型的保守功能域 DUF623 结构, 即 *ovate* 结构域。与拟南芥 *ovate* 家族中其他蛋白、苜蓿未知蛋白和水稻中预测的 *ovate* 类似蛋白在此结构域上也具有相对较低的同源性。*NTovate1* 蛋白与番茄、拟南芥 *ovate* 蛋白都不具有信号肽结构。蛋白信号位点分析发现, *NTovate1* 蛋白具有与番茄的 *ovate* 蛋白类似的核定位信号, 且均无跨膜区, 推测可能为核蛋白, 在拟南芥中已定位的 AtOFP 蛋白也位于核内(Wang 等 2007)。由此可见, 烟草中的 *ovate* 同源基因编码的蛋白质结构特性符合 *ovate* 蛋白的结构特点, 与同科近缘物种番茄具有较高的同源性, 与其他远缘物种在保守域也有一定的序列相似性。

但作为无浆果发育的茄科作物, 烟草的 *ovate* 同源基因也存在不少差异, 烟草的同源基因至少具有 2 个拷贝, 考虑到烟草的基因组比番茄的大约 5 倍, 可能是进化过程中的基因组大片段或基因复制事件的结果。*NTovate1* 蛋白虽与番茄 *ovate* 蛋白结构相似, 但番茄 *ovate* 蛋白经预测是个亲水蛋白(Liu 等 2002), 而 *NTovate1* 蛋白含有一个预测的疏水区域, 并且拟南芥同源蛋白中也不具有

```

                                GCGGGGCATAAGCCAGTTCACACTTCGACACTTTGTGCA -61
AAGGTAGAAAACTCCACTCCCATTATCAGACACAGAGCAGGTGCGGAGAGACAAAAAGA -1
    ──▶
ATGGGAAAAAGCTTTAAGCTTCGGTCTCTAAAGTCATTGCCTCCTCCATTCTGCCGT 60
M G K S F K L R F S K V I A S F H S C R
TCAAAGACTCTTCTACTCTTCTGAAAATCCTGTTCCTTCTCATTCTTCTCAAAATCA 120
S K D S S T L P E N P V P S H F F S K S
TCACATACTAAACCCACTACAAAACCTCATCGCCCTCGATTCCCTGTTGTTATTAECTA 180
S H T K P T T K L I A L D F P V V I N S
CCTCCTTCTTCTTCAAACGCCACGTGTCAGAAAACCGTGATTCTGTTGGCTGTGGGTTT 240
P P S S F K R H V S E T V I S V G C G F
AAGTCGCGCTCACAAGAATTCAGTGGGAGAAGGAAGACAAGTTTCACGTGGTTTCCCTCC 300
K S R S Q E F K W E K E D K F H V V S S
TTTGAAGATTCTGAAGGGGAGAATTTGACCTTAAGACCACCTTCTACTCCTCCAAGATTA 360
F E D S E G E N L T L R P P S T P P R L
GCCGTTGAGAAGAAGAAACGCCGAGCTAAAAAGACGAGTAAAACAAAATCCAGACTCCGA 420
A V E K K K R R A K K T S K T K S R L R
ATGAGCACCTCCTCCGCTGACGACAGTGGGATTTAAGCACTACTGAAACTTTGGATAAT 480
M S T S S A D D S G I L S T T E T L D N
TATAACGAAGAAGAAGATGAAACTGAAACTTTAGTTTCGTCTCCAGAAGCTTCGATTTC 540
Y N E E E D E T E T L V S S S R S F D F
TCAAGTGACTCGTCTACAGATTCAACCCTCAGTTGGAAACCATATGTGAGACCACT 600
S S D D S S T D F N P Q L E T I C E T T
ACTATAAGGCGTCGTAAGAGAGAAACGAAACACCAAGAGGAGAGTGAACATTCCAGA 660
T I R R R K K R N G N T K R R V K H S R
CGAAGCTTCTCGTCTTCAAAGGGTAGAAGGTCGTCAGTTTCTACGTCATCGGACAGCGAG 720
R S F S S S K G R R S S V S T S S D S E
TTGCCAGCGAGGTTGTCGGTGTCCAGAAGCTGATACCGGTAGCGTGAAGGGAAAGTG 780
L P A R L S V F Q K L I P C S V E G K V
AAGGAGAGCTTTGCGATAGTGAAGAAATCACAGAATCCGTACGAGGATTCAAGAGGTCG 840
K E S F A I V K K S Q N P Y E D F K R S
ATGATGGAATGATTTTGAGAAGCAAATGTTGAGAAGAATGAGTTAGACAGCTTTTA 900
M M E M I L E K Q M F E K N E L E Q L L
CAATGTTTCTGTCGTTGAATGGAAAGCATTATCATGGGTTGATTGTTGAGGCTTCTCC 960
Q C F L S L N G K H Y H G L I V E A F S
GAGATTTGGGAGACCTTGTTTTTAGTAATGGTAATGATAAAGCTAGGAGGATGTCAACT 1020
E I W E T L F L G N G N D K A R R M S T
CATCCCCGCCACGTAAGTACAATGCAAAACCCCTTGAATCGAGTTTCTAGTTTAAATG 1080
H P R P R T V Q C K T L *
ATTCCTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1119

```

图4 *NTovate1* 的cDNA序列及编码的氨基酸序列

Fig.4 The nucleotide sequence and amino acid sequence for *NTovate1*
框内为 *ovate* 保守结构域, 星号表示终止密码子。

该疏水区域, 有研究表明疏水区的螺旋结构可能与一定的功能相关(Seigneurin-Berny 等 1999), 因此 *ovate* 基因的功能在烟草中可能具有一定的特异

性。目前, *ovate* 基因的功能仅在番茄和拟南芥中有研究, 在番茄中 *ovate* 是一个控制果实形状从圆形到梨形转变的数量性状基因, 也是多效发育

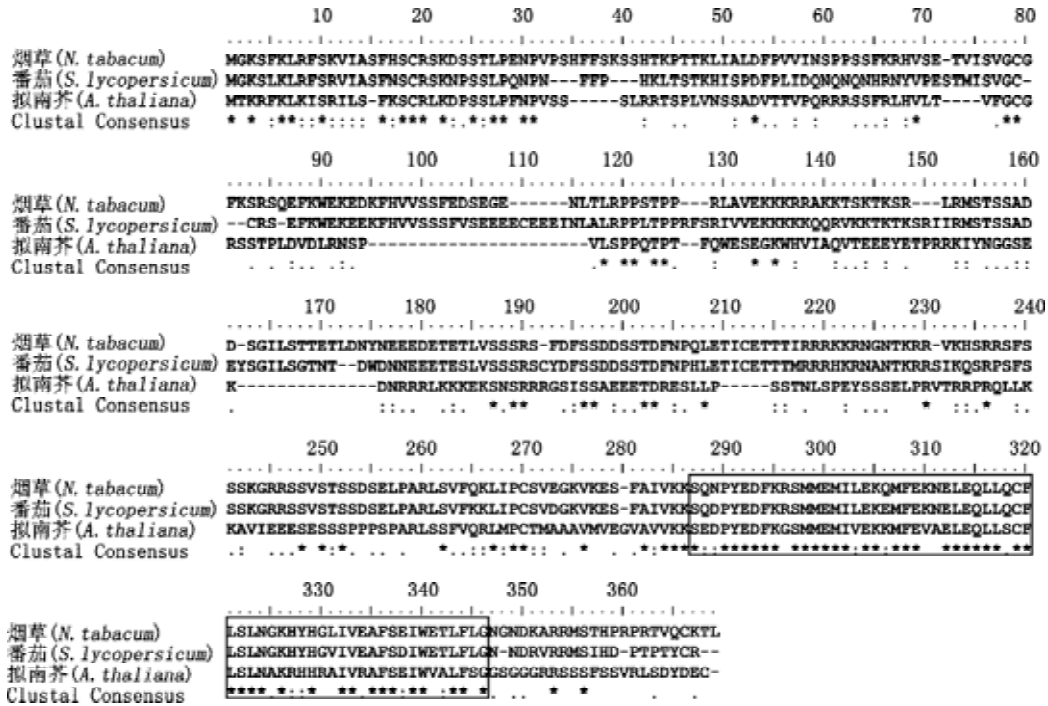


图5 *Ntovate1* 编码的氨基酸序列与番茄和拟南芥的同源氨基酸序列的比较

Fig.5 Alignment of amino acid sequences of *Ntovate1* and *ovate* orthologs from tomato and *Arabidopsis* 加框表示60个氨基酸的 *ovate* 保守结构域, GenBank 登陆号为 EU043369 (烟草)、AAN17752 (番茄)、NP_179440 (拟南芥)。

调控基因, 在植物生长发育中有抑制作用(Liu 等 2002)。拟南芥 *ovate* 蛋白家族中的 *AtOPF1* 是一个转录因子抑制子, 调节一种控制细胞延伸的赤霉素合成酶基因表达, 它的过表达会减小细胞长度, 从而发生胚轴、子叶、叶片、花器官、角果均变短的表型变化, *AtOPF2* 与 *AtOPF7* 也有相似的功能(Wang 等 2007)。本文选择茄科中无浆果的烟草 *ovate* 基因为对象, 该基因全长 cDNA 的获得对于正确解析其基因在非浆果作物中的功能及其表达调控的规律是非常重要的。而要真正了解 *ovate* 同源基因在烟草发育过程中的功能、是否存在特殊的疏水功能结构域、基因功能的进化机制等问题, 则尚需对其进行转基因研究。

参考文献

薛勇彪(2001). 国家重点基础研究发展规划项目: 水稻重要性状的功能基因组学研究项目简介. 植物学报, 43: 769~770
 Bernacchi D, Tanksley SD (1997). An interspecific backcross of *Lycopersicon esculentum*×*L. hirsutum*: linkage analysis and a QTL study of sexual compatibility factors and floral traits. Genetics, 147: 861~877
 Cong B, Liu J, Tanksley SD (2002). Natural alleles at a tomato fruit

size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci USA, 99: 13606~13611
 Eshed Y, Zamir D (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics, 141: 1147~1162
 Frary A, Fulton TM, Zamir D, Tanksley SD (2004). Advanced backcross QTL analysis of a *Lycopersicon esculentum*×*L. pennellii* cross and identification of possible orthologs in the Solanaceae. Theor Appl Genet, 108: 485~496
 Grandillo S, Zamir D, Tanksley SD (1999). Genetic improvement of processing tomatoes: a 20 years perspective. Euphytica, 110: 85~97
 Hackbusch J, Richter K, Müller J, Salamini F, Uhrig JF (2005). A central role of *Arabidopsis thaliana* *ovate* family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. Proc Natl Acad Sci USA, 102: 4908~4912
 Knapp S (2002). Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the Solanaceae. J Exp Bot, 53: 2001~2022
 Ku HM, Liu J, Doganlar S, Tanksley SD (2001). Exploitation of *Arabidopsis*-tomato synteny to construct a high-resolution map of the *ovate*-containing region in tomato chromosome 2. Genome, 44: 470~475
 Ku HM, Vision T, Liu J, Tanksley SD (2000). Comparing sequenced

- segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proc Natl Acad Sci USA*, 97: 9121~9126
- Kyte J, Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157: 105~132
- Liu J, Van Eck J, Cong B, Tanksley SD (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci USA*, 99: 13302~13306
- Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z et al (2005). CDD: a Conserved Domain Database for protein classification. *Nucl Acid Res*, 33: D192~D196
- Nesbitt TC, Tanksley SD (2002). Comparative sequencing in the genus *Lycopersicon*. Implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics*, 162: 365~379
- Opperman CH, Lommel SA, Sosinski B, Burke M, Lakey N, He L, Brierley R, Salstead A, Gadani F, Hayes A (2003). The tobacco genome initiative. *Plant and Animal Genome XI Conference*, San Diego, CA, 2003
- Rogers SO, Bendich AJ (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol Biol*, 5: 69~76
- Seigneurin-Berny D, Rolland N, Garin J, Joyard J (1999). Technical advance: differential extraction of hydrophobic proteins from chloroplast envelope membranes: a subcellular-specific proteomic approach to identify rare intrinsic membrane proteins. *Plant J*, 19: 217~228
- Wang S, Chang Y, Guo J, Chen JG (2007). *Arabidopsis* Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. *Plant J*, 50: 858~872
- Wilkinson JQ, Lanahan MB, Yen HC, Giovannoni JJ, Klee HJ (1995). An ethylene-inducible component of signal transduction encoded by *never-ripe*. *Science*, 270: 1807~1809
- Yen HC, Lee S, Tanksley SD, Lanahan MB, Klee HJ, Giovannoni JJ (1995). The tomato *never-ripe* locus regulates ethylene-inducible gene expression and is linked to a homolog of the *Arabidopsis Etr1* Gene. *Plant Physiol*, 107: 1343~1353