

# 芥菜 *LEAFY* 同源基因的克隆与进化分析

周贤龙<sup>1,5</sup>, 王小兰<sup>2</sup>, 王武源<sup>3</sup>, 葛学军<sup>4\*</sup>

(1. 中国科学院华南植物园, 广州 510650; 2. 广州大学功能基因与芯片研究中心, 广州 510405; 3. 江西农业工程职业学院,樟树 331200; 4. 中山大学生命科学学院, 广州 510275; 5. 中国科学院研究生院, 北京 100039)

**摘要:** *LEAFY* 同源基因是高等植物花的分生组织分化的重要调节基因。根据已发表的 *LEAFY* 同源基因序列保守区设计引物, 以芥菜 (*Capsella bursa-pastoris* (L.) Medic.) 基因组 DNA 序列为模板, 克隆了一条长 2 866 bp 的 *LEAFY* 同源基因。序列分析表明, 该基因含有 3 个外显子和 2 个内含子, 外显子编码 424 个氨基酸组成的多肽。其单个外显子核苷酸序列与拟南芥 (*Arabidopsis thaliana*) *LEAFY* 基因同源性在 90% 以上, 氨基酸序列同源性为 86%, 而与琴叶拟南芥 (*Arabidopsis lyrata*) 的氨基酸序列同源性高达 90%。不同植物物种的 *LEAFY* 同源氨基酸序列在 C 端高度保守, 而 N 端则有较大程度的变异。3 个外显子进化速率不同可能是由于所受选择压力不同所致。存在于芥菜 *CapLFY* 基因 346 位上的精氨酸突变可能是造成芥菜两种生态型花期不同的原因。

**关键词:** 芥菜; *LEAFY* 基因; 克隆; 进化分析

中图分类号: Q349.5 文献标识码: A 文章编号: 1005-3395(2006)02-0113-07

## Cloning and Evolutionary Analysis of Homologous Sequences of *LEAFY* Gene in *Capsella bursa-pastoris* (Cruciferae)

ZHOU Xian-long<sup>1,5</sup>, WANG Xiao-lan<sup>2</sup>, WANG Wu-yuan<sup>3</sup>, GE Xue-jun<sup>4\*</sup>

(1. South China Botanical Garden, the Chinese Academy of Sciences, Guangzhou 510650, China; 2. Centre for Functional Genomics and Microarray, Guangzhou University, Guangzhou 510405, China; 3. Jiangxi Agriculture Engineering Vocational College, Zhangshu 331200, China; 4. School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China; 5. Graduate University of Chinese Academy of Sciences, Beijing 100039, China)

**Abstract:** *LEAFY* is an important regulator of floral meristems genes in higher plants. To investigate its origin and evolution, a *LEAFY* homologs with length of 2 866 bp was cloned from *Capsella bursa-pastoris* genome using the conserved primers from *Arabidopsis thaliana* *LEAFY* gene. Sequence analysis indicated three exons with the length of 457 bp, 407 bp and 411 bp. These exons encoded a protein of 424 amino acids. The single exon was more than 90% homologous to *A. thaliana* *LEAFY* gene, and the amino acid homology was 86%. When compared to *Arabidopsis lyrata* *LEAFY* gene, there was 93 nucleotide substitution in the exons of *C. bursa-pastoris*, and 74.7% of these mutations located at the third position of codon, the amino acid homology reached 90%. Among the three exons, Ds/Dn ratio was 1.9, 2.3 and 6.0. Homologous comparison indicated that the C-terminal amino acids were highly conserved among different plant species while the great variation existed in N-terminal. In addition, the Arg mutation located at 346 in *CapLFY* gene was speculated to be one of main reasons that caused the ecotypes with different flowering time in *C. bursa-pastoris*.

**Key words:** *Capsella bursa-pastoris*; *LEAFY* gene; Cloning; Evolutionary analysis

收稿日期: 2005-09-09 接受日期: 2005-12-12

基金项目: 广东省自然科学基金项目(05006754), 广东省自然科学基金博士启动项目(04300465)资助

\* 通讯作者 Corresponding author

*LEAFY* 基因是 Weigel 等首先从拟南芥 (*Arabidopsis thaliana*) 中分离出来的控制花的分生组织分化的特征基因, 是金鱼草 (*Antirrhinum majus*) *FLORICAULA* (*FLO*) 基因的同源基因<sup>[1]</sup>, 编码一个 DNA 结合域, 作为转录因子通过激活花器官特征基因来启动花的分化<sup>[2]</sup>。它在花序形态和花期的调节等过程中扮演重要角色<sup>[3]</sup>, 它的突变将影响植物的花期<sup>[3]</sup>以及花序形态<sup>[4]</sup>, 这可能是导致莲座状花序进化的直接原因<sup>[5,6]</sup>。*LEAFY* 同源基因存在于所有陆地植物中, 经历了 4 亿年的进化, 在 C 端和 N 端形成包含两个保守功能域的保守序列<sup>[7]</sup>。现已从金鱼草<sup>[1]</sup>、烟草<sup>[8]</sup>、花椰菜<sup>[9]</sup>、水稻<sup>[10]</sup>、辐射松<sup>[11]</sup>、蓝桉<sup>[12]</sup>、紫水芹<sup>[13]</sup>等植物中分离得到了 *LEAFY* 同源基因, 所有这些都是单拷贝或低拷贝基因家族。

芥菜 (*Capsella bursa-pastoris* (L.) Medic.) 为十字花科 1 a 生或 2 a 生草本植物, 广布于全世界温带地区。作为植物研究模式种拟南芥的近缘种, 可以很好地借鉴拟南芥的分子、遗传、生理学数据及对基因调控和功能研究方面的方法和经验, 开展相应研究, 为有关理论研究和应用提供参考依据。

目前尚未见到芥菜 *LEAFY* 同源基因的相关报道。本实验应用 *LEAFY* 同源基因保守引物, 从芥菜中克隆 *LEAFY* 同源基因, 旨在对 *LEAFY* 基因同源序列进行克隆和序列分析, 以期为十字花科系统发育分析提供分子系统学资料。

## 1 材料和方法

**材料** 芥菜 (*Capsella bursa-pastoris*) 新鲜幼叶取自生长于华南植物园的野生个体。

**基因组 DNA 提取** 采用改进的 CTAB 微量抽提法<sup>[13]</sup>提取芥菜基因组 DNA。

**PCR 扩增** 根据 Shu 等<sup>[14]</sup>设计的一对引物: *LEAFY-F*: ATGGATCCTGAAGGTTTCACG; *LEAFY-R*: ACAGCTAATACCGCCAACCAA, 在 MJ Research 公司的 PTC-200 型 PCR 仪上进行 100  $\mu$ L 体系的扩增反应, 反应体系组成: 20 mmol/L Mg<sup>2+</sup>, 10  $\times$  Buffer 10.0  $\mu$ L, 10 mmol/L dNTP 2.0  $\mu$ L, 10  $\mu$ mol/L *LFY-F* 1.5  $\mu$ L, 10 mmol/L *LFY-R* 1.5  $\mu$ L, 50 ng 模板 DNA, 2.5 U Taq DNA 聚合酶, 加双蒸水至 100  $\mu$ L。扩增程序: 94℃ 5 min, 94℃ 30 s, 63.4℃ 45 s, 72℃ 1.5 min, 30 cycles, 72℃ 10 min。PCR 扩增产物于

1% 琼脂糖凝胶中电泳分离, 切胶后采用上海博亚生物技术有限公司生产的 DNA 回收试剂盒进行回收纯化。

**目的片段的亚克隆** 纯化的 DNA 产物与等摩尔的 pGEM-T Easy 载体 (Promega 公司) 在 T<sub>4</sub>DNA 连接酶作用下连接, 连接产物转化大肠杆菌 Top 菌株感受态细胞, 在 LB/Amp/IPTG/X-gal 平板上筛选重组子。重组质粒用 PCR 扩增和 EcoRI 限制性内切酶酶切进行双重鉴定。

**DNA 序列测定与分析** 将含有目的片段的重组质粒委托上海英俊生物技术有限公司使用 PE 公司的 ABI3730 型自动测序仪进行序列测定。用 GENSCAN 软件 (<http://genes.mit.edu/GENSCAN.html>) 预测所得 DNA 序列的外显子及其所编码的氨基酸序列, 并与 Bioedit 软件翻译所得的氨基酸序列进行比较校正。然后使用 BLAST 软件在基因库中 (<http://www.ncbi.nlm.nih.gov/>) 搜索所得 DNA 序列和氨基酸序列的同源序列, 鉴定其是否为 *LEAFY* 同源基因, 对其中相似性最高的两个种的 *LEAFY* 基因序列和氨基酸序列用 ClustalX<sup>[14]</sup> 软件进行进一步的比较分析。通过 PAUP<sup>[15]</sup> 软件对外显子 1、2 构建芥菜和十字花科中已知 *LEAFY* 同源基因序列的几个种的简约树。同时用 1 000 次重复抽样的自展值 (bootstrap) 分析检验分支可信度。所有性状按无序和非加权处理, 缺失和插入作为 “missing” 处理。

## 2 结果

### 2.1 芥菜 *LEAFY* 同源基因的分离与克隆

利用 *LEAFY-F* 和 *LEAFY-R* 引物对芥菜基因组 DNA 进行 PCR 扩增, 结果得到单一的特异电泳条带。经回收、纯化后的 PCR 产物与 pGEM-T Easy 载体连接, 连接产物转化并在含氨苄青霉素 (Amp)、X-gal 及 IPTG 的 LB 平板上涂布。挑取阳性菌落进行 PCR 鉴定, 从电泳图谱可以看出酶切所得片断与 PCR 产物和质粒大小完全吻合, 证明该克隆子确实插入了目的片段 (图 1)。图 2 显示了该基因的基因组序列与 GENSCAN 软件预测的氨基酸序列的分析结果, 表明芥菜中此基因长 2 866 bp, 有 3 个外显子, 2 个内含子, 其中外显子长度分别为 457 bp、407 bp 和 411 bp, 共同编码长度为 424 个氨基酸的蛋白质。

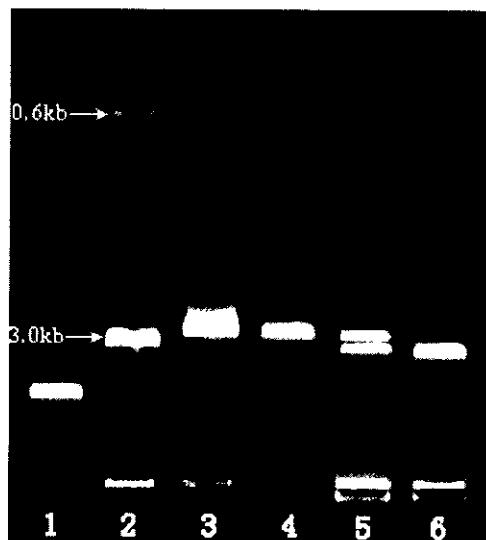


图 1 荠菜 LEAFY 同源基因的鉴定

Fig.1 Analysis of the *LEAFY* homologue clone of *Capsella bursa-pastoris*

1. 带有目的片段的重组 pGEM-T Easy 质粒 pGEM-T Easy plasmid including the target gene; 2. DNA 分子量标准物 SD015-2 SD015-2 DNA Marker; 3. 基因组 DNA 的 PCR 扩增产物 PCR products amplified from genomic DNA; 4. *CapLFY* 克隆的 PCR 产物 PCR products amplified from *CapLFY* clone; 5. EcoRI 酶切产物 The products of *CapLFY* clone digested by EcoRI; 6. pGEM-T Easy 质粒载体 pGEM-T Easy vector.

## 2.2 十字花科 *LEAFY* 同源基因序列的比较

BLAST 结果显示 3 个外显子所在 DNA 区段与十字花科其它已公布的 *LEAFY* 同源基因相应区段的同源性达 87% 以上, 其中与琴叶拟南芥 (*Arabidopsis lyrata*)<sup>[16]</sup> 的同源性高达 93%, 这说明所克隆序列确为拟南芥 *LEAFY* 基因和金鱼草 *FLO* 基因的同源基因, 暂命名为 *CapLFY*。GENSCAN 预测两侧的外显子分别为起始外显子和终止外显子, 山此推断我们得到的是 *CapLFY* 基因的全长序列。将 *CapLFY* 的氨基酸序列与其它已公布的 *LEAFY* 同源基因编码的氨基酸序列进行比对, 发现在所编码的氨基酸序列的 N 端和 C 端分别有长度为 122 和 159 个氨基酸残基的保守区域, 其中 C 端保守区高度保守, 由外显子 2 和 3 共同编码, N 端保守区的变异则相对较大, 由外显子 1 编码。*CapLFY* 所编码的氨基酸序列与琴叶拟南芥的同源性最高, 达 90% (图 3), 对二者的 DNA 序列进一步分析比较的结果表明: 外显子共有碱基替换 93 处, 其中有 74.7% 发生在三联体密码子的第三位。3 个外显子中的碱

基替换率没有显著差异, 同义替换 (ds) 与非同义替换 (dn) 的比值 ds/dn 分别为 1.9、2.3 和 6.0, 转换和颠换的比值分别为 0.71、2.08 和 2.11, 而在 2 个内含子中则分别为 0.91 和 0.66。在整个外显子区域中共有 4 处插入缺失突变, 且均为三碱基的联体密码子的同时缺失或插入。

## 2.3 *CapLFY* 基因在十字花科几个种中的演化关系分析

由于 *CapLFY* 基因外显子 3 高度保守, 导致建树信息过少, 故选用保守性相对较低的外显子 1、2 构建系统发育树。用 PAUP 软件分析十字花科 7 个种, 得到唯一一个简约性树 (图 4), 其一致性指数 (CI) 为 0.8506, 保留性指数 (RI) 为 0.5043。

从聚类树中可以看出, *Capsella bursa-pastoris* 与 *Leavenworthia crassa* 的关系较近, 其次为 *Ionopsisidium acaule*, 与 *Idahoa scapigera* 和 *Brassica oleracea* 姐妹枝最近。拟南芥属的两个种 *A. thaliana* 和 *A. lyrata* 之间由于信息量过少而与 *Capsella bursa-pastoris* 等构成的单系群形成多歧分枝。

## 3 讨论

*LEAFY* 同源基因广泛存在于植物界。Alexis 等<sup>[7]</sup> 对从低等植物苔藓到高等被子植物 7 个种的 *LEAFY* 同源基因的研究结果表明, 从苔藓到被子植物长达 4 亿年的进化过程中, 虽然 *LEAFY* 基因的功能有所变化, 但其蛋白质的 N 端和 C 端则存在两个非常保守的序列, 其中 C 端保守区构成结合 DNA 的功能域, 在这个区域甚至一个氨基酸残基的微小变化即能引起整个基因功能的变化, N 端保守区变异较大, 所构成的功能域能够调节 *LEAFY* 蛋白质结合 DNA 的能力。我们对十字花科 7 个种的 *LEAFY* 同源基因的比较研究同样表明该基因是非常保守的, 但十字花科这几个物种的 N 端保守区比 Alexis 等<sup>[7]</sup> 所发现的 N 端保守区多出 43 个氨基酸残基 (图 3), 可能这一段较为保守的序列在十字花科中执行着特定的功能。在 C 端这个极为保守的区域中, *CapLFY* 基因在 346 位 (共同序列编号) 是由非极性不带电荷的甘氨酸突变为极性带正电的碱性的精氨酸, 而该位置在其它已知 *LEAFY* 同源基因中, 从苔藓到被子植物均为甘氨酸<sup>[7]</sup>。荠菜存在 1 a 生和 2 a 生两种不同的生态型, 二者最主要的差

图 2 莴苣 *LEAFY* 同源基因核苷酸序列及其推导的氨基酸序列

Fig. 2 Nucleic acid sequence and deduced amino acid sequence of *Capsella bursa-pastoris* LEAFY homologue clone

图 3 不同植物 *LEAFY* 同源基因氨基酸序列比对

**Fig. 3** Sequence alignment of *LEAFY* like proteins

“\*”表示相同的氨基酸残基，空格表示有差异，- 表示氨基酸残基缺失，■部分表示十字花科特有的氨基酸保守区域

Asterisk, space and short dash indicate amino acid residues identical, different and absent, respectively, ■ indicates the 43 amino acids conserved in Cruciferaceae.

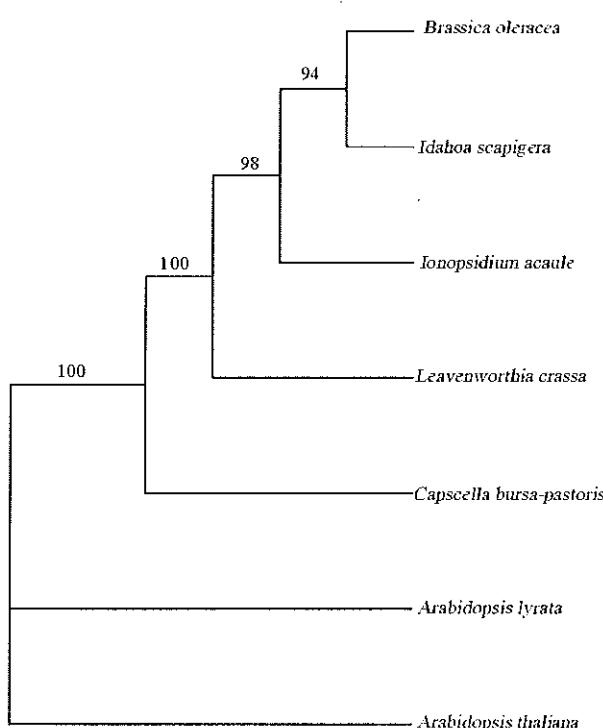


图 4 基于芥菜 *CapLFY* 基因外显子 1、2 建立的严格一致树  
Fig. 4 Strict consensus tree based on sequences of exon 1 and 2  
of *CapLFY* gene from *Capsella bursa-pastoris*

异就在于花期,该突变是否在芥菜花期变化中起作用,还有待于进一步的实验验证。

我们对 *CapLFY* 和 *A. lyrata LEAFY* 同源基因的进一步比较分析也得到了一些有意义的结果。许多研究都发现在碱基替换中转换多于颠换<sup>[17]</sup>,可能是由于三联体密码子的简并多为第三位碱基的 UC 简并或 AG 简并,所以自然选择淘汰了多数由于密码子第三位碱基的颠换所导致的异义替换,以尽量少影响蛋白质的结构和功能。而在本文的研究中,外显子 2 和 3 中转换比颠换多 2 倍,在外显子 1 中转换却少于颠换,这和 3 个外显子中同义替换与非同义替换的比值是一致的。这一结果可能和我们只比较一对基因而引起的取样误差有关,但也可能说明了这 3 个外显子经历着不同的进化过程,其中外显子 1 编码的 N 端保守区可以调节 *LEAFY* 蛋白质结合 DNA 的能力,变异较大,所承受的选择压力较小,而外显子 2 和 3 共同编码的 C 端保守区构成结合 DNA 的功能域,是 *LEAFY* 同源基因作为花分化过程中一个重要的激活因子所必须的功能域,承受着较大的选择压力,极为保守<sup>[7]</sup>,不同的选择压力可

能促成 3 个外显子的不同进化过程。

十字花科大部分物种的花着生于茎的伸长部分,这一部分的叶被潜在的花所抑制。十字花科系统发育证明长花序花比较原始,一些花结构在进化的过程中逐渐被修饰,最终形成莲座状花<sup>[18]</sup>。对几个模式种植物花序结构的研究已经表明 *LEAFY* 基因在植物花序进化中起关键作用<sup>[19]</sup>。Yoon 等<sup>[5]</sup>通过对莲座状花的进化研究,认为 *IscLFY1* 和 *LcrLFY* 基因是导致莲座状花进化的直接原因。Baum 等<sup>[6]</sup>进一步证明了基本功能氨基酸的替换为莲座状花序进化提供了原材料。从本研究的进化树中可以看出,具莲座状花的 *Leavenworthia crassa*、*Ionopsidium acaule* 和 *Idahoa scapigera* 亲缘关系较近。这在一定程度上支持了他们的观点。

**致谢** 感谢华南植物园李忠超先生和穆宏平女士在实验和论文形成中所给予的支持和帮助。

## 参考文献

- [1] Coen E S, Romero J M, Doyle S, et al. *FLORICAULA*: a homeotic gene required for flower development in *Antirrhinum majus* [J]. Cell, 1990, 63:1311–1322.
- [2] Weigel D, Alvarez J, Smyth D R, et al. *LEAFY* controls floral meristem identity in *Arabidopsis* [J]. Cell, 1992, 69:843–859.
- [3] Blázquez M A, Green R, Nilsson O, et al. Gibberellins promote flowering of *Arabidopsis* by activating the *LEAFY* promoter [J]. Plant Cell, 1998, 10:791–800.
- [4] Shu G P, Amaral W, Hileman L C, et al. *LEAFY* and the evolution of rosette flowering in violet cress (*Ionopsidium acaule*, Brassicaceae) [J]. Amer J Bot, 2000, 87:634–641.
- [5] Yoon H S, Baum B A. Transgenic study of parallelism in plant morphological evolution [J]. P Natl Acad Sci USA, 2004, 101: 6524–6529.
- [6] Baum B A, Yoon H S, Oldham, R L. Molecular evolution of the transcription factor *LEAFY* in Brassicaceae [J]. Mol Phylogene Evol, 2005, 37:1–14.
- [7] Alexis M, Maxinilian A B, Takako T, et al. The floral regulator *LEAFY* evolves by substitutions in the DNA binding domain [J]. Science, 2005, 308:260–263.
- [8] Kelly A J, Bonnlander M B, Meeks-Wagner D R. *NFL*, the tobacco homolog of *FLORICAULA* and *LEAFY*, is transcriptionally expressed in both vegetative and floral meristems [J]. Plant Cell, 1995, 7:225–234.
- [9] Anthony R G, James P E, Jordan B R. Cloning and sequence analysis of a *flo/lfy* homologue isolated from cauliflower (*Brassica oleracea* L. var. *botrytis*) [J]. Plant Mol Biol, 1993, 22:1163–1166.
- [10] Junko K, Saeko K, Keisuke N, et al. Down-regulation of *RFL*, the

- FLO/LFY* homolog of rice, accompanied with panicle branch initiation [J]. P Natl Acad Sci USA, 1998, 95:1979–1982.
- [11] Aidyn M, Tina G, Britt H, et al. *NEEDLY*, a *Pinus radiata* ortholog of *FLORICAULA/LEAFY* genes, expressed in both reproductive and vegetative meristems [J]. P Natl Acad Sci USA, 1998, 95:6537–6542.
- [12] Southerton S G, Strauss S H, Olive M R, et al. Eucalyptus has a functional equivalent of the *Arabidopsis* floral meristem identity gene *LEAFY* [J]. Plant Mol Biol, 1998, 37:897–910.
- [13] Doyle J. DNA protocols for plants-CTAB total DNA isolation [A]. In: Hewitt G M, Johnston A. Molecular Techniques in Taxonomy [M]. Berlin, 1991:283–293.
- [14] Thompson J D, Gibson T J, Plewniak F, et al. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. Nucl Acid Res, 1997, 24:4876–4882.
- [15] Swofford D L. PAUP\*. Phylogenetic Analysis Using Parsimony (\*and other methods). 2002, Version 10. Associates, Sunderland, MA.
- [16] Olsen K M, Womack A, Garrett A R, et al. Contrasting evolutionary forces in the *Arabidopsis thaliana* floral developmental pathway [J]. Genetics, 2002, 160:1641–1650.
- [17] Lei C Z (雷初朝), Chen H (陈宏), Yang G S (杨公社), et al. Study on mitochondrial DNA genetic diversity of some cattle breeds in China [J]. Acta Gene Sin (遗传学报), 2004, 31(1):57–62. (in Chinese)
- [18] Koch M, Haubold B, Mitchell-Olds T. Molecular systematics of the Brassicaceae: evidence from coding plastidic *matK* and nuclear *CHS* sequences [J]. Amer J Bot, 2001, 88:534–544.
- [19] Weigel D, Nilsson O. A developmental switch sufficient for flower initiation in diverse plants [J]. Nature, 1995, 377:495–500.