

DOI: 10.3724/SP.J.1005.2013.01117

## 植物 LTR-反转座子中 *Orf1* 基因的分子进化

刘静, 杜建厂

江苏省农业科学院经济作物研究所生物信息项目组, 农业部长江下游棉花和油菜重点实验室, 南京 210014

**摘要:** LTR-反转座子是植物基因组的主要组成部分。它们在结构上非常保守, 通常含有 *gag* 和 *pol* 两个基因, 是完成其转座过程所必需的。在前期研究中, 本项目组对大豆基因组 *SARE* 转座子家族进行了详细的分析。结果表明, 该家族的拷贝中还存在第 3 个基因——*Orf1*。文章借助生物信息学的研究方法, 对 33 个已测序的基因组进行了全基因组注释。结果发现, 在 7 个植物基因组(桉树、杨树、棉花、大豆、百脉根、亚麻和苜蓿)中, 部分 LTR-反转座子元件在 *gag* 基因的上游存在约 1~2 kb 未知的 *Orf1* 基因或基因片段。这类转座子多数在 0~3 百万年内插入到其所在的寄主基因组中, 但它们在不同物种中的分子结构、发生的频率、扩增的强度和活跃的时期等方面差异较大。系统进化树分析表明, 这类具有特殊结构的转座子较整齐的聚类到双子叶植物的一个进化分支上, 表明它们可能是部分双子叶植物在进化过程中所产生的。不同物种间的相对保守性、大量拷贝的转录活性以及可能存在的多个功能结构域, 提示 *Orf1* 基因可能具有一定的生物学功能。

**关键词:** *Orf1* 基因; LTR-反转座子; 植物基因组; 起源; 进化

## Molecular evolution of *Orf1* gene in plant LTR-retrotransposons

LIU Jing, DU Jian-Chang

Bioinformatics Group, Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture, Nanjing 210014, China

**Abstract:** LTR-Retrotransposons are the major DNA components in plant genomes. They usually contain *gag* and *pol*, two genes necessary for transpositional process. Our previous study on soybean genome annotation identified a *SARE* LTR-Retrotransposon family, which carries the third gene, *Orf1*. Using a bioinformatics approach, we here reported that 7 out of 33 sequenced genomes have some LTR-Retrotransposons with an extra *Orf1* gene/gene fragment (~1–2 kb) in the region between 5' LTR and *gag* gene, including *Eucalyptus grandis*, *Populus trichocarpa*, *Gossypium raimondii*, *Glycine max*, *Lotus japonica*, *Linum usitatissimum*, and *Medicago truncatula*. The majority of these elements were inserted into the genomes they reside within the last 3 million years, but their structures, frequencies, intensity, and activity in different host genomes are quite different. Phylogenetic analysis indicated that these unusual elements were clustered in a eudicot branch, suggesting that they may be generated in the evolution of some eudicot species. The relative conservation, transcriptional

收稿日期: 2013-04-22; 修回日期: 2013-05-22

基金项目: 国家人社部留学回国人员启动基金项目(编号: 人社厅函[2012]258号), 江苏省六大人才高峰项目(编号: 2011NY030), 中国农科院棉花所棉花生物学国家重点实验室开放课题项目(编号: CB2013B06)和江苏省农业科技自主创新基金项目(编号: CX(12)5036)资助

作者简介: 刘静, 硕士, 实习研究员, 研究方向: 生物信息学。Tel: 025-84392767; E-mail: liujing200809@163.com

通讯作者: 杜建厂, 博士, 研究员, 研究方向: 进化基因组学。E-mail: dujianchang@hotmail.com

网络出版时间: 2013-6-13 17:55:11

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20130613.1755.001.html>

activity, and the presence of multiple potential conserved motifs suggest that *Orf1* gene may still be functional.

**Keywords:** *Orf1* gene; LTR-retrotransposons; plant genomes; origin; evolution

转座元件或转座子是植物基因组的主要组成部分<sup>[1-5]</sup>。根据转座方式的不同,植物中的转座子可分为反转座子和 DNA 转座子<sup>[6,7]</sup>。前者以 RNA 为媒介,在逆转录酶的作用下,把 RNA 分子逆转录成双链 DNA 分子,插入到基因组中的其他位置<sup>[8]</sup>。反转座子在保留原来拷贝的同时,形成一个新的拷贝。因此,反转座子的复制通常被称为复制-粘贴模式<sup>[9]</sup>。而 DNA 转座子的转座过程是直接从 DNA 到 DNA,没有经过 RNA 分子的逆转录过程。因此,其复制过程通常被称为剪切-复制模式<sup>[9]</sup>。

根据结构的不同,反转座子又可以分为 LTR-反转座子和非 LTR-反转座子<sup>[10]</sup>。后者如 LINE 和 SINE,在人类和哺乳动物基因组中较为常见<sup>[9]</sup>。而在植物基因组中,反转座子通常以 LTR-反转座子的形式存在<sup>[9,11]</sup>。LTR-反转座子具有典型的结构特征,包括 2 个末端正向重复序列(含有转录所必需的启动子及转录起始位点)、GAG 胞衣蛋白(完成逆转录过程的场所)、POL 多聚蛋白(包含转录过程所必需的逆转录酶和整合酶等)。LTR-反转座子通常以‘TG’开头,‘CA’结尾<sup>[11]</sup>。并且在其寄主基因组中的插入位点附近形成 4~6 bp 的正向重复序列(Target site duplications, TSDs)<sup>[9,11]</sup>。此外,LTR-反转座子还含有与转录起始和终止有关的 tRNA 结合位点(Primer binding site, PBS)和多聚嘌呤序列(polypurine tract, PPT)。

随着基因组数据的日益增多,人们逐渐发现,在有些 LTR-反转座子中,除 *gag* 和 *pol* 基因外,在其上游和/或下游往往存在其他类型的编码蛋白<sup>[12,13]</sup>。如大豆基因组中最大的 LTR-反转座子家族 SARE 含有 4 个编码蛋白的基因,包括 *Orf1*、*gag*、*pol* 和 *envelope-like* 基因。这使得 SARE 家族的转座子长度 >20 kb,其 DNA 含量约占大豆整个基因组的 10%<sup>[12]</sup>。由于 *Orf1* 基因不是 LTR-反转座子所必需的,因此,研究其起源及潜在的功能对于研究植物基因组中 LTR-反转座子的进化具有重要意义。

本研究通过生物信息学方法,以大豆基因组中

SARE 家族的 *Orf1* 基因编码的氨基酸序列和豌豆中 *Ogre* 家族 *Orf1* 的氨基酸序列(序列号 AAQ82034.1)为搜索对象,对 33 个已测序的基因组进行同源搜索,并对含有 *Orf1* 基因的 LTR-反转座子(简称 *Orf1* LTR-反转座子)进行结构鉴定、频率统计、插入时间分析及系统进化树分析等。这些研究为进一步探讨 *Orf1* 基因的进化起源和潜在的功能提供了数据支持和研究基础。

## 1 材料和方法

### 1.1 数据来源

本研究涉及 33 个已测序的基因组数据。其中,木薯(*Manihot esculenta*)、蓖麻(*Ricinus communis*)、亚麻(*Linum usitatissimum*)、杨树(*Populus trichocarpa*)、苜蓿(*Medicago truncatula*)、菜豆(*Phaseolus vulgaris*)、大豆(*Glycine max*)、桃(*Prunus persica*)、拟南芥(*Arabidopsis thaliana*)、琴叶拟南芥(*Arabidopsis lyrata*)、芥菜(*Capsella rubella*)、大白菜(*Brassica rapa*)、小盐芥(*Thellungiella halophila*)、番木瓜(*Carica papaya*)、甜橙(*Citrus sinensis*)、柚子(*Citrus clementina*)、桉树(*Eucalyptus grandis*)、葡萄(*Vitis vinifera*)、猴面花(*Mimulus guttatus*)、蓝花耧斗菜(*Aquilegia coerulea*)、高粱(*Sorghum bicolor*)、玉米(*Zea mays*)、谷子(*Setaria italica*)、水稻(*Oryza sativa*)、短坪草(*Brachypodium distachyon*)、棉花(*Gossypium raimondii*)、卷柏(*Selaginella moellendorffii*)、小立碗藓(*Physcomitrella patens*)、莱茵衣藻(*Chlamydomonas reinhardtii*)、团藻(*Volvox carteri*)等 30 个基因组数据来源于 Phytozome 网站(<http://www.phytozome.net>)。百脉根(*Lotus japonica*)的基因组数据来自于 <http://www.plantgdb.org/LjGDB> 网站,番茄(*Solanum lycopersicum*)的基因组数据从 [http://solgenomics.net/organism/Solanum\\_lycopersicum/genome](http://solgenomics.net/organism/Solanum_lycopersicum/genome) 网站上获得,梨(*Pyrus bretschneideri*)的基因组数据来自于 Wu 等<sup>[14]</sup>。

## 1.2 方法

### 1.2.1 LTR-反转座子的鉴定

本研究采用基于结构特征从头寻找的 LTR\_STRUC 程序<sup>[15]</sup>和基于同源搜索的 CROSS\_MATCH 程序 (<http://www.phrap.org/phrap.docs/general.html>) 进行 LTR-反转座子的鉴定。LTR\_STRUC 程序采用默认参数设置。通过 Perl 脚本语言的编程, 对程序给出的结果进行信息提取、展示和校对。

### 1.2.2 *Orf1* LTR-反转座子的鉴定

以大豆基因组中 *SARE* 家族 *Orf1* 基因编码的蛋白序列和 *Ogre* 家族编码的蛋白序列为搜索对象, 对获得的 33 个基因组的 LTR-反转座子数据库进行 Tblastn 同源搜索。搜索的参数设定为  $E\text{-value} < 10^{-6}$ 。同时, 对 *Orf1* LTR-反转座子在不同寄主基因组的个数和发生频率进行统计分析。

### 1.2.3 *Orf1* LTR-反转座子插入时间的计算

将来自于同一个 LTR-反转座子的 2 个 LTR 进行序列提取, 用 MUSCLE<sup>[16]</sup>程序对其进行序列比对。所有运算在本实验室建立的 Linux 系统 IBM3500 服务器上运行。最后, 将得到的序列差异度转换为进化时间(百万年)。突变速率按  $1.3 \times 10^{-8}$ /碱基/年计算<sup>[2]</sup>。

### 1.2.4 序列比对与系统进化树分析

序列比对采用 MUSCLE 软件<sup>[16]</sup>, 在 EMBL-EBI 服务网站上进行 (<http://www.ebi.ac.uk/Tools/msa/muscle>)。系统进化树用 MEGA4<sup>[17]</sup>进行分析。采用 Neighbor-Joining 距离, Jukes-Cantor 模型进行系统进化树构建。重复抽样(Bootstrap)10 000 次。

### 1.2.5 转录组与小 RNA 分析

转录组 EST 数据来自于 NCBI 公共服务网站

(<http://www.ncbi.nlm.gov/nucest>)。用 CROSS\_MATCH 程序搜索 *Orf1* 基因与 EST 数据的匹配。参数设定为匹配长度 > 100 bp, 相似度 > 99%。小 RNA 数据库下载于 <http://www.mirbase.org/>和 <http://bioinformatics.cau.edu.cn/PMRD/>。小 RNA 对 *Orf1* 基因的正义链和反向互补链同时进行搜索。搜索参数设定为相似度 = 100%。

## 2 结果与分析

### 2.1 *Orf1* LTR-反转座子的物种、频率及结构特征

以大豆 *SARE* 家族编码的 ORF1 蛋白序列和豌豆中报道的 *Ogre* 反转座子家族编码的 ORF1 蛋白序列(AAQ82034.1)为探针, 对 33 个已测序的植物基因组和低等生物基因组进行本地 NCBI BLAST 同源搜索。结果显示, 包括大豆在内, 共有 7 个物种的基因组(亚麻、杨树、苜蓿、大豆、棉花、桉树和百脉根)含有 *Orf1* 基因(表 1)。其他 26 个基因组与 *Orf1* 基因没有明显的同源性。

运用基于结构特征开发的 LTR\_STRUC 程序和 NCBI 同源搜索的方法, 对含有 *Orf1* 基因的植物基因组进行 LTR-反转座子的全基因组鉴定与注释。结果表明, 尽管 7 个基因组中都含有 *Orf1* LTR-反转座子, 但它们在各自基因组中所占的比例差异很大。例如, 在苜蓿基因组中鉴定出的 845 个 LTR-反转座子中, 有 154 个转座元件含有 *Orf1* 基因, 占总拷贝数的 18.2%(表 1)。而在百脉根基因组中, 仅有 1 个 LTR-反转座子含有 *Orf1* 基因, 仅占总拷贝数的 0.2%(表 1)。*Orf1* LTR-反转座子在其他 5 个基因组中占的比例分别为 0.4%(亚麻)、0.8%(杨树)、1.2%(桉树)、5.8%(棉花)和 7.9%(大豆)(表 1)。除苜蓿外, 其他 6 个基因组中 *Orf1* LTR-反转座子所占的比例与其基因组大小成正相关( $r = 0.89$ ), 提示这类特殊结构的

表 1 植物基因组中含有 *Orf1* 基因的 LTR-反转座子的物种及其发生频率

物种名称	含 <i>Orf1</i> 的拷贝数	总拷贝数	百分比(%)	基因组大小(Mb)
亚麻( <i>Linum usitatissimum</i> )	1	289	0.4	318
杨树( <i>Populus trichocarpa</i> )	6	762	0.8	370
苜蓿( <i>Medicago truncatula</i> )	154	845	18.2	241
大豆( <i>Glycine max</i> )	2550	32453	7.9	975
棉花( <i>Gossypium raimondii</i> )	156	2668	5.8	761
桉树( <i>Eucalyptus grandis</i> )	36	2976	1.2	691
百脉根( <i>Lotus japonicus</i> )	1	585	0.2	472

转座子在相对复杂的基因组中更容易获得扩增(表 1)。

为比较 *Orf1* LTR-反转座子的结构特征, 本研究对其 LTR 序列及编码蛋白的基因进行了进一步注释。如图 1 所示, 尽管多个基因组中均包含 *Orf1* 基因, 但是, 能检测到的该基因编码的氨基酸的长度差异很大, 表明该基因的祖先拷贝分化后在不同物种中存在较大的遗传差异。除大豆外, 在其他基因组中均没有检测到类病毒(*envelope-like*)基因的存在, 提示这类 *envelope-like* 的基因可能只存在于豆类或其亲缘关系较近的物种中(图 1)。此外, 不同物种中 *Orf1* LTR-反转座子的 LTR 长度和其转座元件的全长也存在一定差异(图 1)。这些结果表明, *Orf1* LTR-反转座子在进化过程中产生了大量的结构变异。

## 2.2 *Orf1* LTR-反转座子在不同物种基因组中的插入时间、扩增规律和转录活性

尽管绝大多数完整的 LTR-反转座子在 <5 百万年插入到其寄主基因组中<sup>[21]</sup>, 但是它们在进化上是非常古老的, 至少可以追溯到单子叶和双子叶植物产生分化之前<sup>[18]</sup>, 从而使植物基因组的进化处于一个长期动态变化的过程。为了解 *Orf1* LTR-反转座子

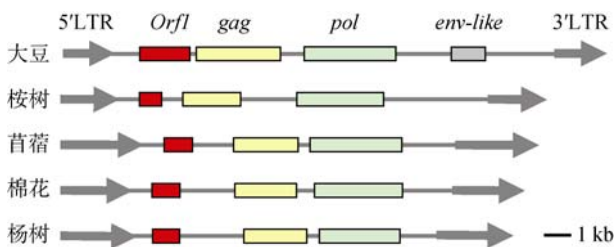


图 1 植物基因组中 *Orf1* LTR-反转座子的特征结构

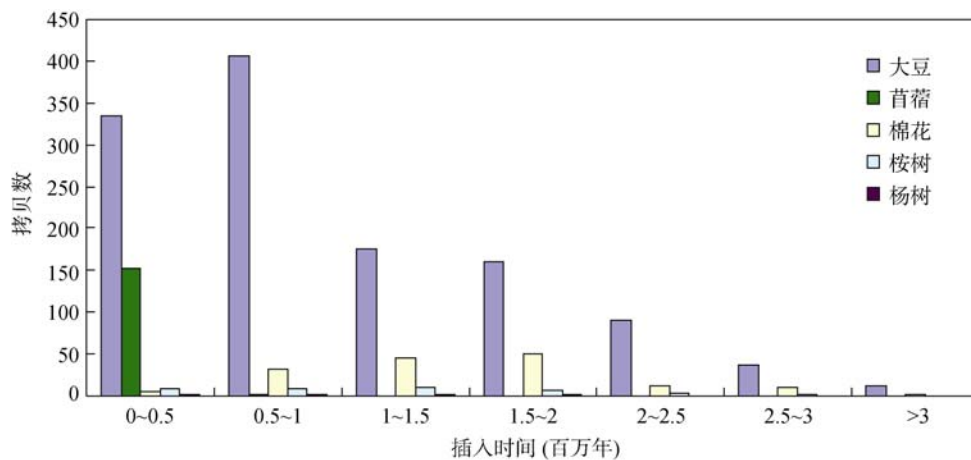


图 2 植物基因组 *Orf1* LTR-反转座子的插入时间

在不同寄主基因组中的插入时间和转座活性, 本研究对 5 个主要的基因组进行了进一步的注释和分析。如图 2 所示, 大豆中 *Orf1* LTR-反转座子在 0~3 百万年时间内均有活性, 但在 0~1 百万年内活性最高。而苜蓿中绝大多数的 *Orf1* LTR-反转座子都是在 <0.5 百万年内插入到其寄主基因组的。与大豆和苜蓿不同, 棉花中 *Orf1* LTR-反转座子在 1.5~2.0 百万年内活性最高。桉树和杨树中 *Orf1* LTR-反转座子的拷贝数较低, 扩增规律不明显(图 2)。这些结果表明, 尽管不同基因组中都含有 *Orf1* LTR-反转座子, 但是它们的扩增频率和转座最活跃的时期差异很大, 这可能与寄主基因组的进化历史和基因组特性有关。

为进一步探讨 *Orf1* LTR-反转座子的转录状况及受小 RNA 调控的情况, 本研究对 *Orf1* 基因进行了转录活性及受小 RNA 调控的生物信息学分析(见材料与方法)。结果显示, 在杨树、苜蓿和大豆基因组中, 分别有 1 个、6 个和 145 个 *Orf1* LTR-反转座子的 *Orf1* 基因与相应的 EST 序列相匹配(相似度 >99%)。提示这些拷贝现在可能仍然具有转录活性。*Orf1* 基因和小 RNA 的比对结果表明, 在苜蓿基因组中, 至少有 4 个 *Orf1* LTR-反转座子的 *Orf1* 基因与 1 个小 RNA 相匹配(mtr-miR5218:UGAGACUUGGU-AGUAAGAUGAU)(相似度=100%), 表明它们的转录过程可能受到该小 RNA 的分子调控。

## 2.3 *Orf1* LTR-反转座子的进化起源及在系统进化树中的位置

由于 *Orf1* 基因不是 LTR-反转座子所必需的, 因此研究它的进化起源是非常有意义的。为了明确

*Orf1* LTR-反转座子在进化过程中所处的位置, 我们以拟南芥 LTR-反转座子中反转座酶最保守的功能结构域(motif)为探针, 对 33 个基因组的 LTR-反转座

子数据库进行同源搜索, 重建了这些物种的系统发生树。如图 3 所示, 该系统发生树大致反映了 33 个物种的进化关系(图 3)。代表进化早期的团藻、莱茵

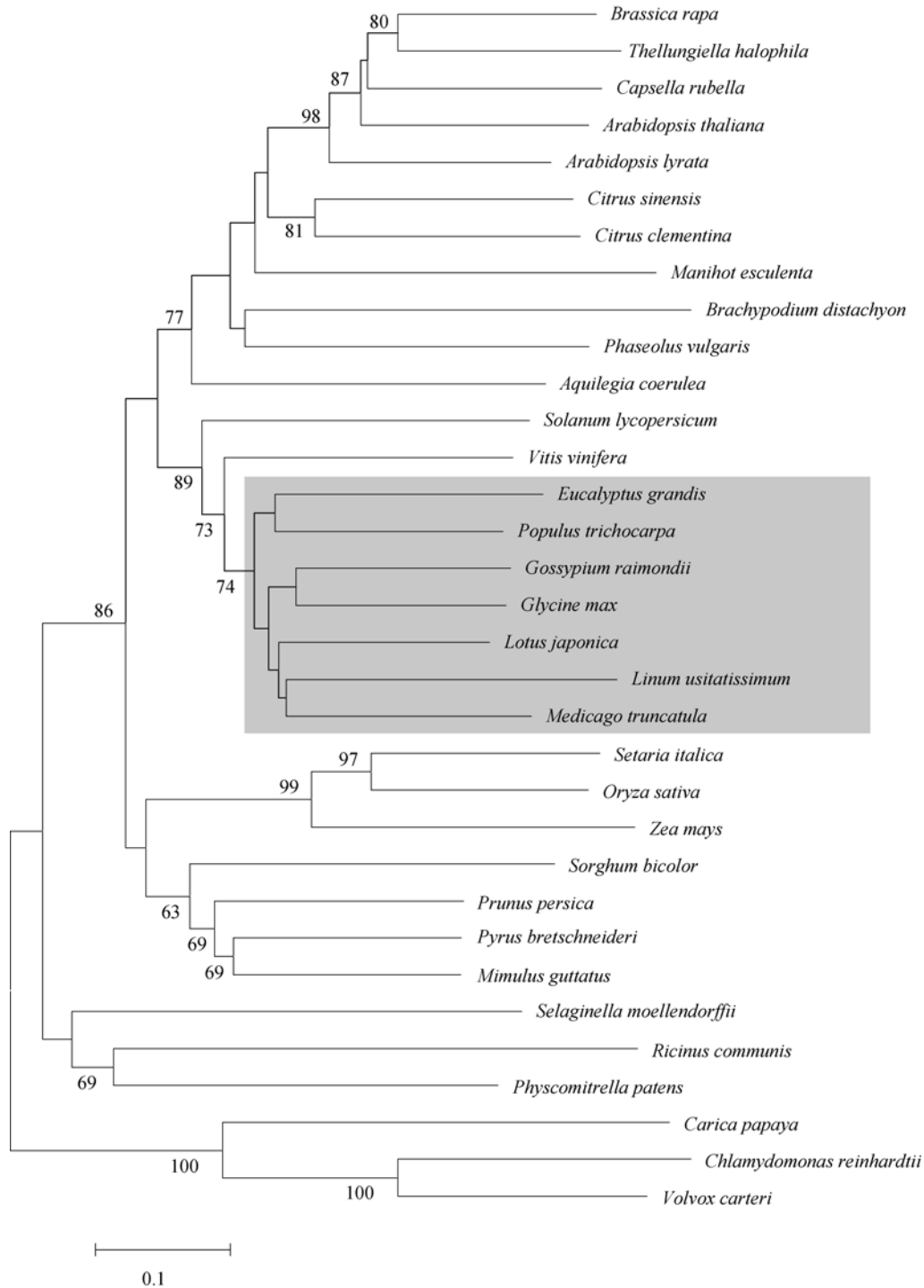


图 3 33 个已测序物种的系统进化树

使用 LTR-反转座子的反转座酶的保守区段进行系统进化树分析。采用 NJ 方法, Jukes-Cantor 模型, Bootstrap 重复 10 000 次。阴影标示的是含 *Orf1* 基因的物种。

衣藻处于系统进化树的最外侧，表明它们进化最为古老。以水稻(*Oryza sativa*)为代表的单子叶植物和以拟南芥(*Arabidopsis thaliana*)为代表的双子叶植物基本分成两个大的分支。值得注意的是，*Orf1* LTR-反转座子所在的 7 个物种(桉树、杨树、棉花、大豆、百脉根、亚麻和苜蓿)整齐的聚类在一起，形成一个较为独立的分支(图 3 中阴影部分所示)。而与它们亲缘关系最近的物种番茄和葡萄则没有被检测到 *Orf1* 基因的存在。这表明，LTR-反转座子中的 *Orf1* 基因可能是部分双子叶植物所特有的，并在其进化过程中形成的。具体的形成机制还不太清楚，有待于进一步研究。

### 2.4 LTR-反转座子中 *Orf1* 基因的保守性及可能的生物学功能

如前所述，由于 *Orf1* 基因很可能是 LTR-反转座子在进化过程中产生的，所以它是否具有功能，以及具有何种功能是值得关注和探讨的问题之一。本研究以大豆 *SARE* LTR-反转座子中 *Orf1* 基因编码的氨基酸序列为探针，对 NCBI 中非冗余的蛋白数据库进行同源搜索。结果显示，在给出的 85 条注释结果中( $E\text{-value} < 10^{-6}$ )，绝大多数都给出了“未知蛋白”或“假定的蛋白”。表明该基因可能还未被准确注释。其可能的生物学功能还不清楚。

LTR-反转座子中包含与转座过程相关的基因，如胞衣蛋白基因 *gag* 和逆转录酶基因 *rt*。这些基因

中往往包含 1 个或多个功能结构域<sup>[18,19]</sup>。由于这些功能结构域在不同物种中非常保守，因此可以用作探针进行系统发生树的重建。为了探讨 *Orf1* 基因是否有功能，我们对来自于不同物种的 *Orf1* 基因编码的氨基酸序列进行了多序列比对。如图 4 所示，来自不同物种比对的约 100 个氨基酸中，至少可以检测到 2 个可能的保守结构域(图 4，如黑框内所示)，提示 *Orf1* 基因可能或者至少在进化过程的某个阶段具有一定的生物学功能。

### 3 讨论

众所周知，*gag* 和 *pol* 是 LTR-反转座子在转座过程中所必需的基因<sup>[9]</sup>。因此，*Orf1* 基因的发现使得 LTR-反转座子的长度增加，结构更为复杂。在之前的研究中，有关 *Orf1* 基因仅有零星的报道，如豌豆中 *Ogre* 家族<sup>[13]</sup>和大豆中 *SARE* 家族<sup>[12]</sup>。本研究首次系统的对 33 个测序的植物基因组和低等生物基因组进行了全基因组分析，发现 7 个植物基因组的 LTR-反转座子含有 *Orf1* 基因或其基因片段。它们在系统进化树上形成一个相对独立的分支，提示可能来自于同一个祖先。并且发现该基因所处的位置相对保守，即通常位于 *gag* 基因的上游和 5'LTR 的下游。对不同物种 *Orf1* LTR-反转座子插入时间的计算表明，尽管它们可能来自于同一个祖先(图 3)，但它们在分化后在不同寄主基因组中发生的频率、扩增的强度，以及转座最活跃的时期等方面均存在着明

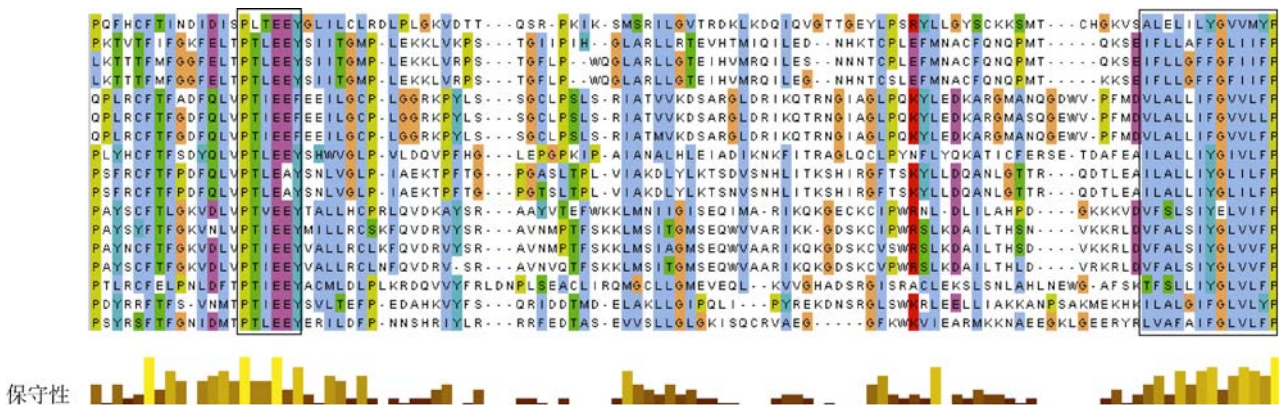


图 4 不同物种 LTR-反转座子中 *Orf1* 基因的保守性

黑框内表示保守性很强，提示可能为 2 个功能结构域。Lu：亚麻；Eg：桉树；Gm：大豆；Mt：苜蓿；Gr：棉花；Lj：百脉根；Pt：杨树。从上到下分别为 Lu\_1、Eg\_1、Eg\_2、Eg\_3、Gm\_1、Gm\_2、Gm\_3、Mt\_1、Mt\_2、Mt\_3、Gr\_1、Gr\_2、Gr\_3、Gr\_4、Lj\_1、Pt\_1 和 Pt\_2。

显差异。这表明 *Orf1* LTR-反转座子家族的扩增可能与其寄主基因组的特性和进化历史有关。

目前, 有关 LTR-反转座子中 *Orf1* 基因的功能还不清楚。但是, 已有的证据暗示, 该基因可能是有功能的。这主要是因为: (1)大豆 *SARE* 家族、豌豆 *Ogre* 家族和苜蓿的很多拷贝都是在近期(<0.5 百万年)扩增, 并且具有转录活性; (2)不同物种 LTR-反转座子的 *Orf1* 基因具有一定的同源性, 并且包含至少 2 个类似功能结构域的区域; (3)*Orf1* 基因在 LTR-反转座子中所处的位置相对固定, 并且多个物种比较时有一定的序列相似性。但是由于现有蛋白数据库中还没有对该基因进行注释, 也没有发现与它具有同源性的相似基因。因此, *Orf1* 基因的功能还不得而知。由于 *Orf1* 基因是在双子叶植物的一个分支在进化过程中产生的, 因此, 它可能是一种新产生的未知的新基因。该基因是 LTR-反转座子中所特有的, 表明它可能参与了 LTR-反转座子的转录、扩增过程, 其具体的生物学功能还有待于进一步研究。

#### 参考文献(References):

- [1] Du JC, Grant D, Tian ZX, Nelson RT, Zhu LC, Shoemaker RC, Ma JX. SoyTEdb: a comprehensive database of transposable elements in the soybean genome. *BMC Genomics*, 2010, 11(1): 113. [DOI](#)
- [2] Ma J, Devos KM, Bennetzen JL. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res*, 2004, 14(5): 860–869. [DOI](#)
- [3] Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddelloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK. The B73 maize genome: complexity, diversity, and dynamics. *Science*, 2009, 326(5956): 1112–1115. [DOI](#)
- [4] Wang KB, Wang ZW, Li FG, Ye WW, Wang JY, Song GL, Yue Z, Cong L, Shang HH, Zhu SL, Zou CS, Li Q, Yuan YL, Lu CR, Wei HL, Gou CY, Zheng ZQ, Yin Y, Zhang XY, Liu K, Wang B, Song C, Shi N, Kohel RJ, Percy RG, Yu JZ, Zhu YX, Wang J, Yu SX. The draft genome of a diploid cotton *Gossypium raimondii*. *Nature Genet*, 2012, 44(10): 1098–1103. [DOI](#)
- [5] Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang HB, Wang XY, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otiillar RP, Penning BW, Salamov AA, Wang Y, Zhang LF, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 2009, 457(7229): 551–556. [DOI](#)
- [6] 程旭东, 凌宏清. 植物基因组中的非 LTR 反转录转座子 SINEs 和 LINEs. *遗传*, 2006, 28(6): 731–736. [DOI](#)
- [7] 侯小改, 张曦, 郭大龙. 植物 LTR 类反转录转座子序列分析识别方法. *遗传*, 2012, 34(11): 1491–1500. [DOI](#)
- [8] Havecker ER, Gao X, Voytas DF. The diversity of LTR retrotransposons. *Genome Biol*, 2004, 5(6): 225–230. [DOI](#)
- [9] Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH. A unified classification system for eukaryotic transposable elements. *Nat Rev*

- Genet*, 2007, 8(12): 973–982. [DOI](#)
- [10] 陈志伟, 吴为人. 植物中的反转录转座子及其应用. *遗传*, 2004, 26(1): 122–126. [DOI](#)
- [11] Kumar A, Bennetzen JL. Plant retrotransposons. *Annu Rev Genet*, 1999, 33: 479–532. [DOI](#)
- [12] Du JC, Tian ZX, Bowen NJ, Schmutz J, Shoemaker RC, Ma JX. Bifurcation and enhancement of autonomous-nonautonomous retrotransposon partnership through LTR Swapping in soybean. *Plant Cell*, 2010, 22(1): 48–61. [DOI](#)
- [13] Neumann P, Požárková D, Macas J. Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. *Plant Mol Biol*, 2003, 53(3): 399–410. [DOI](#)
- [14] Wu J, Wang ZW, Shi ZB, Zhang S, Ming R, Zhu SL, Khan MA, Tao ST, Korban SS, Wang H, Chen NJ, Nishio T, Xu X, Cong L, Qi KJ, Huang XS, Wang YT, Zhao X, Wu JY, Deng C, Gou CY, Zhou WL, Yin HT, Qin GH, Sha YH, Tao Y, Chen H, Yang YA, Song Y, Zhan DL, Wang J, Li LT, Dai MS, Gu C, Wang YZ, Shi DH, Wang XW, Zhang HP, Zeng L, Zheng DM, Wang CL, Chen MS, Wang GB, Xie L, Sovero V, Sha SF, Huang WJ, Zhang SJ, Zhang MY, Sun JM, Xu LL, Li Y, Liu X, Li QS, Shen JH, Wang JY, Paull RE, Bennetzen JL, Wang J, Zhang SL. The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res*, 2013, 23(2): 396–408. [DOI](#)
- [15] McCarthy EM, McDonald JF. LTR\_STRUC: a novel search and identification program for LTR retrotransposons. *Bioinformatics*, 2003, 19(3): 362–367. [DOI](#)
- [16] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 2004, 32(5): 1792–1797. [DOI](#)
- [17] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol*, 2007, 24(8): 1596–1599. [DOI](#)
- [18] Du JC, Tian ZX, Hans CS, Laten HM, Cannon SB, Jackson SA, Shoemaker RC, Ma JX. Evolutionary conservation, diversity and specificity of LTR-retrotransposons in flowering plants: insights from genome-wide analysis and multi-specific comparison. *Plant J*, 2010, 63(4): 584–598. [DOI](#)
- [19] Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J*, 1990, 9(10): 3353–3362. [DOI](#)

## •综合信息•

### 欢迎订阅 2014 年《中国生态农业学报》

《中国生态农业学报》由中国科学院遗传与发育生物学研究所和中国生态经济学会主办, 中国科学院主管, 科学出版社出版。系中国期刊方阵双效期刊、中国科技精品期刊、中文核心期刊、RCCSE 中国权威学术期刊, 为中国学术期刊综合评价数据库、中国期刊全文数据库、中国学术期刊文摘、中国科学引文数据库、中国科技论文与引文数据库、CNKI 中国期刊全文数据库源刊, 并被国际农业生物学文摘(CABI)、美国化学文摘(CA)、哥白尼索引(IC)、美国乌利希国际期刊指南等国际数据库及检索单位收录。荣获第三届、四届全国农业优秀期刊一等奖和首届北方优秀期刊奖, 被评为 2009 年中国北方优秀期刊, 连续多届获得河北省优秀期刊奖。

《中国生态农业学报》主要报道全球环境变化与农业、农业生态系统与生态农业理论基础、农田生态系统与农业资源、生态农业模式和技术体系、农业生态经济学、农业环境质量及环境保护、农业有害生物的综合防治等领域创新性研究成果。适于从事农业生态学、生态学、生态经济学以及环境保护等领域科技人员、高等院校有关专业师生, 农业及环境管理工作者和基层从事生态农业建设的技术人员阅读与投稿。

《中国生态农业学报》国内外公开发行, 国内刊号 CN13-1315/S, 国际刊号 ISSN1671-3990。月刊, 国际标准大 16 开本, 128 页, 每期定价 35 元, 全年 420 元。邮发代号: 82-973, 全国各地邮局均可订阅。漏订者可直接汇款至编辑部补订(另加邮资 50.00 元)

地址: (050022) 河北省石家庄市槐中路 286 号 中科院遗传发育所农业资源中心 《中国生态农业学报》编辑部

电话: (0311) 85818007 传真: (0311) 85815093

网址: <http://www.ecoagri.ac.cn> E-mail: [editor@sjziam.ac.cn](mailto:editor@sjziam.ac.cn)