

水稻两个 *frr* 基因的结构、转录特性及同源性分析

胡欣¹ 胡昊¹ 洪国藩^{1,2} 韩斌^{1*}

(¹中国科学院上海生命科学研究院国家基因研究中心, 上海200233; ²中国科学院上海生命科学研究院生物化学与细胞生物学研究所, 上海200031)

摘要: 对水稻中两个核糖体再循环因子同源基因 *OsfrA* 和 *OsfrB* 进行了鉴定与分析。这两个单拷贝基因分别位于水稻的4号和7号染色体上,且在细胞器基因组中未发现同源序列。它们在不同组织中的转录特性及其蛋白质产物的N端特征提示其翻译产物会被各自转运并分别定位于线粒体和叶绿体中,而序列上的保守性及构成性的表达则表明它们在植物生长中扮演着重要的角色。它们与其它原核及真核RRF之间在基因结构及编码序列上的异同为内共生学说提供了新的证据,也揭示了RRF在分子进化研究方面所具有的潜在价值。

关键词: 核糖体再循环因子;水稻;终止后复合体;半自主细胞器
中图分类号: Q74

核糖体再循环因子 (ribosome recycling factor, RRF) 最早在大肠杆菌 (*Escherichia coli*) 中被发现 (Hirashima 和 Kaji 1970) 和克隆 (Ichikawa 和 Kaji 1989), 一度也被称为核糖体释放因子 (ribosome releasing factor) 和释放因子4 (releasing factor 4, RF4)。近期的研究发现, RRF 在原核生物的蛋白质合成体系中起着不可或缺的作用, 这主要包括以下两个方面。

首先, 体内实验业已证明, RRF 负责将终止后复合体 (post termination complex) 解聚成 mRNA、tRNA 和核糖体 (Janosi 等 1998)。在体外实验中, 这一过程需要 GTP 以及 EF-G 和 RF3 两个蛋白质合成相关因子中的一个参与 (Hirashima 和 Kaji 1973)。而当核糖体的浓度较低时, RRF 能够推动核糖体从一轮到另一轮翻译反应的循环, 从而刺激不同蛋白质的体外合成, 使其效率提高 3-8 倍 (Janosi 等 1996, Ryoji 等 1981)。

其次, RRF 能够避免翻译错误的发生。缺少 RRF 会导致核糖体越过终止密码子开始新一轮翻译反应, 这种计划外的翻译随机选择阅读框, 不

依赖于任何的起始信号 (Ryoji 等 1981)。RRF 对细菌的生长是必要的, 一旦缺少, 会使细菌死亡或失去繁殖能力 (Janosi 等 1994, 1998)。因此, 除古细菌外, 目前所有基因组已被测序的原核生物中都有编码 RRF 的基因 (*frr*) 或其同源基因也就不足为奇了。

RRF 作用的分子机制仍不十分清楚, 只知道它可能与核糖体的 A 位点结合 (Janosi 等 1996, Pavlov 等 1997)。而对 *Thermotoga maritima* RRF 晶体结构的研究发现, 其空间结构几近完美的模拟了 tRNA 的构象 (tRNA mimic), 这样的构象模拟还从未发现过, 这可能提示了 RRF 与翻译后复合体的作用方式 (Selmer 等 1999)。

目前, 在酿酒酵母 (*Saccharomyces cerevisiae*) (Kanai 等 1998) 疟原虫 (*Plasmodium yoelii yoelii*) 秀丽线虫 (*Caenorhabditis elegans*) 拟南芥 (*Arabidopsis thaliana*) 和人 (*Homo sapiens*) (Zhang 和 Spremulli 1998) 的基因组中也发现了 *frr* 的同源基因或序列。对菠菜 (*Spinacia oleracea*) 叶绿体 RRF (RRFHCP) 的研究结果表明, 真核的 RRF 虽然由核基因编码合成, 但却会被转运到细胞器中发挥功能 (Rolland 等 1999)。不过迄今为止, 尚未有对真核生物中 *frr* 同源基因的结构、数目、起源及 RRF 确切功能的综合报告。

在对日本晴 *Oryza sativa* subsp. *japonica* cv. *Nipponbare*) 4 号染色体序列进行注解的工作中, 我们发现了一段 *frr* 同源序列; 通过同源数据搜索的方法继而在 7 号染色体上找到了另一段同源序列。RT-PCR 测序、Southern blotting 等实验的结

2003-07-04 收到, 2003-11-13 接受。

本研究由中国科技部、中科院、上海市科委共同资助。

*通讯作者 (E-mail: bhan@ncgr.ac.cn; Tel: 86-21-64825260)。

果证明这是两个单拷贝的表达基因,它们所具有的不同转录特性及其所编码的蛋白质的N端特点(Emanuelsson等2000)等为其细胞定位提供了证据。而对37个原核及真核RRF进行同源比较的结果则从一个侧面验证了真核生物叶绿体和线粒体的内共生学说,同时也揭示了RRF作为分子进化标尺的潜在价值。

1 材料与amp;方法

1.1 植物材料

本文中所使用的植物材料均为粳稻日本晴。稻种在37℃浸泡萌发后,转移至30℃的培养箱中水培,绿苗的光照时间为12 h/d,黄化苗则在黑暗条件下培养。收取5种不同的组织(10 d左右的黄化苗和绿苗、30 d的绿苗以及10 d的黄化苗和绿苗的根组织)待用。

1.2 软件与软件包

Annotation system (NCGR)是由本中心自主研发的一套基因组序列注解平台,它将4种基因预测软件fGeneSH、GeneMark、GeneScan和Glimmer的预测结果整合到一起以便于进一步的人工校验(manual check),同时还提供tRNA基因及重复序列等相关信息。

BLAST(<http://www.ncbi.nih.gov>)是一个同源序列搜索软件包,由blastn(核酸-核酸比较)、blastp(核酸-蛋白质比较)、tblastn(蛋白质-核酸比较)和blast2(两段序列相互比较)等多个软件组成。

CD-Search(<http://www.ncbi.nih.gov>)是在一段蛋白质序列中寻找已知的保守性结构域(conserved domains)的软件。TargetP(<http://www.cbs.dtu.dk>)是一种基于神经网络算法(neural network-based)的、根据蛋白质的N端序列是否具有转运肽或信号肽的特征来预测其细胞定位的软件。这些软件的预测结果往往能够为预测基因的功能研究提供线索。

CLUSTALW(<http://www.biosino.org>)是一个多重序列同源比较软件,软件TreeView1.6.6(<http://darwin.zoology.gla.ac.uk>)则可根据CLUSTALW的结果画出相应序列的进化树。

PrimerSelector(<http://www.biosino.org>)是

GCG软件包中的引物设计软件,xprimer-online(<http://alces.med.umn.edu/rawprimer.html>)则是另一个在线的引物设计软件。

1.3 后台数据库

本文中用于基因注解及比较分析的日本晴序列来源于本中心(<http://www.ncgr.ac.cn>)和GenBank(<http://www.ncbi.nih.gov>)的数据库,粳稻93-11序列来自北京华大基因研究中心(<http://www.genomics.org.cn>)拟南芥及其它物种的RRF相关数据则是从GenBank中收集得到的。

1.4 基因组测序及注解

OSJNB库是由CUGI(Clemson University Genomics Institute)提供的日本晴的BAC文库。结合指纹法、锚标定位及末端序列比较等方法筛选出其中4号染色体上的克隆并加以定位(Zhao等2002)。挑选其中相互重叠最小者进行测序。拼装完成的基因组序列进行了软件预测和人工注解(Feng等2002)。同时收集相关的其它日本晴和水稻另一亚种粳稻93-11的序列并加以分析。

1.5 RT-PCR测序

用RNeasy Plant Mini Kit(QIAGEN)抽提日本晴不同组织的总RNA并用RQ1 RNase-free DNase(Promega)纯化后作为RT-PCR的模板。根据基因注解的结果,用PrimerSelection和xprimer-online设计多对引物并由上海生工生物工程技术服务有限公司合成(表1)。RT-PCR使用RNA PCR Kit(AMV)ver.2.1(TaKaRa),同时扩增 β -actin基因片段作为阳性对照。RT-PCR产物电泳鉴定后加以纯化(Wizard PCR Preps DNA Purification System, Promega)和测序(Big-Dye Terminator Cycle Sequencing Ready Reaction, ABI),测序结果经人工校验和拼装后,与基因预测结果比对,以修正可能存在的基因组测序和预测错误。

1.6 Southern blotting

抽提日本晴黄化苗的总DNA并用6种不同的限制性内切酶(Roche)分别将其完全酶切。每个反应中加入大约10 μ g的总DNA。酶切产物分为两组电泳分离(0.8% Agarose, 0.5 \times TBE, 4 \times , 1 V/cm):A组包括HindIII、Sacl、XhoI的产物,以OsfrrA的RT-PCR产物为正对照,B组则是EcoRV、PstI、SmaI的产物,以OsfrrB

表1 OsfrrA与OsfrrB的RT-PCR引物

Table 1 The primers used forOsfrrA andOsfrrB RT-PCR

	OsfrrA	OsfrrB
5'UTR-f1	5' CAAAACCCATCAAACCCCTCG3'	5' CAATCAGCAACAAAGGACCTG3'
5'UTR-f2	5' CCTGAACCCTCCGGTAAGC3'	5' GGCGTCAAGCAAACCAAGC3'
5'UTR-f3		5' GTATCCTGTCTCAGTCCCATG3'
Exon1-f1	5' AAGCTGTTCTCCTGGAGGC3'	5' TGCATGGTCTCACCAGAGG3'
Exon2-f1		5' CCACTATTGAAGAAATTGAGGC3'
3'UTR-r1	5' TGAGTGCCTTGCAACCGAATC3'	5' GCTCAAGAAATCTGTGACCG3'

的RT-PCR产物为正对照;随后分别印迹到Hybond-N⁺尼龙膜(Amersham)上。纯化了的OsfrrA和OsfrrB RT-PCR产物用随机引物标记后作为探针分别杂交A组与B组样品(ECL Random Primer Labeling and Detection System, Amersham)。

2 结果

2.1 OSJNBa0084K20和OSJNBa0076N16的注解结果

在对日本晴4号染色体上的两个已经完全测序的克隆OSJNBa0084K20和OSJNBa0076N16(GenBank登录号为AL606613和AL731617,分别被分子标记R896和R1849定位于74.5 cM和75.9 cM处)的注解中,我们在两者的重叠区域发现一个可能存在的假定基因OsfrrA(putative OsfrrA),它所编码蛋白质即为OsRRFA。fGeneSH、GeneMark和Glimmer均对其做出预测,虽然三者的预测结果有较大的差异。该段序列有局部的水稻EST与之匹配(alignment);其可能的翻译产物与已知蛋白RRFHP的后半段有62%的同源(blastp结果),该段同源序列是一个完整的RRF保守结构域(COG0233, pfam01765)(CD-Search结果);其N端序列则具有线粒体转运肽的特征(TargerP的预测)。

进一步的同源搜索发现,日本晴基因组中至少还有一段序列可能编码一个RRF同源因子。该段序列位于七号染色体克隆OJ1121_A05(AP005908)中,但由于是第二阶段测序的结果(phase II sequence),几种软件预测结果分歧很大,我们暂且称其为假定基因OsfrrB。

在北京华大基因研究中心的籼稻93-11框架序

列数据库中,contig3482包含与OsfrrA完全相同的序列,contig444则与上述七号染色体的相关序列相当一致,由于contig444是完全序列(finished sequence),预测和人工注解的结果较好。

图1直观的显示了日本晴的假定基因OsfrrA和OsfrrB的软件预测结果及用RT-PCR产物测序的方法证实的真实的基因结构(具体描述详见2.2),横坐标代表序列长度,上三排是三种软件各自的预测结果,而“manual check”一栏则是基因的真实结构。93-11中两个基因的注解结果这里就不赘述了。

2.2 OsfrrA与OsfrrB是两个表达基因

我们根据2.1所述注解结果设计了引物,以日本晴30 d绿苗的总RNA为模板,用RT-PCR的方法对两个假定基因加以验证。之所以设计和使用了对引物,一方面是考虑到软件预测的局限性以及几种软件预测结果的不完全一致,另一方面也是为了使此后的基因测序中数据能够达到一定的冗余度,从而保证序列的准确性;此外,对于假定基因OsfrrB,我们更是选择在日本晴与93-11序列完全匹配的区域设计引物,以避免因基因组序列的误差而造成实验失败。图2展示的就是其中OsfrrA的RT-PCR结果,扩增产物特异性很好,且片段大小也与预测值基本吻合,只是随引物序列在预测基因中的位置的不同而略有差别。OsfrrB的RT-PCR结果与此相似。这证明了OsfrrA与OsfrrB的确是水稻中的两个表达基因。

将两组RT-PCR产物纯化后用相应的引物双向测序并分别组装,得到1 023 bp的OsfrrA序列重叠群(sequence contig)和1 026 bp的OsfrrB序列重叠群,两段序列都包含了完整的编码区和部分的UTR。将其与预测结果比较发现:

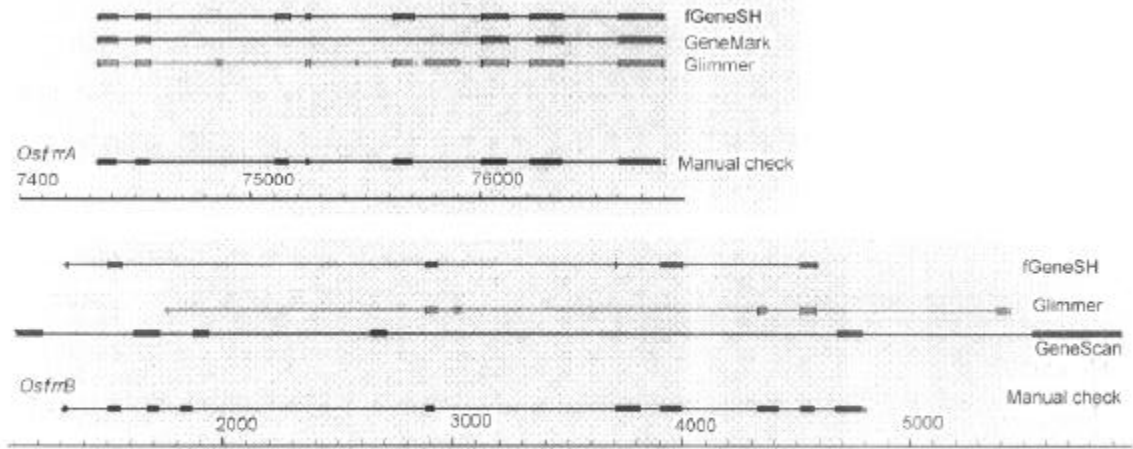


图 1 假定基因 *OsfrrA* 和 *OsfrrB* 的注解结果
 Fig.1 Annotation results of putative *OsfrrA* and *OsfrrB*

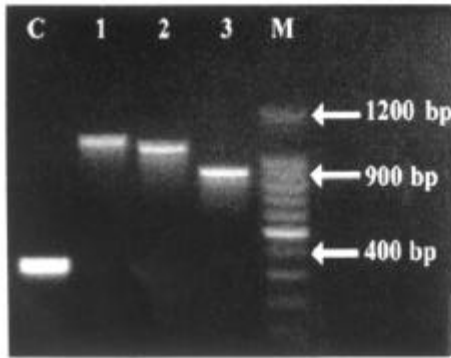


图 2 日本晴 *OsfrrA* 的 RT-PCR 结果
 Fig.2 RT-PCR results of *OsfrrA* in Nipponbare
 M: 100 bp ladder marker; C: Amplified actin gene fragment as positive control; 1: 5' UTR-f1+3' UTR-r1; 2: 5' UTR-f2+3' UTR-r1; 3: Exon1-f1+3' UTR-r1.

(1) 表达基因 *OsfrrA* 的结构, 包括外显子的数目、大小以及外显子与内含子间的剪接位点, 与 fGeneSH 的预测结果完全吻合。该基因跨越一段 2 448 bp 的基因组序列, 由 8 个外显子组成, 其中外显子 5 是一个 22 bp 的微外显子。它的编码序列 (coding DNA sequence, CDS) 全长 783 bp, 编码 260 个氨基酸残基的 OsRRFA。其 5' -UTR 不短于 107 bp, 3' -UTR 全长 239 bp。除了 3' -UTR 中有两个测序错误外, *OsfrrA* 的序列重叠群与对应的基因组序列完全吻合。

(2) 表达基因 *OsfrrB* 的结构与日本晴的基因预测结果有很大的差异, 这主要是由于前面所提

到的基因组序列准确度的不足而引起的; 93-11 的 fGeneSH 软件预测结果除了第一个外显子的位置及大小有误差, 其它大体正确。该基因跨越一段 3 532 bp 的基因组序列, 由 10 个外显子组成, 其中外显子 10 是一个 18 bp 的微外显子。它的 CDS 全长 729 bp, 编码 242 个氨基酸残基的 OsRRFB。其 5' -UTR 不短于 227 bp, 3' -UTR 不短于 116 bp。

(3) 日本晴与 93-11 中的这两个基因的 CDS 及已证实的 UTR 序列均完全一致。

2.3 *OsfrrA* 与 *OsfrrB* 是两个单拷贝基因

日本晴总 DNA 以 1.3 中所述 6 种不同的酶完全酶切 (A 组的 HindIII、SacI 和 XhoI 在 *OsfrrA* 内部、B 组的 EcoRV、PstI 和 SmaI 在 *OsfrrB* 内部均无切点) 并电泳印迹后, 将 *OsfrrA* 与 *OsfrrB* 的 RT-PCR 产物 (即两组电泳样品中各自的阳性对照) 标记后作为探针对两组样品分别进行了杂交。从图 3 显示的杂交结果中可以看到, 无论 A 组还是 B 组, 在 3 种酶切条件下各自都只得到了一条阳性条带, 这说明 *OsfrrA* 与 *OsfrrB* 很可能是两个单拷贝基因。需要特别说明的是, 虽然这两个基因在蛋白质水平有较强的局部同源, 但其基因序列并无明显同源 因此 Southern blotting 的结果并不会相互的干扰。

此外, 用与这两个基因相应的基因组序列、CDS 序列和蛋白质序列搜索 98% 覆盖率的日本晴基因组序列数据库和 70% 以上覆盖率的 93-11 基因组数据库, 均无其它同源匹配。

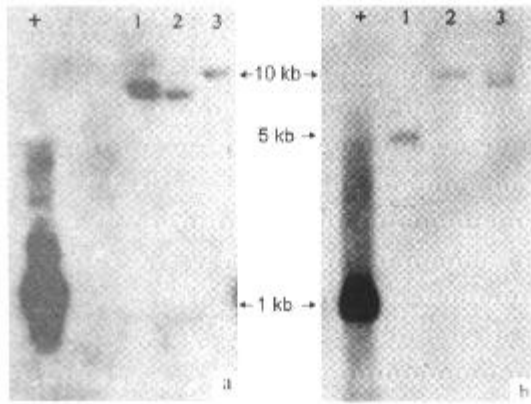


图3 *OsfrrA* 及 *OsfrrB* 的 Southern blotting 结果

Fig.3 Southern blotting results of *OsfrrA* and *OsfrrB*

a: Result of *OsfrrA*. +: *OsfrrA* RT-PCR fragment as positive control; 1: Nipponbare total DNA/*Xho*I fragments; 2: Nipponbare total DNA/*Sac*I fragments; 3: Nipponbare total DNA/*Hind*III fragments.

b: Result of *OsfrrB*. +: *OsfrrB* RT-PCR fragment as positive control; 1: Nipponbare total DNA/*Pst*I fragments; 2: Nipponbare total DNA/*Eco*RV fragments; 3: Nipponbare total DNA/*Sma*I fragments.

这两方面的证据共同指出, *OsfrrA* 与 *OsfrrB* 是水稻基因组中的两个单拷贝基因。

2.4 水稻与拟南芥 *frr* 基因的结构比较

在对水稻四号染色体进行基因注解的过程中我们发现, 一些高度保守的、高等植物中共有的基因以及一些多拷贝基因不仅在其编码的蛋白质的序列上有很高的同源性, 其基因结构的异同也有规律可循: 这些基因相互间外显子数目相同或相近(相近往往意味着有外显子融合或断裂), 而相

应位置上的外显子大小也相同或相差 $3n$ (n 通常都很小)

为验证这一规律, 我们搜索了拟南芥的基因组数据库, 找到并分析了位于其 3 号染色体上的两段 *frr* 同源序列: 基因 F28J7.13 (对应蛋白质 AAF03454.1) 和基因 F16M2_40 (对应蛋白质 CAB86420.1)。两个 *frr* 均有全长匹配的 EST 数据, 根据定位软件的预测, 前者的蛋白质产物位于线粒体, 而后者的位于叶绿体, 故称之为新基因 (novel gene) *AtfrrMT* 和 *AtfrrCP*。

图 4 是水稻与拟南芥 *frr* 的结构比较 (水稻的相关定位证据将在 2.5 中详加描述)。由图可见, *OsfrrB* 后 9 个外显子中的 7 个与 *AtfrrCP* 相应的外显子大小完全相同, 另一个也只相差 6 bp ($n = 2$), 5' 外显子的差异可能是由两者的 N 端转运肽的差异造成的; *OsfrrA* 与 *AtfrrMT* 的结构也有相似的特点, 所不同的在于 *OsfrrA* 的第 4、第 5 个外显子融合在一起才相当 *AtfrrMT* 的第 4 外显子 (相差 9 bp, $n=3$), 另外第 6 个外显子比 *AtfrrMT* 中相应的那个明显要小。

不排除这样的现象是与长时期的基因组进化耦联的, 但其普遍性如何、与基因的保守程度是否成正比、能提示怎样的分子进化机制等等, 还需要通过更大量的数据分析来验证和推测。另一方面, 两对基因的结构相似程度上的差异是否表明叶绿体与线粒体在进化过程中采取了不同的路线, 也有待于进一步的研究。

2.5 *OsRRFA* 与 *OsRRFB* 的亚细胞定位

较之通常大小为 180 个残基左右的原核生物

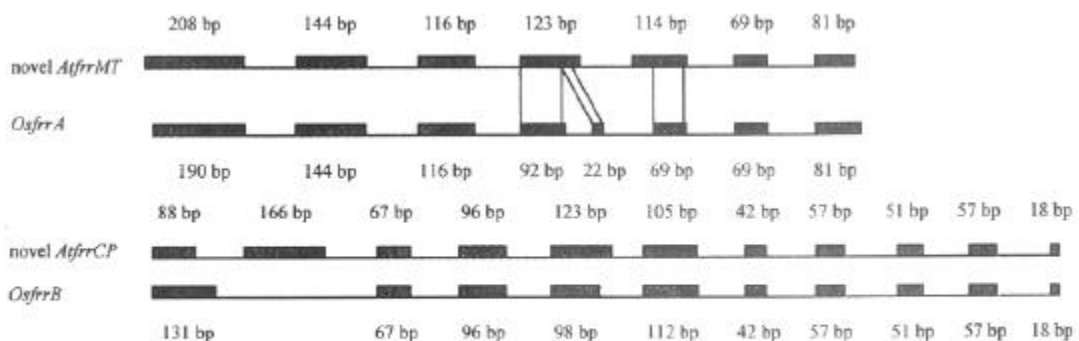


图 4 水稻与拟南芥 *frr* 基因的结构比较

Fig.4 Structural analysis of *frr* genes from *O. sativa* and *A. thaliana*

RRF 因子，水稻中的两个因子都多了一段N 端序列。我们对 *OsfrA* 和 *OsfrB* 的表达进行了分析，并根据以下三方面的证据推测其表达产物可能的亚细胞定位。

首先，我们以来自日本晴的 5 种不同组织（10 d 左右的黄化苗和绿苗、30 d 的绿苗以及 10 d 的黄化苗和绿苗的根组织）的总 RNA 为模板，分别反转录扩增 *OsfrA* 和 *OsfrB*（每种组织所取用的总 RNA 量基本相同，且每种组织用于扩增 3 个不同基因的第一链模板是在同一个反应中合成的，以尽可能的保证其结果的可比性），发现这两个基因有各自不同的转录特点，如图 5 所示。

OsfrA 在所有 5 种组织中均有转录活性，且其转录强度在几种组织间无明显的差异，这与线粒体作为生物体不同组织中的细胞的能量提供者的角色是相符的；*OsfrB* 虽然也在所有 5 种组织中都有转录，但在根——无论是绿苗的还是黄化苗的——中的转录强度大大低于其在苗中的转录，而在幼苗、成苗和黄化苗中的转录则无明显的差

异，这与已知的 RRFHCP 基因的转录特性也完全吻合——虽然不受光的诱导（在绿苗和黄化苗中同样转录），但在光合组织中表达活性更高（苗中的表达活性明显高于根中），且在光合组织中持续表达（幼苗与成苗有相近的转录强度）。

其次，两个相应的 RRF 其 N 端序列都具有明显的转运肽特性。TargetP 的预测（表 2）表明：*OsRRFA* 定位于线粒体中的可能性是它作为叶绿体蛋白、分泌蛋白及其它定位的 3.8、28.4 和 426 倍，其最可能的转运肽切点在第 57 个氨基酸后；*OsRRFB* 则更倾向于定位在叶绿体中，其可能性是线粒体的 1.1 倍、分泌蛋白的 32 倍和其它定位的 2.9 倍，而最可能的切点位置在第 48 个氨基酸后。上述预测的可能的切点位置与已知的 RRFHCP 的切点位置及 RRF 结构域的位置也是相吻合的（图 6）：三者的切点前序列相互之间不具有同源性，而切点后序列在几乎全长的范围内都具有相当高的同源性，并且都包含了完整的 RRF 结构域。

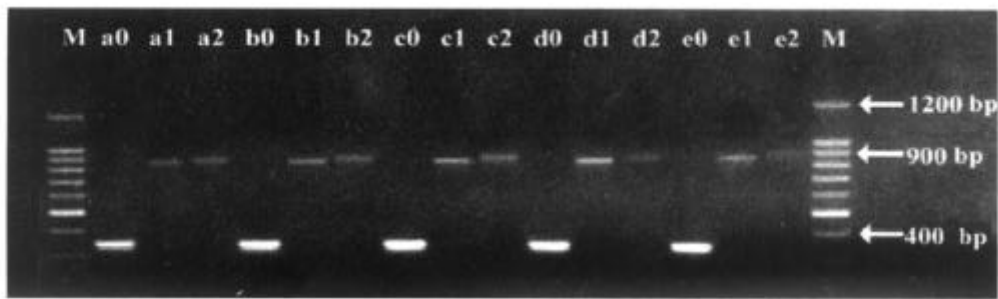


图 5 *OsfrA* 和 *OsfrB* 在不同组织中的转录活性比较

Fig.5 Transcription analysis of *OsfrA* and *OsfrB*

0: Amplified actin gene fragment as positive control; 1: RT-PCR product of *OsfrA*; 2: RT-PCR product of *OsfrB*; a: 10 days' yellow shoot; b: 10 days' green shoot; c: 30 days' green shoot; d: Root of 10 days' yellow shoot; e: Root of 10 days' green shoot; M: 100 bp ladder marker.

表 2 TargetP 对 *OsRRFA* 与 *OsRRFB* 的亚细胞定位预测

Table 2 The subcellular location prediction of *OsRRFA* and *OsRRFB* by TargetP

Name	Length	cTP	mTP	SP	Other	Loc	TPLen
<i>OsRRFA</i>	260	0.221	0.852	0.030	0.002	M	57
<i>OsRRFB</i>	242	0.418	0.384	0.013	0.145	C	48

Length: Length of protein; cTP: Chloroplast transit peptide; mTP: Mitochondrion transit peptide; SP: Signal peptide; Loc: Subcellular location; TPLen: Transit peptide length.

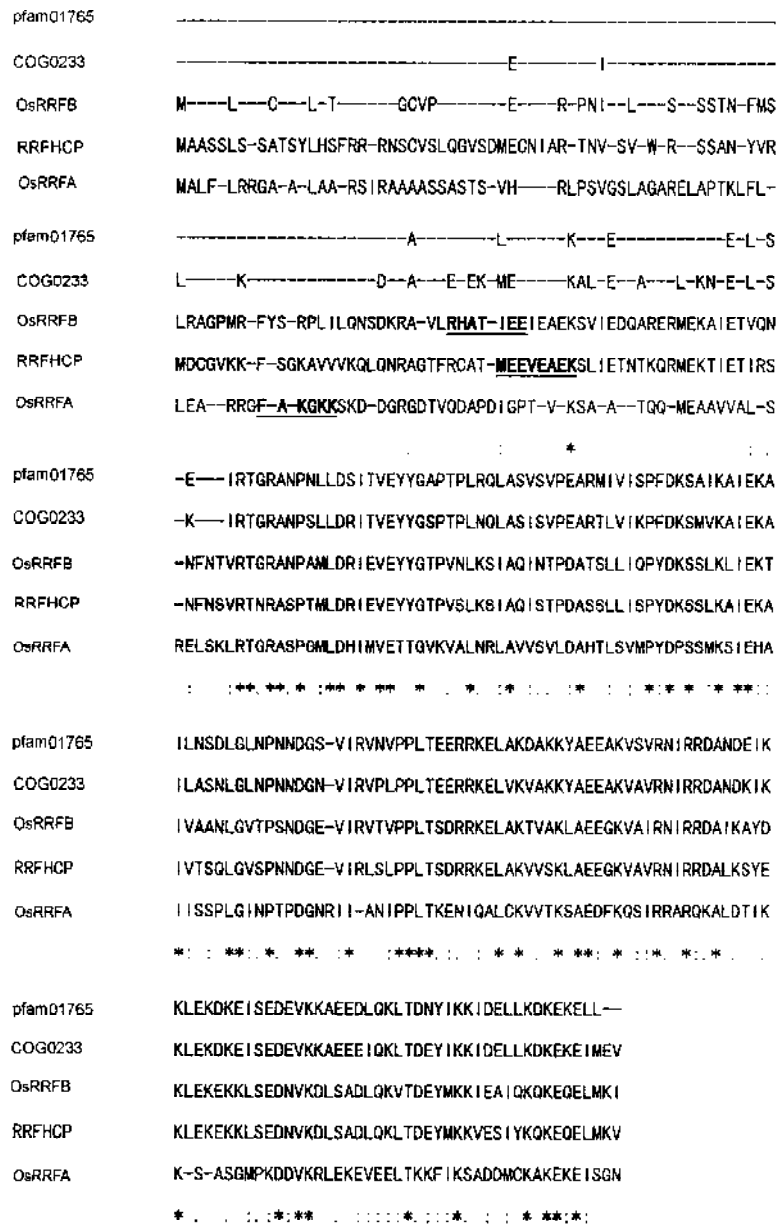


图 6 RRFHCP、OsRRFA、OsRRFB、RRF domain 的同源比较

Fig.6 Homology comparison between RRFHCP, OsRRFA, OsRRFB and RRF domains

“N”: Alternative cleavage site; “*”: Identical amino acids; “: ” & “.”: Similar amino acids.

第三，分子进化的分析结果也认为两个因子具有不同的来源。无论是在对几种真核 RRF 的 CDS 水平上的进化分析还是在对我们收集并整理的包括真核与原核的 35 个物种的 37 段 RRF 蛋白质序列的同源比较中，OsRRFA 都被认为与人、线虫、酵母、支原体、衣原体及预测的拟南芥线粒体等的 RRF 具有更近的亲缘关系，而 OsRRFB 则与 RRFHCP、属于蓝藻门的念珠藻及预测的拟

南芥叶绿体的 RRF 更为相似（图 7）。

已有充分的实验证据说明原核生物的蛋白质合成体系必须循环使用 RRF 和 tRNA，叶绿体和线粒体在进化上和蛋白质合成系统的运作上都具有一定的原核特征，而水稻叶绿体和线粒体的基因组中都没有 frr 的同源序列，再结合上述种种证据，我们推测，水稻 OsRRFA 和 OsRRFB 虽然是由细胞核基因组编码的，但合成后它们会被分别

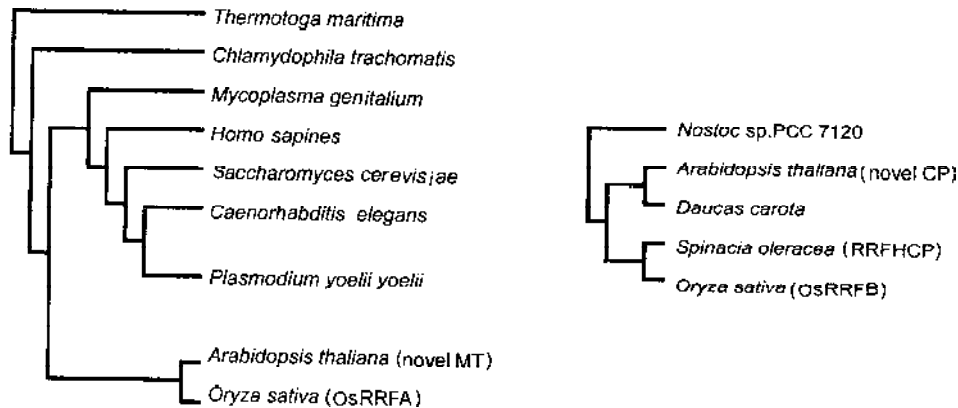


图7 OsRRFA 与 OsRRFB 的进化树

Fig.7 Phylogenetic tree of OsRRFA and OsRRFB

转运至线粒体和叶绿体中发挥作用。

3 讨论

在所有已做过基因组测序的原核生物（除古细菌外）中都发现了 *frr* 同源序列，而在基因组测序已完成或基本完成的真核生物中也发现了 *frr* 的踪迹，这一切都表明 *frr* 在真核生物中也扮演着重要的角色。但除人以外，真核 *frr* 大都还只是基因组注释的结果而未经实验证实。此前在高等植物中唯一已有过报告的是菠菜的 RRFHCP 的基因 (Rolland 等 1999)，但仅仅知道 RRFHCP 定位于叶绿体中，并不了解它是否还有其它的同源基因。

我们根据水稻的基因组序列注解 (图 1) 及同源搜索的结果，用 RT-PCR 测序和 Southern blotting 的方法证明了 *OsRRFA* 和 *OsRRFB* 是两个具有转录活性的单拷贝基因 (图 2、图 3)，它们分别位于水稻 4 号和 7 号染色体上，且在水稻的两个亚种中是完全相同的。根据两个基因在不同组织中的转录特性 (图 5)、所编码蛋白质的 N 端特性 (表 2、图 6) 及与其它 RRF 的亲缘关系远近 (图 7)，可以初步推测，*OsRRFA* 是一个线粒体因子而 *OsRRFB* 则是叶绿体因子。

虽然还没有对其功能的直接证据，但考虑到作为双子叶模式植物的拟南芥基因组中也有两个 *frr* 同源新基因，它们不仅和 rice 的两个基因在结构上相似 (图 4)，而且其编码的蛋白质与 rice 的两个 RRF 在序列上高度同源，在结构上相似

(N 端分别具有线粒体和叶绿体转运肽的特征，其后的序列则包含完整的 RRF 保守结构域) (图 6)，因此我们有理由认为：在高等植物中，可能普遍存在两个而不是以前发现的仅仅一个 RRF 因子，它们由不同的基因编码，分别在线粒体和叶绿体这两种半自主细胞器中发挥与原核 RRF 相似的作用，且由于两者都由核基因编码，在细胞器基因组中不存在同源，因此也可能充当着细胞对半自主细胞器在翻译后水平上进行调控的信号分子。

此外，水稻和拟南芥的 RRF 在相应基因结构上的异同 (图 4) 及其与其它原核及真核生物 RRF 的高度同源 (图 7) 也提示了其在分子进化研究中的潜在价值：RRFHCP 可以作为植物 (包括蓝细菌) 的分子进化标尺 (在蛋白质水平甚至 CDS 水平上)，而 RRFMT 则可作为动物的分子进化标尺。

近年来，对蛋白质的翻译机理及其调控的研究已不仅仅局限于翻译的起始和肽链的延伸，研究人员已越来越意识到翻译的终止反应中也暗藏玄机 (Nakamura 和 Ito 1998)。首先，释放因子 RF 帮助识别终止密码子，使已合成的多肽链从翻译复合体中释放出来；其后，核糖体 - mRNA - tRNA 复合体 (即翻译后复合体) 解聚。已经发现了两大类释放因子：一类是密码子特异的，包括原核生物中的 RF1 和 RF2，以及真核生物中对应的 eRF1；另一类是非特异性的，包括原核中的 RF3 及真核中的 eRF3。然而，对于翻译后复

合体的解聚了解得还很少。RRF 是唯一一个已知与此有关的原核因子, 虽然二十年前就已被发现, 但对其功能研究的突破仅仅是近几年的事, 而且真核生物的核基因表达体系中尚未发现有类似的因子。因此, *OsfrfA* 和 *OsfrfB* 以及其它几个真核生物中 *frr* 的发现就给我们提出了如下几个问题。

首先, 真核的 RRF 是否仅仅在细胞器中发挥作用, 抑或还具有不同于原核 RRF 的其它作用? 如果是前者, 那么为什么它们由核基因编码却在细胞器基因组中没有同源序列? 真核生物的细胞核基因表达体系中翻译后复合体的解聚是否需要其它辅助因子的参与, 而是一个自然发生的、不耗能的过程? 抑或只是相关辅助因子在序列上与原核 RRF 完全没有同源性? 如果是后一种情况, 那么何以 eRF 与 RF 却有着很高的相似性? 毕竟多肽链的释放与翻译后复合体的解聚是两个密切相关的过程, 在原核生物中甚至可能共用某些辅助因子。

其次, 已有证据表明半自主细胞器的蛋白质表达在转录和翻译水平受细胞核的调控以适应不同的发育阶段和环境条件 (Stern 等 1997, Giege 和 Brennicke 2001), 那么类似的调控是否也存在于翻译后水平呢? RRF 是否就是这样的一个信号分子呢? 这样的假设也许可以解释为什么水稻及其它真核 RRF 与原核 RRF 有一定的进化上的同源关系, 并且 *OsRRFA* 与支原体和衣原体、*OsRRFB* 与蓝藻及嗜热菌在进化上更相似——与内共生学说中的细胞器起源完全吻合——而编码它们的基因却并不存在于细胞器基因组中: 作为核对细胞器的调控分子, 其基因在进化过程中从细胞器基因组中转移到了核基因组中, 并且失去了原来的拷贝, 而不像多数核编码的细胞器因子那样发生了基因的扩增, 单拷贝的基因无疑具有更强的调控能力。

第三, 高等植物中保守基因在结构上 (包括外显子数目和长度上) 所具有的相似性规律, 其普遍性如何、在分子进化上具有怎样的意义、是否可以借助于基因结构上的特性来判别基因的性质和功能? 就水稻与拟南芥的两对 *frr* 而言, 其相似程度的差异是否还与线粒体和叶绿体基因可能经历了不同的进化路线有关? 要证实这些推测还需要积累更大量的实验证据并更深入的加以分析归纳。

上述种种都将指导对 RRF 结构与功能进一步研究的方向。此外, 我们选择 RRF 作为研究对象不仅仅是因为其在蛋白质合成中的重要地位以及在分子进化研究中的潜在价值, 同时也是为了对我们的基因注解系统的可靠性和准确性的评估积累数据, 并积累一些基因鉴定的经验, 为将来以基因组测序为平台的全基因组或是整条染色体水平上的基因鉴定和结构功能研究奠定基础。

参 考 文 献

- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*, 300 (4): 1005–1016
- Feng Q, Zhang YJ, Hao P, Wang SY, Fu G, Huang YC, Li Y, Zhu JJ, Liu YL, Hu X et al. (2002). Sequence and analysis of rice chromosome 4. *Nature*, 420 (6913): 316–320
- Giege P, Brennicke A (2001). From gene to protein in higher plant mitochondria. *CR Acad Sci*, 324 (3): 209–17
- Hirashima A, Kaji A (1970). Factor dependent breakdown of polysomes. *Biochem Biophys Res Commun*, 41 (4): 877–883
- Hirashima A, Kaji A (1973). Role of elongation factor G and a protein factor on the release of ribosomes from messenger ribonucleic acid. *J Biol Chem*, 248 (21): 7580–7587
- Ichikawa S, Kaji A (1989). Molecular cloning and expression of ribosome releasing factor. *J Biol Chem*, 264 (33): 20054–20059
- Janosi L, Mottagui-Tabar S, Isaksson LA, Sekine Y, Ohtsubo E, Zhang S, Goon S, Nelken S, Shuda M, Kaji A (1998). Evidence for in vivo ribosome recycling, the fourth step in protein biosynthesis. *EMBO J*, 17 (4): 1141–1151
- Janosi L, Ricker R, Kaji A (1996). Dual function of ribosome recycling factor in protein biosynthesis: Disassembling the termination complex and preventing translational errors. *Biochimie*, 78 (11–12): 959–969
- Janosi L, Shimizu I, Kaji A (1994). Ribosome recycling factor (ribosome releasing factor) is essential for bacterial growth. *Proc Natl Acad Sci USA*, 91 (10): 4249–4253
- Kanai T, Takeshita S, Atomi H, Umemura K, Ueda M, Tanaka A (1998). A regulatory factor, Fil1p, involved in derepression of the isocitrate lyase gene in *Saccharomyces cerevisiae*—a possible mitochondrial protein necessary for protein synthesis in mitochondria. *Eur J Biochem*, 256 (1): 212–220
- Nakamura Y, Ito K (1998). How protein reads the stop codon and terminates translation. *Genes Cells*, 3 (5): 265–278

- Pavlov MY, Freistoffer DV, Heurgue-Hamard V, Buckingham RH, Ehrenberg M (1997). Release factor RF3 abolishes competition between release factor RF1 and ribosome recycling factor (RRF) for a ribosome binding site. *J Mol Biol*, 273 (2): 389–401
- Rolland N, Janosi L, Block MA, Shuda M, Teyssier E, Miede C, Cheniclet C, Carde JP, Kaji A, Joyard J (1999). Plant ribosome recycling factor homologue is a chloroplastic protein and is bactericidal in *Escherichia coli* carrying temperature-sensitive ribosome recycling factor. *Proc Natl Acad Sci USA*, 96 (10): 5464–5469
- Ryoji M, Karpen JW, Kaji A (1981). Further characterization of ribosome releasing factor and evidence that it prevents ribosomes from reading through a termination codon. *J Biol Chem*, 256 (11): 5798–5801
- Selmer M, Al-Karadaghi S, Hirokawa G, Kaji A, Liljas A (1999). Crystal structure of *Thermotoga maritima* ribosome recycling factor: a tRNA mimic. *Science*, 286 (5448): 2349–2352
- Stern DB, Higgs DC, Yang J (1997). Transcription and translation in chloroplasts. *Trends Plant Sci*, 2 (8): 308–315
- Zhang Y, Spremulli LL (1998). Identification and cloning of human mitochondrial translational release factor 1 and the ribosome recycling factor. *Biochim Biophys Acta*, 1443 (1–2): 245–250
- Zhao Q, Zhang Y, Cheng Z, Chen M, Wang S, Feng Q, Huang Y, Li Y, Tang Y, Zhou B et al. (2002). A fine physical map of the rice chromosome 4. *Genome Res*, 12 (5): 817–823

The Structural, Transcriptional and Homology Analysis of Two *frr* Genes in Rice

HU Xin¹, HU Hao¹, HONG Guo-Fan^{1,2}, HAN Bin^{1*}

¹National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200233, China; ²Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China)

Abstract : Two rice (*Oryza sativa* subsp. *japonica* cv. Nipponbare) ribosome recycling factor genes—*OsfrrA* and *OsfrrB* had been identified and characterized (Figs.1, 2 & 4) in this study. The gene *OsfrrA* is located on chromosome 4 while *OsfrrB* on chromosome 7. No other homologue is found in rice organelle genomes. Both genes are unique in rice genome (Fig.3) and constitutively expressed (Fig. 5). The N-terminal character of their encoded protein products suggests that the proteins are transferred to mitochondrion and chloroplast respectively (Table 2, Fig.6) and carry out their functions. The sequence conservation and the

constitutive expression profile of the two genes strongly imply their indispensable role in plant growth. In addition, these sequences share phylogenetic homology to some extent with other prokaryotic and eukaryotic RRFs (Fig.7), providing further evidence for the endosymbiotic theory, and implying the potential value of RRFs in molecular evolution research.

Key words: ribosome recycling factor (RRF); *Oryza sativa*; post termination complex; semi-autonomous organelle

*Corresponding author (E-mail: bhan@ncgr.ac.cn; Tel: 86-21-64825260).