

口蹄疫病毒 R 株基因组序列分析研究

林小敏¹, 孙彦伟², 余业东², 罗满林¹, 林绍荣¹, 贺东生¹

(¹华南农业大学兽医学院, 广东广州 510642; ²广东省动物防疫监督总所, 广东广州 510230)

摘要: 根据 GenBank 上发表的口蹄疫 (foot-and-mouth disease, FMD) 全基因序列数据, 设计了多对引物, 采用 RT-PCR 的方法, 分别对 R 株编码区和非编码区基因进行扩增, 将各片段克隆至 pMD18-T 载体, 进行核苷酸序列测定。按顺序拼接各基因的序列, 将结果序列与 GenBank 上各 FMDV 毒株的序列进行比较和同源性分析。序列的比较和分析表明, R 株编码区全长为 6 966 个核苷酸, 共编码 2 322 个氨基酸; 非编码区 5' 端内部核糖体结合位点 (internal ribosome entry site, IRES) 长约 450 个核苷酸, 3' 非翻译区 (untranslatable region, urR) 从 TA 始至 Poly(A) 前长度为 94 个核苷酸, 所测出 Poly(A) 尾长度为 18 个核苷酸。与 GenBank 世界流行毒株的序列比较, 编码区核苷酸序列同源性为 88%—90%, 推导氨基酸序列同源性为 88%~95%。

关键词: 口蹄疫病毒; 基因组; 序列分析

中图分类号: S852.659.6 文献标识码: A 文章编号: 1001-411X(2006)02-0118-03

Genome Sequence Analysis of Foot-and-Mouth Disease Virus R Strain

LIN Xiao-min¹, SUN Yan-wei², YU Ye-dong², LUO Man-lin¹, LIN Shao-rong¹, HE Dong-sheng¹

(¹ College of Veterinary Medicine, South China Agric. Univ., Guangzhou 510642, China;

² Guangdong Provincial Veterinary Station of Epidemic Prevention & Supervision, Guangzhou 510230, China)

Abstract: Several pairs of primers were designed to amplify the complete genome of foot-and-mouth disease virus R strain by RT-PCR. Each fragment was cloned to pMD18-T vector and sequenced. Analysis showed that the length of the encoding region of R strain was 6 966 bp, encoding for 2 322 amino acid. The 5' IRES region and 3' UTR were approximately 450 bp and 94 bp in length. Eighteen nucleotide acid was detected in Poly (A) region by 3' RACE. Compared with other FMDV strains' sequences in GenBank, the encoding region of R strain had 88% - 90% similarities in nucleotide sequence and 88% ~ 95% similarities in deduced amino acid sequence.

Key words: foot-and-mouth disease virus; genome; sequence analysis

口蹄疫 (foot-and-mouth disease, FMD) 是偶蹄动物的一种急性、高度接触性、发热性传染病, 传播迅速、感染率极高, 并能形成大范围流行, 被国际兽疫局 (OIE) 列为 A 类动物疫病之首。其病原为口蹄疫病毒 (foot-and-mouth disease virus, FMDV), 属于小 RNA 病毒科 Picornaviridae 口蹄疫病毒属 *Aphthovirus*。基因组为单股正链 RNA, 全长约 8.5 kb, 由 5' 非翻译区 (untranslatable region, UTR)、3' UTR 和一个大的开放阅读框 (open reading frame, ORF) 架组成。本

研究通过测定 R 株全基因组的核苷酸序列及推导其氨基酸序列, 从基因水平上检验该毒株变异情况, 为 FMD 的综合防控打下基础, 对规模化养猪场疫苗的使用有一定的指导作用。

1 材料与方法

1.1 材料

1.1.1 毒株 FMDV R 株来源于猪, 经传代致弱, 由华南农业大学家畜传染病教研室保存。

收稿日期: 2005-08-31

作者简介: 林小敏 (1980-), 女, 硕士研究生; 通讯作者: 贺东生 (1968-), 副教授, 博士, E-mail: dhe@scau.edu.cn

基金项目: 广东省科技计划项目 (2003A20403); 广东省农业厅动植物科研项目粤农 [2004] 314 号

1.1.2 工程菌和分子生物学试剂 大肠杆菌 DH5 α 由华南农业大学家畜传染病教研室保存. 总 RNA 抽提试剂 Trizol 为北京赛百盛生物工程有限公司产品; 反转录酶 M-MLV 为日本东洋坊公司产品; PCR 所用试剂、限制性内切酶和 pMD18-T 载体均为 Takara 生物工程公司产品; 胶回收、质粒回收试剂盒均为 Omega 公司产品.

1.1.3 引物 参考 GenBank 上发表的 FMDV 全基因组序列, 应用 Oligo 6.24、Primer 5.0 等软件, 设计覆盖了整个基因组编码区、非编码区 5' IRES 和 3' UTR 及 Poly(A) 的多对特异性引物(序列略), 2 个相邻片段之间均有部分重叠, 以便拼接获得完整的序列. 引物均由上海博亚生物技术有限公司合成.

1.2 方法

1.2.1 病毒总 RNA 的提取 具体操作按试剂说明书步骤进行, 最后用 50 μ L 无 RNA 酶的水溶解, 直接进行反转录反应.

1.2.2 RT-PCR 获得各基因片段 在同一反应管中加入 RNA 溶解液、5 $\times$ Buffer、dNTPs、M-MLV、RNase Inhibitor 和引物, 最后补充无 RNA 酶的水至总体积 20 μ L, 42 $^{\circ}$ C 进行反应 1 h 获得 cDNA. PCR 反应在同一管中加入 cDNA、10 $\times$ Buffer、dNTPs、引物和 Ex Taq, 补充蒸馏水至总体积 100 μ L, 在 PCR 仪运行程序, PCR 产物经琼脂糖凝胶电泳, 用试剂盒回收各特异片段.

1.2.3 各基因片段的克隆、鉴定和测序 回收片段经加 A 尾反应后与 pMD18-T 载体连接, 转化至大肠杆菌 DH5 α 中, 用氨苄抗性筛选出带有阳性重组质粒的单菌落, 增殖后菌液以通用引物 M13 进行 PCR 鉴定, 为阳性的用试剂盒抽提质粒后用 HindIII、BamHI 双酶切进一步鉴定. 经 PCR 鉴定和双酶切鉴定均为阳性的菌液送到上海博亚生物技术公司进行测序.

1.2.4 序列的拼接和分析 将各基因片段的序列按顺序拼接, 划分出编码区及非编码区各个基因和区域(LP、VP4、VP2、VP3、VP1、2A、2B、2C、3A、3B、3C、3D、5' IRES、3' UTR), 检查片段的大小与位置是否一致. 用软件 DNASTAR 5.0 推导编码区相应的氨基酸序列. 用 RNAdraw 1.1 软件推测计算非编码区的二级结构图谱.

2 结果

测序结果经拼接, R 株的编码区从起始密码子 ATG 到终止密码子 TAA 前止有 6 966 个核苷酸, 共编码 2 322 个氨基酸, 非编码区 5' IRES 区长约 450

个核苷酸, 3' UTR 从 TAA 始至 Poly(A) 前长度为 94 个核苷酸, 所测出 Poly(A) 尾长度为 18 个氨基酸. 通过 NCBI Blast 在线分析系统, 将各基因序列分别与 GenBank 收录的所有基因作最大同源性比较, 结果表明各基因与世界各地流行毒株序列差异不大, 同源性比较高, 各项比较数据如表 1 所示.

表 1 各基因序列在 GenBank 的最大同源性比较

Tab. 1 Comparison of gene sequences of FMDV R strain with those available in GenBank

编号 No.	核苷酸 长度 length of nucleotide acid	核苷酸 序列 同源性 homology of nucleotide acid sequence	推导 氨基酸 长度 length of deduced amino acid	氨基酸 序列 同源性 homology of deduced amino acid sequence
L	603	82% ~ 90%	201	63% ~ 92%
VP4	255	86% ~ 90%	85	87% ~ 89%
VP2	654	85% ~ 90%	218	97% ~ 98%
VP3	660	84% ~ 90%	220	49% ~ 98%
VP1	633	86% ~ 98%	211	71% ~ 97%
2A	54	91% ~ 97%	18	88% ~ 100%
2B	462	90% ~ 91%	154	83% ~ 92%
2C	954	88% ~ 90%	318	88% ~ 98%
3A	429	86% ~ 97%	143	68% ~ 95%
3B	213	91% ~ 99%	71	84% ~ 99%
3C	639	88% ~ 93%	213	84% ~ 97%
3D	1 410	90% ~ 91%	470	91% ~ 98%
ORF	6 966	88% ~ 90%	2 322	88% ~ 95%
IRES	450	82% ~ 96%	...	...
3' UTR	94	82% ~ 92%	...	...

3 讨论

3.1 结构蛋白结构和功能的关系

O 型 FMDV 有 5 个抗原位点, 其中 3 个位于 VP1 基因上, 抗原位点 1 由 1D GH 环(133 ~ 157 位)和 200 ~ 213 位氨基酸的线性表位组成, 此位点是 FMDV 最重要的抗原位点, 其中的关键氨基酸为 144(V)、148(L)、154(K)、208(P); 抗原位点 3 位于 VP1 的 β B-C 环上, 其关键氨基酸为 43(T) 和 44(P); 抗原位点 5 位于 VP1 G-H 环内, 其关键氨基酸为 149(Q). 另外, 在 145-148 位有保守的 RGDL 细胞表面受体分子识别基序^[1]. 从测序结果推导的氨基酸序列分析, R 株这几个位点的氨基酸都没有发生变化, 而 VP1 中 145 ~ 148 位氨基酸为保守的 RGDL.

3.2 非结构蛋白结构和功能的关系

测序结果表明 R 株 2 个起始密码子的位置相距 84 个核苷酸,具有真核核糖体识别有效起始翻译的特征性序列 AXX AUG G(本试验毒株表现为 AAA AUG G);L 蛋白酶的活性中心第 51、148 和 164 位氨基酸分别为 C、H 和 D,与病毒的切割有关的,在所有的 FMDV 中均高度保守,位于 L 蛋白最后 5 个氨基酸和 VP4 蛋白前 5 个氨基酸分别为 VQK(KorR)LKLGAGQS(本毒株表现为 VQKRLK 和 GAGQS)^[2].

有学者研究表明 2C 的中央区包含核苷三磷酸(NTP)结合蛋白的特征性基序,AGXXXGKS/T 参与和 NTP 磷酸基的结合^[3]. R 株 2C 在中央区并没有找到该特征性基序,而在第 110 位有 GKS GQ GKS 序列.

P3 区中,FMDV 所在科属病毒的 3B 蛋白都是通过保守的第 3 位酪氨酸残基连接到 RNA 上的;而 3D 中有一个高度保守的以 2 个天冬氨酸残基为中心的氨基酸疏水区基序 YGDD^[4]. 本研究的结果分析表明 R 株仍保持有这 2 个特点,YGDD 基序位于 3D 蛋白的第 336~339 位.

3.3 5' IRES 区域结构和功能的关系

本次试验应用 3'RACE 原理成功扩增出 3'UTR 的完整序列,也曾经尝试用 5'RACE 原理扩增 5'端 IRES 外的序列,但并没有成功,有待进一步的试验进行.

本研究的测序结果表明 R 株的 RNA 复制顺式作用元件(cis-acting replicative element, cre)结构区域具有保守的 AAACA 序列,位于第一个 AUG 上游 474 个核苷酸处;而 IRES 区域长约 450 个核苷酸,在该区中找到多个与长距离 RNA-RNA 相互作用有关的 GNRA(N 是任意核苷酸,R 是嘌呤)基序^[5]. 另外从软件计算推测的二级结构图可以看出其有 5 个结构域,各个结构域中又有不同的分支,与其他文献报道的结构相似^[6].

3.4 3'UTR 结构和功能的关系

Chatterjee 等^[7]报道 FMDV 基因组 3'末端具有一个长达 60~80 个核苷酸的 Poly(A)尾. 刘光清

等^[8]用 3'RACE 通过基因克隆方法得到的长度为 56 个核苷酸. 本试验由于试验操作、病料采集和保存等各种因素影响,所测出的 Poly(A)数目可能与实际相比有所差异,测出 Poly(A)的长度只有 18 个核苷酸. 另外从 2 级结构图可看出其有 3 个结构域,每个结构域基本上没有分支,与其他文献报道的结构相似^[6].

参考文献:

- [1] VERDAGUER N. A similar pattern of interaction for different antibodies with a major antigenic site of foot-and-mouth disease virus: implications for intratypic antigenic variation[J]. *Virology*, 1998, 72: 739-748.
- [2] SANGAR D V. All foot and mouth disease virus serotypes initiate protein synthesis at two separate AUGs[J]. *Nucleic Acids*, 1987, 15: 3 305-3 315.
- [3] DOEDENS J R. Inhibition of cellular protein secretion by poliovirus proteins 2B and 3A[J]. *EMBO*, 1995, 14: 894-907.
- [4] HANSEN J L. Structure of the RNA dependent RNA polymerase of poliovirus [J]. *Structure*, 1997, 5: 1 109-1 122.
- [5] de QUINTO L S. Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain are required for internal initiation of translation[J]. *Virology*, 1997, 71: 4 171-4 175.
- [6] CLARKE B E, BROWN A L, CURREY K M, et al. Potential secondary and tertiary structure in the genomic RNA of foot-and-mouth disease virus[J]. *Nucleic Acids Research*, 1987, 15(17): 7 067-7 079.
- [7] CHATTERJEE N K. Foot-and-mouth disease virus RNA. Presence of 3' terminal polyribadenylic acid and absence of amino acid binding ability[J]. *Virology*, 1976, 69: 369-377.
- [8] 刘光清, 刘在新, 谢庆阁. 用 3'RACE 方法扩增并克隆口蹄疫病毒基因组 3'末端序列[J]. *中国兽医学报*, 2003, 23 (6): 524-526.

【责任编辑 柴 焰】