

SARS 冠状病毒基因结构中未知蛋白 Orf7 的研究*

赵琳¹, 涂亦娴¹, 于杨¹, 金华利¹, 陈则², 王宾^{1**}

(1. 中国农业大学农业生物技术国家重点实验室和生物学院微生物与免疫学系, 北京 100094; 2. 中国科学院武汉病毒研究所, 湖北武汉, 430071)

Role of ORF7 in the Genome of SARS-CoV

ZHAO Lin¹, TU Yi-xian¹, YU Yang¹, JIN Hua-li¹, CHEN Ze², WANG Bin^{1**}

(1. State Key Laboratory for Agro-Biotechnology and Department of Microbiology and Immunology, College of Biological Science, China Agricultural University, Beijing 100094, China; 2. Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China)

Abstract: The aim of this study was to characterize functions of the unknown open-reading frames, Orf7, Orf8 and Orf9 within the Severe acute respiratory syndrome coronavirus (SARS-CoV) genome. We cloned the cDNAs of Orf7, Orf8, Orf9 from the SARS-CoV BJ01 strain were cloned by RT-PCR and constructed into five expression vectors designated as pEGFP-Orf7, pEGFP-Orf8, pEGFP-Orf9, pcDNA3-Orf7 and pcDNA3-Orf8. HeLa cells were transfected with the constructs and analyzed for transient expressions by fluorescent analysis. We observed that the expression of pEGFP was enhanced when pcDNA3-Orf7 and pEGFP were co-transfected into HeLa cells. Also the expression of pDsRed was increased with pEGFP-Orf7 and pDsRed. More cells which expressed green and red fluorescence at the same time were detected. Based on these results, we demonstrated for the first time that the 63 amino acid protein encoded by Orf7 served as a transcriptional transactivator in SARS-CoV.

Key words: SARS-CoV; Transcriptional transactivator; ORF7; ORF8; ORF9

摘要: 为探讨SARS-CoV未知蛋白在基因表达调控中的作用, 本研究以鉴定的SARS-CoV BJ01毒株cDNA为材料, 采用PCR方法扩增SARS-CoV未知蛋白基因Orf7, Orf8, Orf9, 引入GFP和pDsRed作为报告基因, 构建出5种表达质粒: pEGFP-Orf7, pEGFP-Orf8, pEGFP-Orf9, pcDNA3-Orf7, pcDNA3-Orf8。进行体外转染实验后, 采用荧光酶标仪测量其细胞中荧光蛋白的表达量, 再采用流式细胞仪进一步进行瞬时表达的分析比较。结果发现, pcDNA3-Orf7与pEGFP共转染时, pcDNA3-Orf7能够增强绿色荧光蛋白(GFP)的表达, pEGFP-Orf7与pDsRed共转染时pEGFP-Orf7可以增强红色荧光蛋白(RFP)的表达, 且同时表达红色和绿色荧光的细胞数增加。本文首次证明了ORF7编码的63个氨基酸的蛋白具有反式激活作用, 即为SARS-CoV自身的反式激活因子。

关键词: SARS冠状病毒(SARS-CoV); 反式激活因子; ORF7; ORF8; ORF9

中图分类号: R373

文献标识码: A

文章编号: 1003-5125(2006)04-0343-05

严重急性呼吸综合征(Severe acute respiratory syndrome, SARS)的病原体能够引发以呼吸道肺上皮细胞损害为主的全身性疾病, 抑制免疫系统而导致全身感染, 使得血管病变, 是一类传染性很强的新型冠状病毒, 被命名为 SARS coronavirus (SARS-CoV), 属于巢状病毒目(Nidovirales), 冠状病毒科(Coronaviridae), 冠状病毒属(Coronavirus), 基因组大小约30kb, 属于单链正义RNA病毒^[1,2]。在对

目前其它已知冠状病毒的研究中, 发现其病毒包膜主要包括三种糖蛋白: S蛋白(Spike Protein), 是冠状病毒表面最重要的膜蛋白; M蛋白(Membrane Protein), 一种跨膜糖蛋白, 是冠状病毒颗粒组装过程中的一个关键信号; N蛋白(Nucleocapsid phosphoprotein), 处于病毒颗粒的核心部分, 对于病毒基因组RNA特征性序列的识别、与其他结构蛋白的相互作用及病毒粒子的准确组装都有着重要意义^[3,4]。目前, 研究人员

收稿日期: 2006-01-20, 修回日期: 2006-02-27

* 基金项目: 中国农业大学优秀人才引进基金资助项目。

作者简介: 赵琳(1979-), 女, 内蒙古籍, 硕士, 分子免疫学研究方向。

** 通讯作者. Corresponding author. Tel: 010-62733055; Fax: 010-62732012; E-mail: bwang3@cau.edu.cn

已对这些重要的结构基因开展了大量的研究工作,然而尚缺乏对结构蛋白编码区可能的开放读码框(Open Reading Frame, ORF)中未知蛋白(Predicted Unknown Protein, PUP)^[5,6]的深入研究。对这些未知蛋白的研究具有相当重要的意义,能够有助于了解疾病的传染过程、病毒的毒力变化以及控制该疾病的传播。例如,牛泡沫病毒(bovine foamy virus, BFV), Orf-1蛋白是BFV的转录反式激活因子,可极大地激活病毒5'长末端重复序列(long termination repeat LTR)起始基因的表达^[7]。顺式作用元件, DNA序列, 分子内相互作用, 是DNA与DNA作用; 反式激活因子, 蛋白质因子, 是蛋白质对DNA的作用, 是不同分子间作用。本文为了研究SARS-CoV中的未知蛋白, 特别是Orf7, Orf8, Orf9蛋白在病毒基因表达调控中, 是顺式作用元件或者反式作用因子。以分离并鉴定的BJ01毒株的cDNA为材料, 采用PCR方法克隆SARS-CoV的orf7, orf8, orf9各基因片断, 构建各种表达质粒, 引入pDsRed作为报告基因, 体外转染Hella细胞进行瞬时表达分析, 荧光酶标仪读取细胞中荧光蛋白的表达量后, 取适量细胞采用流式细胞仪进行分析, 发现pcDNA3-Orf7与pEGFP共转染时pcDNA3-Orf7能够增强pEGFP的表达, pEGFP-Orf7与pDsRed共转染时pEGFP-Orf7可以增强pDsRed的表达。最终, 结果首次证明了SARS-CoV编码的63个氨基酸的Orf7蛋白为SARS-CoV的反式激活因子。

1 材料与方法

1.1 实验材料

Hela细胞为本室保藏, SARS-CoV BJ01毒株的cDNA来自中国科学院武汉病毒研究所。细胞转染试剂Lipofectamine™ 2000购自Invitrogen (CA, USA), DMEM和胎牛血清均购自GIBCO/BRL公司。PCR试剂、常用的内切酶购自TaKaRa大连公司。DNA片断回收试剂盒购自北京鼎国生物技术有限责任公司。

1.2 引物

实验中使用的引物均参照已发表的序列^[13] (GenBank AY278488) 设计, 由奥科生物公司合成, 引物在正链基因组中的位置(均为5'→3'方向), 具体如下, (P1:表示上游引物; P2:表示下游引物; 斜体字划线为引物中所引入的酶切位点; 黑体字为原有的起始密码子ATG:

Orf7P1: 5'-AAGAAATTCATGTTTCATCTTGTTGAC-3'(26 mer, 27055-27072, EcoR I)

Orf7P2: 5'-AAGTCGACTGGATAATCTAACTCCA-T-3'(26 mer, 27221-27246, Sal I)

Orf8P1: 5'-AAGAATTCATGAAAATTATTCTCTTC-3'(26 mer, 27254-27271, EcoR I)

Orf8P2: 5'-AAGTCGACTTCTGTCTTTCTCTTAAT-3'(26 mer, 27605-27622, Sal I)

Orf9P1: 5'-AAGAATTCATGAATGAGCTCACT-3'(23mer, 27619-27633, EcoR I)

Orf9P2: 5'-AAGTCGACGACTTTGGTACAAGG-3'(23mer, 27739-27753, Sal I)

1.3 PCR

按照 TaKaRa PCR Kit 操作指南, 各个基因的PCR 条件经过多次预反应确定, 在退火温度上稍有区别。循环数一般设为 35 个, 结果经 2%的琼脂糖凝胶电泳检测。其中以 PCR orf7 的条件为例: 94℃ 4min后, 进行下列循环: 95℃ 30s, 56℃ 35s, 72℃ 30s, 共 35 次循环, 最后 72℃再延伸 10min, 终止反应。

1.4 质粒构建

以赠送的 SARS-CoV BJ01 毒株的 cDNA 为模板, 采用 PCR 方法扩增出 orf7, orf8, orf9 各个片断, 然后用回收试剂盒回收 PCR 产物, 最后用 T4 DNA 连接酶连接到 pMD18-T (TakaRa, Dalian, China) 载体上。酶切鉴定正确后, 序列分析证明所得的片段为目的片段, 再将各个片断分别亚克隆到 pEGFP 和 pcDNA3(均为本室保藏)载体上。质粒的提取、连接、酶切、及转化的操作方法均参照文献^[8]。

1.5 细胞培养与转染

用含 10%胎牛血清的 DMEM, 37℃ 5% CO₂ 条件下培养 Hela 细胞, 用于转染。转染时, Hela 细胞以 2.5×10⁵ mL⁻¹ 接种于 24 孔板, 每孔 500L。按照文献^[8]方法, 对各质粒进行纯化后, 通过阳离子脂质体介导转染培养好的 Hela 细胞。pcDNA3-Orf7, pcDNA3-Orf8, pcDNA3-Orf9 分别以相同的质量梯度与 pEGFP 共转染 Hela 细胞, pEGFP-orf7, pEGFP-orf8, pEGFP-orf9, pEGFP 分别单独转染 Hela 细胞作为对照。同时, pEGFP-orf7 与 pDsRed 共转染 Hela 细胞, 以 pEGFP, pDsRed, pEGFP-orf7, 单独转染和 pEGFP 与 pDsRed 共转染作为对照。每种转染都设置三个重复, 具体转染步骤参见 Lipofectamine™ 2000 产品说明书 (Invitrogen, CA, USA)。

1.6 荧光值测定

在细胞转染后的 24h, 48h, 72h 各时间点, 分别在荧光显微镜下观察细胞的状态及各质粒的表达, 进行拍照记录后, 并使用荧光酶标仪测其细胞中荧光蛋白的表达量。

1.7 流式检测表达红色和绿色荧光蛋白的细胞数量

引入 pDsRed 后, 转染细胞后 24h, 48h, 72h 各时间点, 在荧光显微镜下观察细胞状态后, 取一定比例的细胞, 0.25% 的胰蛋白酶消化 2min 后制成单细胞悬液, 磷酸缓冲液 (PBS) 洗涤三次后, 流式细胞仪检测表达红色和绿色荧光的细胞的数量。

2 结果

2.1 质粒构建及鉴定

以 cDNA 为模板, PCR 扩增得到各目的片段 *orf7*, *orf8*, *orf9*, 大小分别为 192bp, 369bp, 135bp, 与预期大小相等。将各目的片段分别克隆到 pMD18-T 载体上, 经序列分析结果正确后, 用限制性内切酶 *Eco*R I 和 *Sal* I 切下各目的片段, 低熔点琼脂糖凝胶电泳回收各目的基因片段, 再将回收得到的 *orf7*, *orf8*, *orf9* 基因片段亚克隆到 pEGFP 载体上, 当 *Eco*R I 和 *Sal* I 酶切鉴定正确, 都会产生两条带, 各片段大小分别为: pEGFP4700bp, *orf7* 192bp, *orf8* 369bp, *orf9* 135bp (如图 1a 所示)。同时, 用限制性内切酶 *Xba* I 和 *Eco*R I 将 *orf7* 和 *orf8* 由 pMD18-T 载体上切下, 按照同样的方法亚克隆到 pcDNA3 载体上, 当 *Xba* I 和 *Eco*R I 酶切鉴定正确, 也都会产生两条带, 各片段大小分别为: pcDNA3 5400bp, *orf7* 192bp, *orf8* 369bp (如图 1b 所示)。

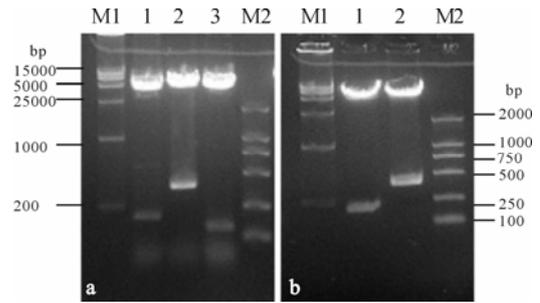


图 1 重组质粒的酶切鉴定

Fig.1 Identification of recombinant plasmid by digestion M1, DL 15000; 1, pEGFP-Orf7; 2, pEGFP-Orf8; 3, pEGFP-Orf9; 4, pcDNA3-Orf7; 5, pcDNA3-Orf8; M2, DL 2000.

2.2 转染及表达荧光细胞数量的检测

为了探讨 SARS-CoV 未知蛋白在转录调控过程中的作用, 即顺式作用因子或反式作用因子。我们克隆了各基因片段 *orf7*, *orf8*, *orf9*, 成功构建出表达质粒 pEGFP-Orf7, pEGFP-Orf8, pEGFP-Orf9, pcDNA3-Orf7, pcDNA3-Orf8。然后将 pEGFP 转染培养好的 Hela 细胞, 作为对照, 并将 pEGFP-Orf7, pEGFP-Orf8, pEGFP-Orf9 三种质粒分别单独转染 Hela 细胞, 研究 Orf7, Orf8, Orf9 蛋白是否有顺式调控作用。同时, 将 pcDNA3-Orf7, pcDNA3-Orf8 分别与 pEGFP 共转染 Hela 细胞, 混合时都设有相同的梯度, 研究 Orf7, Orf8 蛋白是否有反式调控作用。最终实验得出瞬时表达分析结果如图 2 和图 3 所示。

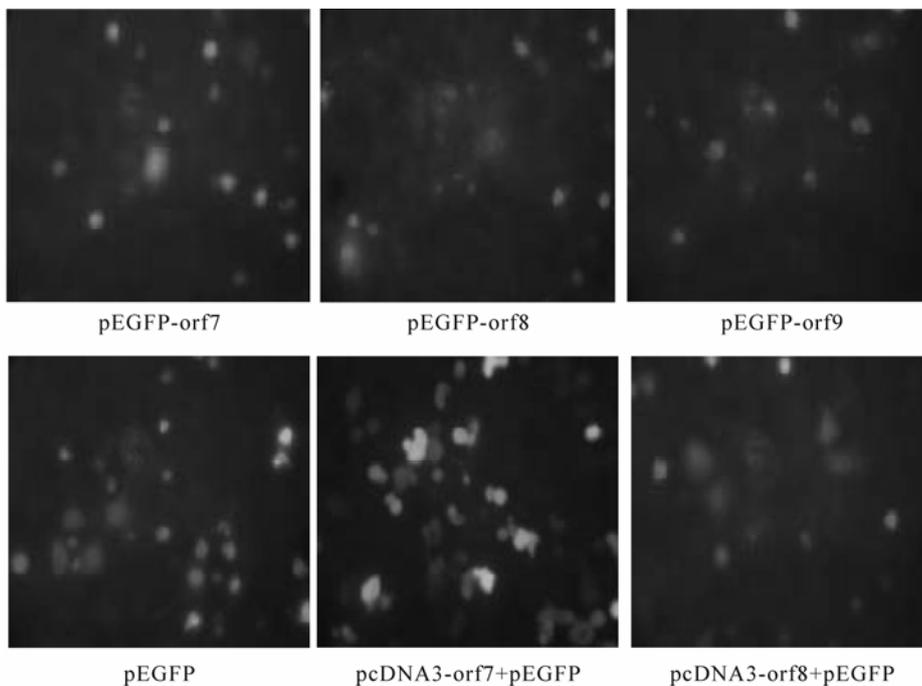


图 2 细胞转染后绿色荧光蛋白的瞬时表达结果

Fig.2 Results of green fluorescence protein expressed in cell after transfection Pictures at 72h after transfection.

从图 2 的结果可以看出, pcDNA3-Orf7 与 pEGFP 共转染时, 细胞中荧光较其它质粒单独转染或共转染的亮度更强, 即说明 pcDNA3-Orf7 使 pEGFP 在细胞中的表达量增加。荧光酶标仪测其细胞中荧光蛋白的表达量得到荧光值(图 3), 这一结果进一步证实, 随着 pcDNA3-Orf7 量的增加, 荧光值成线性升高, 且经 t 检验, pcDNA3-Orf7 与 pEGFP 共转染相比 pEGFP 单独转染, 细胞中的荧光值有显著差异性 ($P < 0.05$)。而 pcDNA3-Orf8 与 pEGFP 共转染时细胞中的荧光值低于 pEGFP 单独转染细胞中的荧光值。pEGFP-Orf7 单独转染的荧光值虽高于 pEGFP 单独转染的荧光值, 但差异不显著。pEGFP-Orf8 和 pEGFP-Orf9 单独转染的荧光值均低于 pEGFP 单独转染的荧光值。引入 pDsRed 转染后, 图 4 的结果表明, pDsRed 与 pEGFP-Orf7 共转染 Hella 细胞时, 表达红色荧光的细胞数量较 pDsRed 单独或 pDsRed 与 pEGFP 共转染时更多, 且随着 pEGFP-Orf7 量的增加表达红色荧光的细胞数量也增多。pDsRed 与 pEGFP-Orf7 共转染时, 同时表达红色和绿色荧光的细胞数量较 pEGFP 与 pDsRed 共转染时有明显增加, 差异是极显著的 ($P < 0.05$), 同样随着 pEGFP-Orf7 量的增加而增多。这些结果表明由 Orf7 编码的蛋白具有反式调控的作用, 因而可以确定该蛋白是 SARS-CoV 基因表达调控的反式激活因子。

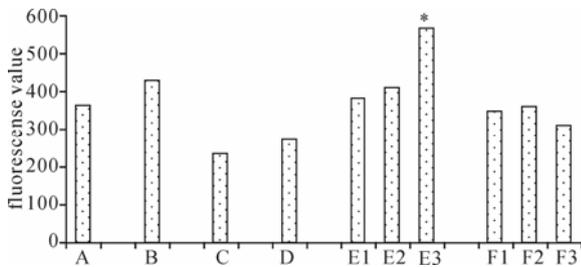


图 3 转染后细胞中绿色荧光值的测定

Fig.3 Detection of fluorescence value from cells after transfection

A: pEGFP(5 μ g); B: pEGFP-ORF7(5 μ g); C: pEGFP-ORF8(5 μ g); D: pEGFP-ORF9(5 μ g); E1: pcDNA3-ORF7(2 μ g)+pEGFP(3 μ g); E2: pcDNA3-ORF7(3 μ g)+pEGFP(3 μ g); E3: pcDNA3-ORF7(4 μ g)+pEGFP(3 μ g); F1: pcDNA3-ORF8(2 μ g)+pEGFP(3 μ g); F2: pcDNA3-ORF8(3 μ g)+pEGFP(3 μ g); F3: pcDNA3-ORF8(4 μ g)+pEGFP(3 μ g)

3 讨论

冠状病毒的基因组结构一般较为紧凑, 很少存在冗余的核苷酸, 而不同的冠状病毒都确实存在部分非保守的 ORF, 这些 ORF 的功能都有待进一步的实验证实^[9]。从进化树分析可以看出, SARS-CoV 明显不同于其他三个冠状病毒群^[10]。USCCDC^[11]、

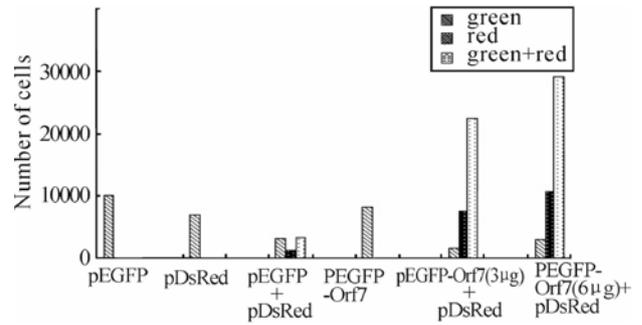


图 4 表达绿色和红色荧光蛋白细胞数量的分析

Fig.4 Analysis the number cells expressing green or red fluorescence protein after transfection

Single-cell suspensions were prepared at 72h after transfection then washed and run on a flow cytometer.

加拿大 BCCA 基因组研究所^[12]、北京华大基因组研究中心^[13]均研究得出结论: 在结构蛋白编码区可能的开放读码框(Open Reading Frame, ORF)中, 存在已有蛋白质序列数据库中未找到任何同源序列的未知蛋白(Predicted Unknown Protein, PUP)。加拿大和美国的研究组, 均对冠状病毒基因组的 5'端非翻译区(UTR)和各个 ORF 起始密码子上游的区域进行了分析, 以期待寻找到类似于其他冠状病毒“UCU AAAC”的, 参与“不连续转录”识别的特征性转录调控序列。两个研究组均在 orf1a、S 蛋白、ORF3、M 蛋白、ORF8、ORF9 和 N 蛋白上游, 发现了一个保守的 8 个核苷酸的“AAACGAAC”。另外, 加拿大的研究还显示在其他几个 ORF 上游, 也有类似的转录调控序列。上述发现, 进一步支持了冠状病毒的不连续转录模型, 也为 ORF 预测的准确性提供了强有力的证据。本文通过体外转染实验, 只对 Orf7、Orf8、Orf9 进行了研究, 首次证明了 Orf7 是 SARS-CoV 的反式转录激活因子。

尚需指出的是, SARS-CoV 与其它冠状病毒同源性不高, 核酸序列相似性分数是 0.56~0.63, 氨基酸序列的相似性分数是 0.57~0.74^[10]。这表明 SARS-CoV 与其它冠状病毒的基因表达调控机理有所不同, Orf7 蛋白的反式转录激活作用机理也明显不同, 暗示 SARS-CoV 的复制、基因表达调控有其独到之处, 还需要进一步研究。加拿大研究组应用已有的工具对 ORFs 这类蛋白质的结构及功能进行了分析, 作出了初步的推断, 但是还需要实验的验证。清楚的了解这些非结构蛋白的功能, 也许有助于 SARS 的预防及治疗。

GFP 是一种生物发光物质, 其荧光的强弱取决于蛋白的含量; GFP 分子量小, 无细胞毒性, 因此可在活细胞体内长期存在。由于稳定表达的细胞株

生长状态正常, 荧光没有衰变, 可在胞内动态观察所研究蛋白质, 并推测其转运过程和代谢通路以及对细胞的影响。因此 GFP 是一种理想的生物分子标签^[14]。本研究构建融合基因, 利用 GFP 表达量的变化观察 ORF 在基因表达调控中的作用, 便于观察且取得良好效果。pDsRed 是 Clontech 公司的第 1 代红色荧光蛋白, 在细胞中通常以四聚体的形式存在, 可溶性不及 EGFP, 毒性也较大。但是在本实验中没有观察到 pDsRed 对细胞的活力有明显的影 响, 只是等量的 pDsRed 和 pEGFP 转染细胞时, pDsRed 较 pEGFP 得到较少的阳性的细胞, 荧光强度也较弱, 并不影响本实验结果的比较。

目前研究表明, SARS 病毒很可能是原先寄生在野生动物体内, 由于与人发生频繁接触后, 传染到人身^[15]。虽然 SARS 近期已从人们视线中消失, 但是这并不能说明它不会卷土重来, 再次引发新一轮的疾病流行。由于冠状病毒的结构和序列变异的非常快, 因此使得找到广谱的疫苗非常困难。靶向一种病毒株的疫苗可能无法有效抵抗另外一种病毒株。人们对病原体进行深入了解是预防和治疗 SARS 的必要手段, 仅仅靠已知的 SARS 病毒蛋白是不足以构成这样凶猛的病毒的, 本文主要着眼于未知的开放阅读框编码的蛋白在转录调控中作用的研究具有重要意义, 其机理及其它的开放阅读框的研究都是都非常重要且有待进行的。彻底弄清楚 SARS 病毒的蛋白组成, 不仅具有极高的科学研究价值, 也将为医学上早日找到预防治疗 SARS 的药物奠定理论基础。

References

- [1] Drosten C, Gunther S, Preiser W, *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome [J]. *N Engl J Med*, 2003, 348 (20): 1967-1976.
- [2] Escors D, Otego J, Laude H, *et al.* The membrane M protein carboxy terminus binds to transmissible gastroenteritis coronavirus core and contributes to core stability [J]. *J Virol*, 2001, 75 (3): 1312-1324.
- [3] Nelson G W, Stohlman S A, Tahara S M. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA [J]. *J Gen Virol*, 2000, 81 (1): 181-188.
- [4] Chen H, Wurm T, Britton P, *et al.* Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell [J]. *J Virol*, 2002, 76 (10): 5233-5250.
- [5] Holmes K V. SARS-Associated Coronavirus [J]. *N Engl J Med*, 2003, 348 (20): 1948-1951.
- [6] van Vliet A L W, Smits S L, Rottier P J M, *et al.* Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus [J]. *The EMBO J*, 2002, 21 (23): 6571-6580.
- [7] Liu J J, Liu S H, Chen Q M, (刘佳建, 刘淑红, 陈启民) *et al.* ORF-1 encoded by bovine foamy virus serve as transactivator itself [J]. *Chin Sci Bull (科学通报)*, 1999, 44 (8): 844-847.
- [8] Sambrook J, Fritsch E, Maniatis T. *Molecular cloning: a laboratory manual* [M]. New York: Cold Spring Harbor Laboratory Press; 1989.
- [9] Holmes K V, Enjuanes L. *Virology. The SARS coronorvirus: a postgenomic era* [J]. *Science*, 2003, 300 (5624): 1377-1378.
- [10] Wu S F, Du C J, Wan P, (吴松锋, 杜春娟, 万平) *et al.* The compare and analysis of genome from SARS-CoV and other coronorvirus [J]. *Hereditas (遗传)*, 2003, 25 (4): 373-382.
- [11] Rota P A, Oberste M S, Monroe, S S. *et al.* Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome [J]. *Science*, 2003, 300: 1394-1399.
- [12] Marra M A, Jones S J M, Astell C R, *et al.* The Genome Sequence of the SARS-Associated Coronavirus [J]. *Science*, 2003, 300: 1399-1404.
- [13] Qin E D, Zhu Q Y, Wang J, *et al.* A complete sequence and comparative analysis of a SARS-associated virus (Isolate BJ01) [J]. *Chin Sci Bull*, 2003, 48 (10): 941-948.
- [14] Scales S J, Pepperkok R, Kreis T E. Visualization of ER to Golgi transport in living cells reveals a sequential mode of action for cop II and cop I [J]. *Cell*, 1997, 90: 1137-1146.
- [15] Li W D, Shi Z, Yu M, *et al.* Bats Are Natural Reservoirs of SARS-like Coronaviruses [J]. *Science*, 2005, 310: 676-679.