

SARS-CoV Sars7a 和 EGFP 融合蛋白真核表达载体构建及其表达*

何贤辉¹, 徐丽慧^{1,2}, 刘毅^{1,4}, 蔡小嫦³, 曾耀英^{1**}(1. 暨南大学组织移植与免疫教育部重点实验室, 广东广州 510632; 2. 暨南大学生物工程研究所, 广东广州 510632;
3. 暨南大学第一附属医院皮肤科, 广东广州 510632; 4. 郑州大学第一附属医院皮肤科, 河南郑州 450052)Construction of an Eukaryotic Expression Vector for SARS-CoV Sars7a
and EGFP Fusion Protein ExpressionHE Xian-hui¹, XU Li-hui^{1,2}, LIU Yi^{1,4}, CAI Xiao-chang³, ZENG Yao-ying^{1**}

(1. Key Laboratory of Tissue Transplantation and Immunology, Ministry of Education, Jinan University, Guangzhou 510632; 2. Institute of Bioengineering, Jinan University, Guangzhou 510632; 3. Department of Dermatology, the First Affiliated Hospital, Jinan University, Guangzhou 510632, China; 4. Department of Dermatology, the First Affiliated Hospital, Zhengzhou University, Zhengzhou 450052, China)

Abstract: Partial overlapping primers were designed based on the severe acute respiratory syndrome coronavirus (SARS-CoV) *sars7a* gene and were chemically synthesized. The *sars7a* gene fragment was obtained by two-round of PCR and this fragment was used as template for a further round of PCR by using a pair of primers to introduce Kozak sequence and to delete the stop codon. The mammalian cell expression vector for Sars7a and enhanced green fluorescent protein (EGFP) fusion protein was generated by inserting the PCR product into pEGFP-N1 vector, with *sars7a* gene upstream to EGFP gene. K562 cells were transfected by the expression vector and the green fluorescence of EGFP could be detected with flow cytometry and confocal microscopy, indicating the expression of Sars7a-EGFP fusion protein. This fusion protein was distributed in the whole cells, which suggested that Sars7a was probable a cytosolic protein rather than a membrane protein. Besides, the expression of Sars7a had no significant effect on the apoptotic cell death of K562 cells.

Key words: Severe acute respiratory syndrome; Coronavirus; Sars7a; Green fluorescent protein; Fusion protein; Expression

摘要: 根据 SARS-CoV *sars7a* 基因设计并化学合成部分重叠引物, 经二轮 PCR 获得 *sars7a* 基因片段, 以此片段为模板并利用一对带有 Kozak 序列及删除终止密码的引物进行 PCR, 获得产物与 pEGFP-N1 载体连接, 使 *sars7a* 基因位于 EGFP 的基因上游, 得到含编码 Sars7a-EGFP 融合蛋白基因的哺乳动物细胞表达载体。采用细胞核转染技术将重组表达载体转染 K562 细胞, 以流式细胞仪和共聚焦显微镜分析, 可检测到 EGFP 的绿色荧光, 表明 Sars7a-EGFP 得到表达, 该蛋白分布于整个细胞, 提示 Sars7a 并非膜蛋白, 更可能是胞浆蛋白。此外, 该蛋白的表达对 K562 细胞凋亡无明显影响。

关键词: 严重急性呼吸综合征; 冠状病毒; Sars7a; 绿色荧光蛋白; 融合蛋白; 表达

中图分类号: Q78

文章标识码: A

文章编号: 1003-5152(2004)03-0209-05

严重急性呼吸综合征 (severe acute respiratory syndrome, SARS) 是一种由病毒引起的急性呼吸系

统感染性疾病, 其病原体为 SARS 冠状病毒 (SARS coronavirus, SARS-Cov) [1-3], 属于一种过去未知的

收稿日期: 2003-11-10, 修回日期: 2004-02-16

* 基金项目: 国家自然科学基金重点项目 (No. 30230350); 国家重点基础研究发展规划“973”项目 (No. G200057006); 暨南大学 SARS 攻关基金和暨南大学自然科学基金资助

作者简介: 何贤辉 (1965-), 男, 浙江宁海籍, 副研究员, 博士, 分子免疫学。

** 通讯作者: 曾耀英 (1946-), 男, 广东台山籍, 研究员, 博士生导师。Corresponding author. Tel: (020)85220732, E-mail: ozms@jnu.edu.cn

冠状病毒^[4,5]。目前已有 30 多个 SARS-Cov 分离株的基因组完成测序,其编码的大多数蛋白质功能都通过生物信息学分析得到初步确立^[4,5],也有一些蛋白在原核获得表达^[6-9],其中一些蛋白的功能得到初步验证^[6-9],但尚有少数几个基因编码的蛋白质的功能无法确定,因为这些推定蛋白 (putative protein) 的序列与所有已知蛋白都无相似性, Sars7a (又称 ORF8 或 X4) 即是其中一个功能未知的推定蛋白^[4,5]。为了研究 Sars7a 蛋白的潜在功能以及其亚细胞定位,我们通过化学合成结合聚合酶链反应 (PCR), 成功构建 *sars7a* 基因, 并与增强型绿色荧光蛋白 (EGFP) 基因融合, 构建 Sars7a-EGFP 融合蛋白表达载体, 将表达载体转入 K562 细胞后, Sars7a-EGFP 融合蛋白得到快速表达, 以 EGFP 的自发绿色荧光为标志, 分析了 Sars7a 在哺乳动物细胞内的表达定位及其对细胞凋亡的影响。

1 材料和方法

1.1 材料

大肠杆菌 DH5 α 购自 Novagen 公司; 真核表达载体 pEGFP-N1 购自 ClonTech; HLA-A*0201 与 EGFP 融合蛋白表达载体 pEGFP-A0201 为本实验室构建; 限制性内切酶 *EcoR* I 和 *BamH* I、T4 DNA 连接酶、高保真 DeepVent DNA Taq 聚合酶等均为美国 New England Biolabs 产品; QIAquick Gel Extraction Kit、QIAprep Spin Miniprep Kit 和 Qiagen Plasmid Midi Kit 为德国 QIAGEN 公司产品; 碘化丙锭 (propidium iodide)、7-氨基放线菌素 D (7-aminoactinomycin D, 7-AAD) 购自 Sigma 公司; RPMI-1640 培养基及胎牛血清 (FBS) 为美国 Gibco 产品; Annexin-V-PE 为 Pharmingen 产品。

1.2 *sars7a* 引物设计

参照 Urbani 株 SARS-Cov cDNA 全序列^[4]设计引物, 所有引物均由上海博亚 (Bioasia) 公司合成, 引物序列如下:

P1: 5'atgaaaattattcttctctgacattgattgtattacatcttgcgagctatatact
atcagagtggttagaggtacgactgtactactaaaagaacctgcccacagg3'
P2: 5'caagcaaaagcaaaagtggtgctagtgcaagttagtgcaaaattattgtcagca
agagggtgaaatggtgaattgccctgctatgttctgatggcgaagttcttttag3'
P3: 5'gcacacacttgcctttgctgtgctgacggtactgcacatacctatcagctgcgtg
Caagatcagttccacaaaacttttcatcagacaagaggagttcaacaagag3'
P4: 5'tcattctgtcttcttcttaattggtgaagcaaaagtattaaaatactagagcagcaa
caatgagaaaaagtgcgagtagagctcttggtaacctctcttctgtctg3'
P5: 5'actagaattcgcaccatgaaaattattcttctctgacattgattg3'
P6: 5'atctgtaccgtttctgtcttcttcttaattggtgaagca3'

1.3 PCR 扩增 *sars7a* 基因及其克隆

为扩增 *sars7a* 基因, 需进行三轮 PCR。第一轮

2 次 PCR 反应: 引物 P1+P2 和 P3+P4, 不加模板; 第二轮 PCR 以第一轮 2 个 PCR 产物为模板, 引物为 P1+P4; 第三轮 PCR 以第二轮 PCR 产物为模板, 引物为 P5+P6。三轮 PCR 反应条件均为: PCR 反应在 50 μ L 总体积中进行, 反应条件为在 94 $^{\circ}$ C 变性 2 min 后开始循环, 然后 94 $^{\circ}$ C 变性 30s, 55 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 30 s, 共 30 个循环后, 于 72 $^{\circ}$ C 延伸 10 min。第三轮 PCR 产物以 *EcoR* I + *Kpn* I 双酶切并以 QIAquick Gel Extraction Kit 从凝胶中回收, 与 pEGFP-N1 (*EcoR* I + *Kpn* I 双酶切) 在 16 $^{\circ}$ C 连接 16 h 后, 转化 *E. coli* DH5 α 感受态细胞, 以碱裂解法从转化子中提取质粒, 双酶切法筛选接入正确外源基因的转化子, QIAprep Spin Miniprep Kit 少量制备质粒, 以 ABI 377 DNA 自动测序仪测定克隆的基因序列 (由 Takara 公司测定)。经测序验证的质粒在 DH5 α 中扩增, 以超纯质粒抽提试剂盒 (Qiagen Plasmid Midi Kit) 大量制备无内毒素转染用载体。

1.4 K562 细胞培养和基因转染

人慢性骨髓性白血病细胞株 K562 细胞 (购自中国典型培养物保藏中心) 以含 10% FBS 的 RPMI-1640 培养液 (加 100 \times 10³ U/L 的青霉素及 100 mg/L 链霉素) 在 37 $^{\circ}$ C 及 5% CO₂ 的湿化培养箱中培养, 每 3 d 传代一次。转染前 2d 以 1 \times 10⁵/mL 细胞接种, 每个转染反应细胞数为 1 \times 10⁶, 质粒 DNA 用量为 3 μ g, 以德国 Amaxa 公司的 K562 细胞优化核转染试剂盒 (Cell line Nucleofector Kit V) 和细胞核转染仪 (Nucleofector Device), 并按其提供的转染程序将表达载体转入细胞核, 转染后的细胞在 24 孔板培养 5 h 和 24 h 分析 EGFP 表达水平。

1.5 流式细胞仪分析

离心收集转染的 K562 细胞, 死亡细胞以碘化丙锭染色 (阳性) 排除, 以 Becton Dickinson 公司的流式细胞仪 FACSCalibur 和软件 CellQuest 获取参数及进行分析。激发光为 $\lambda_{488\text{nm}}$ 激光, 于 $\lambda_{530\text{nm}}$ 检测 EGFP 荧光。所有测定都在相同的仪器设定下进行, 每个样品收集 10000 个细胞的参数。

1.6 激光共聚焦显微镜观察

将转染的 K562 细胞在 24 孔板上培养 24 h 后, 离心收集细胞, 将细胞封于载玻片上, 以激光共聚焦显微镜 (High Speed Confocal Microscope, Perkin Elmer) 进行观察, 激发光为 $\lambda_{488\text{nm}}$ 激光, 细胞产生的荧光以冷 CCD 成像, 并以伪彩色 (绿色) 显示荧光。

1.7 细胞凋亡分析

离心收集细胞, 以 Annexin-V-PE 和 7-AAD 染色分析细胞凋亡。坏死细胞以 7-AAD 染色 (阳性)

排除, 以流式细胞仪进行分析。

1.8 生物信息学分析

以 SignalP^[10] (网址 <http://www.cbs.dtu.dk/services/SignalP/>) 分析潜在的信号肽序列, 以 Tmpred (网址 http://www.ch.embnet.org/software/TMPRED_form.html) 鉴定跨膜区。

2 结果

2.1 sars7a 基因片段扩增及其克隆和鉴定

根据 Sars-Cov sars7a 序列^[4]设计一系列部分序列重叠的引物 (图 1)。首先分别以引物 P1-P2 及 P3-P4 分别配对进行第一轮 PCR, 均扩增出与预期大小 (约 200 bp) 相符的 DNA 片段。回收这二个片段, 作为第二轮 PCR 模板, 并以 P1+P4 为引物进行扩增, 得到与预期大小 (370 bp) 相符的片段。回收 370 bp 片段作为模板, 进行第三轮 PCR 反应, 以带有 Kozak 序列及删除终止密码的一对引物 (P5+P6) 扩增 sars7a 基因, 产物约 400 bp。该片段经 *EcoRI* 和 *KpnI* 双酶切后, 与 pEGFP-N1 载体连接, 转化大肠杆菌 DH5 α , 挑取转化子抽提质粒, 经酶切鉴定表明已接入预计大小的外源片段 (图 2), DNA 序列分析表明, 获得的 sars7a 基因序列与原 SARS-CoV Urbani 株的该段基因编码区序列完全一致, 且该序列 (包括 Kozak 序列) 已正确接至 pEGFP-N1, 位于 EGFP 基因的上游, 其阅读框与 EGFP 基因一致, 获得的表达载体称为 pEGFP-sars

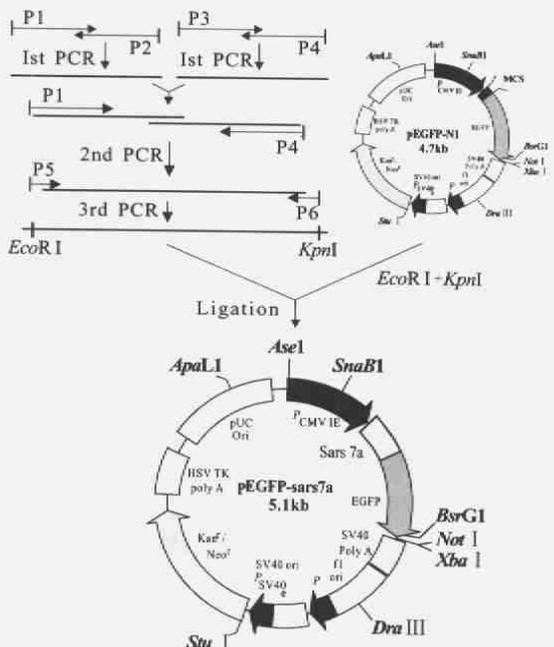


图 1 PCR 扩增 sars7a 基因及表达载体构建图

Fig.1 PCR amplification of sars7a gene and construction of the expression vector

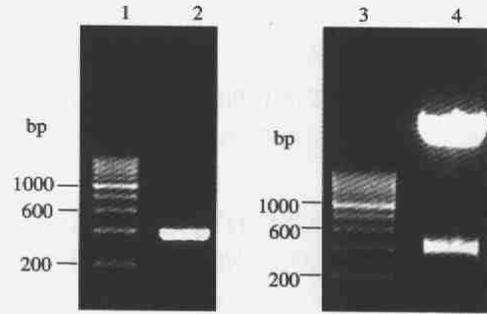


图 2 第三轮 PCR 扩增 sars7a 基因片段及其克隆

Fig.2 Third round PCR amplification of sars7a gene

1/3, 200 bp DNA ladder; 2, Third round PCR product; 4, Recombinant plasmid cut with *EcoRI*+*KpnI*.

7a (图 1)。该合成基因已在 GenBank 登记, 登录号为 AY444813。

2.2 Sars7a-EGFP 融合蛋白在 K562 细胞表达分析

利用细胞核转染仪及其配套试剂盒将 pEGFP-sars7a 转入 K562 细胞, 通过对 EGFP 绿色荧光的检

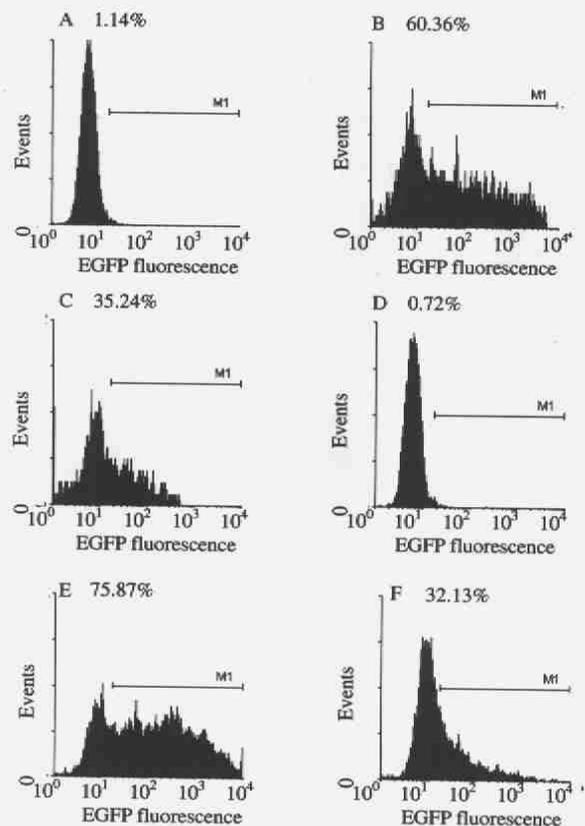


图 3 转染 pEGFP-sars7a 载体的 K562 细胞表达 EGFP 绿色荧光水平

Fig.3 Analysis of the expression levels of EGFP fluorescence in K562 cells transfected with pEGFP-sars7a vector

Numbers are the percentages of EGFP positive cells (M1) in live cells. Data represent one of three independent experiments. Panels A, B, C were 5 h and D, E, F 24 h. A/D, control; B/E, pEGFP-N1; C/F, pEGFP-sars7a.

测表明, Sars7a-EGFP 融合蛋白表达百分率在 5 h 和 24 h 分别为 35% 和 32%, 随时间延长变化不大; 而转入空载体 pEGFP-N1 的细胞 5 h 和 24 h 绿色荧光表达率分别为 60% 和 75%, 随时间延长表达水平有所提高 (图 3)。

为分析 Sars7a-EGFP 的亚细胞分布, 以共聚焦显微镜观察绿色荧光的分布特征, 表明该融合蛋白在大多数细胞中分布于整个细胞, 但在表达水平较高的细胞中可以观察到细胞核中荧光亮度较胞浆高 (彩版图 4C); 作为对照的 EGFP 则较均匀地分布于整个细胞 (彩版图 4B), 未转染的细胞无任何荧光 (彩版图 4A)。作为比较, 表达于 K562 细胞的 HLA-A*0201-EGFP 融合蛋白主要分布于细胞膜和高尔基体等部位 (彩版图 4D)。

2.3 表达 Sars7a-EGFP 融合蛋白的 K562 细胞凋亡分析

在 pEGFP-sars7a 转染 K562 细胞 24h 后, 以 Annexin-V-PE 染色分析细胞凋亡情况。未转染的细胞自发凋亡率 6%, 转染空载体 pEGFP-N1 和 pEGFP-sars7a 的 K562 细胞凋亡率分别为 53% 和 43% (图 5)。进一步对 EGFP 绿色荧光阴性和阳性的细胞的凋亡情况分别分析, 结果显示这些凋亡细胞主要集中于 EGFP 阴性的细胞, 无论是表达 EGFP 或 Sars7a-EGFP 融合蛋白的细胞凋亡水平均较低, 且较接近, 分别为 9% 和 6% (图 6)。

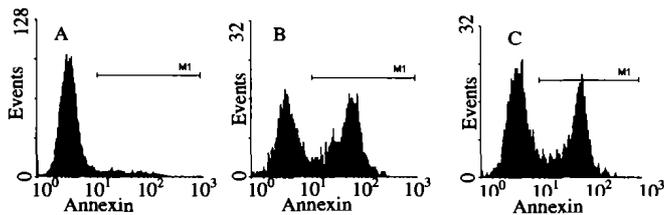


图 5 转染 pEGFP-sars7a 载体的 K562 细胞凋亡水平分析
Fig.5 Analysis of apoptotic levels in K562 cells transfected with pEGFP-sars7a vector

Numbers are the percentages of apoptotic cells with dead cells excluded by 7-AAD staining. Data represent one of three independent experiments. A, control; B, pEGFP-N1; C, pEGFP-sars7a.

2.4 Sars 7a 生物信息学分析

Signalp 分析表明 Sars 7a 的 N-端存在可能的信号肽序列(1-15), TM pred 分析显示序列 99-117 为潜在的跨膜区。

3 讨论

Sars7a (又称 X4 和 ORF8) 是 SARS-CoV 基因

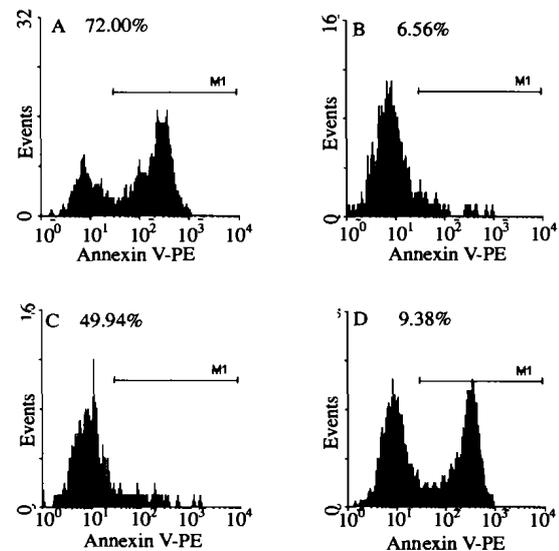


图 6 转染 pEGFP-sars7a 载体的 K562 细胞中 EGFP⁻ 和 EGFP⁺ 细胞凋亡水平比较

Fig.6 Analysis of apoptotic levels in K562 cells transfected with pEGFP-sars7a vector

Numbers are the percentages of apoptotic cells. Data represent one of three independent experiments. A/B, pEGFP-N1; C/D, pEGFP-sars7a. A/D, EGFP⁻ cells; B/C, EGFP⁺ cells.

组编码的一个功能未知的推定蛋白^[4,5], 在 Vero 细胞中表达的 SARS-CoV 的 mRNA 中可检测到编码 sars7a 和 sars8b (X5 或 ORF11) 的转录本 (2.5 kb)^[4], 说明这些蛋白有可能在感染细胞中表达, 但到目前为止还未有该蛋白表达的直接证据。生物信息学分析表明, Sars7a 由 122 个氨基酸组成^[4,5], N-端有潜在的信号肽 (1-15), 并存在一个潜在的跨膜区 (99-117), 提示可能为一个膜蛋白。Marra 等^[5]推测, sars7a 为 I 型膜蛋白, 根据膜定位不同, 其 N-端朝向于内质网/高尔基体腔内, 或位于细胞膜或病毒颗粒表面。我们通过部分化学合成结合 PCR, 获得 sars7a 基因编码区全长, 并成功构建 Sars-7a-EGFP 融合蛋白哺乳动物细胞表达载体, 以绿色荧光蛋白为标志研究了 sars7a 在真核细胞的表达及其表达产物的亚细胞定位。从 Sars7a-EGFP 在 K562 细胞中的表达情况看, 该融合蛋白分布于整个细胞中, 而不呈膜蛋白的分布特征, 说明该蛋白的潜在信号肽并不能有效地将新合成的蛋白质导入内质网; 由于不能进入内质网, 该蛋白的跨膜区不能整合至生物膜, 这表明 Sars7a 并非膜蛋白而更可能是一个胞浆蛋白。这一结果与生物信息学预测及 Marra 等^[5]的推测为 I 型膜蛋白不一致, 说明预测方

法可能存在不足。值得一提的是, 从该蛋白在细胞核区分布较多这一现象来看, 它似乎易于进入细胞核, 推测对细胞核功能有潜在影响。

传统的哺乳动物细胞转染方法转染细胞, 一般需要 2-3 d 后外源基因才能表达, 而利用新型哺乳动物细胞核转染系统进行转染, 被转入的基因在数小时内就有较高水平的表达, 且不依赖于细胞增殖。Sars7a-EGFP 融合蛋白虽然在 5 h 就有明显表达, 但在 24 h 后并未有任何提高, 并且其表达率和表达水平均较单独 EGFP 的低, 这和其它与 EGFP 的 N-端融合的外源蛋白情形相似。同时, 我们尝试分析该蛋白表达与细胞凋亡关系, 表达 Sars7a-EGFP 的细胞的凋亡水平和表达 EGFP 的细胞无明显差别, 说明 Sars7a 对细胞凋亡无明显影响。它对细胞其它功能的影响值得进一步探讨。此外, 我们还构建了原核表达载体, 尝试在大肠杆菌中表达 Sars7a, 但没有观察到明显的表达条带 (未发表数据), 目前正尝试在酵母中进行表达。

总之, 我们成功构建 Sars7a-EGFP 融合蛋白的哺乳动物细胞表达载体, 初步研究表明它并非膜蛋白, 更可能是一个胞浆蛋白, 并且对细胞凋亡无明显影响。这些结果为进一步研究 Sars7a 的功能打下基础。

参考文献

- [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, 348:1967-1976.

- [2] Ksiazek T G, Erdman D, Goldsmith C S, *et al.* A novel coronavirus associated with severe acute respiratory syndrome [J]. *N Engl J Med*, 2003, 348:1953-1966.
- [3] Peiris J S, Lai S T, Poon L L, *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome [J]. *Lancet*, 2003, 361:1319-1325.
- [4] 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.
- [5] Marra M A, Jones S J, Astell C R, *et al.* The Genome sequence of the SARS-associated coronavirus [J]. *Science*, 2003, 300:1399-1404.
- [6] Yang H, Yang M, Ding Y, *et al.* The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor [J]. *Proc Natl Acad Sci USA*, 2003, 100:13190-13195.
- [7] Campanacci V, Egloff M P, Longhi S, *et al.* Structural genomics of the SARS coronavirus: cloning, expression, crystallization and preliminary crystallographic study of the Nsp9 protein [J]. *Acta Crystallogr D Biol Crystallogr*, 2003, 59:1628-1631.
- [8] 俞浩, 杨勇, 张伟, 等. 重组 SARS 冠状病毒刺突蛋白的表达和分离纯化[J]. *生物化学与生物物理学报*, 2003, 35: 774-778.
- [9] Tanner J A, Watt R M, Chai Y B, *et al.* The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases [J]. *J Biol Chem*, 2003, 278: 9578-39582.
- [10] Nielsen H, Engelbrecht J, Brunak S, *et al.* Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites [J]. *Protein Eng*, 1997, 10: 1-6.

彩版

何贤辉, 等. SARS-CoV Sars7a和EGFP融合蛋白真核表达载体构建及其表达 (正文见第209页)

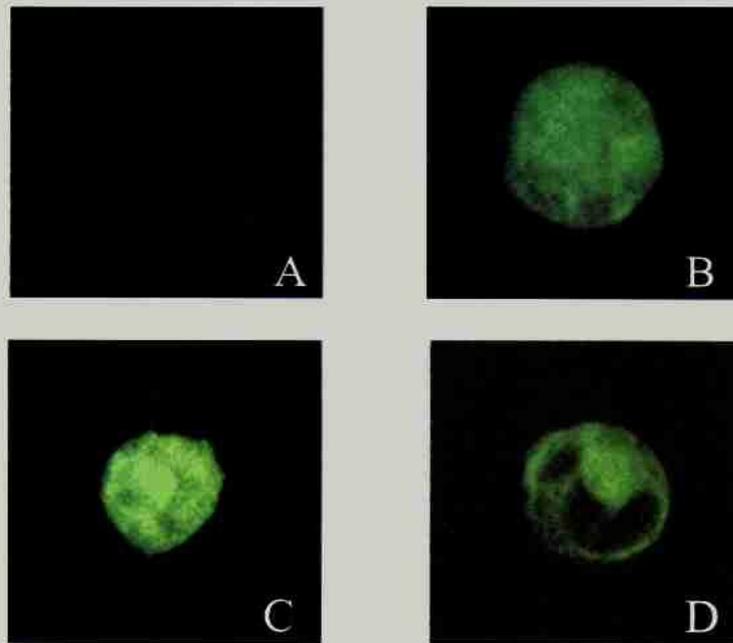


图4 激光共聚焦显微镜观察转染pEGFP-sars7a载体的K562细胞中绿色荧光分布

Fig.4 Laser scanning confocal microscopy analysis of the green fluorescence distribution in pEGFP-sars7a transfected K562 cells

A, Control; B, pEGFP-N1; C, pEGFP-sars7a; D, pEGFP-A0201.

罗保君, 等. 汉滩病毒囊膜糖蛋白G1、G2腺病毒载体的表达及免疫分析 (正文见第224页)

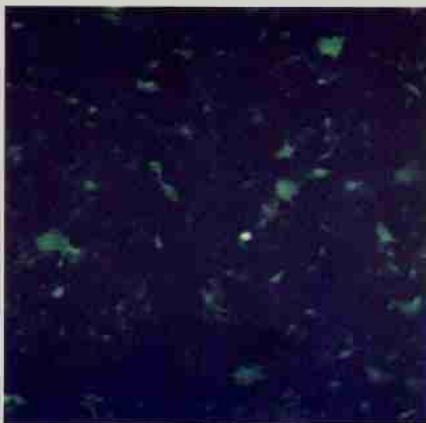


图4 荧光显微镜下重组病毒感染293细胞

Fig.4 293 cells infected by recombinant virus

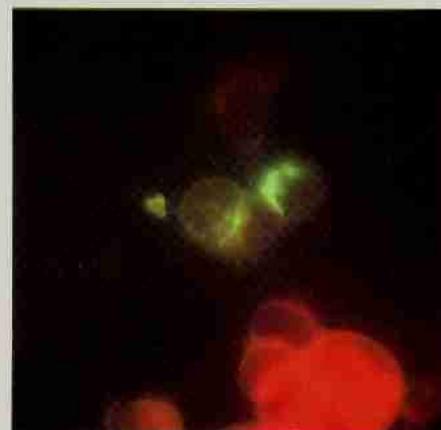


图5 阳性免疫鼠血清间接免疫荧光检测结果

Fig.5 IFA results of immunized mouse sera