

## SARS冠状病毒 S2 基因变异性分析

周浩<sup>1</sup>, 龙北国<sup>1</sup>, 张文炳<sup>1</sup>, 江丽芳<sup>2</sup>, 陈丽丹<sup>1</sup>, 贡树基<sup>1</sup>, 赵卫<sup>1</sup> ( <sup>1</sup>南方医科大学微生物教研室, 广东广州 510515; <sup>2</sup>中山大学中山医学院微生物学教研室, 广东广州 510080 )

**摘要:**目的 测定 SARS 冠状病毒 GD322 株 S2 基因序列并对 SARS-CoV S2 基因进行系统进化树分析。方法 用 RT-PCR 法从 SARS-CoV GD322 株基因组中扩增 S2 基因片段并克隆至 pGEM-T 载体, 转化大肠杆菌 DH5 $\alpha$  株后测序。利用 Lasergene、Clustal X、DNAMAN 和 Treeview 等软件 将基因序列翻译成氨基酸序列后 与早、中、晚期各地暴发的 SARS-CoV S2 蛋白序列进行比对 构建系统进化树 并在 Internet 网数据库中对 S2 蛋白氨基酸序列进行 T 细胞抗原表位预测。结果 随着 SARS-CoV 的传播流行, S2 基因趋于稳定, 晚期的 SARS-CoV S2 基因的同源性达到 99.9%。T 细胞抗原表位预测发现, SARS-CoV S2 蛋白第 57 位氨基酸突变对 T 细胞抗原表位有影响。结论 SARS 流行过程中 SARS-CoV S2 基因较为稳定, S2 蛋白 57 位氨基酸的突变可影响病毒 T 细胞抗原表位。

**关键词:** SARS-CoV; S2 基因; 序列分析

中图分类号: Q754 文献标识码: A 文章编号: 1673-4254(2006)04-0463-03

### Variability analysis of S2 gene of SARS-CoV

ZHOU Hao<sup>1</sup>, LONG Bei-guo<sup>1</sup>, ZHANG Wen-bing<sup>1</sup>, JIANG Li-fang<sup>2</sup>, CHEN Li-dan<sup>1</sup>, GONG Shu-ji<sup>1</sup>, ZHAO Wei<sup>1</sup>

<sup>1</sup>Department of Microbiology, Southern Medical University, Guangzhou 510515, China; <sup>2</sup>Department of Microbiology, Zhongshan Medical College, Sun Yat-sen University, Guangzhou 510080, China

**Abstract: Objective** To determine the sequence of S2 gene of SARS-associated coronavirus (SARS-CoV) GD322 and analyze the phyletic evolution of S2 gene. **Method** S2 gene fragment was amplified from SARS-CoV GD322 genome with RT-PCR and ligated to pGEM-T vector for sequence analysis after transformation of the plasmid into *E.coli* DH5a. The variability of S2 genes and S2 proteins from 12 strains isolated in the early, intermediate and advanced stages of the SARS outbreak were analyzed and the phylogenetic tree was constructed with Lasergene, Clustal X, DNAMAN and Treeview. T cell antigen epitopes of S2 protein were predicted on the basis of Internet database. **Result** With the epidemic spread of SARS-CoV, the S2 genes of the virus tended to become stable. Homology of S2 genes of SARS-CoV isolated in advanced stage of the outbreak reached 99.9%. Prediction of T cell antigen epitope showed that mutation at the 57th amino acid effected T cell antigen epitope. **Conclusion** S2 gene of GD322 SARS-CoV is relatively stable during the epidemic spread of the virus, and mutation at the 57th amino acids of S2 protein may affect the T cell antigen epitope.

**Key words:** Coronavirus, severe acute respiratory syndrome-associated; S2 gene; sequences analysis

严重急性呼吸综合征(SARS)又称传染性非典型肺炎,是一种新发现的传染病,其病原体为 SARS 冠状病毒(SARS-CoV)<sup>[1]</sup>。SARS-CoV S2 基因全长 1 605 bp, S2 蛋白在病毒与宿主细胞的融合、病毒的出芽、胞外分泌以及介导细胞免疫等方面具有重要作用<sup>[2]</sup>。本研究对 SARS-CoV GD322 株 S2 基因测序,并收集早、中、晚期不同地区流行株的 S2 基因序列,利用生物信息学软件,分析了 S2 基因的变异规律。

### 1 材料与方法

收稿日期:2005-09-13

基金项目:广东省科技攻关项目(A2004378)

Supported by Key Science Research Project of Guangdong Province (A2004378)

作者简介:周浩(1980-),男,在读硕士研究生, E-mail: zhou\_hao120@163.com

通讯作者:赵卫, E-mail: zhaowei@fimmu.com

#### 1.1 SARS-CoV GD322 株 S2 基因序列测定

根据 GenBank 提供的 SARS-CoV BJ01 株基因组全序列(AY278488),利用 Lasergene -Primer Select 软件设计引物。上游引物序列是“5'-CCT GTTTC-TATG GCTAAAACC -3'”;下游引物序列是 5'-CTG-GCTGTGCAGTAATTGATC -3'。由上海博亚生物技术有限公司合成。采用 RNeasy Mini Kit(QIAGEN 公司),SARS-CoV GD322(广东地区 SARS 流行早期分离株<sup>[4]</sup>)Vero E6 细胞培养上清中提取 RNA。

提取 RNA 后立即进行逆转录。总体积 20  $\mu$ l,其中 RNase Inhibitor(TaKaRa) 1  $\mu$ l, 20 mmol/L 的 dNTPs 1  $\mu$ l, AMV (TaKaRa) 1  $\mu$ l, 5 $\times$ 缓冲液 4  $\mu$ l, 下游引物 ST25240 1  $\mu$ l, 水 2  $\mu$ l, RNA 模板 10  $\mu$ l。反应条件: 42  $^{\circ}$ C 1 h, 95 $^{\circ}$ C 5 min。

使用 TaKaRa 公司 ExTaq 酶进行 PCR 反应。反应液包括 2.5 mmol 的 dNTPs 3  $\mu$ l, 10 $\times$ PCR buffer 3  $\mu$ l,

上游引物 1  $\mu\text{l}$ 、下游引物 1  $\mu\text{l}$ 、ExTaq 0.5  $\mu\text{l}$ 、cDNA 模板 5  $\mu\text{l}$ 、水 16.5  $\mu\text{l}$ 、总体积 30  $\mu\text{l}$ 。反应条件: 94  $^{\circ}\text{C}$ , 2 min; 94  $^{\circ}\text{C}$ , 30 s; 55  $^{\circ}\text{C}$ , 1 min; 72  $^{\circ}\text{C}$ , 2 min, 30 个循环; 72  $^{\circ}\text{C}$ , 8 min; 4  $^{\circ}\text{C}$  保存。

使用申能博采公司 DNA 纯化试剂盒回收 PCR 产物, 将回收产物与 Promega 公司的 pGEM-T 载体连接, 4  $^{\circ}\text{C}$  放置 12 h。将重组载体转化 DH5 $\alpha$  感受态细胞, 于含氨苄青霉素的固体 LB 培养基上 37  $^{\circ}\text{C}$  培养 16 h, 利用 X-Gal 和 IPTG 筛选阳性重组克隆。获得的阳性克隆菌分别经 PCR 和双酶切初步鉴定, 送 TaKaRa 公司测序。

### 1.2 序列分析

采用 DNAMAN 软件对测序结果与 BJ01 株对应基因核苷酸序列进行同源性比较。另从 GenBank 下载 11 株不同时期、不同地点, 并在流行过程中有联系的 SARS-CoV 的全基因组序列, 分别输入 Lasergene 软件包中的 Editseq 软件中。截取 S2 基因片段, 翻译成氨基酸序列, 利用 Clustal X 软件分析不同毒株间 S2 蛋白氨基酸序列差异。使用 TreeView 软件构建系统进化树。

### 1.3 T 细胞抗原表位预测

在 Internet 网数据库 ([http://bimas.dcrn.nih.gov/molbio/hla\\_bind](http://bimas.dcrn.nih.gov/molbio/hla_bind)) 中对 GD322 株 S2 蛋白的氨基酸序列进行 T 细胞抗原表位预测。

## 2 结果

### 2.1 S2 基因扩增结果

从 GD322 株 Vero E6 细胞培养上清中提取 RNA, 用设计的引物进行 RT-PCR, 经 7% 琼脂糖凝胶电泳可见一约 1 620 bp 的片段, 与预计的 S2 基因大小相符 (图 1)。

### 2.2 S2 基因克隆

将上述 PCR 产物纯化回收后, 与 pGEM-T 载体

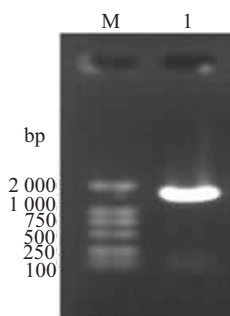


图 1 RT-PCR 法扩增 SARS-CoV GD322 S2 基因  
Fig.1 RT-PCR amplified S2 gene of SARS-CoV GD322

M: DNA Marker DL2000; Lane 1: Amplified S2 gene from SARS-CoV GD322 isolate

连接, 转化 DH5 $\alpha$  感受态细胞, 利用蓝白斑菌落筛选阳性重组克隆。对阳性克隆质粒进行 PCR 和双酶切初步鉴定, PCR 产物经琼脂糖凝胶电泳分析大小约 1 620 bp, 阳性克隆质粒双酶切后有两条带, 一条大小约 1 620 bp, 另一条约 3 000 bp 左右 (图 2)。

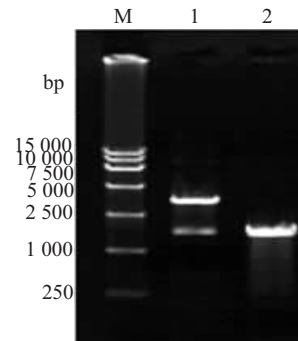


图 2 SARS-CoV GD322 S2 基因重组克隆的鉴定

Fig.2 Identification of recombinant clone of SARS-CoV GD322 S2 gene

M: DNA Marker DL15000; Lane 1: pGEM-T S2 recombinant plasmid after double digestion by *EcoRI/XhoI*;  
Lane 2: Identification by PCR

### 2.3 序列测定及分析

测序结果表明, 所获基因与 BJ01 株 S2 基因核苷酸序列同源性达 99.6%。将其和另外 11 株不同时期不同地区分离的 SARS-CoV S2 基因进行核酸和氨基酸比对 (表 1)。发现与 GD322 株相比, 11 株 S2 基因突变位点有 12 个, 其中非同义突变 8 个。所有的 11 株在相同的变异分别是 106 (D-N)、139 (I-V)、231 (L-F) 和 332 (K-Q)。有 3 株第 57 位氨基酸变异 (Y-D), 分别是早期分离株 GD01、GZ02、GZ50。有 1 株 427 位氨基酸变异 (L-F), 是中期分离株 HZS2-C。

### 2.4 系统进化树分析

使用 Clustal X 软件对 12 株不同时期不同地区分离的 SARS-CoV S2 基因进行比对分析, 并使用 TreeView 软件构建系统进化树 (图 3)。早期分离株 GD322、GZ50、GD01、GZ02 之间存在的差异大于 0.3%, 中期分离株 BJ01、CUUHK-W1、HZS2-Cc 之间的差异小于 0.1%, 晚期分离株同部分中期分离株之间的差异小于 0.1%。

### 2.5 T 细胞表位预测

通过 Internet 网数据库对 GD322 株 S2 蛋白氨基酸序列进行 T 细胞抗原表位预测。结果发现, 只存在一处 T 细胞抗原表位的改变。早期分离株 GD01、GZ02、GZ50 第 57 位氨基酸为 D, 其余各株均为 Y, 该氨基酸的改变引起 T 细胞抗原表位 B<sub>3501</sub> 的改变。

表 1 11 株 SARS-CoV 与 GD322 株 S2 蛋白核酸和氨基酸序列比较  
**Tab.1 Comparison of the nucleic acid and amino acid sequences of GD322 S2 protein with those of another 11 SARS-CoVs**

SARS-CoV isolates	Mutation of S2 gene	Mutation of S2 protein	Origin of isolation	Phase*
GD322			Guangzhou	E
GD01	120(A-T) 169(T-G) 217(C-T)244(C-T) 316(G-A) 415(A-G)691(C-T) 912(T-C) 944(A-C)	57(Y-D) 70(P-S) 106(D-N) 139(I-V) 231(L-F) 332(K-Q)	Zhongshan	E
GZ02	120(A-T) 169(T-G) 244(C-T)316(G-A) 415(A-G) 691(C-T)912(T-C)944(A-C) 1323(A-G)	57(Y-D) 106(D-N)139(I-V) 231(L-F) 332(K-Q) 442(K-E)	Shenzhen	E
GZ50	120(A-T) 169(T-G) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	57(Y-D) 106(D-N)139(I-V) 231(L-F) 332(K-Q)	Guangzhou	E
BJ01	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Beijing	M
HZS2-Cc	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)1279(C-T)	106(D-N)139(I-V) 231(L-F) 332(K-Q) 427(L-F)	Guangdong	M
CUHK-W1	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T) 332(K-Q)944(A-C)	106(D-N)139(I-V) 231(L-F)	Hongkong	M
TOR2	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)1342(A-G)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Toronto	L
PUMC01	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Beijing	L
Sin2500	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Singapore	L
TaiwanTC1	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Taiwan	L
CUHK-AG01	120(A-T) 217(C-T) 244(C-T)316(G-A) 415(A-G) 691(C-T)944(A-C)	106(D-N)139(I-V) 231(L-F) 332(K-Q)	Hongkong	L

\*E: Early stage of SARS epidemic; M: Intermediate stage of SARS epidemic; L: Late stage of SARS epidemic

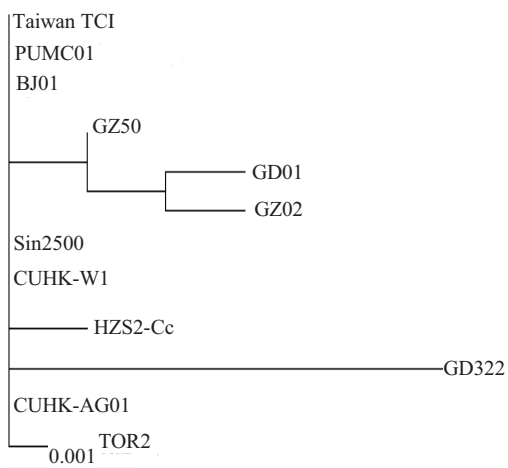


图 3 系统进化树分析

Fig.3 Analysis of phyletic evolution

### 3 讨论

SARS-CoV 是一种新型冠状病毒。各国多个实验室相继从病人标本中分离出 SARS-CoV 并对基因序列进行测定<sup>[3]</sup>。S 蛋白在病毒侵入宿主细胞时发挥重要作用。可分为 S1 和 S2 两部分<sup>[4]</sup>，其中 S1 识别敏感细胞表面的受体，S2 介导病毒包膜与宿主细胞膜融合。通过对包括本实验室分离的 GD322 株在内的 12 株 SARS-CoV S2 基因进行同源性比较和系统进化树

的分析，发现 S2 蛋白在 SARS 流行早期各功能域均有非同义突变，但总体而言 S2 基因的变异早期仅有 12 个碱基，中期 S2 基因间的差异只有 2 个碱基，晚期的 S2 基因同源性则高达 99.9%。可见 S2 基因早期变异明显，中晚期则变异很小。

系统进化树的分析可将不同流行时期和不同流行地域的毒株进行分类。第一类为早期广东省境内流行的 SARS 分离株 GD322、GZ50、GD01、GZ02，它们之间的非同义突变有 7 个，占有非同义突变的 78%。相近地区的早期流行分离株变异很大，其共同特点是 57 位氨基酸的变异 (Y-D)。从 T 细胞抗原表位分析可以看出，57 位氨基酸 (Y-D) 的变异是唯一对 T 细胞抗原产生影响的位点。该突变可能有利于 SARS-CoV 适应人体环境。第二类是由广州传播至香港 Hotel M，再传播至多伦多的 HZS2-Cc、TOR2 分离株<sup>[5, 6]</sup>，这二株之间的变异小于 0.05%。第三类是从世界各地分离的晚期 SARS-CoV BJ01、CUHK-W1、PUMC01、Sin2500、TaiwanTC1、CUHK-AG01。这些毒株之间没有差异，说明随着时间的推移，SARS-CoV S2 蛋白趋于稳定。

对普通冠状病毒 S 蛋白的研究发现，尽管 S2 蛋  
 (下转 471 页)

### 3 讨论

登革病毒为非逆转录 RNA 病毒,在复制周期中不形成 DNA 中间体,直接进行分子生物学操作存在一定困难<sup>[4-6]</sup>。反向遗传学技术为 RNA 病毒分子生物学研究开辟了新途径。主要包括病毒基因组全长 cDNA 克隆技术,以及由质粒 DNA 转录 RNA 的感染性转录体制备技术<sup>[7]</sup>。现在通常先采用长链 RT-PCR 技术扩增出覆盖病毒基因组全长的几个 cDNA 片段,构建亚克隆,再进行体外连接,最终体外转录获得感染性转录体。Srburi 等<sup>[8]</sup>利用类似方法已成功获得了一株 DV2 的感染性克隆。

在研究中,我们首先应用长链 RT-PCR 法一次性扩增出 DV2 NGC 株病毒全长基因组片段,试图与 pCR-XL-TOPO 载体连接,构建全长 cDNA 克隆。但经过多次连接后发现,病毒全长基因组片段不能克隆至该载体。主要原因可能是由于 DV2 病毒基因组较大,超过了 pCR-XL-TOPO 载体的承受范围,一般 pCR-XL-TOPO 载体插入片段长度为 7~10 kb。因此,我们将 DV2 NGC 株病毒基因组分为 2 个片段进行亚克隆的构建,利用病毒基因组序列上所含的酶切位点(*Acl* I)设计了 2 对引物,并在第一条引物 5' 端引入 Sp6 RNA 聚合酶启动子核心序列。采用长链 RT-PCR 技术扩增出覆盖病毒基因组全长的 2 个 cDNA 片段,并克隆至 pCR-XL-TOPO 载体。最终构建出 DV2 NGC 株病毒基因组的 2 个 cDNA 亚克隆,经酶切鉴定和序列测定表明,获得的 cDNA 克隆是 DV2 NGC 株病毒特异的。

DV2 NGC 株病毒基因组全长 cDNA 亚克隆的成功构建,为我们下一步构建出全长 cDNA 克隆及体外转录获得恢复性病毒奠定基础,由此获得 DV2 NGC 株基因库,为开展其它登革病毒分子生物学研究提供方便。

### 参考文献:

- [1] 贡树基,赵卫,曹虹.登革病毒感染实验室诊断的研究进展[J].中国人兽共患病杂志,2005,21(12):1103-5.
- [2] 赵卫,张文炳,蔡鄂德.登革病毒进化遗传学与毒力关系研究进展[J].微生物学免疫学进展,2003,31(2):54-6.
- [3] Vaughn DW, Green S, Kalayanarooj S, et al. Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity[J]. J Infect Dis, 2000, 181(1): 2-9.
- [4] Dar A, Munir S, Vishwanathan S, et al. Transcriptional analysis of avian embryonic tissues following infection with avian infectious bronchitis virus[J]. Virus Res, 2005, 110(1-2): 41-5.
- [5] Khatri M, Palmquist JM, Cha RM, et al. Infection and activation of bursal macrophages by virulent infectious bursal disease virus[J]. Virus Res, 2005, 113(1): 44-50.
- [6] 赵卫,胡志军,杨佩英,等.登革 2 型内嵌合病毒全长 cDNA 克隆的构建[J].微生物学免疫学进展,2001,29(2):8-12. Zhao W, Hu ZJ, Yang PY, et al. Construction of the full-length cDNA clone of intratypic chimeric dengue 2 virus[J]. Prog Microbiol Immunol, 2001, 29(2): 8-12.
- [7] 范宝昌,杨佩英.RNA 病毒的反向遗传学操作[J].军事医学科学院院刊,2001,25(3):223-7. Fan BC, Yang PY. Reverse genetic manipulation on RNA virus[J]. Bull Acad Mil Med Sci, 2001, 25(3): 223-7.
- [8] Suiburi R, Keelapang P, Duangchinda T, et al. Construction of infectious dengue 2 virus cDNA clones using high copy number plasmid[J]. J Virol Methods, 2001, 9(2): 71-2.

(上接 465 页)

白只是组成 S 蛋白柄的亚单位,并不参与病毒与受体最初的识别,但是在 S2 蛋白上具有介导病毒与宿主细胞膜融合的亮氨酸拉链结构。在体内及体外试验中,识别冠状病毒 S2 蛋白的抗体都可以阻断病毒与宿主细胞的融合,使病毒不能感染细胞,是一种具有中和活性的抗体<sup>[7]</sup>。从 T 细胞抗原表位分析看出,57 位氨基酸变异(Y-D)对 T 细胞抗原表位产生影响。研究发现,T 细胞抗原表位 B<sub>3501</sub> 结构的精细差别影响 MHC I 类分子与 CD8<sup>+</sup> 的 T 细胞的结合<sup>[8]</sup>。S2 蛋白某个碱基的突变就有可能改变病毒对宿主细胞的侵袭能力<sup>[9]</sup>。同时 S2 蛋白在病毒进化的后期非常稳定,是抗病毒药物及疫苗研制的理想靶位。

### 参考文献:

- [1] Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus associated with severe acute respiratory syndrome[J]. N Engl J Med, 2003, 348(20): 1953-66.
- [2] Krueger DK, Kelly SM, Lewicki DN, et al. Variations in disparate regions of the murine coronavirus spike protein impact the initiation

of membrane fusion[J]. J Virol, 2001, 75(6): 2792-802.

- [3] Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome[J]. Science, 2003, 300(5624): 1394-9.
- [4] Marra MA, Jones SJ, Astell CR, et al. The genome sequence of the SARS associated coronavirus[J]. Science, 2003, 300(5624): 1399-404.
- [5] He JF, Peng GW, Min J, et al. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China[J]. Science, 2004, 303(5664): 1666-9.
- [6] Yeh SH, Wang HY, Tsai CY, et al. Characterization of severe acute respiratory syndrome coronavirus genomes in Taiwan: molecular epidemiology and genome evolution[J]. Proc Natl Acad Sci USA, 2004, 101(8): 2542-7.
- [7] Fagchi F. The S2 subunit of the murine coronavirus spike protein is not involved in receptor binding[J]. J Virol, 1995, 69(10): 7260-3.
- [8] Gao GF, Willcox BE, Wyer JR, et al. Classical and nonclassical class I major histocompatibility complex molecules exhibit subtle conformational differences that affect binding to CD8alphaalpha[J]. J Biol Chem, 2000, 275(20): 15232-8.
- [9] Guan Y, Zheng BJ, He YQ, et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China[J]. Science, 2003, 302(5643): 276-8.