

# SARS 冠状病毒基因片段变异分析

王义权<sup>1</sup>, 姚冠华<sup>2</sup>, 牛建军<sup>2</sup>, 黄磊<sup>1</sup>, 纪念念<sup>1</sup>, 李奇渊<sup>1</sup>, 黄建炜<sup>2</sup>,  
邵寒娟<sup>1</sup>, 陈亮<sup>1</sup>, 章军<sup>1</sup>, 宋思扬<sup>1</sup>, 彭宣宪\*

(1. 厦门大学生命科学学院, 细胞生物学与肿瘤细胞工程教育部重点实验室, 福建 厦门 361005; 2. 厦门市卫生防疫站, 福建 厦门 361004)

**摘要:** 采用基因序列和生物信息分析法, 研究了 SARS-CoV 复制酶基因中 4 个片段. 结果发现, 片段 I 与已报道的 SARS 冠状病毒 ZJ01 有 2 个位点不同, 片段 III 与 CUHK-W1、CUHK-Su10、HKU-39849 和 Hong Kong 各有 1 个位点的差别; 片段 IV 与 GZ01 间有 1 个位点的变化. 同时对已公布的 17 个全基因组序列进行比对分析, 可找到共 137 个变异位点, 其中仅出现 1 次的变异位点 119 个, 间约信息位点 18 个.

**关键词:** SARS-CoV; 病毒; 变异; 序列分析

**中图分类号:** Q 933.44

**文献标识码:** A

急性严重呼吸道综合征(severe acute respiratory syndrome, SARS) 是一种由 SARS 病毒引起的新急性呼吸道传染病, 流行于全球<sup>[1, 2]</sup>. 目前认为, SARS 病毒是一种新型冠状病毒, 单列为冠状病毒第 4 组, 缩写为 SARS-CoV<sup>[3-5]</sup>. 对从患者体内分离出的 14 株 SARS-CoV 的全基因分析发现, 病毒存在 129 处变异, 提示 SARS-CoV 的基因序列变异性较高<sup>[6]</sup>. SARS-CoV 变异及其意义是学术界高度重视的问题. 本文对 SARS-CoV 的 4 个基因片段进行序列分析, 并与 GenBank 公布的序列进行了比较.

## 1 材料和方法

### 1.1 病毒核酸提取

SARS-CoV 来源于厦门市卫生防疫站保存的材料. 采用 Trizol 一步法在厦门市卫生防疫站 P3 实验室提取, 并经 DNase I 处理以去除其中的 DNA.

### 1.2 RT-PCR 步骤

用 TaKaRa 公司 RNA LAMP<sup>TM</sup> 试剂盒, 按其

说明书进行操作.

### 1.3 PCR 产物的克隆

1) PCR 产物的回收: 将 PCR 产物用 1.5% 琼脂糖, 120 V 电泳 30 min 后用 Watson's 小量胶回收试剂盒回收目的 DNA 片段.

2) 连接: 用 TaKaRa 公司 pMD 18-T Vector 连接试剂盒将回收的 DNA 片段与 pMD 18-T Vector 相连. 连接体系 15  $\mu$ L, 16  $^{\circ}$ C 反应过夜.

3) 转化: 用 15  $\mu$ L 的连接产物, 转化 100  $\mu$ L 经 CaCl<sub>2</sub> 处理后的 JM101 感受态细胞, 在含有 Amp 的 LB 琼脂平板培养基上培养, 形成单菌落.

4) 克隆子的 PCR 检测: 挑取平板上的单菌落, 接种于含 Amp 的 LB 液体培养基中, 37  $^{\circ}$ C 摇床培养 3 h; 直接吸取 0.5  $\mu$ L 菌液作为 PCR 的模板, 用 M13 Primers 引物, BioAsia PCR 试剂盒, 15  $\mu$ L PCR 反应体积, 进行“菌落 PCR”检测; 扩增产物经 1.5% 琼脂糖凝胶电泳, 检测扩增片段大小与预期产物相符的克隆用于测序.

### 1.4 测序反应

取上述单克隆菌落的含 Amp LB 菌液(菌体浓度较高时) 3 mL, 用 Watson's 小量质粒抽提纯化试剂盒提取 PCR 鉴定为阳性的重组质粒, 再经 1.5% 琼脂糖电泳估测质粒 DNA 浓度. 测序反应采用

收稿日期: 2003-06-18

基金项目: 厦门市科技计划资助课题

作者简介: 王义权(1957-)男, 教授, 博士生导师.

\* Corresponding author

CEQ™ Dye Terminator Cycle Sequencing (DTCS) 试剂盒, 分别用引物 M13 Primers M4 和 M13 Primers RV 双向测序. 反应体系为 DTCS Premix 6 μL, 质粒 DNA 模板 120 ng, 引物 5 pmol, 补水至 10 μL, 测序反应循环为 96℃ 20 s, 50℃ 20 s, 60℃ 4 min. 反应结束后用反应终止液(NaAc 3mol/L, pH 5.2, 1 μL; EDTA 100 mmol/L, pH 8.0, 1 μL; Glycogen 20 mg/mL, 0.5 μL) 终止反应, 然后加入-20℃预冷的 95% 乙醇 30 μL, 4℃ 12 000 r/min 离心 20 min; 弃上清后, 再用-20℃预冷的 70% 冰乙醇 200 μL 洗涤沉淀, 4℃ 12 000 r/min 离心 5 min, 弃上清, 重复洗涤沉淀 2 次. 沉淀物经真空抽干后用 30 μL 甲酰胺溶解. 测序用 Beckman 公司 CEQ 8 000 遗传分析仪进行.

## 2 结 果

本研究得到的 4 个片段的序列共 730 bp, 这些片段均位于 SARS-CoV 基因组中复制酶 1AB(replicase 1AB) 蛋白编码区, 与 GenBank 中公布的序列经 Blast 分析发现, 片段 I 除了与 ZJ01 相比, 在 15534 位点有一个 A 插入(位点编号与 TOR2 一致, 下同)

和在 15568 处有一个 T 缺失外, 与其余已公布的序列完全一致. 片段 II 与全部已公布的该段序列完全一致. 片段 III 与 CUHK-W1 和 CUHK-Su10 相比, 在 17846 处有一个 T → C 的转换; 与 HKU-39849 和 Hong-Kong 相比在 18065 外有一个 G → C 的转换, 其它各处序列与已公布的序列相同. 片段 IV 与绝大多数 SARS-CoV 序列完全一致, 仅在 18 167 的位置, 与 GZ01 比较有一个 A → G 的转换(表 1).

## 3 讨 论

自 SARS-CoV 被发现以来, 迄今已有 17 个病毒株的全基因序列在 GenBank 中公布. 对这些 SARS-CoV 基因组序列的研究表明, 该病毒存在较多的变异位点. 分析比较不同来源的病毒株序列, 有助于疾病传播源的调查<sup>[6]</sup>, 我们对已公布的这 17 个全基因组序列进行比对分析, 共找到 137 个变异位点, 其中仅出现 1 次的变异位点 119 个, 间约信息位点 18 个. BJ01 和 BJ02 在 8573 位置出现了 G → T 的颠换; 在 9405 位点 BJ01, BJ02, BJ03, GZ01 和 CUHK-W1 出现了 T → C 的转换; GZ01 和 CU-

表 1 4 个片段的序列与 GenBank 中已知序列 Blast 结果\*  
Tab. 1 Results of the sequence blast with SARS-CoV in GenBank

| 毒株名              | 片段 I                             | 片段 II     | 片段 III       | 片段 IV        |
|------------------|----------------------------------|-----------|--------------|--------------|
| TOR2             | Identical                        | Identical | Identical    | Identical    |
| BJ01             | Identical                        | Identical | Identical    | Identical    |
| BJ02             | Identical                        | Identical | Identical    | Identical    |
| BJ03             | Identical                        | Identical | Identical    | Identical    |
| BJ04             | Identical                        | Identical | Identical    | Identical    |
| CUHK-W1          | Identical                        | Identical | T → C(17846) | Identical    |
| CUHK-Su10        | Identical                        | Identical | T → C(17846) | Identical    |
| HKU-39849        | Identical                        | Identical | G → A(18065) | Identical    |
| Hong Kong        |                                  |           | G → A(18065) | Identical    |
| GZ01             | Identical                        | Identical | Identical    | A → G(18167) |
| ZJ01             | - → 插入 A(15534)<br>T → 缺失(15568) | Identical | Identical    | Identical    |
| TW1              | Identical                        | Identical | Identical    | Identical    |
| Taiwan JG 2003   | Identical                        |           |              |              |
| Taiwan RN A dire | Identical                        |           |              |              |
| Urbani           | Identical                        | Identical | Identical    | Identical    |
| isolate SIN2500  | Identical                        | Identical | Identical    | Identical    |
| isolate SIN2677  | Identical                        | Identical | Identical    | Identical    |
| isolate SIN2679  | Identical                        | Identical | Identical    | Identical    |
| isolate SIN2748  | Identical                        | Identical | Identical    | Identical    |
| isolate SIN2774  | Identical                        | Identical | Identical    | Identical    |
| Frankfurt        | Identical                        | Identical | Identical    | Identical    |
| Vietnam strain   | Identical                        |           |              |              |

\*. 空白处为 GenBank 中该毒株缺少相应片段序列的数据, 括弧内数字表示在 TOR2 中的碱基位置

HK-W1 还在 9480 位再次出现 T<sup>→</sup>C 的转换; BJ01 ~ BJ04 在 9855 位出现 C<sup>→</sup>T; BJ01 和 GZ01 在 10588 位还出现了 A<sup>→</sup>C 的颠换; 在 17570 位点 BJ01~ 04, GZ01 和 CUHK-W1 出现了 T<sup>→</sup>G 的转换; 在 17852 位置 CUHK-Su10 和 CUHK-W1 均为 C<sup>→</sup>C; 在 19070 位置 CUHK-W1 和 Urbani 由 A<sup>→</sup>G; 在 19090 处, 5 个 insulate 样品中, 除了 SIN2679 外, 其它样品全部由 C<sup>→</sup>T; 在 19844 位点, BJ01~ BJ04 和 GZ01 均由 A<sup>→</sup>G; 在 21727 处, BJ01~ BJ03, GZ01 和 CUHK-W1 出现了 G<sup>→</sup>A 的颠换. 在 22228 位点 BJ01~ 04, GZ01 和 CUHK-W1 均由 T<sup>→</sup>C; 在 22428 处 BJ02 和 GZ01 出现 G<sup>→</sup>A; 在 22523 处这两株序列又由 A<sup>→</sup>G; 在 25305 处 BJ02 和 BJ03 出现了 G<sup>→</sup>A; 在 26056 处 BJ01 和 BJ03 出现了 A<sup>→</sup>C; 在 27249 处, BJ01~ BJ04 和 GZ01 均出现 C<sup>→</sup>T; 在 27834 处 BJ01~ BJ04, GZ01 和 CUHK-W1 出现 T<sup>→</sup>C. 此外, GZ01 还在 27 891~ 27 919 间有一 39 bp 的插入.

本文的测序与既往研究不同的是, 我们选择的是该病毒的保守区, 以考查在这样的区段中可能发生的变异情况. 结果表明, 所分析的 3 个基因片段均与目前国际报道的 SARS-CoV 序列没有明显差别, 其中仅片段 I 与 ZJ01 有 2 个位点不同; 片段 III 与 CUHK-W1、CUHK-Su10、HKU-39849 和 Hong-

Kong 各有 1 个位点的差别; 片段 IV 与 GZ01 间有 1 个位点的变化. 我们还将对该株病毒的基因序列进行深入研究.

### 参考文献:

- [1] Nicholls J M, Poon L L M, Lee K C, et al. Lung pathology of fatal severe acute respiratory syndrome [J]. *Lancet*, 2003, 361: 1 773- 1 778.
- [2] Peiris J S M, Chu C M, Cheng V C C, et al. Clinical progression and viral load in a community outbreak of coronavirus associated SARS pneumonia: a respective study [J]. *Lancet*, 2003, 361: 1 767 1 772.
- [3] Peiris J S M, Lai S T, Poon L L M, et al. Coronavirus as a possible cause of severe acute respiratory syndrome [J]. *Lancet*, 2003; 361: 1 319- 1 324.
- [4] Marra M A, Jones S J M, Astell C R, et al. The Genome Sequence of the SARS Associated Coronavirus [J]. *Science*, 2003, 300: 1 399- 1 404.
- [5] 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: 1 394- 1 399.
- [6] Ruan Y J, Wei C L, Ee L A, et al. Comparative full length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection [J]. *Lancet*, 2003, 361: 1 779- 1 780.

## Analysis of Variations in Gene Fragments of SARS-CoV

WANG Yi quan<sup>1</sup>, YAO Guang hua, NIU Jian jun<sup>2</sup>, HUANG Lei<sup>1</sup>, JI Niannian<sup>1</sup>,  
LI Qi yuan<sup>1</sup>, HUANG Jian wei<sup>2</sup>, SHAO Hair juan<sup>1</sup>, CHEN Liang<sup>1</sup>,  
ZHANG Jun<sup>1</sup>, SONG Si yuan<sup>1</sup>, PENG Xuannian<sup>1</sup>

(1. The Key Laboratory of Education Ministry for Cell Biology and Tumor Cell Engineering, School of Life Sciences, Xiamen University, Xiamen 361005, China; 2. The Center for Diseases Control of Xiamen, Xiamen 361004, China)

**Abstract:** Four gene fragments from replication enzyme gene of SARS-CoV, namely fragment I, fragment II, fragment III and fragment IV, were investigated with combination of sequence analysis and bioinformatics. Our results showed that there were several variations existing in three of the four fragments, showing that fragment I had two variance against GZ01, fragment III had 1 nuclear acid change comparing with CUHK-W1, CUHK-Su10, HKU-39849 and Hong-Kong respectively, and fragment IV had 1 point mutation against GZ01. In addition, alignment analysis of 17 complete sequences downloaded from GeneBank revealed 137 variable sites, of which 18 sites were of parsimony informative. Our findings are valuable for further study on variations of SARS-CoV.

**Key words:** SARS-CoV; Virus; Variation; Sequence