

## 我国森林脑炎病毒森张株编码区序列测定

马新英<sup>1</sup>, 司炳银<sup>2</sup>, 高 轩<sup>1\*\*</sup>, 彭文明<sup>2</sup>, 祝庆余<sup>2</sup>

(1. 河北农业大学动物科技学院, 河北保定, 071000; 2. 军事医学科学院微生物学流行病学研究所, 北京, 100071;)

**Nucleotide Sequence of the Polyprotein of the *Tick-borne encephalitis virus* Senzhang strain Isolated in China**MA Xin-ying<sup>1</sup>, SI Bing-yin<sup>2</sup>, GAO Xuan<sup>1\*\*</sup>, PENG Wen-ming<sup>2</sup>, ZHU Qing-yu<sup>2</sup>

(1, Institute of Animal Science and Technology, Agricultural university of Hebei, Baoding 071000, China; 2, Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, China)

**Abstract :** To assign the *Tick-borne encephalitis virus*(TBEV) Senzhang strain isolated in China to one of three TBEV subtypes and understand the relation between the polyprotein structure and biological function and aid in the development of new TBEV vaccines, the nucleotide sequence of the polyprotein of Senzhang strain was determined. Nine pairs of primers were designed according to the nucleotide sequence of Sofjin-HO, Oshima 5-10 strain. By RT-PCR, cDNA fragments of Senzhang strain were acquired from infected BHK-21 cells. Then the cDNA fragments was cloned into the vector pGEM<sup>®</sup>-T and then transformed to competent DH5  $\alpha$  cells. Positive clones were screened and identified by *Eco*R I cleavage and PCR. The sequences of inserted fragment were determined by 3100 Genetic Analyzer. Sequence analysis showed that Senzhang strain contains a sigle open reading frame (ORF) of 10245 nucleotides (nt) which encodes a polyprotein of 3414 amino acids(aa). Senzhang strain is more similar to Oshima5-10, Sofjin-HO, Sofjin, and assigned to the Far Eastern Subtype.

**Key words:** *Tick-borne encephalitis virus*; Polyprotein; Sequence analysis

**摘要:** 为鉴定我国森林脑炎病毒亚型, 了解基因组结构与生物功能的关系, 同时为森林脑炎病毒新型疫苗研制打下基础, 对森林脑炎病毒森张株编码区序列进行测定。根据已发表的 Sofjin-HO、Oshima5-10 株序列设计 9 对重叠引物, 通过 RT-PCR 扩增不同的 cDNA 片段, 分别克隆至 pGEM<sup>®</sup>-T 载体, 转化 DH5  $\alpha$  菌, 挑阳性克隆 PCR、酶切鉴定后测序。结果表明森张株编码区全长 10 245nt, 编码 3 414 个氨基酸。我国森林脑炎病毒森张株与 Oshima5-10、Sofjin-HO、Sofjin 同源性最近, 属于远东亚型。

**关键词:** 森林脑炎病毒; 多聚蛋白; 序列分析

中图分类号: R373

文章标识码: A

文章编号: 1003-5125 (2003) 04-0322-04

森林脑炎病毒(*Tick-borne encephalitis virus*, TBEV)是黄病毒科黄病毒属成员, 重要的致人疾病病毒之一。感染后可引起发热、循环系统障碍、神经系统障碍。根据系统进化学森林脑炎病毒分为远东亚型、西伯利亚型、欧洲亚型, 此种分类与 E 蛋白同源性相一致<sup>[1]</sup>。远东亚型毒力最强, 临床症状

重, 脑神经症状明显, 病人恢复期长, 常留有麻痹性后遗症, 病死率达 20%。欧洲亚型毒力最弱, 临床症状轻, 以脑膜炎为主, 较少出现后遗症, 死亡率为 1%~5%。TBEV 流行于欧洲、亚洲的苏联及日本、中国等地, 在我国主要分布于东北、云南、新疆, 西藏林芝地区血清抗体阳性率很高<sup>[2,3]</sup>。

收稿日期: 2002-12-26, 修回日期: 2003-03-07

作者简介: 马新英 (1977-), 女, 河北保定籍, 硕士研究生, 主要从事人畜共患病研究。E-mail: ndmxy@163.net

\*\* 通讯作者。Correspondence author.

本文基因序列已登录 GenBank, 序列号为 AY182009。

森林脑炎病毒基因组为长约 11kb 的单股正链 RNA, 由单一阅读框架编码三种结构蛋白 (C、M (prM)、E)、七种非结构蛋白 (NS1、NS2a、NS2b、NS3、NS4a、NS4b、NS5)。E 蛋白是最重要的病毒蛋白, 与病毒的受体结合、融合活性、血凝活性有关, 诱导宿主保护性免疫应答<sup>[4]</sup>。Mandl 等根据晶体衍射结果将其病毒膜外部分分成三个结构域: I、II、III, 并对其跨膜区功能元件进行了研究<sup>[5,6]</sup>。1997 年侯宗柳等对我国东北株 HLJ-1 的 E 蛋白核苷酸序列进行了测定<sup>[7]</sup>, 但对我国分离 TBEV 其它蛋白基因核苷酸序列测定未见报道。TBEV 森张株于 1953 年从我国东北某林场病尸脑中分离, 现对其整个编码区进行了测定, 结果报告如下。

## 1 材料与方法

### 1.1 病毒株

森张株 TBEV Senzhang, 1953 年从东北某林场病尸脑中分离, 鼠脑传第 37 代, 由军事医学科学院微生物流行病学研究所病毒室保存。

### 1.2 试剂及工具酶

TRIZOL、SUPERSRIPT™ II 及 Taq DNA Polymerase (GIBCOBRL 公司); PCR Fragment Recovery Kit (宝生物工程(大连)有限公司); pGEM®-T Easy Vector System I 及 Wizard® Plus SV Minipreps DNA Purification System (Promega 公司); 5' /3' RACE Kit (Roche 公司)。

### 1.3 引物设计与合成

根据 GenBank 中已公布的 Sofjin-HO、oshima5-10 株核苷酸序列设计引物, 由上海生物工程公司合成。

### 1.4 总 RNA 的提取

将  $10^{-3}$  浓度森张株乳鼠脑悬液接种于 BHK-21 细胞单层 (3.5cm×9cm 空斑瓶) 上, 待细胞病变至+时, 弃细胞上清, 每瓶细胞加入 2mL TRIZOL 试剂。按 TRIZOL 法提取总 RNA, 详细操作见 TRIZOL 说明书。

### 1.5 RT-PCR 扩增

取总 RNA 8μL, 加入 2μL 10μmol/L 的反向引物, 70℃ 水浴 10min。取出立即置冰上, 加入 4μL 5× buffer、2μL 10μmol/L 的 dNTP、2μL 0.1mol/L 的 DTT、1μL SUPERSRIPT™ II、1μL Rnasin, 42℃ 水浴 60min。随后 70℃ 水浴 10min。取 2-3μL cDNA, 用 GIBCOBRL 公司的 Taq DNA Polymerase 进行 PCR 扩增。

### 1.6 PCR 产物克隆及测序

将 PCR 产物经 1% 的琼脂糖电泳后, 用 PCR Fragment Recovery Kit 回收, 与 pGEM®-T Easy Vector 连接, 转化感受态的 DH5α 菌, 经 PCR 和酶切鉴定后, 阳性克隆交给军事医学科学院八所测定。采用荧光标记双脱氧测序法利用 Applied Biosystems 公司生产的 3100 Genetic Analyzer 进行序列测定<sup>[8]</sup>。

### 1.7 序列结果处理

序列用计算机软件 DNASTar 进行编辑、校对与拼接。在 GenBank 上进行了同源性比较, 并从中检索 6 株森林脑炎病毒编码区序列、12 株森林脑炎病毒 E 蛋白序列, 利用 DNASTar 进行氨基酸比较并根据 E 蛋白构建系统发生树。

## 2 结果

### 2.1 森张株编码区序列测定结果

整个编码区由 10245 个核苷酸组成, 四种碱基含量分别为 A 2544 (24.8%)、C 2326 (22.7%)、G 3235 (31.6%)、T 2140 (20.9%)。编码三种结构蛋白、七种非结构蛋白, 共 3414 个氨基酸。其中组氨酸含量最高, 为 9.5%。密码子 GCC、AUG 使用频率最高, 为 3.0%; UUA 使用频率最低, 为 0.25%。GenBank 登录号为 AY182009。

### 2.2 核苷酸、氨基酸同源性比较

将森张株核苷酸序列在 Transeq 上翻译成氨基酸, 并在 DDBJ 上进行了同源性比较, 与 TBE 病毒 Oshima5-10 株、Sofjin-HO 株、Sofjin 株、Vasilchenko 株、Neudoerfl 株、263 株、Hypr 株的核苷酸同源性分别为 94%、93%、93%、85%、83%、83%、83%, 氨基酸同源性分别为 95%、95%、94%、92%、90%、90%、90%, 而与 Louping ill 病毒核苷酸、氨基酸同源性分别为 82%、88%。

同 Oshima5-10、Sofjin-HO、Vasilchenko、Neudoerfl、263、Hyp 相比, 森张株存在 16 个特异性氨基酸变异, 包括 Thr<sub>72</sub>→Met、Lys<sub>77</sub>→Arg、Val<sub>254</sub>→Ile、Val<sub>271</sub>→Ala、Ala<sub>443</sub>→Pro、Asn<sub>646</sub>→Thr、Lys<sub>893</sub>→Arg、Leu<sub>922</sub>→Pro、Phe<sub>1100</sub>→Ile、Val<sub>1493</sub>→Aln、Ile<sub>1631</sub>→Thr、Val<sub>2103</sub>→Ile、Arg<sub>2671</sub>→Lys、Val<sub>2963</sub>→Ala、Ile<sub>3102</sub>→Val、Lys<sub>3390</sub>→Arg, 这些氨基酸仅森张株发生变异, 其余六株均保守。通过比较还发现 13 个亚型特异氨基酸, 见表 1。

### 2.3 系统进化树分析

根据 E 蛋白序列采用 DNASTar 中的 clustal 方法进行系统进化树分析, 结果见图 1。TBE 有三个主要分支: 远东亚型分支、西伯利亚亚型分支、欧洲

亚型分支。森张株位于远东亚型分支内。

表 1 亚型特异氨基酸  
Table 1 Subtype-special amino acids

位置 Position	远东亚型 Far Eastern Subtype.	西伯利亚亚型 Siberian Subtype.	欧洲亚型 European Subtype.
3	Gly	Arg	Lys
486	Ser	Leu	Val
597	Ile	Thr	Ala
1061	Arg	Lys	Thr
1228	Asn	Gly	Ser
1302	Met	Ile	Val
1303	Leu	Ile	Val
1615	Ile	Met	Leu
1747	Val	Met	Ala
1865	Ile	Val	Ala
2280	His	Gln	Arg
2355	Ala	Ser	Thr
3182	Val	Ile	Leu

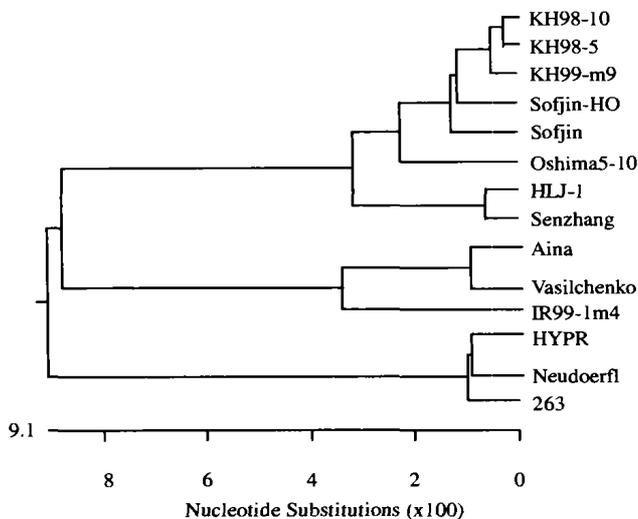


图 1 TBE 的系统进化树

Fig.1 TBE phylogenetic tree

Virus strain (GenBank No.): Sofjin(X07755), Sofjin-HO(AB022703), Senzhang(AY182009), HYPR(X75286), Aina(AF091006), Vasilchenko (M 97659); Neudoerfl(U27495), 263(U27491), Oshima5-10(AB001026), KH98-5 (AB022296), KH98-10 (AB022297), IR99-1m4 (AB049349), KH99-m9 (AB049346), HLJ-1(No.).

### 3 讨论

本文首次测定我国森林脑炎病毒森张株整个编码区核苷酸序列,全长 10 245bp。同源性比较结果表明:森张株与 Sofjin-HO 株、Sofjin 株同源性较近,与 HLJ-1 株、Oshima5-10 株同源性最近,很可能与 HLJ-1 株、Oshima 5~10 株共同起源于 Sofjin

株,因此属于远东亚型<sup>[1]</sup>。

森张株 C 蛋白内的 72 位、77 位变异发生在假定螺旋 III 上,结构与 Leucine zipper motifs 相似,每圈包含 7 个 aa,可能参与形成卷曲螺旋间相互作用,此处变异可能影响病毒核衣壳的形成<sup>[9]</sup>。254 位、271 位变异发生在 M 蛋白跨膜区内,可能影响病毒装配。E 蛋白有两个变异,一个是发生在 I 区  $\beta$ -折叠 F<sub>0</sub> 上的 443 位变异, F<sub>0</sub> 与 A<sub>0</sub>、C<sub>0</sub>、D<sub>0</sub>、E<sub>0</sub> 同位于 E 蛋白外侧面,参与 I 区、III 区间相互作用,但由于 Ala、Pro 都是非极性氨基酸,疏水性、所带电荷相差不大,故该点变异对此影响可能不大。另一个是 646 位变异,发生在 III 区  $\beta$ -折叠 E、D<sub>x</sub> 的连接处,  $\beta$ -折叠 A、B、E、D 与 I 区相对,并与 II 区的 cd 环相接触, D<sub>x</sub> 与 A<sub>x</sub>、C<sub>x</sub> 一起覆盖 III 区的  $\beta$ -折叠桶顶。此处变异可能影响 I 区与 III 区间的相互作用、低 pH 值诱导的 E 蛋白构象转变、膜融合过程中 cd 环的暴露,从而影响森林脑炎病毒的宿主范围、细胞敏感性及毒力<sup>[5]</sup>。NS1、NS3、NS5 蛋白内分别有 3、3、4 个变异,共同影响病毒复制<sup>[10]</sup>。

氨基酸序列比较时发现的 13 个亚型特异性氨基酸很可能是三个病毒亚型特性不同的原因之一,486 位氨基酸 Ecker.M 等<sup>[1]</sup>曾报道过。如多聚蛋白第三位氨基酸在 Kunjin 病毒中已证实对病毒基因组 RNA 的包装、病毒的复制都很重要,三个病毒亚型此位置上的差异也可能影响到病毒基因组 RNA 的包装、病毒的复制,引起生物学性状、致病性上的差异<sup>[9]</sup>。486 位氨基酸位于 E 蛋白 II 区  $\beta$ -折叠 g、f 间的  $\alpha$  螺旋上,参与低 pH 值下诱导的 E 蛋白构象转变,从而影响病毒毒力。597 位氨基酸位于 E 蛋白 III 区的  $\beta$ -折叠 A 上, Ile、Thr、Ala 疏水性上存在一定差异,可能影响 I 区、III 区间作用力,从而影响 E 蛋白构象转变及病毒毒力<sup>[5]</sup>。

森林脑炎病毒森张株编码区序列的测定,有助于进一步研究病毒之间核苷酸序列差异与病毒致病性关系,了解病毒的复制、装配及发病机理,以及弄清我国流行株的分子基础,为研制和开发相应的森林脑炎疫苗奠定基础。

致谢:感谢黄如统老师审校本文初稿,自登云老师为本文提供资料。

### 参考文献

- [1] Ecker M, Allison S L, et al. Sequence analysis and genetic classification of tick-borne encephalitis viruses from Europe and Asia[J]. J Gen Virol, 1999, 80: 179-185.

- [2] 自登云, 陈伯权, 俞永新. 虫媒病毒与虫媒病毒病[M]. 1995: 192-201.
- [3] 马仕金, 卢登明, 张有植. 西藏林芝地区森林脑炎血清抗体调查[J]. 中华预防医学杂志. 1996, 30: 346.
- [4] Heinz F X, Mandl C W. The molecular biology of tick-borne encephalitis virus[J]. APMS, 1993, 101: 735-745.
- [5] Rey F A, Heinz F X, Mandl C W, *et al.* The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution[J]. Nature. 1995, 375: 291-298.
- [6] Allison S L, Stiasny K. Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E[J]. J Virol, 1999, 73: 5605-5612.
- [7] 侯宗柳, Heinz F X, Ecker M, 等. 蜱传脑炎病毒东北株 E 蛋白的核酸和氨基酸序列[J]. 病毒学报, 1997, 13(1): 47-53.
- [8] 萨姆步鲁克 J, 弗里奇 EF, 曼尼阿蒂斯 T(金东雁, 黎孟枫等译). 分子克隆实验指南[M]. (第二版)北京: 科学出版社, 1992. 34-66.
- [9] Kofler R M, Heinz F X, Mandl C W. Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favorable target for attenuation of virulence[J]. J Virol, 2002, 76: 3534-3543.
- [10] Lindenbach B D, Rice C W. Genetic interaction of flavivirus non-structural proteins NS1 and NS4A as a determinant of replicase function[J]. J Virol, 1999, 73: 4611-4621.

(上接第 314 页)

**作者(Authors):** 采用汉语拼音 (姓氏字母全部大写置前, 双名的前名首字母大写、后名全部小写, 两名中间加“-”符隔开, 姓名均不缩写。如: WANG Da-ming), 作者间用“,”隔开, 国外作者遵从所在国习俗。

**单位英译全称** (单位名称, 省市名 邮编, China)。

**Abstract:** 研究论文和简报均需撰写英文摘要, 英文摘要要求同中文摘要, 可稍详细)。

**Key words:** 3-5 词组, 词组间用“;”隔开, 单词首字母大写。

**2.1.5 中文摘要:** 应简要介绍本文的目的意义、主要方法、结果及结论, 应有独立性和自含性, 一气呵成, 不分段, **不另列小标题**, 不附图表, 不用非公知公用符号、略写语。简报不需要撰写中文摘要, 但需撰写英文摘要。

**2.1.6 中文关键词:** 3-5 个, 词组间用“;”隔开。

**2.1.7 关键词下行标注** “**中图分类号:** …”、“**文献标识码:** …”“**文章编号:** …” (内容由本刊填写)

**2.1.8 文章首页地脚处:**

收稿日期: YYYY-MM-DD, 修回日期: YYYY-MM-DD (由本刊填写)

\* **基金项目:** 基金名称 1 (编号); 基金名称 2 (编号)

**作者简介:** (第一作者) 姓名 (出生年-), 性别 (民族), 籍贯, 职称, 学位, 从事专业或研究方向。

\*\* **通讯作者:** (内容同上, 也可根据通讯作者的要求省略, 但应附电话、电传和 E-mail 地址)。

**2.1.9 正文**

**2.1.10 文章正文内标题层次标号**采用阿拉伯数字分级编号法, 一级通常为“1 材料和方法”, “2 结果”, “3 讨论”; 二级为 1.1、1.2……, 三级为 1.1.1、1.1.2……。

**2.1.11 参考文献集中附于文后。**

### 3. 文稿撰写要求

**3.1 文稿**请用计算机打印, 字体请勿小于 5 号宋体字。文中大小写、正斜体、上下角、以及希腊文、罗马数字均应打印清楚。

**3.2 使用法定的计量单位。**如: 分子量用 **Mr**; 溶液浓度单位不用 **M**、**N** 而用 **mol/L**; 体积单位不用 **l** 而用 **L** (**ml**、**μl** 分别应为 **mL**、**μL**); 时间单位用 **s**、**min**、**h**、**d** 分别表示秒、分、小时、天; 离心机转速用 **r/min** 取代 **rpm**, 或用 **g** (离心力) 表示。

**3.3 蛋白质和核酸名称的英文缩写**一律用正体, 首字母大写或全部大写; **限制性内切酶的前 3 个字母用斜体**, 如 *BamH I*。基因座名称的英文缩写一律用小写斜体, 基因位点用正体大写, 如 *aroG*; 基因表达产物和蛋白名称用正体, 首字母大写, 如 *AroG*。正式定名的病毒的科属及毒株名称为斜体, 首字母大写, 如 *Potato leafroll virus*

(下转第 375 页)