

# 副粘病毒 Tianjin 株 NP、P、M 及 L 蛋白的生物信息学分析

石立莹 李梅 李晓眠\* 何健民 袁立军

(天津医科大学基础医学院微生物学教研室 天津 300070)

**摘要:** 为了进一步明确副粘病毒 Tianjin 株的来源和种系进化地位, 探讨其高致病性的机制。对 Tianjin 株 NP、P、M 及 L 蛋白进行了生物信息学分析。进化树显示: Tianjin 株属于副粘病毒亚科呼吸道病毒属, 且很可能为仙台病毒新的基因型。相似性比较表明, P 蛋白变异最大, 相似性仅为 78.7%~91.9%; L 蛋白相似性最高, 为 96.0%~98.0%。序列比对显示: NP 蛋白氨基酸序列中存在 15 个独特的变异位点, P 蛋白存在 29 个, M 蛋白存在 6 个, L 蛋白存在 29 个。这些独特变异位点的存在很可能是导致 Tianjin 株在宿主来源和致病特点等方面与已知仙台病毒株具有较大差异的原因。

**关键词:** 副粘病毒, 仙台病毒, 狨猴, 生物信息学

## Bioinformatics Analysis of the NP, P, M and L Protein of Paramyxovirus Tianjin Strain

SHI Li-Ying LI Mei LI Xiao-Mian\* HE Jian-Min YUAN Li-Jun

(Department of Microbiology, College of Basic Medical Science, Tianjin Medical University, Tianjin 300070)

**Abstract:** In order to demonstrate the taxonomic position of paramyxovirus Tianjin strain and explore its mechanism of pathogenesis. Bioinformatics methods were used to analyze the deduced amino acid sequences of NP, P, M, and L protein of Tianjin strain. Phylogenetic analysis based on NP, P, M, and L protein sequences demonstrated that Tianjin strain belonged to the genus *Respirovirus*, in the subfamily *Paramyxovirinae* and most likely a new genotype of Sendai virus. Sequence similarities comparisons indicated that Tianjin strain P protein was poorly conserved, sharing only 78.7%~91.9% amino acid identity with 6 known Sendai viruses, while L protein was the most conserved, having 96.0%~98.0% amino acid identity with other Sendai viruses. Multiple-sequence alignments of Tianjin strain NP, P, M, and L protein with those of 6 known Sendai viruses showed that Tianjin strain possessed a lot of unique amino acid substitutions in protein sequences, 15 in NP, 29 in P, 6 in M, and 29 in L. The presence of these unique amino acid substitutions suggests that Tianjin strain maybe has a significant difference in host or pathological characteristics from the known Sendai viruses.

**Keywords:** Paramyxovirus, Sendai virus, Marmoset, Bioinformatics

副粘病毒为有包膜的负链RNA病毒,是一类重要的人类和动物致病病毒,如尼帕病毒(Nipah virus)、亨德拉病毒(Hendra virus)、Menangle病毒等。1999年6月,天津医科大学实验动物中心饲养的普通棉耳狨猴群体中暴发急性呼吸道传染病,发病率几乎100%,病死率为33%<sup>[1]</sup>。为了明确病因,在除外细菌感染的基础上,我们将死亡狨猴的肺组织匀浆后接种鸡胚培养,分离得到一株高血凝效价的病毒。通过电镜观察病毒形态、交叉血凝抑制试验、RT-PCR、互联网Blast服务器联配检索等方法初步证实该分离株可能是一株副粘病毒<sup>[2]</sup>,暂时命名为副粘病毒Tianjin株。

在病原鉴定过程中,我们曾经通过RT-PCR方法扩增得到该毒株一段长为375 bp的片段,并将测序结果发送到互联网Blast服务器进行联配检索,发现其与仙台病毒(Sendai virus, SeV)HN基因同源性较高,但推导出的氨基酸序列还存在较大差异。更重要的是,仙台病毒是啮齿类动物致死性病原体,然而在普通棉耳狨猴呼吸道疾病流行期间,饲养在同一实验中心的各种实验用鼠无一例发病,因此我们推测该毒株可能是仙台病毒发生了某些变异而跨越了种属界限,形成对啮齿类动物低致病性,而对灵长类动物高致病性的毒株;或者是一株未被发现的能引起灵长类动物呼吸道感染的新的副粘病毒。为了明确该毒株的来源,探讨其高致病性的机制,我们对Tianjin株全基因组进行了测序。Tianjin株基因组由15384个核苷酸组成,与已知仙台病毒基因组长度完全相同。从基因组的3'端至5'端依次排列着3'UTR,结构基因NP、P、M、F、HN和L编码区,以及5'UTR。Tianjin株基因组各基因的起始、终止位置及基因间隔区序列与仙台病毒具有较高的相似性。

在本研究中,我们对Tianjin株预测编码的结构蛋白NP、P、M、及L蛋白进行生物信息学分析(关于HN及F蛋白的分析在其他文章中已有论述<sup>[3,4]</sup>),其结果不仅有助于阐明Tianjin株蛋白结构与功能的关系,而且还将为进一步研究基因突变与病毒生物活性乃至致病性的关系提供分子生物学依据。

## 1 材料与方法

### 1.1 材料

副粘病毒 Tianjin 株:登录号 EF679198。用于

分析的其他副粘病毒,来源于副粘病毒科7个属以及目前尚未归类的副粘病毒,共25株,具体包括:(1)Avulavirus属:禽副粘病毒6型(Avian paramyxovirus 6, APMV-6, NC\_003043)、新城疫病毒(Newcastle disease virus, NDV, NC\_002617);(2)Henipavirus属:Nipah virus(NiV, NC\_002728)、Hendra virus(HeV, NC\_001906);(3)Morbillivirus属:犬瘟热病毒(Canine distemper virus, CDV, NC\_001921)、海豚麻疹病毒(Dolphin morbillivirus, DMV, NC\_005283)、麻疹病毒(Measles virus, MeV, NC\_001498)、Peste-des-petits-ruminants virus (PPRV, NC\_006383)、牛瘟病毒(Rinderpest virus, RPV, NC\_006296);(4)Respirovirus属:仙台病毒(Sendai virus Hamamatsu, SeV ha, AB039658)、人副流感病毒3型(Human parainfluenza virus type 3, hPIV-3, NC\_001796)、人副流感病毒1型(Human parainfluenza virus 1, hPIV-1, NC\_003461)、牛副流感病毒3型(Bovine parainfluenza virus 3, bPIV-3, NC\_002161);(5)Rubulavirus属:人副流感病毒2型(Human parainfluenza virus 2, hPIV-2, NC\_003443)、腮腺炎病毒(Mumps virus, MuV, NC\_002200)、猿副流感病毒5型(Simian parainfluenza virus 5, SV5, NC\_006430)、Tioman virus (TiV, NC\_004074)、Menangle virus (MenV, NC\_007620);(6)Pneumovirus属:牛呼吸道合胞病毒(Bovine respiratory syncytial virus, bRSV, NC\_001989)、人呼吸道合胞病毒(Human respiratory syncytial virus, hRSV, NC\_001781);(7)Metapneumovirus属:禽后肺病毒(Avian metapneumovirus, AMPV, NC\_007652)、人后肺病毒(Human metapneumovirus, HMPV, NC\_004148);(8)未分类的副粘病毒:Fer-de-lance virus (FDLV, NC\_005084)、J virus (NC\_007454)、Tupaia paramyxovirus (TPMV, NC\_002199)。

用于分析的6株仙台病毒代表株:BB1株(DQ219803)、Hamamatsu株(AB039658)、Nagoya株(AB195968)、Ohita株(NC\_001552)、pi株(AB195967)、Z株(M30202)。

### 1.2 分子进化分析

利用BioEdit软件采用ClustalW方法,将Tianjin株NP、P、M及L蛋白氨基酸序列与副粘病毒科中25株分别来自7个属及未分类的副粘病毒相

应序列进行联配分析, 用 ProtDist 程序绘制系统进化树, 用 TreeView 软件生成树文件。通过进化树的结构来确定 Tianjin 株在副粘病毒科中的分类地位。然后采用同样方法将 Tianjin 株 NP、P、M 及 L 蛋白与 GenBank 中选取的 6 株仙台病毒代表株相应序列进行列队比较, 并构建系统进化树, 用 TreeView 软件生成树文件。

### 1.3 相似性比较

使用 DNASTar 软件包中的 MegAlign 程序将推导出的 Tianjin 株 NP、P、M 及 L 蛋白氨基酸序列分别与 GenBank 中选取的仙台病毒相应序列进行相似性比较。

### 1.4 变异点分析

通过与已知仙台病毒相应蛋白氨基酸序列进行序列比对, 找出 Tianjin 株 NP、P、M 及 L 蛋白氨基酸序列中独特的变异位点, 分析这些变异位点对蛋白结构及功能可能造成的影响。

## 2 结果

### 2.1 进化树分析

将 Tianjin 株 NP、P、M 及 L 蛋白氨基酸序列与来自于副粘病毒科 7 个属以及尚未分类的 25 株病毒相应序列进行系统进化分析。从得到的进化树来看(图 1 显示出 NP 及 M 蛋白进化树), 除 J-V、TPMV 和 FDLV 三个未分类病毒分别独立成一支外, 其他病毒被划分为 7 个分支, 与已经建立的分类学结果一致, 即 7 个分支分别对应副粘病毒科 7 个已知属。Tianjin 株与 Respirovirus 属的 SeV、hPIV-1、hPIV-3 和 bPIV-3 位于同一个分支, 且与仙台病毒亲缘关系较近, 表明 Tianjin 株属于副粘病毒科、副粘病毒亚科、呼吸道病毒属, 很可能为仙台病毒种系中新的成员。

将 Tianjin 株 NP、P、M 及 L 蛋白氨基酸序列分别与 GenBank 中选取的 6 株仙台病毒相应序列进行系统进化分析。结果显示, 7 株病毒被分为 4 个分支, 其中 Hamamatsu 株和 Ohita 株为一个分支, 具有近似来源的 Z 株、Pi 株及 Nagoya 株为一个分支, Tianjin 株和北京登录的 BB1 株各自独立成一个分支。而且 NP、P、M 和 L 蛋白的进化树基本一致, 仅有一些细微的差别, 在 M 蛋白进化树中, Tianjin 株与 Ohita 株和 Hamamatsu 株代表的分支最接近, 而在其他蛋白中, Tianjin 株均是是与 BB1 株亲缘关系最

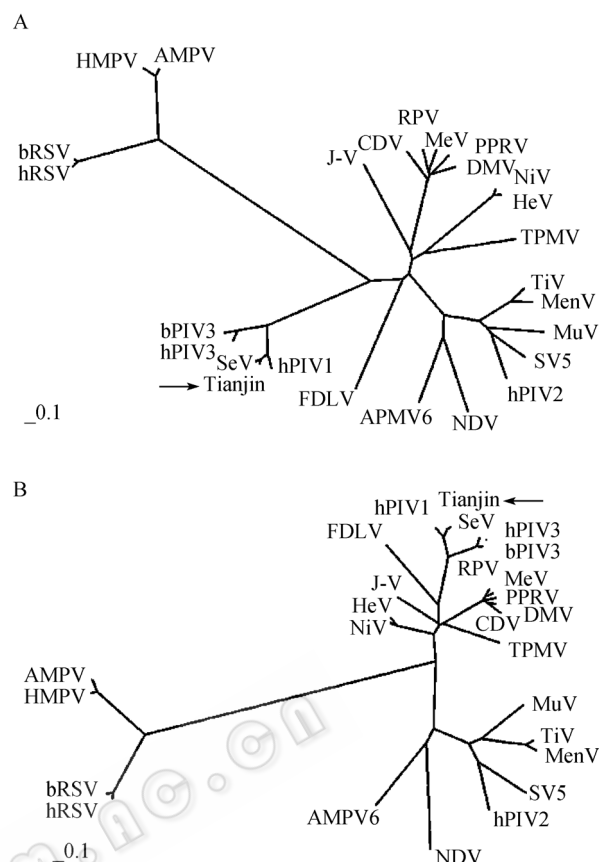


图 1 副粘病毒 Tianjin 株与 25 株副粘病毒 NP 蛋白(A)及 M 蛋白(B)的进化树

Fig. 1 Phylogenetic trees based on the complete NP (A) and M protein (B) sequences of Tianjin strain and other Paramyxoviridae

接近(图 2 显示出 NP 及 M 蛋白进化树)。

### 2.2 相似性比较

将 Tianjin 株 NP、P、M 及 L 蛋白氨基酸序列与 6 株仙台病毒相应序列进行相似性比较, 结果显示(见表 1), NP 蛋白氨基酸序列相似性在 92.4%~96.0%, P 蛋白相似性在 78.7%~91.9%, M 蛋白相似性在 93.7%~97.1%, L 蛋白相似性在 96.0%~98.0%。说明 L 蛋白较为保守, 而 P 蛋白差异较大。

### 2.3 变异位点分析

通过与已知仙台病毒 NP、P、M 及 L 蛋白氨基酸序列的序列比对发现, Tianjin 株 NP 蛋白氨基酸序列中存在 15 个独特的变异位点, P 蛋白存在 29 个独特的变异位点, M 蛋白存在 6 个独特的变异位点, L 蛋白存在 29 个独特的变异位点, 而且 M 蛋白在 83 位、106 位、158 位、251 位和 295 位共含有 5 个保守的 Cys 残基(见表 2、3、4、5)。

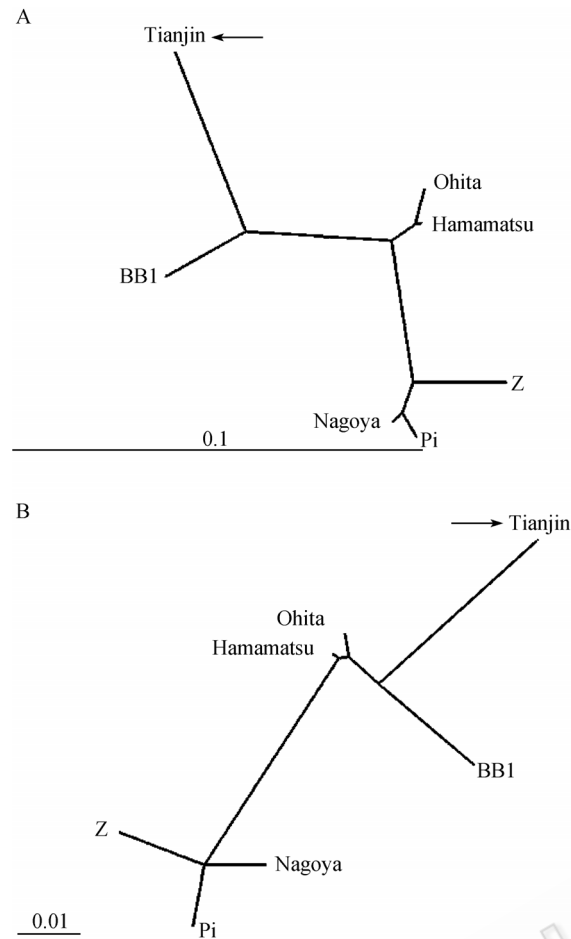


图2 副粘病毒 Tianjin 株与 6 株仙台病毒 NP 蛋白(A)及 M 蛋白(B)的进化树

Fig. 2 Phylogenetic trees based on the complete NP (A) and M protein (B) sequences of Tianjin strain and 6 known SeVs

表 1 副粘病毒 Tianjin 株与 6 株仙台病毒 NP、P、M 及 L 蛋白氨基酸序列相似性比较(%)				
Table 1 Similarities comparisons of amino acid sequences of NP, P, M, and L protein between Tianjin strain and 6 known SeVs (%)				
病毒株 Strain	NP	P	M	L
Ohita	94.1	83.7	97.1	97.1
Hamamatsu	94.5	84.9	97.1	97.4
Pi	92.8	78.9	94.0	96.1
Nagoya	93.0	78.7	94.0	96.2
Z	92.4	80.0	93.7	96.0
BB1	96.0	91.9	96.6	98.0

3 讨论

副粘病毒 Tianjin 株基因组中含有 6 个结构基因,可能编码 NP、P、M、F、HN 和 L 六种结构蛋白。

表 2 副粘病毒 Tianjin 株 NP 蛋白独特的氨基酸变异位点		
Table 2 Unique amino acid substitutions in NP proteins sequence of Tianjin strain		
变异位点 Substitution sites	已知仙台病毒 The known SeVs	Tianjin 株 Tianjin strain
14、481	Arg	Lys
121	Lys	Arg
150	Asp	Glu
156	Thr/Ala	Val
221	Ser	Ala
293	Ile	Leu
372	Ala	Thr
446	Gly	Asp
472	Asn/Lys	Ser
485	Arg/Met	Lys
493	Glu/Asp	Asn
503	Arg	Leu
508	Asp	Asn
515	Ala/Thr	Val

表 3 副粘病毒 Tianjin 株 P 蛋白独特的氨基酸变异位点		
Table 3 Unique amino acid substitutions in P protein sequence of Tianjin strain		
变异位点 Substitution sites	已知仙台病毒 The known SeVs	Tianjin 株 Tianjin strain
21	Gly	Glu
25、137、239	Leu	Phe
71	Glu	Lys
72、166	Gly	Arg
86、557	Glu	Asp
88	Ser	Ile
103	Ala	Val
119	Ala/Thr	Ile
123	Ser/Ala	Val
208	Val	Leu
231、304	Pro	Ser
242	Val/Ala	Glu
255	Asp/Tyr	Asn
258、322、500	Thr	Ile
263	Arg/Gly	Glu
264	Lys	Thr
294、443	Ser	Pro
387	Asp	Glu
491	Arg	Cys
534	Ala	Ser
540	Val	Ile

表 4 副粘病毒 Tianjin 株 M 蛋白独特的氨基酸变异位点 Table 4 Unique amino acid substitutions in M protein sequence of Tianjin strain		
变异位点 Substitution sites	已知仙台病毒 The known SeVs	Tianjin 株 Tianjin strain
17	Thr	Ala
38	Val	Ile
139	Gly	Ser
165	Ile	Thr
244	Lys	Arg
323	Phe	Leu

表 5 副粘病毒 Tianjin 株 L 蛋白独特的氨基酸变异位点 Table 5 Unique amino acid substitution in L protein sequence of Tianjin strain		
变异位点 Substitution sites	已知仙台病毒 The known SeVs	Tianjin 株 Tianjin strain
6	Ser	Phe
115	Ala	Ser
125、629	Ser	Ala
186	Met	Ile
271	Glu/Arg	Lys
278	Ser	Pro
306、1987	Ile	Met
326	Ser/Asp	Asn
333	Pro/Ala	Leu
337	Ala/Thr	Val
514	Glu	Lys
623	Ser/Gly	Asp
718	Met	Thr
1090、1516、2108	Val	Ile
1404	Asp	Asn
1550	Val	Leu
1604、2173	Arg	Lys
1653	Val/Ile	Met
1722、2212	Thr	Ile
1861	Gly	Ser
1924	Ala	Thr
1998	His/Leu/Asn	Lys
2002	Lys	Arg

我们推测, 和仙台病毒相似, HN 和 F 蛋白是两种跨膜糖蛋白, 位于囊膜外表面, 分别形成大、小纤突, 是重要的宿主保护性抗原。M 蛋白是非糖基化蛋白, 构成囊膜内表面的支撑物。NP、P 及 L 蛋白可包裹病毒基因组 RNA 形成 RNP 复合体, 共同参与病毒 RNA 的转录与复制。

为了进一步明确副粘病毒 Tianjin 株的分类地位, 我们绘制了包括 Tianjin 株在内的副粘病毒科 26 株病毒 NP、P、M 及 L 蛋白的系统进化树。从得到的

进化树来看, Tianjin 株与 SeV、hPIV-1、hPIV-3 和 bPIV-3 位于同一个分支, 且与仙台病毒亲缘关系最接近。进一步与 GenBank 中 6 株仙台病毒代表株相应序列构建系统进化树, 进化树显示, Tianjin 株不属于现有的 3 个分支, 而是独立为一个分支, 且与北京登录的仙台病毒 BB1 株亲缘关系较为接近, 我们考虑可能是地域对毒株进化的影响。各蛋白氨基酸序列相似性比较同样支持这一结果, Tianjin 株与 BB1 株相似性较高, 但二者之间的差异明显大于同一分支内各病毒之间的差异, 说明不属于同一分支。因此, Tianjin 株在分类上属于副粘病毒科、副粘病毒亚科、呼吸道病毒属, 而且很可能为仙台病毒新的基因型。

Tianjin株NP蛋白编码区为 1575 nt, 编码 524 个氨基酸残基。研究证实, 仙台病毒NP蛋白除了与病毒基因组RNA结合外, 还与多种病毒蛋白作用, 如 NP0-P, NP-NP。缺失分析显示<sup>[5-7]</sup>, 高度保守的aa 258~357 区域与NP-NP相互作用有关; aa 472~491 影响NP0-P复合物的形成而影响复制; aa 462~471 是核衣壳与P-L复合物中的P蛋白结合所必需的, 而aa 440~461 是NP自身装配形成核衣壳所必需的。Tianjin株NP蛋白氨基酸序列中存在 15 个独特的变异位点, 其中G446D、N(K)472S、R481K和R(M)485K与NP蛋白以上功能区有关, 是否影响到这些区域的功能还有待研究。

Tianjin株P蛋白编码区为 1707 nt, 编码 568 个氨基酸残基, 为高度磷酸化蛋白。研究表明, P蛋白可与L蛋白结合, 形成P-L RNA多聚酶复合体, 结合位点位于aa 412~478<sup>[8]</sup>; 另外, P蛋白两个不连续区域 aa 345~411 及aa 479~568 为P-L复合体与核衣壳的结合位点<sup>[9,10]</sup>。Tuckis等<sup>[11]</sup>通过基因定点突变技术对仙台病毒NP结合区, 即P蛋白N端 88 个氨基酸进行了研究, 发现aa 506~509、aa 524~526 和aa 560~562 是这一区域的关键氨基酸, 突变或缺失后均影响到RNA聚合酶活性。Tianjin株P蛋白氨基酸序列中存在 29 个独特变异位点, 其中有 5 个位于NP结合区(aa 479~568), 但是在 aa 506~509、aa 524~526 和 aa 560~562 位置均未发生变异, 提示这些区域可能具有重要的功能, 同时也揭示了Tianjin株可能具有同样的RNA聚合酶活性。

Tianjin株M蛋白编码区为 1047 nt, 编码 348 个氨基酸残基。氨基酸序列比对分析发现, Tianjin株M蛋白在 83 位、106 位、158 位、251 位和 295 位含

有5个Cys残基,与已知仙台病毒相同。而且这些Cys残基在副粘病毒亚科中也是保守的,说明它们可能具有重要的生物学功能。Sakaguchi等<sup>[12]</sup>曾经利用基因突变方法证实,M蛋白中单个Cys残基突变可破坏M蛋白空间构象的稳定性,进而影响到仙台病毒形态和核衣壳的形成,说明这些Cys残基对于M蛋白维持病毒形态的功能起重要作用。

Tianjin株L基因全长为6690 nt,共编码2229个氨基酸残基。通过对不分节段的负链RNA病毒L蛋白序列比对分析发现<sup>[13]</sup>,这些病毒L蛋白中均存在6个高度保守的区域,即~区,且由可变的非保守区相连接,提示可能具有重要的功能。仙台病毒L蛋白也具有6个保守区,经证实,这些保守区均与RNA聚合酶活性有关,突变后会不同程度地影响病毒复制及转录,提示可能在RNA合成的不同阶段发挥作用<sup>[14-17]</sup>。比较重要的是~区,~区包括4个保守的基序A~D,其中C基序内保守的五肽Q-G-D-N-Q(aa 771~775)被认为是RNA聚合酶的活性位点,突变后完全丧失RNA合成能力。本研究中,通过氨基酸序列比对发现,Tianjin株L蛋白氨基酸序列中出现了29个独特的变异位点,其中C基序内保守的五肽Q-G-D-N-Q未发生变异。

本研究对副粘病毒Tianjin株NP、P、M及L蛋白进行了初步分析,发现Tianjin株很可能为仙台病毒新的基因型,且其氨基酸序列中存在较多独特的变异位点,说明Tianjin株有可能在生物学特性、致病性、流行病学等方面与已知仙台病毒株具有很大的区别。为了确定这些独特的变异位点与病毒生物学特性乃至致病性之间的关系,将来有望通过反向遗传学技术构建重组病毒来证实。

## 参 考 文 献

- [1] 张国际,李晓眠,金孟珏,等. 48只普通棉耳绒猴死亡病因的初步分析. 天津医科大学学报, 2000, 6(2): 208-209.
- [2] 石建党,李晓眠,刘凤勇,等. 一株引起普通棉耳绒猴死亡病毒的分离与鉴定. 中国病毒学, 2003, 18(4): 357-361.
- [3] 李梅,石建党,石立莹,等. 一株致普通棉耳绒猴严重下呼吸道感染的副粘病毒种系进化分析. 中华微生物学和免疫学杂志, 2007, 27(7): 603-606.
- [4] 石立莹,李晓眠,李梅,等. 副粘病毒Tianjin株F基因克隆及序列分析. 中华微生物学和免疫学杂志, 2007,

27(8): 583-589.

- [5] Myers TM, Moyer SA. An amino-terminal domain of the Sendai virus nucleocapsid protein is required for template function in viral RNA synthesis. *J Virol*, 1997, 71(2): 918-924.
- [6] Cevik B, Kaesberg J, Smallwood S, et al. Mapping the phosphoprotein binding site on Sendai virus NP protein assembled into nucleocapsids. *Virology*, 2004, 325(2): 216-224.
- [7] Curran J, Boeck R, Kolakofsky D. The Sendai virus P gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *EMBO J*, 1991, 10(10): 3079-3085.
- [8] Smallwood S, Ryan KW, Moyer SA. Deletion analysis defines a carboxyl-proximal region of Sendai virus P protein that binds to the polymerase L protein. *Virology*, 1994, 202(1): 154-163.
- [9] Ryan KW, Morgan EM, Portner A. Two noncontiguous regions of Sendai virus P protein combine to form a single nucleocapsid binding domain. *Virology*, 1991, 180(1): 126-134.
- [10] Ryan KW, Portner A. Separate domains of Sendai virus P protein are required for binding to viral nucleocapsids. *Virology*, 1990, 174(2): 515-521.
- [11] Tuckis J, Smallwood S, Feller JA, et al. The C-terminal 88 amino acids of the Sendai virus P protein have multiple functions separable by mutation. *Virology*, 2002, 76(1): 68-77.
- [12] Sakaguchi T, Uchiyama T, Huang C, et al. Alteration of Sendai virus morphogenesis and nucleocapsid incorporation due to mutation of cysteine residues of the matrix protein. *Virology*, 2002, 76(4): 1682-1690.
- [13] Poch O, Blumberg BM, Bougueleret L, et al. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *Journal of General Virology*, 1990, 71(Pt 5): 1153-1162.
- [14] Feller JA, Smallwood S, Horikami SM, et al. Mutations in conserved domains IV and VI of the large (L) subunit of the sendai virus RNA polymerase give a spectrum of defective RNA synthesis phenotypes. *Virology*, 2000, 269(2): 426-439.
- [15] Chandrika R, Horikami SM, Smallwood S, et al. Mutations in conserved domain I of the Sendai virus L polymerase protein uncouple transcription and replication. *Virology*, 1995, 213(2): 352-363.
- [16] Cortese CK, Feller JA, Moyer SA. Mutations in domain V of the Sendai virus L polymerase protein uncouple transcription and replication and differentially affect replication *in vitro* and *in vivo*. *Virology*, 2000, 277(2): 387-396.
- [17] Smallwood S, Hovel T, Neubert WJ, et al. Different substitutions at conserved amino acids in domains II and III in the Sendai L RNA polymerase protein inactivate viral RNA synthesis. *Virology*, 2002, 304(1): 135-145.