

乙型流感病毒 Victoria 系和 Yamagata 系 HA1 基因的分子进化研究

金青青 茅海燕 孙逸 卢亦愚 冯燕 徐昌平 莫世华

【摘要】 目的 探讨乙型流感病毒两大谱系的进化特征和进化规律。方法 从 GenBank 数据库下载 1940—2012 年乙型流感病毒流行株共 126 条,采用贝叶斯-马尔科夫链-蒙特卡洛 (Bayesian-MCMC) 和分子钟方法,对乙型流感病毒的 HA1 基因进行系统发育学分析,计算乙型流感病毒两大谱系可能的起源时间与分化时间。结果 1978—2010 年乙型流感病毒 Victoria 系与 Yamagata 系的 aa 平均差异率为 5.4%~10.2%,两谱系的 aa 差异和组间遗传距离随时间推移呈逐渐增大的趋势。与 Victoria 系毒株相比, Yamagata 系全部毒株的 163 位 aa 及部分毒株的 166 位 aa 缺失,但是这些年来的乙型流感病毒 HA1 基因除个别位点外,尚未受到明显的正向选择压力。每年乙型流感病毒 HA1 基因的碱基替换速率为 2.138×10^{-3} (95%HPD: $1.833 \times 10^{-3} \sim 2.437 \times 10^{-3}$) 替代/位点,推算乙型流感病毒 Victoria 系和 Yamagata 系的最近共同祖先出现在 1971 年 (95%HPD: 1969—1972 年),两大谱系的分化时间点分别为 1973 年 (95%HPD: 1971—1974 年) 和 1977 年 (95%HPD: 1975—1978 年)。结论 乙型流感 Victoria 系和 Yamagata 系均较以往发生大的变异,且两大谱系的差异日趋增大,将来有可能分化为不同的亚型,在流感监测中应密切关注这一变化及其流行病学意义。

【关键词】 乙型流感病毒; HA1 基因; 贝叶斯-马尔科夫链-蒙特卡洛方法; 分子钟方法; 进化

Molecular evolution of two lineages related to influenza B virus based on HA1 gene JIN Qing-qing¹, MAO Hai-yan², Sun Yi², LU Yi-yu², FENG Yan², XU Chang-ping², MO Shi-hua². 1 Medical School of Ningbo University, Zhejiang, Ningbo 315211, China; 2 Zhejiang Provincial Centre for Disease Control and Prevention

Corresponding author: LU Yi-yu, Email: luyiyuzjh@yahoo.com.cn

This work was supported by grants from the Zhejiang Provincial Natural Science Foundation (No. Y2080987) and Key Disciplines of Zhejiang Province Public Health Inspection (No. XKQ-009-003).

【Abstract】 **Objective** To study the evolutionary characteristics and rules of two lineages on influenza B virus. **Methods** A total of 126 HA1 sequences of strains isolated during 1940 to 2012 were downloaded from the GenBank. Time of the most recent common ancestor (TMRCA) and divergence of the two lineages were calculated based on the data from phylogenetic analysis of HA1 gene, using Bayesian Markov Chain Monte Carlo (Bayesian-MCMC) and molecular clock method. **Results** The average amino acid variant ratios were ranged from 5.4% to 10.2% within the strains of influenza B virus isolated during 1978 to 2010. Compared with the Victoria-like strains, all Yamagata-like strains showed an amino acid deletion at 163th site, while some of them showing a deletion at position 166. HA1 gene of influenza B virus seemed not have been affected by positive selection except a few sites. The evolutionary average rate on HA1 gene was 2.138×10^{-3} substitutions/site/year (95%HPD: $1.833 \times 10^{-3} \sim 2.437 \times 10^{-3}$ substitutions/site/year). The estimated dates for TMRCA of the two lineages of influenza B virus could be dated back to 1971 (95% HPD: 1969–1972), while the divergence times of the two lineages were 1973 (95% HPD: 1971–1974) and 1977 (95% HPD: 1975–1978) respectively. **Conclusion** Significant differences were found on HA1 gene between earlier and recent identified strains of Victoria and Yamagata lineage. Differences between the two lineages increased and showing the potential of dividing themselves into different subtypes in the future. More attention should be paid to these trends and the related epidemiological significance.

【Key words】 Influenza B virus; HA1 gene; Bayesian-MCMC; Molecular clock; Evolution

DOI: 10.3760/cma.j.issn.0254-6450.2013.04.013

基金项目:浙江省自然科学基金项目(Y2080987);浙江省公共卫生检验检测重点学科群(XKQ-009-003)

作者单位:315211 宁波大学医学院(金青青);浙江省疾病预防控制中心(茅海燕、孙逸、卢亦愚、冯燕、徐昌平、莫世华)

通信作者:卢亦愚, Email: luyiyuzjh@yahoo.com.cn

乙型流感病毒的 HA 基因片段包括 HA1 和 HA2 两部分,是流感病毒主要的抗原区域,该区域的氨基酸(aa)变异将导致流感病毒的抗原性改变,从而引起乙型流感的暴发与流行。根据抗原性和基因特征的不同,乙型流感病毒分为 Victoria 和 Yamagata 两大谱系^[1-3],然而国内外学术界关于两大谱系的分化以及起源时间的研究结论并不完全一致^[4,5]。为探讨乙型流感病毒两大谱系的进化特征及进化规律,本研究采用贝叶斯-马尔科夫链-蒙特卡洛(Bayesian-MCMC)方法,对现有乙型流感病毒的 HA1 基因进行系统发育学分析,并通过分子钟方法推算出乙型流感病毒两大谱系可能的起源及分歧时间。

材料与与方法

1. 样本来源:从 GenBank 数据库下载 1940—2012 年不同地区、不同来源的乙型流感病毒流行株的 HA1 基因序列共 126 条,用 Clustal W 程序进行比对分析。

2. 选择压力分析:根据乙型流感病毒的 HA1 基因序列,用 HyPhy 在线软件([http:// www. datamonkey.org](http://www.datamonkey.org))的 maximum likelihood SLAC (Single Likelihood Ancestor Counting)方法对乙型流感病毒及其两大谱系分别计算进化压力 ω 值($\omega = dN/dS$)^[6],其中 dN 表示非同义替代率, dS 表示同义替代率。aa 位点在检验水准 α 取 0.05 时, $\omega < 1$ 表示负向选择位点; $\omega = 1$ 表示中性选择位点; $\omega > 1$ 表示正向选择位点。

3. 遗传距离分析及系统进化树的构建:利用 Mega 4.0 软件进行遗传距离的计算,并构建 NJ 树(Neighbor-Joining),PAUP 软件构建 MP 树(Maximum Parsimony)。MP 树各分支单位位置信度通过自展值(bootstrap)法进行检验,共 1000 个循环^[7,8]。

4. 核苷酸替代模型的选择、进化速率及分化时间的估算:根据 Modeltest 3.7 软件的 Akaike 信息准则(Akaike's Information Criterion, AIC)选择本研究核苷酸序列矩阵的最佳替换模型为 GTR+I+G。利用 Beast 软件对乙型流感病毒的 HA1 基因序列进行分析,并且计算核苷酸的进化速率和谱系之间的

最近共同祖先(the most recent common ancestor, TMRCA)^[9,10]。采用不相关松散对数分子钟模型(uncorrelated relaxed lognormal)结合种群指数增长模型(exponential growth),共运行 3.2×10^7 代,每 1000 代进行抽样。用 Tracer 分析 Beast 抽样产生的数据, burn-in 值为 3.2×10^6 。同时利用后验概率法,推算出的最大分支可信树(maximum clade credibility, MCC),用 Figtree 软件分析构建的进化树。

结 果

1. 两谱系 HA1 基因编码 aa 的差异:Victoria 系流行株 HA1 基因共编码 340 个 aa, Yamagata 系流行株共编码 338~339 个 aa。与 Victoria 系毒株相比, Yamagata 系全部毒株的 163 位 aa 及部分毒株的 166 位 aa 缺失。1978—2010 年乙型流感病毒 Victoria 系与 Yamagata 系的 aa 平均差异率为 5.4%~10.2%,其中 1978—1988 年的 aa 平均差异数最小(18.4 ± 4.4), 2000—2010 年的 aa 平均差异数最大(34.7 ± 2.0), aa 差异总体呈现出随时间推移逐渐增大的趋势(表 1)。分化初期和当前乙型流感两大谱系流行株 HA1 基因的 aa 差异数分别为 13~14 个和 34~36 个(表 2、3)。

表 1 1978—2010 年乙型流感病毒 Victoria 系与 Yamagata 系 HA1 基因编码 aa 的差异

年份	总差异数	平均差异数	平均差异率(%)
1978—1988	13~24	18.4 ± 4.4	5.4
1989—1999	25~31	27.6 ± 2.2	8.1
2000—2010	32~38	34.7 ± 2.0	10.2

2. 乙型流感病毒选择压力分析:通过 SLAC 方法计算,乙型流感病毒 HA1 基因的选择压力平均 ω 值为 0.278,谱系分化前的乙型流感病毒为 0.312, Victoria 系为 0.234, Yamagata 系为 0.280。在 $\alpha=0.05$ 的水平下,乙型流感病毒 Victoria 系和 Yamagata 系 HA1 基因的 199 位受到明显的正向选择压力,为两大谱系的共同正向选择位点;同时分化前乙型流感病毒、Victoria 系和 Yamagata 系的负向选择位点数分别为 5、7 和 7 个(表 4)。

3. 乙型流感病毒遗传距离分析:将 1978—2012 年乙型流感病毒 Victoria 系和 Yamagata 系流行株按照时间顺序(间隔 6 年)分成 5 组,并进行遗传距离分

表 2 乙型流感病毒 Victoria 系与 Yamagata 系分化初期流行株 HA1 区的 aa 差异

谱系	流行株	56	75	122	129	137	141	163	164	166	183	197	199	233	235	255	262	267
Victoria 系	B/Shanghai/10/1980	K	N	H	T	V	R	N	N	N	K	N	T	S	T	S	T	I
	B/Xuanwu/23/1982	K	T	H	T	V	G	N	N	N	E	N	I	N	A	S	T	I
Yamagata 系	B/Singapore/222/1979	N	T	R	R	I	G	-	D	-	K	T	T	N	T	P	V	V
	B/Oregon/5/1980	N	T	R	R	I	G	-	D	-	K	N	A	N	T	P	V	V

表 3 乙型流感病毒 Victoria 系与 Yamagata 系当前流行株 HA1 区的 aa 差异

谱系	流行株	15	40	48	56	58	71	73	75	76	80	81	116	122	126	129	136	137	146	148
Victoria 系	B/Georgia/07/2010	I	H	E	K	L	K	T	K	I	R	V	H	H	N	N	K	I	I	N
	B/Ningbo/49/2010	V	H	E	K	P	K	T	K	I	R	V	H	H	N	N	K	I	V	N
Yamagata 系	B/Pennsylvania/13/2010	V	Y	R	D	L	M	V	T	T	K	A	N	Q	D	K	R	L	A	S
	B/Wisconsin/01/2010	V	Y	R	D	L	M	V	T	T	K	A	N	Q	D	K	R	L	A	S

谱系	流行株	149	150	163	164	165	166	168	175	197	198	202	203	209	230	233	252	255	262	267
Victoria 系	B/Georgia/07/2010	G	N	N	D	K	N	T	I	N	E	A	K	K	G	N	V	S	T	I
	B/Ningbo/49/2010	G	N	N	D	K	N	T	I	D	E	A	K	K	G	N	V	S	T	I
Yamagata 系	B/Pennsylvania/13/2010	K	I	D	-	N	Y	N	V	N	K	K	N	N	D	D	V	P	V	V
	B/Wisconsin/01/2010	K	I	D	-	N	Y	N	V	N	K	K	S	N	D	D	M	P	V	V

注: - 表示该 aa 位点缺失

表 4 谱系分化前乙型流感病毒、Victoria 系和 Yamagata 系流行株 HA1 基因的选择压力分析

毒株	株数	$\omega = dN/dS$	正向选择位点数 (aa 位置)	负向选择位点数
谱系分化前毒株	17	0.312	0	5
Victoria 系毒株	49	0.234	1(199)	7
Yamagata 系毒株	60	0.280	1(199)	7

析。各时间段内毒株组内距离在 0~0.025 之间,均小于各组的组间距离 0.039~0.120。除 1985—1991 年外,其余年份的 Victoria 系的组内距离均小于 Yamagata 系的组内距离,两大谱系的组内距离差距在 1999—2005 年最大,其中 Yamagata 系 0.025, Victoria 系 0.016,二者间差异达 0.009。此外,Victoria 系和 Yamagata 系的组间距离随时间推移逐渐增大,其中 1978—1984 年与 1985—1991 年间的增加幅度最大,为 0.027,1992 年之后的增加幅度在 0.017~0.019 之间,保持在相对稳定的水平(表 5)。

4. 乙型流感病毒系统进化树分析:用 Mega 4.0 软件构建的 NJ 树与 PAUP 软件构建的 MP 树,以及 Bayesian-MCMC 法构建的 MCC 树三者结果基本一致,均显示乙型流感病毒在毒株 B/Hongkong/5/72 后形成了 Victoria 系和 Yamagata 系两大分支。另外,从系统进化树上看,1940—2012 年乙型流感病毒总体呈现以时间为序进化的趋势(图 1)。

5. 乙型流感病毒进化速率及共同祖先的计算:本研究用 Bayesian-MCMC 法,计算得出每年乙型流感病毒 HA1 基因核苷酸平均碱基替代率为 2.138×10^{-3} (95% HPD: $1.833 \times 10^{-3} \sim 2.437 \times 10^{-3}$) 替代/位点,推算乙型流感病毒的最近祖先出现在 1928 年 (95% HPD: 1921—1934 年),Victoria 系和 Yamagata

系的最近共同祖先出现在 1971 年 (95% HPD: 1969—1972 年),Victoria 系的分化时间点为 1973 年 (95% HPD: 1971—1974 年),Yamagata 系的分化时间点为 1977 年 (95% HPD: 1975—1978 年)(图 1)。

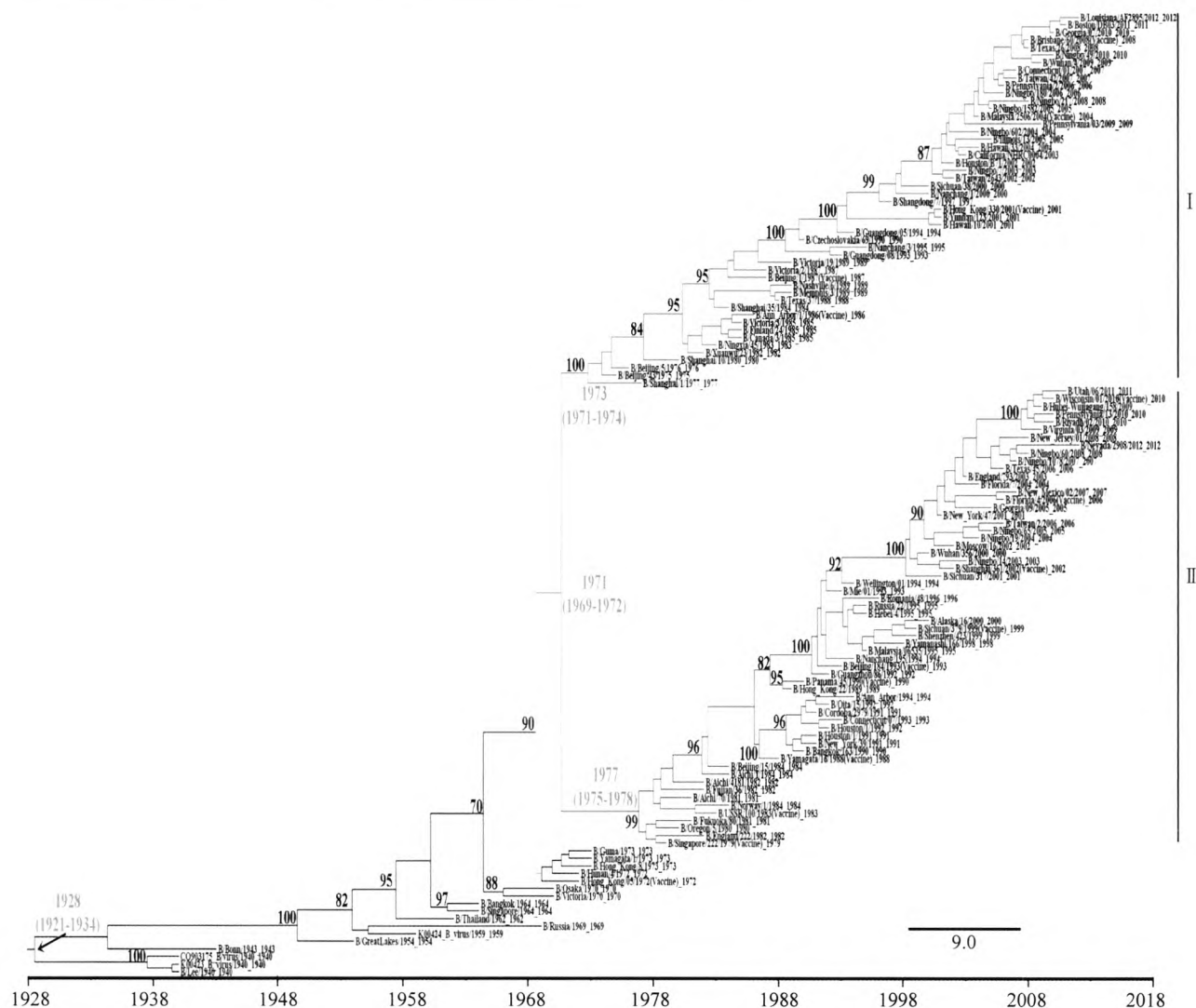
讨 论

一种病毒各粒子间的核苷酸序列并不完全相同,存在一个差异谱,但仍以某一种序列为主,即病毒的主序列,这种主序列与变异谱之间随着时空的转移而不断变化的过程就是病毒的进化^[11]。20 世纪 70 年代以前乙型流感病毒尚未分化成 Victoria 和 Yamagata 两大谱系,推测该时期内的流行株为当时乙型流感病毒的主序列,此后分离到的毒株才发现与早年毒株具有较大的差异^[2,3]。本研究结果显示,乙型流感病毒自分化为 Victoria 系和 Yamagata 系以来,两个谱系间 HA1 基因的 aa 差异随时间推移逐渐增大,1978—2010 年谱系间的 aa 平均差异数从 18.4 个增加至 34.7 个,aa 的平均差异率则由 5.4% 扩大至 10.2%,到 2010 年两者间的 aa 差异率已达 10%。同时,从遗传距离的分析中也表明了两大谱系的组间距离随时间推移逐渐增大 (0.039→0.120),结合 Victoria 系和 Yamagata 系分别于 1973、1977 年开始分化,提示乙型流感两大谱系沿着各自的进化途径发展,差异逐年增大,随着时间积累将来可能由谱系分化成不同的亚型^[12-15]。

通过选择压力分析发现,谱系分化前乙型流感病毒、Victoria 系和 Yamagata 系的 HA1 基因的 ω 值均在 0.3 以下,乙型流感两个谱系在 HA1 基因上的正向选择位点仅第 199 位 aa,提示乙型流感病毒 HA1 区

表 5 1978—2012 年乙型流感病毒 Victoria 系和 Yamagata 系流行株遗传距离分析

分组	1978—1984 年		1985—1991 年		1992—1998 年		1999—2005 年		2006—2012 年	
	Victoria 系	Yamagata 系	Victoria 系	Yamagata 系	Victoria 系	Yamagata 系	Victoria 系	Yamagata 系	Victoria 系	Yamagata 系
Victoria 系	0.014	-	0.019	-	0.018	-	0.016	-	0.013	-
Yamagata 系	0.039	0.019	0.066	0.013	0.084	0.019	0.103	0.025	0.120	0.020



注: I (绿色)表示 Victoria 系, II (蓝色)表示 Yamagata 系;关键节点显示 Bootstrap 值 > 70% 及祖先序列出现的时间 (TMRCa), 括号内数据为 95% CI, 水平分支表示分化时间

图 1 乙型流感病毒 HA1 基因的 MCC 树

除个别位点外, 尚未受到明显的正向选择压力^[16]。此外, 本研究发现, 近 10 年来 (1990—2012 年) 的乙型流感 HA1 区的 4 个抗原决定簇 (A 区: 116~137; B 区: 141~150; C 区: 162~167; D 区: 194~202) 上共有 17 个 aa 位点发生改变^[17], 这些位点的变异可能造成乙型流感抗原性的改变。与早年的乙型流感流行株相比, 两大谱系流行株在 163~166 位 aa 插入或缺失比较频繁, 1980 年以后的 Victoria 流行株均在 163 位插入一个 aa, 并发生多次变异 (Asn→Asp→Glu→Asp), 而 Yamagata 系的 166 位 aa 在 1993 年前缺失, 直到 1994 年之后才保留下来 (Asn→Tyr→Asn), 提示这两个位点的 aa 变异可能在两大谱系的适应与进化上存在重要意义。

从乙型流感病毒的进化速率研究上显示, 每年乙型流感 HA1 基因的碱基替换速率为 $1.833 \times$

$10^{-3} \sim 2.437 \times 10^{-3}$ 替代/位点, 与 Rubing 和 Edward^[18] 的研究结果相近, 略小于每年甲型流感病毒的碱基替换速率 ($2.68 \times 10^{-3} \sim 12.50 \times 10^{-3}$ 替代/位点)^[19]。据此速率推算的 Victoria 系和 Yamagata 系的最近共同祖先应出现在 20 世纪 70 年代初, 稍早于国内学者提出的两大谱系起源于 20 世纪 70 年代中期的结果^[5]。造成这种差异的原因可能有以下几方面: 首先, 本研究所选毒株为 1940—2012 年期间分离的乙型流感流行毒株, 选择的毒株较多, 毒株分离的年份、地区跨度较大, 能够更为全面地分析乙型流感 Victoria 系和 Yamagata 系的自然动态变化; 其次, 两大谱系的起源时间, 以往是根据乙型流感毒株的 HA1 系统进化树来进行推算的。由于病毒变异与谱系分化在前, 流行毒株分离在后, 尤其是早年分离到的乙型流感毒株本身就少, 由此推算的两大谱系共

同祖先在时间上会产生一定的滞后现象。本研究采用的 Bayesian-MCMC 法是根据碱基的平均替换速率来推算乙型流感两大谱系的分歧时间,能够充分利用流行株的分离时间及进化信息,据此推算的共同祖先时间可能更接近真实的谱系分化时间。

乙型流感两大谱系之间的转换及病毒本身抗原性的变异,与乙型流感的暴发与流行密切相关^[20],因此加强乙型流感病毒的进化研究,对于了解乙型流感的流行与进化规律,开展流感的预测预警工作,具有重要的流行病学意义。

参 考 文 献

- [1] Guo YJ, Cheng XW. Influenza virus and experimental technology. Beijing: Three Gorges Press, 1997. (in Chinese)
郭元吉,程小雯. 流行性感病毒及其实验技术. 北京:三峡出版社,1997.
- [2] Nerome R, Hiromoto Y, Sugita S, et al. Evolutionary characteristics of influenza B virus since its first isolation in 1940: dynamic circulation of deletion and insertion mechanism. Arch Virol, 1998, 143(8):1569-1583.
- [3] Wang YP, Li CH, Di B, et al. Analysis of pathogen of influenza B in Guangzhou during 1986-1997. Chin J Health Lab Technol, 1999, 9(1):39-40. (in Chinese)
王艳萍,李钊华,狄飏,等. 广州市1986-1997年乙型流感病原学分析. 中国卫生检验杂志,1999,9(1):39-40.
- [4] Rota PA, Wallis TR, Harmon MWR, et al. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. Virology, 1990, 175:59-68.
- [5] Chen JM, Guo YJ, Wu KY, et al. Study on the origin and evolution characteristics of two lineages of influenza B. Chin J Virol, 2001, 17(4):322-327. (in Chinese)
陈继明,郭元吉,吴昆昱,等. 乙型流行性感病毒两大谱系的起源及其演变特征. 病毒学报,2001,17(4):322-327.
- [6] Kosakovsky Pond SL, Frost SDW. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics, 2005, 21(10):2531-2533.
- [7] Meng SL, Xu GL, Lei YL, et al. Analysis of genetic diversity of rabies virus isolates in China. Chin J Biol, 2010 (50): 449-454. (in Chinese)
孟胜利,徐葛林,雷永良,等. 中国狂犬病病毒遗传多样性分析. 中国生物制品杂志,2010(50):449-454.
- [8] Pan XL. Study on the molecular evolution and recombination of genome of Japanese encephalitis virus. Beijing: Chinese Center for Disease Control and Prevention, 2009. (in Chinese)
潘晓玲. 乙型脑炎病毒全基因组分子进化及基因重组研究. 北京:中国疾病预防控制中心,2009.
- [9] Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol, 2007, 7:214.
- [10] Drummond AJ, Ho SYW, Phillips MJ, et al. Relaxed phylogenetics and dating with confidence. Plos Biol, 2006, 4(5):e88.
- [11] Jin Q. Medical Molecular Virology. Beijing: Science Press, 2001: 2-3. (in Chinese)
金奇. 医学分子病毒学. 北京:科学出版社,2001:2-3.
- [12] Li MH, Zhou M, Mao HY, et al. Analysis of HA1 gene of influenza virus B isolated in Zhejiang province. Chin J Health Lab Technol, 2008, 18(3):429-432. (in Chinese)
李敏红,周敏,茅海燕,等. 浙江省乙型流感病毒 HA1 基因分析. 中国卫生检验杂志,2008,18(3):429-432.
- [13] Qi SX, Han GY, Liu YF, et al. Virological surveillance and molecular characteristics of influenza B viruses in Hebei province during 2004-2008. Chin J Vac Immun, 2009, 15(1): 27-30. (in Chinese)
齐顺祥,韩光跃,刘艳芳,等. 河北省2004-2008年乙型流行性感病毒监测与血凝素蛋白基因序列分析. 中国疫苗和免疫, 2009, 15(1):27-30.
- [14] Lv YN, Peng XM, Liang HJ, et al. Analysis on the characteristics of HA gene of influenza virus B circulated in Beijing area from 2009 to 2010. Chin J Dis Control Prev, 2011, 15(8): 689-694. (in Chinese)
吕燕宁,彭晓敏,梁慧洁,等. 北京地区2009-2010年流感监测季乙型流感病毒 HA 基因特征分析. 中华疾病控制杂志,2011, 15(8):689-694.
- [15] Tapasi R, Anurodh SA, Anupam M, et al. Surveillance and molecular characterization of human influenza B viruses during 2006-2010 revealed co-circulation of Yamagata-like and Victoria-like strains in eastern India. Infection, Gen Evol, 2011 (11):1595-1601.
- [16] Air GM, Gibbs AJ, Laver WG, et al. Evolutionary changes in influenza B are not primarily governed by antibody selection. Proc Natl Acad Sci USA, 1990, (87):3884-3888.
- [17] Zhang DY, Feng Y, Mao HY, et al. Study on the differences of the amino acid sequences in HA1 regions between the vaccine strains and the epidemic strains of influenza virus B. Zhejiang Prev Med, 2011, 23(7): 1-4. (in Chinese)
张冬艳,冯燕,茅海燕,等. 中国南方地区乙型流感流行株与疫苗株的 HA1 区氨基酸差异研究. 浙江预防医学,2011, 23(7): 1-4.
- [18] Rubing C, Edward CH. The evolutionary dynamics of human Influenza B virus. J Mol Evol, 2008, 66(6):655-663.
- [19] Nelson MI, Simonsen L, Viboud C, et al. Stochastic processes are key determinants of short-term evolution in Influenza A virus. PLoS Pathog, 2006, 2:e125.
- [20] Li GX, Zheng GZ, Qiu DH, et al. Phylogenetic analysis of HA1 gene of influenza B viruses isolated and identified in Taizhou city. Chin J Health Lab Technol, 2011, 21(9): 2126-2129. (in Chinese)
李桂霞,郑官增,裘丹红,等. 台州市乙型流感病毒分离、鉴定及 HA1 基因特性研究. 中国卫生检验杂志, 2011, 21(9): 2126-2129.

(收稿日期:2012-11-13)

(本文编辑:张林东)