

甲型流感病毒血凝素基因的进化与分型研究

熊成龙 张志杰 梁 笛 蒋露芳 陆一涵 居丽雯 陈 跃 姜庆五<sup>△</sup>  
(复旦大学公共卫生学院流行病学教研室-公共卫生安全教育部重点实验室 上海 200032)

**【摘要】 目的** 分析甲型流感病毒各亚型血凝素基因的系统发育特征,探讨甲型流感病毒血凝素基因的起源与演化规律。**方法** 从互联网获取甲型流感 16 个亚型的基因序列资料,运用 DNASTar、Clustal X1.83、MEGA 4 及 PHYLIP 3.67 等生物信息学软件对其进行系统发育关系研究。**结果** 16 个亚型的甲型流感病毒的血凝素基因能够按照各自的血清亚型实现聚类,亚型内各基因的遗传距离在 0.10231~0.22516 厘摩(centimorgan, cM, SE 为 0.006~0.011)之间不等,按照系统发育关系,16 个血清亚型可以分为 5 个大的进化簇:亚型 H1、H2、H5 与 H6 构成 A 簇,H8、H9 与 H12 构成 B 簇,H11、H13 与 H16 构成 C 簇,H7、H10 与 H15 构成 D 簇,H3、H4 与 H14 构成 E 簇,其中 A 簇、B 簇与 C 簇可进一步聚类为一个进化群,而 D 簇与 E 簇也可以进一步聚类为另一进化群。**结论** 按血凝素基因的序列信息可将甲型流感病毒分为 5 个进化簇并进而分为两个进化群。**【关键词】** 甲型流感病毒; 血凝素; 基因进化; 基因分型  
**【中图分类号】** R 372 **【文献标志码】** A **doi:** 10.3969/j.issn.1672-8467.2011.04.009

Evolution and classification of haemagglutinin genes of influenza A

XIONG Cheng-long, ZHANG Zhi-jie, LIANG Di, JIANG Lu-fang,  
LU Yi-han, JU Li-wen, CHEN Yue, JIANG Qing-wu<sup>△</sup>  
(Key Laboratory of Public Health Safety, Ministry of Education-Department of Epidemiology,  
School of Public Health, Fudan University, Shanghai 200032, China)

**【Abstract】 Objective** To determine the phylogenetic relationships of haemagglutinin gene sequences of 16 influenza A virus sera subtypes, and to study their genetic origins. **Methods** Haemagglutinin gene sequences of 16 influenza A virus sera subtypes were derived from an Internet database. DNASTar, Clustal X1.83, MEGA 4, and PHYLIP 3.67 softwares were used to measure the genetic distances within and between the influenza A virus sera subtypes, and to construct their phylogenetic trees. **Results** Haemagglutinin gene sequences of 16 influenza A virus sera subtypes were successfully divided into 16 genotypes, with the genetic distances within genotypes being 0.10231 to 0.22516 centimorgan (cM, SE = 0.006 - 0.011). A taxonomy study demonstrated that these 16 influenza A virus sera subtypes could be congregated into 5 clades including: A (subtypes H1, H2, H5 and H6), B (H8, H9 and H12), C (H11, H13 and H16), D (H7, H10 and H15) and E (H3, H4, and H14). They could be further classified into two groups: group I (clades A, B and C) and group II (clades D and E). **Conclusions** Haemagglutinin gene sequences of 16 influenza A virus sera subtypes could be congregated into five clades and further into two groups.  
**【Key words】** influenza A virus; haemagglutinin; gene evolution; gene classification

流感病毒(influenza virus)属正黏病毒科(orthomyxoviridae),有 A、B 和 C 三种型别,分别引起甲、乙和丙型流感,其中甲型流感(influenza A)可以感染人、禽、猪、马等多种动物,是重要的人(禽)畜

<sup>△</sup>Corresponding author E-mail:jiangqw@fudan.edu.cn

共患传染病。流感病毒的基因组由 8 条单股负链的 RNA 组成,变异主要集中在病毒的血凝素(haemagglutinin, HA)和神经胺酸酶(neuraminidase, NA)两种表面结构蛋白上,血凝素基因是其中变异最大的基因。目前已发现 16 种不同的血凝素血清亚型,流感病毒的抗原性和致病性很大程度上取决于该基因的变异情况,使得该基因成为病毒变异、毒力和宿主特异性的主要决定因素<sup>[1-2]</sup>,与流感的流行暴发和疫苗防控有着密切关系。近年来,出于开发具有广泛抗原活性的流感疫苗的需要,对流感病毒血凝素蛋白的研究得到了广泛的重视<sup>[3-5]</sup>。本研究对 16 个亚型的 HA 基因进行分子进化分析,旨在分析甲型流感病毒各亚型的起源进化关系,在病毒的分子遗传学水平上探讨开发具有广泛抗原活性的流感病毒疫苗的可能性。

## 材 料 和 方 法

### 甲型流感病毒各型 HA 基因的预分析与序列选取

序列选取使用美国国立生物技术信息中心网站(<http://www.ncbi.nlm.nih.gov>)Nucleotide 搜索引擎,设置检索式为“Influenza A + hn(n 分别为 1、2、3…16) + hemagglutinin”,H1、H2、H3…H16 分别获得搜索结果分别为 679、151、1072、20、438、69、162、8、220、22、13、17、13、4、5、24(截止日期为中国北京时间 2010 年 6 月 2 日 9:00),考虑到片段过短将导致后续分析受软件空位罚分(Gap penalty)策略影响偏大,仅对片段长度大于 1 000 bp 的各序列一一下载。

使用 DNASTar 软件包 MegAlign 模块之 Clustal V 程序进行快速预分析,选取各亚型内彼此分歧尽可能大的基因序列作为各型的代表株,由于本文考查序列数量庞大,为节省篇幅,按惯例仅列出各基因在 GenBank 中的准入编码(Access No.)<sup>[6-7]</sup>。获得 H1 亚型的 HA 基因序列 22 个,按 h1、h1(2)…h1(22)顺序各基因的准入编码分别是 AF117241、U11857、AB514930、CY038903、CY040600、FJ986618、U38242、AB514937、CY040587、D00839、FJ986622、L25071、AF116576、CY040594、D13574、GU168284、M81707、CY016329、CY040596、FJ966082、FJ984355、FJ998207;H2 亚型 11 个,按 H2、H2(2)…H2(11)排列分别是 L11129、D13580、L11130、L11139、L11133、L20406、L11128、L11136、L20409、AF116201、L11138;H3 亚型 22 个,按 h3、h3(2)…h3(22)分别为 GU215026、M24718、M24721、M24722、M24726、M73774、M73776、M16739、M1674、U08858、AB537470、EU301219、

AB537477、AY531037、CY031843、CY062351、D00930、D00932、EU301215、EU422987、FJ215839、GU143843;H4 亚型 7 个,按 h4、h4(2)…h4(7)分别为 FJ439565、GU168298、M25285、M25286、M25290、M25291、CY020805;H5 亚型 21 个,按 h5、h5(2)…h5(21)分别为 U79452、U79454、U79456、AB241614、AB241615、AB241620、AB241623、AF082042、AF290443、EF619972、AF468837、AM911095、AY296064、CY046150、EF597255、EF597262、EF597268、EF597269、EU099400、GU046743、GU361203;H6 亚型 8 个,按 h6、h6(2)…h6(8)分别为 GQ414867、AF100181、AF474035、AJ507205、GQ117282、GQ117283、GQ414852、GQ414856;H7 亚型 12 个,按 h7、h7(2)…h7(12)分别为 L43915、AF202231、AF202234、AF202237、AF202239、EU684259、AJ704811、AY240892、AY240914、AY240915、EF470585、EU158101;H8 亚型 8 个,按 h8、h8(2)…h8(8)分别为 AB450435、AB450454、AF310987、CY042882、CY043848、D90304、EF061122、GU051909;H9 亚型 17 个,按 h9、h9(2)…h9(17)分别为 FJ807715、AF203008、AF523390、AY206676、AY594194、GQ117285、AF523372、AJ781818、AY206677、HM008885、AF523376、AJ781822、AY206679、FJ434569、AF523386、AY206675、AY428485;H10 亚型 7 个,按 h10、h10(2)…h10(7)分别为 M21647、EU124207、EU747323、EU747324、EU747327、EU747329、HM179251;H11 亚型 6 个,按 h11、h11(2)…h11(6)分别为 AB298283、GU046752、DQ424859、DQ424860、DQ424861、DQ435281;H12 亚型 10 个,按 h12、h12(2)…h12(10)分别为 GQ907286、CY006006、CY006007、CY021293、CY029873、CY038101、CY064065、D90307、EU557489、EU743480;H13 亚型 7 个,按 h13、h13(2)…h13(7)分别为 M26090、AB284988、EU293859、EU293860、EU293862、GQ117287、M26089;H14 亚型 4 个,按 h14、h14(2)…h14(4)分别为 M35996、AB289335、FJ975075、GQ247868;H15 亚型 5 个,按 h15、h15(2)…h15(5)分别为 L43916、L43917、AB295613、CY006032、CY006033;H16 亚型 6 个,按 h16、h16(2)…h16(6)分别为 EU293864、AY684887、AY684891、CY015160、EU148600、FM179755。

甲型流感病毒 HA 基因的系统发育树构建 随机选取 5 株乙型流感(influenza B)病毒的血凝素基因为外群,分别记作 InfluenzaB1~InfluenzaB5,其 GenBank 准入编码(Access No.)分别为 DQ508913、

M58428、M18384、J02093、M10298,序列比对使用 Clustal X1.83 软件(ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/),参数均取默认设置;用 MEGA4 软件(http://bioinfo.weizmann.ac.il/databases/info/mega.sof;)以邻接法(NJ)构建甲型流感病毒 HA 基因的系统发育树,分析重抽样数(replicates)设为 1 000。

不设外群,用 PHYLIP 3.67 软件(http://evolution.genetics.washington.edu/phylip.html)构建甲型流感病毒 H 基因的无根系统发育树。为了尽可能在同一页面清晰展示各基因亚型的进化关系,在以上 MEGA4.0 分析建树的基础上进一步选取各型 HA 基因的代表序列,得到 HA1、HA2、HA3…HA16 基因序列分别为 7、11、7、5、11、6、5、5、8、5、4、5、5、2、3,和 5 个,自展分析(bootstrap)仍采用 1 000 个重抽样数(replicates),运用最大似然法(DNA maximum likelihood, DNAML)建树,以 Tree View 软件(http://taxonomy.zoology.gla.ac.uk/rod/treeview.html)认读和拓扑变形所生成的进化树。

## 结 果

MEGA4 软件构建的甲型流感病毒各血清亚型的 HA 基因系统发育树显示,除 H16 亚型中有一基因序列 h16(2)落入 H13 亚型以外,其余各亚型流感病毒的血凝素均能按其血清亚型很好地实现聚类(见图 1,由于用于分析的序列资料太多,除 H16 亚型与 H13 亚型因有交错需要体现而采用分支展开型以外,其余各型均采用分支压缩型)。由图 1 可见,H1、H2、H5、H6 在系统发育树上处于紧密相邻的位置,显示出这 4 类血清亚型的甲型流感病毒的基因可能有着较近的起源演化关系,可归为一个进化分支,命名为 A 簇(clade A);存在这种亲缘进化关系的还有 H8、H9,与 H12 所构成的 B 簇(clade B);H11、H13,与 H16 所构成的 C 簇(clade C);H7、H10,与 H15 所构成的 D 簇(clade D);H3、H4,与 H14 所构成的 E 簇(clade E),其中 A 簇、B 簇与 C 簇可以进一步聚类为进化群一(group I),D 簇与 E 簇可进一步聚类为进化群二(group II)。此外,MEGA4 软件所构建的甲型流感病毒各血清亚型 HA 基因的系统发育树还显示,早期按照免疫学性质划分的血凝素类型虽然与基因型有着很好的对应关系,但在分类层次上各型别显然存在跨度标准不同的局限性。换言之,若有适当的甲型流感基因分型标准,H2 与 H5、H13 与 H16、H8 与 H12、H7 与 H15,及 H4 与 H14 亚型彼此之间的差别极有可能并未上

升到足以被分为不同亚型的程度,这一现象在 H13 与 H16 的系统发育关系上表现得尤其明显。

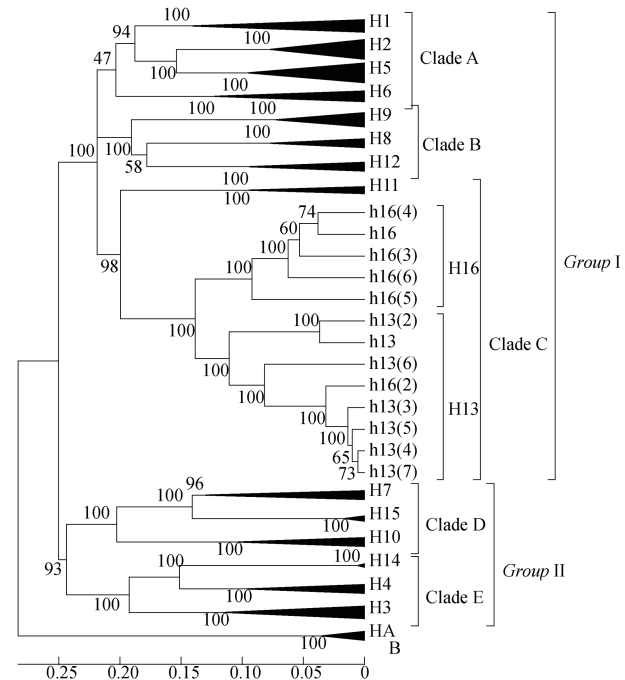


图 1 MEGA4 软件构建的甲型流感病毒各血清亚型 HA 基因系统发育树

Fig 1 Phylogenetic tree of 16 influenza A viruses sera subtypes constructed by MEGA4 software

PHYLIP 3.67 软件所构建的甲型流感病毒 HA 基因的无根系统发育树与上述 MEGA 4 软件的建树结果极其吻合(图 2),简化为图 3 所示的枝状图,可见 16 个亚型的甲型流感血凝素基因也聚类为 5 个分支,即前述的簇,各簇亦能进一步聚类成两个进化

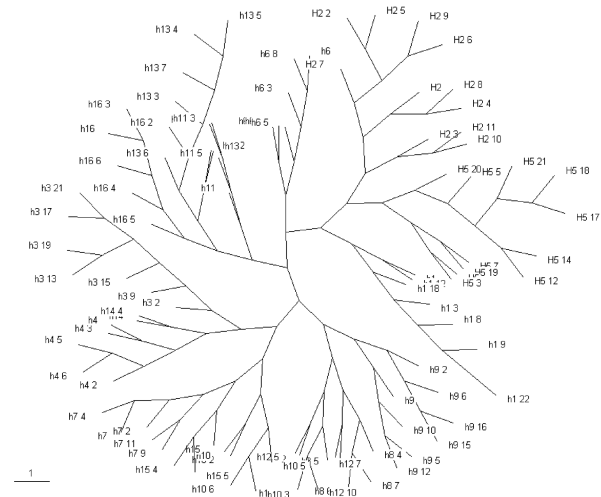


图 2 PHYLIP 3.67 软件构建的甲型流感病毒各血清亚型 HA 基因系统发育树

Fig 2 Phylogenetic tree of 16 influenza A viruses sera subtypes constructed by PHYLIP 3.67 software

群,各簇、群的构成及其在进化树上的位置关系与MEGA 4 软件的建树结果基本一致。其中 H7 与 H15、H8 与 H12、H13 与 H16 在无根树上的交错关系也体现出上述根据免疫学性质划分血凝素类型在分类层次上所存在的问题。

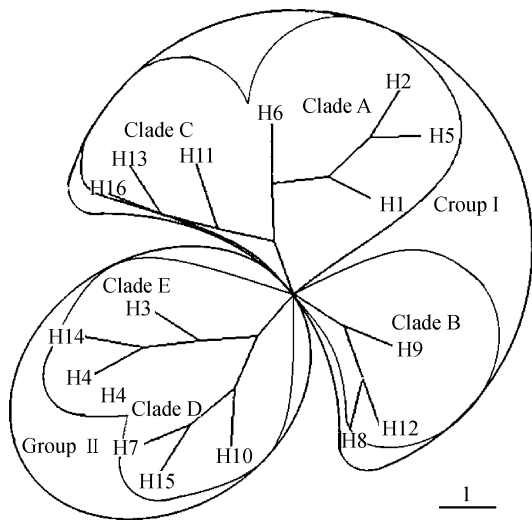


图3 甲型流感病毒各血清亚型聚群关系的枝状图  
Fig 3 Trunk-and-branch-like graph of 16 sera subtypes influenza A viruses

讨 论

甲型流感病毒能够引起包括人、禽、畜在内的多种宿主动物具有高度传染性的呼吸道疾病,是第一个实行全球监测的呼吸道传染病,全球每年有300 000~500 000 人死于甲型流感所致的传染性疾病,其中尤以 H1N1、H1N2、H3N1、H3N2 及 H5N1 等亚型影响为烈,这些亚型常造成全球范围内的暴发流行<sup>[8]</sup>。疫苗接种是预防和控制甲型流感的重要手段,全球每年甲型流感疫苗投入达数十亿美元<sup>[9]</sup>,但由于流感病毒众所周知的基因片断的重排(reassortent)作用,使得病毒的毒力、致病性以及抗原特性变异极大,这种变异不但能导致针对特定型别流感病毒的疫苗免疫相对他种型别的病毒无效或效用降低,而且还为流感的正确诊治带来极大不便<sup>[10-11]</sup>。在流感病毒诸蛋白中,血凝素蛋白与流感的发生与流行关系最为密切,同时也是流感病毒发生抗原变异的分子基础<sup>[12]</sup>。因此,从基因水平了解流感病毒血凝素蛋白各亚型之间的关系,了解病毒基因的变异动态,对流感防治及诊断均具有十分重要的意义。

本研究对甲型流感的 16 个血清亚型的基因序

列进行了比较全面的系统发育分析,运用两种不同的方法构建了甲型流感病毒血凝素基因的系统发育树,研究结果证实,16 个亚型的流感病毒按照各自的血清亚型能够很好地实现聚类,各亚型内部的遗传距离在 0.102 31~0.225 16 厘摩(centimorgan, cM, SE=0.006~0.011)之间不等,按照系统发育关系,可将 16 个血清亚型的流感病毒分为 5 个进化簇并进一步归入 2 个进化群,在进化簇与进化群各自的分类层次上均能得到相应遗传距离的数据支持,说明在基因核苷酸序列水平上的这种分类(分型)是客观合理的。免疫学的证据亦为这一分类(分型)观点提供了支持和佐证,2009 年, Sui 与 Hwang 等<sup>[13]</sup>从来源于不同个体、数量达到 270 亿之巨的人源抗体库中筛选出位于流感病毒血凝素蛋白上的名为 F10 的抗原表位,该研究小组在详细分析了该表位的晶体衍射结构之后,根据这一表位结合特性的不同,将所考查的流感病毒亚型分为两个组(group),其中组 1(group 1)包括 H1、H5 和 H9 亚型,组 2(group 2)则由 H3 与 H7 亚型构成。Varecková 等<sup>[14]</sup>充分研究了流感病毒血凝素蛋白的 HA2 亚单位,发现可以与流感病毒血凝素第 IV 抗原位点结合的人源单克隆抗体 FC12 与 FE1 对 H3 和 H4 亚型的流感具有交叉保护作用,而可以结合该蛋白抗原位点 II 的单抗 IIF4 则具有更为广泛的亚型结合活性,并认为这种广泛的结合特性可以在流感病毒的检测诊断以及设计具有广泛抗原活性的流感病毒疫苗上有着广阔的应用前景。

早在 1999 年, Smirnov 等<sup>[15]</sup>人就明确提出,按照血凝素蛋白的抗原表位特性, H1、H2、H5 以及 H6 血清亚型有着密切的近缘关系,并认为这 4 种血清型似乎应归为一类。更早的相关研究则分别由 Air 等<sup>[16]</sup>在 1981 年以及 Nobusawa 等<sup>[17]</sup>在 1991 年提出,在研究了各亚型甲型流感病毒血凝素蛋白的氨基酸序列同源性之后,研究人员发现流感病毒的 12 (或 13)个血清亚型彼此的同源性在 40%~80%,而亚型内的同源性可达 90%,并最早提出将 12 (或 13)个血清亚型的流感病毒血凝素分为 5 个进化簇(clade)和两个进化群(group)。长期以来,相关的研究结果并没有得到重视,主要原因在于这一成果缺乏相应的理论和应用价值。近年来,随着基因工程在疫苗研发方面的广泛应用,寻找与各亚型流感病毒遗传距离适中的自然分离株,或者通过基因工程改造病毒,使之成为合适的疫苗株已成为可能,相关的基因操作势必要求在病毒基因的核苷酸序列层次、而不仅是病毒蛋白的氨基酸序列层次上进一步探讨流感病毒各亚型的系统发育关系,并试图在此



基础上通过相对简单的操作形成具有针对各亚型甲型流感病毒广泛抗原特性的新疫苗<sup>[14,18]</sup>。此外,一些抗流感病毒药物的研究亦对这种病毒的分型研究提出了类似要求<sup>[19-20]</sup>,这正是开展本研究的初衷与研究意义之所在。

### 参 考 文 献

- [1] Fouchier RA, Munster V, Wallensten A, *et al.* Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls[J]. *J Virol*, 2005, 79(5): 2 814 - 2 822.
- [2] Knossow M, Skehel JJ. Variation and infectivity neutralization in influenza [J]. *Immunology*, 2006, 119(1): 1 - 7.
- [3] Chen J, Ma C, Wong C. Vaccine design of hemagglutinin glycoprotein against influenza[J/OL]. *Trends Biotechnol*, 2011(2011 - 04 - 07) [2011 - 05 - 02]. <http://www.sciencedirect.com/science/article/pii/S0167779911000734>.
- [4] Hifumi E, Fujimoto N, Ishida K, *et al.* Characteristic features of infA-15 monoclonal antibody recognizing H1, H3, and H5 subtypes of hemagglutinin of influenza virus A type [J]. *J Biosci Bioeng*, 2010, 109 (6): 598 - 608.
- [5] Corti D, Suguitan AL, Pinna D, *et al.* Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine [J]. *J Invest*, 2010, 120(5): 1 663 - 1 673.
- [6] Sahini L, Tempczyk-Russell A, Agarwal R, *et al.* Large-scale sequence analysis of hemagglutinin of influenza A virus identifies conserved regions suitable for targeting an anti-viral response [J]. *PLoS ONE*, 2010, 5(2): 9 268 - 9 276.
- [7] Ghedin E, Sengamalay NA, Shumway M, *et al.* Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution [J]. *Nature*, 2005, 437(7 062): 4 239 - 4 243.
- [8] Goñi N, Fajardo A, Moratorio G, *et al.* Modeling gene sequences over time in 2009 H1N1 Influenza A virus populations [J]. *J Virol*, 2009, 6: 215.
- [9] Friesen RHE, Koudstaal W, Koldijk MH, *et al.* New class of monoclonal antibodies against severe influenza: prophylactic and therapeutic efficacy in ferrets [J]. *PLoS ONE*, 2010, 5(2): 1 - 7.
- [10] Adams B, McHardy AC. The impact of seasonal and year-round transmission regimes on the evolution of influenza A virus [J]. *Proc Biol Sci*, 2011, 278 (1 716): 2 249 - 2 256.
- [11] Smith DJ, Lapedes AS, Jong JC, *et al.* Mapping the antigenic and genetic evolution of influenza virus [J]. *Science*, 2004, 305(5 682): 371 - 377.
- [12] Kaverin NV, Rudneva IA, Ilyushina NA, *et al.* Structure of antigenic sites on the haemagglutinin molecule of H5 avian influenza virus and phenotypic variation of escape mutants [J]. *J Gen Virol*, 2002, 83 (Pt10): 2497 - 2505.
- [13] Sui J, Hwang WC, Perez S, *et al.* Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses [J]. *Nat Struct Mol Biol*, 2009, 16(3): 265 - 273.
- [14] Varecková E, Mucha V, Kostolanský F, *et al.* HA2-specific monoclonal antibodies as tools for differential recognition of influenza A virus antigenic subtypes [J]. *Virus Res*, 2008, 132(1 - 2): 181 - 186.
- [15] Smirnov YA, Lipatov AS, Gitelman AK, *et al.* An epitope shared by the hemagglutinins of H1, H2, H5, and H6 subtypes of influenza A virus [J]. *Acta Virol*, 1999, 43(4): 237 - 244.
- [16] Air GM. Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus [J]. *Proc Natl Acad Sci USA*, 1981, 78(12): 7 639 - 7 643.
- [17] Nobusawa E, Aoyama T, Kato H, *et al.* Comparison of complete amino acid sequences and receptorbinding properties among 13 serotypes of hemagglutinins of influenza A viruses [J]. *Virology*, 1991, 182 (2): 475 - 485.
- [18] Ghose T. Single-handed flu combat? [J/OL]. The Scientist-Magazine of the Life Sci, 2009, (2009 - 02 - 23) [2011 - 04 - 30]. <http://classic.the-scientist.com/blog/display/55443/>
- [19] Yoshimoto J, Kakui M, Iwasaki H, *et al.* Identification of a novel HA conformational change inhibitor of human influenza virus [J]. *Arch Virol*, 1999, 144(5): 865 - 878.
- [20] Plotch SJ, Hara BO, Morin J, *et al.* Inhibition of influenza A Virus replication by compounds interfering with the fusogenic function of the viral hemagglutinin [J]. *J Virol*, 1999, 73(1): 140 - 151.

(收稿日期: 2011 - 05 - 16; 编辑: 张秀峰)