

中国甲型流感病毒血凝素基因进化分析

袁彦婷¹, 陈建威², 陆祖宏¹

¹东南大学生物电子学国家重点实验室, 江苏南京 210096; ²深圳华大基因研究院, 广东深圳 518083

摘要: **目的** 探索甲型流感的暴发规律及毒株血凝素 (hemagglutinin, HA) 基因的进化。**方法** 通过分析中国采集并公布的甲型流感病毒基因数据, 对中国大陆所有宿主为人、禽类和哺乳动物的 HA 基因氨基酸序列进行多序列比对, 并用于构建 HA 的系统发育树, 同时对香港地区的 HA 基因蛋白序列同样处理; 分析不同宿主的 HA 基因的选择压力。**结果** 选择中国大陆 1 252 个 HA 基因和香港 485 个 HA 基因, 对其分析发现, 自 2002 年以来, 禽流感的暴发次数和频率多于历史记录; 系统发育树结果显示中国大陆和香港地区的 HA 基因邻接树的拓扑结构一致; 宿主为人的 HA 基因受选择大于禽类宿主的 HA 基因。**结论** 不同宿主的 HA 基因受选择有差异, 提示 HA 基因的快速进化可能会对人类健康产生潜在威胁。

关键词: 甲型流感病毒; 群体遗传学; 血凝素基因; 进化

中图分类号: R 511.7 文献标志码: A 文章编号: 2095-5227(2014)09-0906-03 DOI: 10.3969/j.issn.2095-5227.2014.09.009

网络出版时间: 2014-05-09 17:14 网络出版地址: <http://www.cnki.net/kcms/detail/11.3275.R.20140509.1714.003.html>

Evolution of influenza A virus hemagglutinin gene in China

YUAN Yan-ting¹, CHEN Jian-wei², LU Zu-hong¹

¹State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, Jiangsu Province, China; ²Shenzhen BGI, Shenzhen 518083, Guangdong Province, China

Corresponding author: LU Zu-hong. Email: luzuhong@gmail.com

Abstract: Objective To study the outbreak rules of influenza A virus and the evolution of hemagglutinin (HA). **Methods** A phylogenetic tree of HA gene was constructed respectively for influenza A virus in China mainland and Hong Kong District of China to analyzed the influenza genome data published in China mainland and compare the HA gene sequences in humans, avian, and mammals in China mainland and Hong Kong District of China. The selection pressure of HA gene from different hosts were analyzed. **Results** The analysis of 1 252 and 485 HA genes selected from China mainland and Hong Kong District showed that the outbreak times and frequencies of influenza A virus had been more than their historical records since 2002. The phylogenetic trees of HA gene for China mainland and Hong Kong District showed that their topological structure was identical. The selection pressure of HA gene was greater from humans than from avian. **Conclusion** The selection pressure of HA gene is different from different hosts, suggesting that the rapid evolution of HA gene may pose a potential threat to the health of mankind.

Key words: influenza A virus; population genetics; hemagglutinin gene; evolution

甲型流感病毒可在鸟类和哺乳动物中引起流行性感, 根据其表面蛋白血凝素 (hemagglutinin, HA) 和神经氨酸酶 (neuraminidase, NA) 的不同, 可以将甲型病毒分为多种亚型^[1]。目前甲流病毒的血凝素 (H) 有 18 种亚型, 神经氨酸酶 (N) 有 11 个亚型。禽流感就是甲型流感病毒的亚型。历史上几次世界性流行性感都与禽流感有关, 最早记录为 1878 年在意大利发生的禽流感。1918 年在西班牙发生的禽流感, 被推断是由 H1N1 亚型引起的。自 1997 年中国香港发现人感染 H5N1 禽流感后, 2003 年的东南亚禽流感、2009 年的全球“猪流感” (H1N1)、2012 年的 H5N2 禽流感及始于 2013

年的 H7N9 禽流感, 可以发现禽流感的变异对人的生命和财产安全造成了极大威胁。本研究通过分析中国采集并公布的甲型流感病毒基因数据, 构建了 HA 的系统发育树, 并分析了不同宿主的 HA 基因的选择压力, 探索 HA 基因的进化对于人类健康的影响。

材料和方法

1 材料 本文研究所采用的数据来自于流感研究数据库 (Influenza Research Database, IRD) 在线数据库 (<http://www.fludb.org>)^[2]。通过流感序列搜索, 筛选其中中国的甲型流感病毒的 HA 片段 (宿主种类去除 Lab、Other 及 Unknown), 选择具有完整片段的 HA 基因并去除重复的 HA 序列。最终得到中国大陆 1 252 条 HA 的氨基酸序列, 中国香港 485 条 HA 氨基酸序列。并用同样的条件筛选出香港地

收稿日期: 2014-03-06

作者简介: 袁彦婷, 女, 在读硕士。研究方向: 病原微生物学及生物信息学。Email: yuanyanting@genomics.cn

通信作者: 陆祖宏, 男, 教授, 博士生导师。Email: luzuhong@gmail.com

区的甲型病毒 HA 基因蛋白序列。对筛选的中国大陆和中国香港的 HA 基因, 使用 Perl 语言编程, 把 HA 亚型和采集年代信息使用哈希进行存储和分析, 来汇总采样年代和各年感染的亚型种类。

2 基因多序列比对和系统发育树构建 对 HA 的基因做多序列比对。由于 MUSCLE 软件是一个常用的对蛋白多序列就行比对的软件, 所以我们使用 MUSCLE 软件 (v3.8.31) 对 HA 基因蛋白序列行进行多比对^[3]。并对比对后的 HA 基因构建系统发育树。选择 TreeBeST 软件来建树, 选用邻接法建树, 并设置 Bootstraps 参数为 1 000 对数的可靠性进行检验^[4]。

3 选择压力分析 对宿主为人、鸡、鸭和猪的 HA 基因分别做选择压力检验。由于受正向选择位点可为揭示进化历史和遗传动力提供信息, 所以我们可以选择压力分析来探索 HA 基因在不同宿主内的进化情况。我们选用 MEGA 软件 (Version5.5.2) 来分析 HA 基因的受选择压力^[5]。

4 统计学分析 Tajima's D 检验可以鉴定目标序列在进化过程中是否遵循中性进化模型^[6-8]。该检验通过比较不同方法产生的群体突变率的两 θ 估计值差异来检测正向选择, 即比较 θW ($\theta W=K/a$, 其中 K 为分离位点数, 即指所有序列比对后找出的变异碱基位点数, $a=[1+1/2+1/3+\dots+1/(n-1)]$) 和 θT ($\theta T=\pi$) 的差异值。而 θ 的理论值为 $\theta=4N_e\mu$ (其中 N_e 为有效群体大小, μ 为突变频率), 在中性进化时, θW 和 θT 的值得理论应相等, 此时 Tajima's D 的理论值为 0。若 Tajima's D 值明显偏离 0, 表明实际的等位基因频率相对于中性进化模型的期望出现偏离。所以我们通过 Tajima's D 检验值来推测 HA 基因的进化。

结 果

1 中国甲型流感亚型 对目前中国大陆的所有有完整基因组的毒株年代和亚型统计得出, 我国主要流行的亚型是 H6 (宿主基本为禽类, 如鸡、鸭和野鸭)、H1N1 (猪流感, 宿主为人)、H5N1、H7N9 和 H9N2 (宿主为禽类)。自 2002 年, 禽流感的暴发次数和频率高于历史记录, 且 2004 - 2013 年, 每年均有至少 5 种 HA 亚型出现感染。见图 1。

2 构建系统发育树 对中国大陆和中国香港的甲型流感病毒的 HA 基因氨基酸序列分别就行多序列比对后, 分别构建 NJ 树, 两个树的主要分支 Bootstrap 值均在 900 以上, 说明树的分支是可靠的。中国大陆和中国香港的系统发育树的拓扑结构一

致。从系统发育树结果可以看出, H5 亚型与 H2 亚型距离较近, H7 与 H10 较近缘, H3 与 H4 较近缘, H8、H12 与 H9 的亲缘关系较近。见图 2。

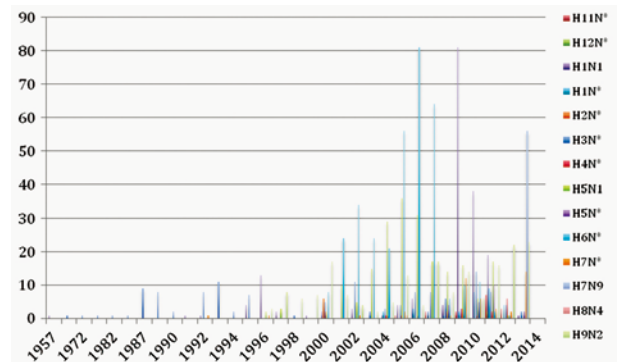


图 1 中国历年甲型流感的亚型及样品数目分布

H10N* 包含 H10N3、H10N5、H10N7、H10N8 和 H10N9; H11N* 包含 H11N2、H11N3 和 H11N9; H12N* 包含 H12N7 和 H12N8; H1N* 包含 H1N2 和 H1N3; H2N* 包含 H2N8 和 H2N9; H3N* 包含 H3N2、H3N3、H3N6、H3N8 和 H3N9; H4N* 包含 H4N1、H4N2、H4N6、H4N8 和 H4N9; H5N* 包含 H5N2 和 H5N5; H6N* 包含 H6N1、H6N2、H6N5、H6N6、H6N8 和 H6N9; H7N* 包含 H7N1、H7N2、H7N3、H7N6、H7N7 和 H7N8

Fig.1 Distribution of influenza A virus subtypes and its samples in China

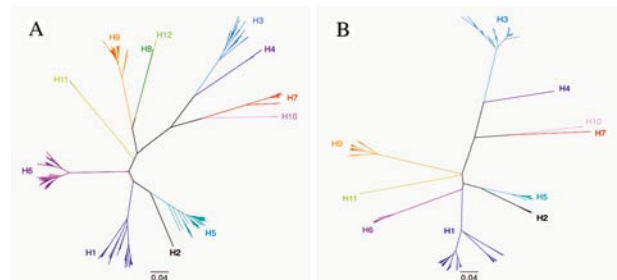


图 2 中国甲型流感病毒 HA 基因系统发育树 A: 中国大陆; B: 中国香港

Fig.2 Phylogenetic tree of HA gene for China mainland (A) and Hong Kong District (B)

表 1 中国甲型流感病毒不同宿主的 HA 基因的 Tajima 检验结果

Tab. 1 Tajima test of HA gene from different hosts of influenza A virus in China

Host	m	S	Ps	Θ	π	D
Human	196	247	0.860627	0.147046	0.440594	6.351911
Chicken	211	125	0.905797	0.152833	0.388152	4.787975
Duck	484	40	0.851064	0.125929	0.356007	4.961876
Swine	149	444	0.814679	0.146057	0.312858	3.740909

m=number of sequences; S=number of segregating sites; Ps=S/m; Θ = π s/a; π =nucleotide diversity; D=Tajima test statistic

3 自然选择分析 宿主为人的 HA 基因受选择作用最显著 (D 值 =6.35), 鸭与鸡为宿主的 HA 受选择作用其次 (D 值分别为 4.96、4.79), 宿主为猪的 HA 基因受选择作用最低 (D 值 =3.74)。虽然宿主为猪的受选择效应较其他宿主低, 但是也偏离了中性选择^[9-10]。见表 1。

讨 论

甲型流感的发生种类、次数与频率近年来都有所提高,其暴发受到多种因素影响,如候鸟迁徙、环境、家禽家畜养殖运输贩卖等^[11-12]。分析其暴发规律及毒株的进化,可以帮助预防控制禽流感。

本研究结合了流感病毒数据库分析、群体遗传学分析、中性进化检验等方法。借助了流感研究数据库,在海量的数据库中筛选出自己需要的信息。

分析中国流感历年发生的时间和亚型,发现中国的流感自2002年以后,不论是亚型种类还是暴发的次数都在提高,已出现同一年多种亚型均出现小范围流行的情况,人群一般对新亚型病毒缺少免疫力,所以如果同时出现多种亚型流行,很可能引起全球性大流行,这会给流感的防控带来不少压力^[13]。

对中国大陆和中国香港的系统发育树拓扑结构分析发现,中国分离的HA基因,在宏观上拓扑结构比较稳定。我们的系统发育树与Chan等^[14]的使用核酸分析出的树在拓扑结构上基本一致。这些都证明了这两个系统发育树的可信度,也反映出目前为止中国地区的HA基因亚型关系在大范围上是比较稳定且一致的。同时由于此树是无根树,所以树枝的远近只能反映出HA亚型间的距离远近,揭示HA基因亚型间的同源性,近缘的HA亚型在进化上更具有相似性。

对不同宿主的HA基因做选择压力分析后发现,不同宿主内的HA基因受选择压力情况有所不同,宿主为人的HA基因受选择比禽类或者猪要显著。这提示我们,HA基因受选择情况还与宿主有关,人体内的HA基因受选择作用较强,可能会促使HA基因快速进化。人宿主内的HA基因受明显的选择作用,这也许与抗生素滥用有一定的关联^[15]。也提示我们流感病毒可能在加速进化中。也有文献报道,病毒在家禽和人中并没有适应,还处于

快速的变异期^[11]。综上所述,甲型流感的发生种类、次数与频率近年来都有所提高,可能对人类造成很大威胁。

参考文献

- 1 Hilleman MR. Realities and enigmas of human viral influenza : pathogenesis, epidemiology and control [J] . *Vaccine*, 2002, 20 (25-26) : 3068-3087.
- 2 Squires RB, Noronha J, Hunt VA, et al. Influenza research database : an integrated bioinformatics resource for influenza research and surveillance [J] . *Influenza Other Respir Viruses*, 2012, 6 (6) : 404-416.
- 3 Edgar RC. MUSCLE : multiple sequence alignment with high accuracy and high throughput [J] . *Nucleic Acids Res*, 2004, 32 (5) : 1792-1797.
- 4 Nandi T, Ong C, Singh AP, et al. A genomic survey of positive selection in *Burkholderia pseudomallei* provides insights into the evolution of accidental virulence [J] . *PLoS Pathog*, 2010, 6 (4) : e1000845.
- 5 Tamura K, Peterson D, Peterson N, et al. MEGA5 : molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J] . *Mol Biol Evol*, 2011, 28 (10) : 2731-2739.
- 6 Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism [J] . *Genetics*, 1989, 123 (3) : 585-595.
- 7 K Wei YL, Xie D. Evolutionary and ecological dynamics of transboundary disease caused by H5N1 virus in Southeast Asia [J/OL] . <http://onlinelibrary.wiley.com/doi/10.1111/tbed.12147/full>.
- 8 Xu J, Zhong HA, Madrahimov A, et al. Molecular phylogeny and evolutionary dynamics of influenza A nonstructural (NS) gene [J] . *Infect Genet Evol*, 2014, 22 : 192-200.
- 9 Nei M, Kumar S. *Molecular evolution and phylogenetics* [M] . Oxford : Oxford University Press, 2000 : 212-213.
- 10 Li W, Shi W, Qiao H, et al. Positive selection on hemagglutinin and neuraminidase genes of H1N1 influenza viruses [J] . *Virology*, 2011, 8 : 183.
- 11 Cui L, Liu D, Shi W, et al. Dynamic reassortments and genetic heterogeneity of the human-infecting influenza A (H7N9) virus [J] . *Nat Commun*, 2014, 5 : 3142.
- 12 Chen Y, Liang W, Yang S, et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry : clinical analysis and characterisation of viral genome [J] . *Lancet*, 2013, 381 (9881) : 1916-1925.
- 13 Lukens S, Clermont G, Swigon D. Sensitivity of human immune response to influenza a virus infection and its dependence on virus and host phenotypes [J] . *J Crit Care*, 2013, 28 (1) : e7-e8.
- 14 Chan JM, Carlsson G, Rabadan R. Topology of viral evolution [J] . *Proc Natl Acad Sci U S A*, 2013, 110 (46) : 18566-18571.
- 15 Hai R, Schmolke M, Leyva-Grado VH, et al. Influenza A (H7N9) virus gains neuraminidase inhibitor resistance without loss of in vivo virulence or transmissibility [J] . *Nat Commun*, 2013, 4 : 2854.