

·综述·

系统发育研究在传染病分子流行病学中的应用

王璐琦 熊海燕 王伟炳

复旦大学公共卫生学院流行病教研室, 上海 200032

通信作者: 熊海燕, Email: haiyanxiong@fudan.edu.cn

【摘要】 测序技术的快速发展导致病原体基因数据急剧增加, 将这些数据与系统发育分析方法相结合, 可用于阐述病原体的起源和进化、流行过程中的时空分布、参数变化以及抗原、毒力、耐药性等表型特征变化规律、病原体传播趋势预测等。本文简述了系统发育研究的目的和系统发育树的构建方法, 阐述了距离法、最大简约法、最大似然法、贝叶斯法等常用系统发育重建方法的优缺点和适用范围, 并重点回顾了系统动力学和系统地理学方法在国内外研究中的应用及主要流行病学参数的估计方法, 通过综述也发现, 病毒基因组数据建立的带有时间和地点注释的系统发育树越来越多地用于传染病疫情暴发调查和常规监测。

【关键词】 传染病; 分子流行病学; 系统发育

基金项目: 国家自然科学基金(81673233); 上海市公共卫生体系建设三年行动计划(GWV-10.1-XK16)

Application of phylogenetic analysis in the molecular epidemiological study of infectious diseases

Wang Luqi, Xiong Haiyan, Wang Weibing

Department of Epidemiology, School of Public Health, Fudan University, Shanghai 200032, China

Corresponding author: Xiong Haiyan, Email: haiyanxiong@fudan.edu.cn

【Abstract】 The rapid development of sequencing technology brings the explosive growth of pathogen genetic data. The combination of genomic data and phylogenetic method is being used to elaborate the origin and evolution of pathogens, the time and space distribution and parameter changes in the prevalence process, and how phenotypes like antigen, virulence, and resistance change over time. This method is also being used to predict pathogen transmission trends. In this study, we described the aim of phylogeny and the process of the phylogenetic construction method. We elaborated the advantages and disadvantages and scope of application of tree-building methods including distance-based, maximum parsimony, maximum likelihood and bayesian methods. We have reviewed the application and the estimation methods of major epidemiological parameters of phylodynamics and phylogeography in domestic and foreign studies. We concluded that the time- and location-scaled phylogenetic trees are increasingly used for outbreak investigation and routine surveillance of infectious diseases.

【Key words】 Infectious disease; Molecular epidemiology; Phylogenetics

Fund programs: National Natural Science Foundation of China (81673233); Three-Year Initiative Plan for Public Health System Construction in Shanghai Municipality (GWV-10.1-XK16)

得益于测序技术成本的下降和分子生物学的发展, 传染病的分子流行病学研究取得了巨大成就, 在传染病暴发调查、日常监测和致病机制研究等方面发挥了重要作用^[1]。

但是, 关于病原体进化与流行之间关系的研究尚处于发展阶段, 而研究两者关系是理解病原体传播模式变化、表型变化、疫苗研发以及新发传染病出现的关键^[2-3]。

DOI: 10.3760/cma.j.cn112338-20210701-00515

收稿日期 2021-07-01 本文编辑 万玉立

引用格式: 王璐琦, 熊海燕, 王伟炳. 系统发育研究在传染病分子流行病学中的应用[J]. 中华流行病学杂志, 2022, 43(2): 282-285. DOI: 10.3760/cma.j.cn112338-20210701-00515.

Wang LQ, Xiong HY, Wang WB. Application of phylogenetic analysis in the molecular epidemiological study of infectious diseases[J]. Chin J Epidemiol, 2022, 43(2):282-285. DOI: 10.3760/cma.j.cn112338-20210701-00515.



迄今为止,国内外已有较多研究运用系统发育理论并结合复杂的进化模型与方法,从病原体的遗传序列中提取流行病学信息,进而重建传染病流行过程中病原体时间、空间甚至表型范围上的进化过程。本文介绍传染病系统发育研究中常见方法的概念、理论发展、目前的应用和局限。

1. 病原体的系统发育研究:系统发育分析(phylogenetic analysis)是描述病原体的起源、传播、进化以及证明病原体多样性的一种核心方法,有助于病原体溯源和中间宿主的判断^[4]。自从对病原体进行分子流行病学研究以来,系统发育分析就发挥了核心作用。与细菌和DNA病毒相比, RNA病毒是系统发育分析的主要病原体,因为其结构简单、突变速度快,可在短时间内观察到病原体的改变。同样在细菌性疾病中,如结核分枝杆菌、大肠埃希菌、霍乱弧菌等的系统发育也已得到广泛研究^[5-6]。

在新发传染病的研究中,系统发育溯源分析显得尤为重要,如新型冠状病毒运用系统发育分析表明,与MERS、SARS等共同属于冠状病毒家族^[7]。系统发育分析还可以帮助鉴定病原体的亚型或跨物种传播,如HIV的系统发育分析表明其是由猿类免疫缺陷病毒(SIV)跨物种传播而来并在进化过程中进行多种循环重组^[8]。除此之外,系统发育分析还可以推断传染病流行期间病原体的传播网络,因为传播网络中的病例通常具有相似的基因型,如结核分枝杆菌的系统发育分析表明其正在全球范围内传播^[9]。当前,系统发育分析一个有前途的应用是推断病原体进化与表型流行的关系,因为突变会导致毒力、耐药性、抗原等表型有较为明显的变化。使用系统发育作为支架,在该支架上绘制关键表型发生转变的位置,可以了解导致表型进化的原因,如位于更深节点的突变意味着更有可能具有选择性优势^[10]。

2. 系统发育树构建:系统发育树也称系统进化树(phylogenetic tree),即利用一种类似树枝状的图形来形象地展示物种间亲缘关系,是系统进化研究的核心。根据根有无可分为有根树和无根树,有根树呈二进制树模样,可以表示进化方向,构建时通常需要引入一个外群;而无根树只能显示节点之间的关系^[11]。树的外节点代表当前物种的序列,内节点代表已灭绝祖先,其序列数据通常无法获得。建树的序列可以是核苷酸序列或者氨基酸序列;主流建树软件众多,其中MEGA属于图形化软件,因界面友好而被广泛使用;方法包括距离法(DM)、最大简约法(MP)、最大似

然法(ML)和贝叶斯法(BP),其中DM又包括最少进化法(ME)和邻位法(NJ)。由于NJ建树极快,通常用于建树尝试阶段,而正式建树常选用可靠性高的ML。理论上应该尝试过所有的方法之后,根据一定的标准选择最可靠的树,但这在现实中几乎很难实现,除非是非常小的序列数据集。树构建的方法和软件可参考张丽娜等^[12]的研究。系统发育树的选择标准、优缺点和适用范围见表1。

随着新发和再发传染病事件近年来呈上升趋势,一种新型的带有时戳的贝叶斯进化树正在兴起^[13],其节点和分支带有病原体可能被引入当地传播的时间,有助于在疾病暴发和流行期间实时管理疫情。常用贝叶斯物种分化时间估计软件可参考朱天琪^[14]的研究。BEAST是最常用的软件之一,通过单击界面导入序列、设置分类群、序列收集日期、核苷酸替代模型、分子钟类型、树先验模型并调整参数的权重,结合马尔科夫链蒙特卡罗算法采样,得到收敛的可靠性高的后验进化树,已广泛用于系统动力学和系统地理学研究^[15-16]。

3. 传染病系统动力学(phyldynamics)研究:传染病系统动力学研究流行病学、免疫学和进化过程如何相互作用以形成病原体谱系,是传染病研究最新和发展最快的领域之一^[17],已被用来重建多种病原体在最近或更长时间范围内的传播动力学。其以系统发育分析为基础,结合序列采样时间,估计病原体的分歧时间、进化速率、种群增长和流行病学参数等^[18]。虽然现有研究主要关注快速进化的RNA病毒,但全基因组测序技术的发展使系统动力学方法扩展到分析缓慢进化的微生物(例如细菌和真菌等)成为可能^[5,19]。

(1)分歧时间估计:疾病何时出现、何时传播一直以来都是流行病学关心的重点问题。在溯祖理论和分子钟假设下,通过对已知的分化时间设置先验分布,生成带有时戳的进化树,帮助我们了解病原体的最近共同祖先出现时间、分歧时间和进化速率等^[20-22]。进行分歧时间估计之前应先进行序列比对,并计算合适的核苷酸替代模型,后在BEAUti软件包中设置一系列参数,最后在BEAST软件中生成时间树,并在Tracer软件中根据收敛程度评价树的可靠性。需要注意的是基于分子钟模型估计出来的分歧时间和进化速率并不一定完全符合过去实际发生的情况,因为病原体基因组的代表性和分子异速进化现象的存在。

(2)有效种群数量(N_e)估计:传染病病原体 N_e 是传播

表1 系统发育树构建方法

方法	系统发育树选择标准	优点	缺点	适用范围
距离法	树长总距离最小的树	计算速度极快	相对准确	尤其适用于进化距离小的序列
最大简约法	进化过程中所需核苷酸(或氨基酸)替换数目最少的树	可分析某些特殊序列如含有插入、缺失等片段	计算所需时间较长	适用于序列间具有显著相似性的序列
最大似然法	给定进化模型,具有最高对数似然值的树	准确性高	计算复杂、所需时间长	尤其适用于有合理的核苷酸替代模型时
贝叶斯法	具有最大后验概率的树	准确性高;与马尔科夫链蒙特卡罗算法结合可估计拓扑结构等各参数值;估计分歧时间	受先验影响;计算所需时间长	适用于大而复杂的数据集

过程中一个重要的流行病学参数,可用于评价防控工作的有效性,当 Ne 呈指数增长时提示疾病正在流行^[23];此外,在一定条件下, Ne 还可用作传染病流行率的估计^[24]。估计 Ne 最早的方法是溯祖模型,但该模型只能描述整个种群在时间上向后发展即过去的动态,且只能基于恒定规模、指数增长等情形^[25-26];鉴于传染病流行过程通常呈随机非线性,研究人员开发了一种更为灵活的天际线模型,其允许种群规模变化函数以分段的方式组合,以适应多变的种群规模^[27-28],但其尚未允许纳入协变量;最近的研究正在试图考虑将影响种群规模的协变量纳入模型中,如季节、温度等^[29]。

(3)基本再生数(R_0)估计:传染病 R_0 是传播动力学研究中一个重要参数,即感染个体在完全易感人群中引起的二次感染的平均数,它可以提示我们疾病流行期间的干预措施是否有效^[30]。虽然 R_0 常常通过传统动力学模型(SIR等)估计,但越来越多的研究开始使用分子序列估计。因为系统发育树根部的序列可以代表传染病暴发阶段,分枝末端的序列可以代表最近的传播。因此,树分支模式提供了传播随时间变化的特点。应用于引起大流行的多种病原体,如埃博拉、流感、新型冠状病毒等^[31-32]。最常用的两种树先验模型是指数增长模型和出生-死亡模型。当疾病暴发初期病原体呈指数增长时,在BEAUti软件包中将树先验设置为指数增长获得可靠的指数增长率后验值 r ,结合人际传染间隔时间 D , $R_0 = rD + 1$ ^[33]。然而,传染病流行特点往往受多种因素影响呈非线性变化,当病例数量不再呈指数增加时,一种更为灵活的估计是将树先验设置为出生-死亡模型,进而获得可靠的出生率和死亡率后验估计值。出生率对应于传染病流行期间的传播率 λ ,死亡率对应于感染者因治愈或死亡等原因转为非感染者率 μ ,结合采样率 ψ , $R_0 = \lambda/(\mu + \psi)$ ^[34]。

4. 传染病系统地理学(phylogeography)研究:传染病系统地理学是了解传染病病原体的起源地、迁移和地理分布的有力工具。该研究的原理是利用系统发育树将遗传序列数据与空间位置相结合,生成带有位置信息的系统发育树,进而分析疾病流行过程中系统地理学模式,已应用于大量威胁人类健康的传染病^[35-36]。

传染病病原体的系统地理学模式主要有混合传播、波状传播、源头输入传播和空间分割传播等^[37]。混合传播在系统发育树上表现为来自不同地理位置的样本随机混合;而波状传播在系统发育树上表现为传播距离越来越远;源头输入传播表现为病原体某一分型是其他流行地区的来源,如全球季节性流感可能源自印度尼西亚等热带地区的抗原漂移流感病毒株,这些病毒株每年可通过空中交通网络迁移到南北半球^[38];空间分割传播则表现为病原体的不同类型或亚型都有不同的地理分布,通常与特定的风险群体有关,如丙型肝炎病毒HC1a和3a亚型主要存在于注射毒品使用者中^[39]。病原体传播过程中可呈单一系统地理学模式或呈多种系统地理学模式共存。

系统地理学研究的主要目标是重现病毒在传播范围内的运动轨迹,通过收集序列的采样日期和地点,设置带有离散特征或连续特征注释的MCC树,在SPREAD或SPREAD3(<https://rega.kuleuven.be/cev/ecv/software/Software>)软件中进行分析。迄今为止,多种病原体传播的系统地理分布谱系图已被绘制^[35-37]。系统地理学的另一个研究目标是结合数学模型,发现与病原体运动相关的社会、地理或环境因素,帮助预测未来的传播,如Dudas等^[40]重建了2014-2016年埃博拉病毒在疫情最严重的3个国家内部和之间的传播情况,并整合空间变量研究地理、气候、环境、经济等因素对预测埃博拉病毒输入风险程度的可能性。证明了国家边界对埃博拉病毒的地理传播有抑制作用,而城市特点(经济产量、人口密度)和气候因素与病毒传播没有显著关联。Njoto等^[41]的研究确定了高致病性禽流感H5N1病毒的传播与野生鸟类飞行路线和家禽贸易路线一致,并与鸡密度高和人类发展指数低的地区有关。

然而,系统地理学研究的准确性与采样覆盖率和代表性有很大关系,在“方便”采样盛行的情况下,局限在特定地理区域取样可能会限制对空间传播动态的全面理解。

5. 总结与展望:从病原体基因组中提取进化和流行病学信息已发展为传染病流行病学研究中重要一部分。然而,到目前为止,大部分研究仍集中在小基因组的RNA病毒上。为了更好地预防人类传染病,需要更多关于DNA病毒、细菌性病原体等的系统动力学和系统地理学研究。尽管基因组测序技术飞速发展带来了海量数据,但同时意味着研究者需要学习更多的基因大数据处理方法以及需要更高的计算机存储和运行环境。此外,由于病原体进化过程十分复杂,统计学方法很难精确模拟,导致结果与事实常存在一定偏倚,需要更加强大和高效的统计计算方法。未来,基于基因序列进行流行病学研究和日常监测,了解病原体进化过程中的传播动力与机制,将会是一个很有前景的领域。

利益冲突 所有作者声明无利益冲突

参 考 文 献

- [1] Riley LW, Blanton RE. Advances in molecular epidemiology of infectious diseases: definitions, approaches, and scope of the field[J]. Microbiol Spectr, 2018, 6(6):AME-0001-2018. DOI:10.1128/microbiolspec. AME-0001-2018.
- [2] Liberles DA, Chang B, Geiler-Samerotte K, et al. Emerging frontiers in the study of molecular evolution[J]. J Mol Evol, 2020, 88(3):211-226. DOI:10.1007/s00239-020-09932-6.
- [3] Geoghegan JL, Holmes EC. The phylogenomics of evolving virus virulence[J]. Nat Rev Genet, 2018, 19(12):756-769. DOI:10.1038/s41576-018-0055-5.
- [4] Geoghegan JL, Holmes EC. Evolutionary virology at 40[J]. Genetics, 2018, 210(4):1151-1162. DOI:10.1534/GENETICS. 118.301556.
- [5] Ingle DJ, Howden BP, Duchene S. Development of phylodynamic methods for bacterial pathogens[J]. Trends Microbiol, 2021, 29(9):788-797. DOI:10.1016/j.tim.2021. 02.008.
- [6] Choi SY, Rashed SM, Hasan NA, et al. Phylogenetic diversity of *Vibrio cholerae* associated with endemic cholera in Mexico from 1991 to 2008[J]. mBio, 2016, 7(2): e02160. DOI:10.1128/mBio.02160-15.

- [7] Sironi M, Hasnain SE, Rosenthal B, et al. SARS-CoV-2 and COVID-19: a genetic, epidemiological, and evolutionary perspective[J]. Infect Genet Evol, 2020, 84: 104384. DOI: 10.1016/j.meegid.2020.104384.
- [8] Zulfiqar HF, Javed A, Sumbal, et al. HIV diagnosis and treatment through advanced technologies[J]. Front Public Health, 2017, 5:32. DOI:10.3389/fpubh.2017.00032.
- [9] Napier G, Campino S, Merid Y, et al. Robust barcoding and identification of *Mycobacterium tuberculosis* lineages for epidemiological and clinical studies[J]. Genome Med, 2020, 12(1):114. DOI:10.1186/s13073-020-00817-3.
- [10] Urbanowicz RA, McClure CP, Sakuntabhai A, et al. Human adaptation of Ebola virus during the west African outbreak[J]. Cell, 2016, 167(4): 1079-1087. e5. DOI: 10.1016/j.cell.2016.10.013.
- [11] 迟文静, 刘宜昕, 王粟, 等. 进化树在细菌亲缘关系分析中的应用研究[J]. 检验医学, 2020, 35(12):1310-1314. DOI: 10.3969/j.issn.1673-8640.2020.12.026.
- [12] Chi WJ, Liu YX, Wang S, et al. Evolutionary tree and its application in the analysis of bacterial kinship[J]. Lab Med, 2020, 35(12):1310-1314. DOI:10.3969/j.issn.1673-8640.2020.12.026.
- [12] 张丽娜, 荣昌鹤, 何远, 等. 常用系统发育树构建算法和软件鸟瞰[J]. 动物学研究, 2013, 34(6): 640-650. DOI: 10.11813/j.issn.0254-5853.2013.6.0640.
- [12] Zhang LN, Rong CH, He Y, et al. A bird's eye view of the algorithms and software packages for reconstructing phylogenetic trees[J]. Zool Res, 2013, 34(6):640-650. DOI: 10.11813/j.issn.0254-5853.2013.6.0640.
- [13] Nascimento FF, dos Reis M, Yang ZH. A biologist's guide to Bayesian phylogenetic analysis[J]. Nat Ecol Evol, 2017, 1(10):1446-1454. DOI:10.1038/s41559-017-0280-x.
- [14] 朱天琪. 使用基因组数据进行贝叶斯物种分化时间估计[J]. 中国科学:生命科学, 2019, 49(4):472-483. DOI:10.1360/N052018-00224.
- [14] Zhu TQ. Bayesian molecular dating with genomic data[J]. Sci Chin: Life Sci, 2019, 49(4): 472-483. DOI: 10.1360/N052018-00224.
- [15] Wright AM. A systematist's guide to estimating Bayesian phylogenies from morphological data[J]. Insect Syst Divers, 2019, 3(3):2. DOI:10.1093/isd/ixz006.
- [16] Darriba D, Flouri T, Stamatakis A. The state of software for evolutionary biology[J]. Mol Biol Evol, 2018, 35(5): 1037-1046. DOI:10.1093/molbev/msy014.
- [17] Lessler J, Azman AS, Grabowski MK, et al. Trends in the mechanistic and dynamic modeling of infectious diseases[J]. Curr Epidemiol Rep, 2016, 3(3): 212-222. DOI: 10.1007/s40471-016-0078-4.
- [18] Vasylyeva TI, Zarebski A, Smyrnov P, et al. Phylodynamics helps to evaluate the impact of an HIV prevention intervention[J]. Viruses, 2020, 12(4): 469. DOI: 10.3390/v12040469.
- [19] Rife BD, Mavian C, Chen XG, et al. Phylodynamic applications in 21st century global infectious disease research[J]. Glob Health Res Policy, 2017, 2: 13. DOI: 10.1186/s41256-017-0034-y.
- [20] Nabil B, Sabrina B, Abdelhakim B. Transmission route and introduction of pandemic SARS-CoV-2 between China, Italy, and Spain[J]. J Med Virol, 2021, 93(1):564-568. DOI: 10.1002/jmv.26333.
- [21] 李伟红, 高志勇, 严寒秋, 等. 2016-2019 年北京市 G II.2 [P16] 型诺如病毒全长基因组系统进化分析[J]. 国际病毒学杂志, 2020, 27(2): 97-101. DOI: 10.3760/cma.j.issn.1673-4092.2020.02.003.
- [21] Li WH, Gao ZY, Yan HQ, et al. Phylogenetic analysis of complete genomes of G II.2 [P16] noroviruses in Beijing, 2016-2019[J]. Int J Virol, 2020, 27(2): 97-101. DOI: 10.3760/cma.j.issn.1673-4092.2020.02.003.
- [22] 金青青, 茅海燕, 孙逸, 等. 乙型流感病毒 Victoria 系和 Yamagata 系 HA1 基因的分子进化研究[J]. 中华流行病学杂志, 2013, 34(4):366-370. DOI: 10.3760/cma.j.issn.0254-6450.2013.04.013.
- [22] Jin QQ, Mao HY, Sun Y, et al. Molecular evolution of two lineages related to influenza B virus based on HA1 gene[J]. Chin J Epidemiol, 2013, 34(4):366-370. DOI:10.3760/cma.j.issn.0254-6450.2013.04.013.
- [23] Volz EM, Koelle K, Bedford T. Viral phylodynamics[J]. PLoS Comput Biol, 2013, 9(3):e1002947. DOI:10.1371/journal.pcbi.1002947.
- [24] Ragonnet-Cronin M, Boyd O, Geidelberg L, et al. Genetic evidence for the association between COVID-19 epidemic severity and timing of non-pharmaceutical interventions[J]. Nat Commun, 2021, 12(1): 2188. DOI: 10.1038/s41467-021-22366-y.
- [25] Müller NF, Rasmussen D, Stadler T. MASCOT: parameter and state inference under the marginal structured coalescent approximation[J]. Bioinformatics, 2018, 34(22):3843-3848. DOI:10.1093/bioinformatics/bty406.
- [26] Alkhamis MA, Perez AM, Murtaugh MP, et al. Applications of Bayesian phylodynamic methods in a recent U. S. porcine reproductive and respiratory syndrome virus outbreak[J]. Front Microbiol, 2016, 7: 67. DOI: 10.3389/fmicb.2016.00067.
- [27] Parag KV, Pybus OG, Wu CH. Are skyline plot-based demographic estimates overly dependent on smoothing prior assumptions? [J]. Syst Biol, 2021, syab037. DOI: 10.1093/sysbio/syab037. Online ahead of print.
- [28] Nie Q, Li XG, Chen W, et al. Phylogenetic and phylodynamic analyses of SARS-CoV-2[J]. Virus Res, 2020, 287:198098. DOI:10.1016/j.virusres.2020.198098.
- [29] Hill V, Baele B. Bayesian estimation of past population dynamics in BEAST 1.10 using the skygrid coalescent model[J]. Mol Biol Evol, 2019, 36(11): 2620-2628. DOI: 10.1093/molbev/msz172.
- [30] Miller D, Martin MA, Harel N, et al. Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel[J]. Nat Commun, 2020, 11(1):5518. DOI: 10.1038/s41467-020-19248-0.
- [31] Lai A, Bergna A, Acciarri C, et al. Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2[J]. J Med Virol, 2020, 92(6): 675-679. DOI: 10.1002/jmv.25723.
- [32] Lam A, Duchene S. The impacts of low diversity sequence data on phylodynamic inference during an emerging epidemic[J]. Viruses, 2021, 13(1): 79. DOI: 10.3390/v13010079.
- [33] Anderson RM, Donnelly C, Hollingsworth TD, et al. Reproduction number (*R*) and growth rate (*r*) of the COVID-19 epidemic in the UK: methods of estimation, sources data, causes of heterogeneity, and use as a guide in policy formulation[EB/OL]. (2020-08-24) [2021-06-30]. <https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf>.
- [34] Stadler T, Kouyos R, von Wyl V, et al. Estimating the basic reproductive number from viral sequence data[J]. Mol Biol Evol, 2012, 29(1): 347-357. DOI: 10.1093/molbev/msr217.
- [35] Reimering S, Muñoz S, McHardy AC. Phylogeographic reconstruction using air transportation data and its application to the 2009 H1N1 influenza pandemic[J]. PLoS Comput Biol, 2020, 16(2):e1007101. DOI:10.1371/journal.pcbi.1007101.
- [36] Sun YK, Han XL, Wei YF, et al. Phylogeography, phylodynamics and the recent outbreak of lineage 3 porcine reproductive and respiratory syndrome viruses in China[J]. Transbound Emerg Dis, 2019, 66(5): 2152-2162. DOI:10.1111/tbed.13269.
- [37] Holmes EC. Evolutionary history and phylogeography of human viruses[J]. Annu Rev Microbiol, 2008, 62:307-328. DOI:10.1146/annurev.micro.62.081307.162912.
- [38] Root ED, Agustian D, Kartasmita C, et al. Demographic and ecological risk factors for human influenza A virus infections in rural Indonesia[J]. Influenza Other Respir Viruses, 2017, 11(5):425-433. DOI:10.1111/irv.12468.
- [39] Husic-Selimovic A, Sofic A, Jahic E, et al. Impact of different sources of infection on therapy response in chronic hepatitis C[J]. Med Arch, 2018, 72(3): 182-186. DOI:10.5455/medarch.2018.72.182-186.
- [40] Dudas G, Carvalho LM, Bedford T, et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic[J]. Nature, 2017, 544(7650): 309-315. DOI: 10.1038/nature22040.
- [41] Njoto EN, Scotch M, Bui CM, et al. Phylogeography of H5N1 avian influenza virus in Indonesia[J]. Transbound Emerg Dis, 2018, 65(5): 1339-1347. DOI: 10.1111/tbed.12883.