

技术方法

DOI: 10.13406/j.cnki.cyxh.002570

基于 VOSviewer 软件的 COVID-19 关系图研究

胡仁旺, 龚建平

(华中科技大学同济医学院附属同济医院胃肠外科, 武汉 430000)

【摘要】目的:寻找新型冠状病毒肺炎(coronavirus disease 2019, COVID-19)的研究方向,理清研究思路。**方法:**通过 PubMed 或中国知网数据库检索 COVID-19 的相关文献,检索时间起点为 2019 年 11 月 1 日,止点时间不限,利用 VOSviewer 软件对纳入的文献关键词构建研究热点关系网络图,通过不同节点的大小和复杂程度,发现不同研究热点之间的关联性,从而不断寻找新的研究热点,为疾病的研究指明方向。**结果:**COVID-19 与严重急性呼吸综合征(severe acute respiratory syndrome, SARS)、中东呼吸综合征(Middle East respiratory syndrome, MERS)、血管紧张素转化酶 2(angiotensin converting enzyme 2, ACE2)、spike 蛋白等研究热点具有明显的相关性。部分研究者已经重视到重症患者细胞因子释放综合征(cytokine release syndrome, CRS)或细胞因子风暴的存在。ACE2 与 at1 基因及其受体、mas 受体、ang-(1-7)、心脏和肺损伤、肥胖或糖尿病患者的相关性较大。spike 蛋白与 T 细胞表面抗原 CD26(dpp4)、s1 基因及其蛋白、弗林蛋白酶(furin)及跨膜丝氨酸蛋白酶 2(transmembrane serine protease 2, TMPRSS2)具有一定相关性。细胞因子风暴与 T 细胞、B 细胞及其相关的 CD19、CD40 和 CD28 受体等相关。同时 CAR-T 疗法(嵌合抗原受体 T 细胞免疫疗法)、白介素-6(interleukin-6, IL-6)与细胞因子风暴也有明显的相关性。**结论:**通过构建 COVID-19 研究热点网络图,可以在很短的时间内了解研究进展,寻找下一步研究思路。

【关键词】新型冠状病毒;新型冠状病毒肺炎;VOSviewer

【中图分类号】R512.99

【文献标志码】A

【收稿日期】2020-03-12

Relation graph of COVID-19 based on VOSviewer software

Hu Renwang, Gong Jianping

(Department of Gastroenterology, Tongji Hospital, Tongji Medical College,
Huazhong University of Science and Technology)

【Abstract】Objective: To find out the research direction of coronavirus disease 2019 (COVID-19) and explore the research ideas. **Methods:** Relevant literatures of COVID-19 were extracted on PubMed or CNKI Databases, with starting point of November 1, 2019. VOSviewer software was used to build the keywords network map of included literatures. Discovered the correlation among different research hot points through the size and complexity of different nodes, so as to constantly look for new research ideas and point out the direction for disease research. **Results:** COVID-19 was significantly correlated with hot points of severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), angiotensin converting enzyme 2 (ACE2) and spike protein. Some researchers have paid attention to the existence of cytokine release syndrome (CRS) or cytokine storm in critically ill patients. There was a strong correlation between ACE2 and at1 gene and its receptor, mas receptor, ang-(1-7), heart and lung injury, and obesity or diabetes. Spike protein was correlated with T cell surface antigen CD26(dpp4), S1 gene and its protein, furin protease and transmembrane serine protease 2 (TMPRSS2). Cytokine storm was associated with T cells, B cells and related CD19, CD40 and CD28 receptors. At the same time, CAR-T therapy (chimeric antigen receptor T cell immunotherapy) and interleukin-6 (IL-6) were also significantly correlated with cytokine storm. **Conclusion:** By constructing the COVID-19 research network map, we can understand the research progress in a very short time and find out the further research ideas.

【Key words】 2019 novel coronavirus; coronavirus disease 2019; VOSviewer

作者介绍: 胡仁旺, Email: huhu1366@sina.com,

研究方向: 分子医学、胃肠外科常见疾病诊疗。

通信作者: 龚建平, Email: jpgong@tjh.tjmu.edu.cn。

优先出版: <https://kns.cnki.net/kcms/detail/50.1046.R.20200618.1632.050.html>

(2020-06-19)

新型冠状病毒肺炎(coronavirus disease 2019, COVID-19)的暴发,已经发展为公共卫生事件。由于病毒的传染性较强,短时间内大量患者被感染,严重威胁人民的生命健康安全^[1]。疾病初期,由于对疾病特点的了解不充分,很难作出完全合理的预防措施。随着确诊患者的增多,以及对疾病研究的加深,广大医务者越来越全面地掌握了 COVID-19 的临床特点及诊疗方法。目前疫情在很大程度上得以控制,也有很多的诊疗经验发表在中外学术期刊。由于医务工作者临床工作繁重,很难在短时间内了解所有的研究进展。如何在短时间内从大量的学术文献中寻找研究热点的相关性、理清下一步研究思路,是极为重要的。本研究基于 VOSviewer 软件,通过对大量学术文献进行聚类分析,构建研究热点网络图,发现研究热点的相关性,从而更好地指导研究方向。本研究是目前中文数据库和英文数据库中第 1 篇关于 COVID-19 研究思路寻找的研究,具有很大的临床意义。

1 材料与方法

通过 PubMed 数据库或 CNKI 数据库,检索 COVID-19 的所有文献,检索时间起点为 2019 年 11 月 1 日,止点时间不限。通过 PubMed 数据库检索 COVID-19 相关文献,检索式为“2019-nCoV[Title/Abstract]” OR “2019 novel coronavirus [Title/Abstract]” OR “Novel coronavirus pneumonia[Title/Abstract]” OR “SARS-CoV-2[Title/Abstract]” OR “Severe acute respiratory syndrome coronavirus 2[Title/Abstract]” OR “COVID-19[Title/Abstract]”。通过中国知网数据库检索 COVID-19 相关文献,检索式为“新型冠状病毒肺炎”或“COVID-19”。然后利用 VOSviewer 软件对所有文献的研究热点,即文献关键词进行聚类分析,构建研究热点网络图,通过不同节点的大小和复杂程度,发现研究热点的关联性。通过对 COVID-19 的热点分析,抽取目前研究的突破点,然后再以突破点为关键词,重新检索中、英文文献,再次构建研究热点网络图,以此类推,不断寻找研究热点的关联性,构建研究热点网络图。从而发现研究思路,指明研究方向。

2 结果

2.1 COVID-19 研究热点网络图(一级网络图)

通过筛选和删除重复后,检索到英文相关文章 277 篇,利用 VOSviewer 软件对关键词进行网络构建,发现 387 个关键词,纳入 472 篇中文期刊文章,发现 1 114 个关键词,选取英文和中文数据库频数表前 15 位关键词,绘制表 1。然后将 PubMed 数据库所有关键词纳入,由软件自动聚类分析,分为 35 类(聚类信息只能通过 VOSviewer 软件进行查看,原理为通过关键词之间的关联性进行聚类,无具体的聚类名称,只有纳入聚类的关键词条目),绘制研究热点网络图如图 1 所示。将中国知网数据库关键词出现 2 次以上的纳入,由软件自动聚类分析,分为 24 类,绘制研究热点网络图如图 2 所

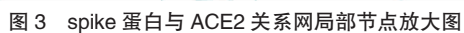
示。通过中英文关键词研究热点网络图构建,重点关注网络图中节点较大、复杂程度较高的关键词,因为节点的大小代表关键词出现的频率,节点的复杂程度代表不同研究方向的交叉,节点越复杂说明目前研究的范围越广,程度越深。同时结合目前已经报告出来的研究热点,进一步缩小研究方向。研究发现,新型冠状病毒与严重急性呼吸综合征(severe acute respiratory syndrome, SARS)、中东呼吸综合征(Middle East respiratory syndrome, MERS)^[2]、血管紧张素转化酶 2(angiotensin converting enzyme 2, ACE2)、spike 蛋白^[3]等具有明显的相关性。由于图片网络节点太多,本文只能展示部分内容。网络图可以很好地显示研究方向的发生发展变化,此处以最近的研究热点 spike 蛋白与新型冠状病毒受体 ACE2 为例,放大其中一个局部节点,如图 3 所示。所有的关键词节点均可以放大,从而了解目前的研究热点^[4],寻找研究思路;同时还可以明显地看到很多研究者关注此次肺炎的流行病学特点、CT 表现^[5]、预测模型^[6],关注疾病的预防和治疗方法等^[7]。总之,通过构建关系网络图,能对疾病的认识带来宏观感受,最重要的是指明下一步研究方向。本文后面章节中(2.2~2.7)的网络图构建是通过构建新型冠状病毒肺炎研究热点网络图后,结合目前已经报告出来的研究热点进行的二级网络图构建,在 VOSviewer 软件中有更加直观的表现,因此不再描述关键词频数表等信息。

表 1 COVID-19 研究热点频数表(前 15 位)

关键词 (PubMed 数据库)	频数表	关键词 (中国知网数据库)	频数表
COVID-19	197	新型冠状病毒肺炎	245
coronavirus	59	新型冠状病毒	218
SARS	23	肺炎	43
epidemiology	22	冠状病毒感染	21
outbreak	19	中医药	20
pneumonia	17	冠状病毒肺炎	16
China	15	疫情防控	15
humans	12	感染防控	14
Wuhan	12	冠状病毒	12
MERS	11	诊疗方案	12
child	7	儿童	11
infection	7	感染	11
ACE2	6	抗病毒药物	11
diagnosis	5	疫病	10
public health	4	防护	10

2.2 SARS 研究热点网络图

COVID-19 的研究与 SARS 的研究具有明显相关性^[8]。通过 PubMed 数据库检索 SARS 相关文献,检索式为“SARS[Title/Abstract]” OR “SARS-CoV[Title/Abstract]”,检索到相关文献 8 284 篇,利用 VOSviewer 软件对关键词进行网络构建,发现 5 217 个关键词,研究热点网络图如图 4 所示。本研究发现,SARS 研究热点主要围绕动物模型、疫苗、c-met 基因、nsp14 非结构蛋白^[9-10]、mda5 抗体、ACE2 等各种指标。



蛋白^[11]、受体结合域^[12-13]、自身免疫反应、基因组 k-mer 分析等各种方向,同时疾病的诊疗、预防方案也是最主要的研究热点之一。

2.4 ACE2 研究热点网络图

COVID-19 的研究与 ACE2 的研究具有明显相关性, ACE2 目前也被认为是新型冠状病毒的受体。通过 PubMed 数据库检索 ACE2 相关文献,检索式为“angiotensin-convert-
ing enzyme 2[Title/Abstract]” OR “ACE2[Title/Abstract]”,检索

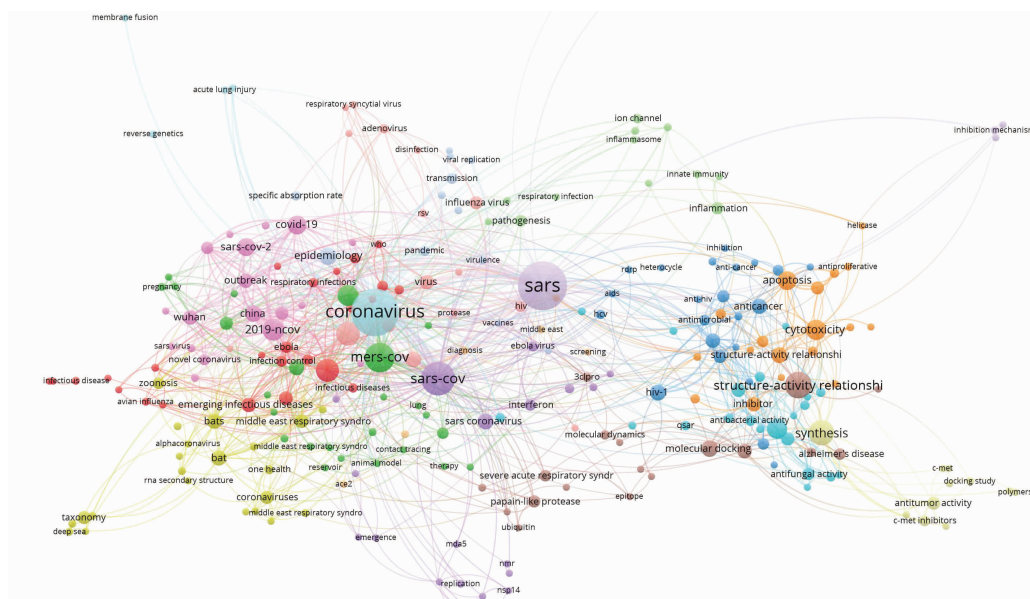


图 4 SARS 研究热点网络图

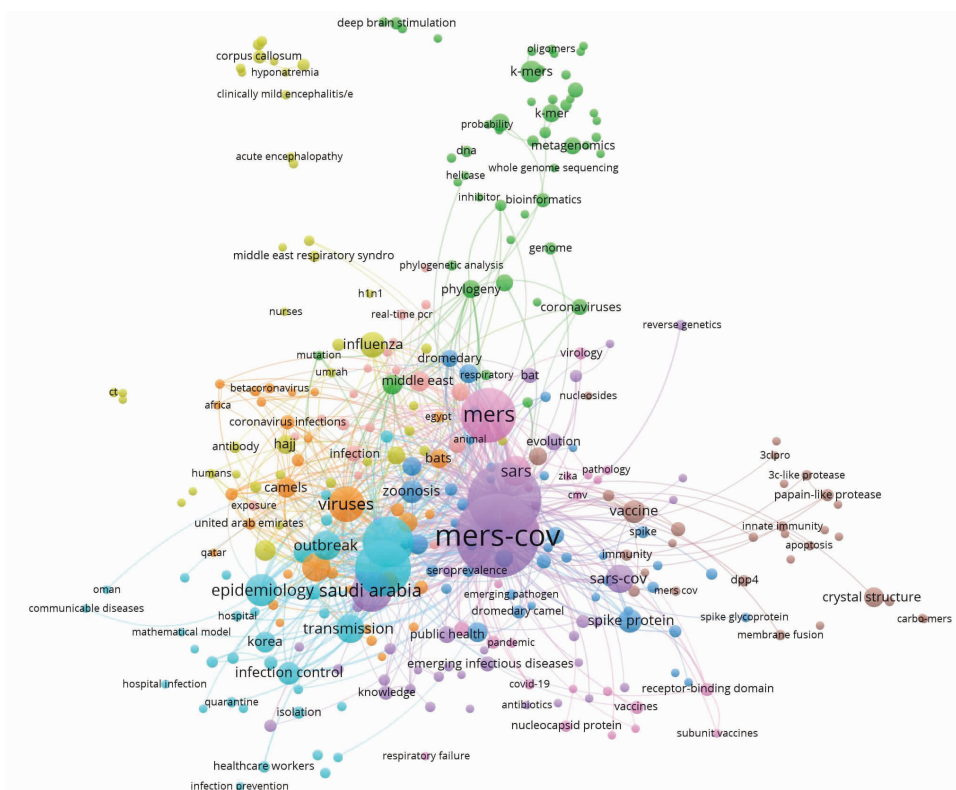


图 5 MERS 研究热点网络图

PCR 检测^[19]、spike 蛋白、模型构建等仍然是研究的热点。

3 讨论

近些年来,互联网不断发展,各种公开数据库不断出现,获取文献更为简便。但是,由于文献量巨大,研究者很难在短时间内了解某一疾病的研究热点或各种研究热点之间的内在关系。VOSviewer 软件是一款文献计量软件,专为探索某一领域研究热点的内在关系而设计。通过聚类分析,绘制研究

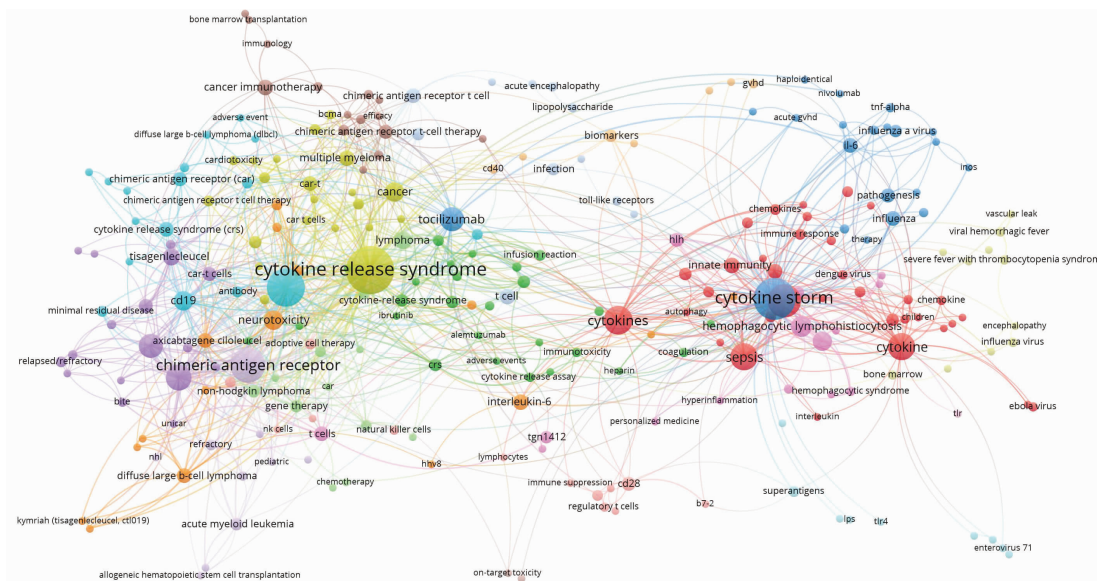


图 8 细胞因子释放综合征或细胞因子风暴研究热点网络图

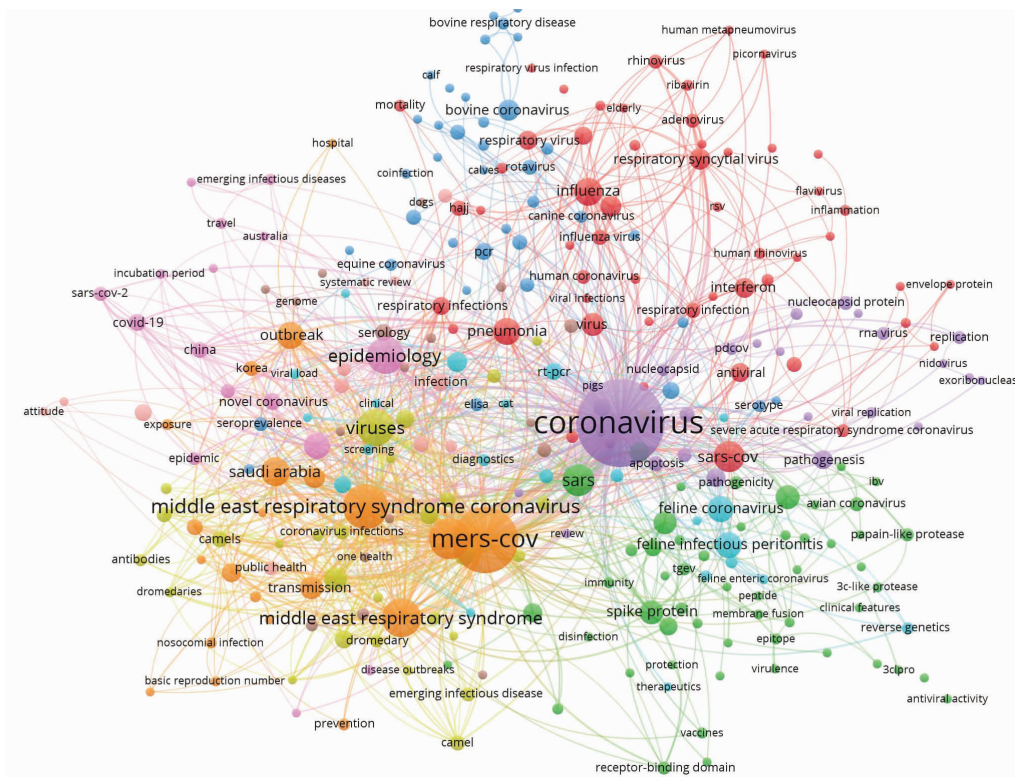


图 9 冠状病毒研究热点网络图

热点网络图,可以很清楚地找到研究热点的内部联系。绘制的网络图信息量很大,每一个节点都与其他节点具有千丝万缕的联系,在动态图中可以得到很好的体现,为研究者发现研究热点、寻找研究方向提供了很大的便利。

由于文献量巨大,短时间内阅读完所有文献是不可能的。例如,本研究对冠状病毒研究热点的探索,从 PubMed 数据库检索到文献 9 902 篇,若将这么多的文献浏览完是一个很大的工作量。通过聚类分析发现仍然有 5 943 个关键词,但此时的研究热点已经可以得到大致了解,这些研究热点的内部关系如何呢?通过设置节点阈值,构建研究热点网络图,便可很清晰地发现其内部关系,同时也能够更好地寻找研究思路。比如,在构建网络图的过程中发现之前的 SARS 研究与 nsp14 非结构蛋白^[20]、spike 蛋白等热点有关,而 spike 蛋白又与 T 细胞表面抗原 CD26(dpp4)^[21]、furin 及 TMPRSS2^[22-23]明显相关性。这种研究热点之间的不断串联、不断进行网络图的构建,可以帮助我们探索研究热点之间的内部关系,同时也能了解世界各国研究者的关注重点。

COVID-19 的暴发对医务工作者是一个很大的考验。面对一个未知的病毒,在初期阶段缺乏很多的确诊手段和治疗方法。随着疾病的进展和诊疗经验的积累,基础研究者也会从各个方面对病毒的发病机制进行研究。此时,构建研究热点网络图、发现其内部关系显得尤为重要。本研究的另一个重要作用是从目前仅有的研究成果入手,例如新型冠状病毒受体 ACE2 和其作用相关的 spike 蛋白。通过对这 2 个研究点进行文献检索,检索到几千篇文献,然后对其研究热点构建网络图,逐步追踪其研究方向,为下一步研究指明道路。特别是对这次暴发型的 COVID-19,尽早找到下一步的研究思路,对挽救更多的生命、减少病毒的感染人数至关重要。

参 考 文 献

- [1] Tang B, Wang X, Li Q, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions[J]. J Clin Med, 2020, 9(2): E462.
- [2] Chen J. Pathogenicity and transmissibility of 2019-nCoV: a quick overview and comparison with other emerging viruses[J]. Microbes Infect, 2020, 22(2): 69-71.
- [3] Xia S, Zhu Y, Liu M, et al. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein[J]. Cell Mol Immunol, 2020[Epub ahead of print]. DOI: 10.1038/s41423-020-0374-2.
- [4] Luo H, Tang QL, Shang YX, et al. Can Chinese medicine be used for prevention of corona virus disease 2019(COVID-19)? A review of historical classics, research evidence and current prevention programs[J]. Chin J Integr Med, 2020, 26(4): 243-250.
- [5] Chung M, Bernheim A, Mei X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV)[J]. Radiology, 2020, 295(1): 202-207.
- [6] Zhao S, Musa SS, Lin Q, et al. Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak[J]. J Clin Med, 2020, 9(2): E388.
- [7] Yu F, Du L, Ojcius DM, et al. Measures for diagnosing and treating infections by a novel coronavirus responsible for a pneumonia outbreak originating in Wuhan, China[J]. Microbes Infect, 2020, 22(2): 74-79.
- [8] Wan Y, Shang J, Graham R, et al. Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus[J]. J Virol, 2020, 94(7): e00127.
- [9] Ma Y, Wu L, Shaw N, et al. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex[J]. Proc Natl Acad Sci USA, 2015, 112(30): 9436-9441.
- [10] Eckerle LD, Becker MM, Halpin RA, et al. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing[J]. PLoS Pathog, 2010, 6(5): e1000896.
- [11] Sisk JM, Frieman MB, Machamer CE. Coronavirus S protein-induced fusion is blocked prior to hemifusion by Abl kinase inhibitors[J]. J Gen Virol, 2018, 99(5): 619-630.
- [12] Alsaadi EAJ, Neuman BW, Jones IM. A fusion peptide in the spike protein of MERS coronavirus[J]. Viruses, 2019, 11(9): E825.
- [13] Jaimes JA, Millet JK, Goldstein ME, et al. A fluorogenic peptide cleavage assay to screen for proteolytic activity: applications for coronavirus spike protein activation[J]. J Vis Exp, 2019(143): 58892.
- [14] Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation[J]. Science, 2020, 367(6483): 1260-1263.
- [15] Coutard B, Valle C, de Lamballerie X, et al. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade[J]. Antiviral Res, 2020, 176: 104742.
- [16] Tetro JA. Is COVID-19 receiving ADE from other coronaviruses? [J]. Microbes Infect, 2020, 22(2): 72-73.
- [17] Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV[J]. J Med Virol, 2020, 92(5): 495-500.
- [18] Chen Y, Guo Y, Pan Y, et al. Structure analysis of the receptor binding of 2019-nCoV[J]. Biochem Biophys Res Commun, 2020[Epub ahead of print]. DOI: 10.1016/j.bbrc.2020.02.071.
- [19] Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR[J]. Euro Surveill, 2020, 25(3): 2000045.
- [20] Chen Y, Cai H, Pan J, et al. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase[J]. Proc Natl Acad Sci USA, 2009, 106(9): 3484-3489.
- [21] Song W, Wang Y, Wang N, et al. Identification of residues on human receptor DPP4 critical for MERS-CoV binding and entry[J]. Virology, 2014, 471-473: 49-53.
- [22] Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor[J]. Cell, 2020[Epub ahead of print]. DOI: 10.1016/j.cell.2020.02.052.
- [23] Iwata-Yoshikawa N, Okamura T, Shimizu Y, et al. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection[J]. J Virol, 2019, 93(6): e01815.

(责任编辑:冉明会)