Contents lists available at ScienceDirect
Journal of Infection
journal homepage: www.elsevier.com/locate/jinf

Letter to the Editor

Metagenomic sequencing with spiked-in internal control to monitor cellularity and diagnosis of pneumonia

Dear editor

We read with interest the study by Peng and colleges showing the performance of metagenomic Next-Generation Sequencing (mNGS) in detecting pulmonary pathogens was not superior to conventional microbiological tests (CMT) in a cohort of 101 immunocompromised adults.[1] Indeed, although mNGS enables untargeted "pan-pathogen" detection that covers a broad array of microorganisms with known genomic sequences, clinical application of this test has encountered challenges. For instance, the diagnostic sensitivity is affected by the quantity of host DNA, which varies considerably from sample to sample.[2] Host cell depletion techniques have been used to improve the sensitivity of mNGS but may lead to unspecific removal of pathogens.[3] In our independent study, we developed a spike-in internal control to assess the abundance of host and microbial DNA in bronchoalveolar lavage fluid (BALF) and evaluated the analytical and diagnostic performance of mNGS with and without host depletion in a cohort of 205 patients suspected of lower respiratory tract infections (**Supplementary Table 1**).

As shown by our data (Fig. 1**A, B**) and previous studies,[4] the microbial reads decreased with elevating concentrations of host cells in the sample, independent of whether PCR-based or PCR-free library preparation was used. It has been proposed that a nucleic acid internal control spiked at constant concentrations into all specimens could be utilized (Fig. 1**C**) to monitor the cellularity and microbial abundance.[2] Therefore, we designed double-stranded DNA to serve as spiked-in internal control (hereinafter referred to as spike) that shared no significant homology to genomes of any known organisms. We tested five spike molecules of varying length and nucleotide sequences, which were added into artificial samples containing different concentrations of human cells. As expected, spike RPM was positively correlated with input DNA quantity, but inversely correlated with the amount of host cells (Fig. 1**D**). We then added spike 1 to each BALF sample and carried out mNGS testing to evaluate the relationship between spike RPM and nucleated cell count. Both spike RPM and host index (**Supplementary Materials**) were inversely correlated with cell count (R^2 of 0.6278 and 0.6331, respectively) (Fig. 2**A, B**). Notably, cell-free DNA (cfDNA) derived from host was also present in BALF,[5] which cannot be quantified by cell counting.

We then examined the effect of saponin-based differential lysis for removing host cells. We discovered a total of 8 cases in which initial mNGS missed microorganisms that were detected after host depletion (8/205, 3.90%). These included 4 cases of *Mycobacterium tuberculosis* (MTB), 2 cases of *Aspergillus fumigatus* (AF) and 2 cases of *Candida albicans* (CA) (Fig. 2**C**). Saponin treatment has led to both increases and decreases of microbial reads (Fig. 2**D, Supplementary Fig. 1**). The enrichment in bacterial and mycobacterial reads was more apparent as compared to fungi and viruses. The main pathogen loss due to host depletion was *Pneumocystis jirovecii* (two cases, RPM decreased by 7695.87 and 8397.47, respectively). Pre-depletion mNGS was in agreement with post-depletion mNGS in 94.79% of samples (Fig. 2**E**).

To evaluate the analytical performance of mNGS in microbial detection, we used a combination of conventional tests as orthogonal methods. These included culture, acid-fast bacillus (AFB) stain, galactomannan antigen test (GM test), 1–3-β-D-glucans antigen test (*G* test), glucuronoxylomannan antigen test (GXM test), and PCR assays such as GeneXpert.TB, *Mycoplasma pneumoniae, Chlamydia pneumoniae*, adenovirus and cytomegalovirus (**Supplementary Fig. 2**). Overall, the traditional tests were positive in 64/205 (31.22%) samples, in which 51 had matching mNGS results (79.69%) (**Supplementary Table 2**). We discovered one case of *Cryptococcus* and 4 cases of MTB via glucuronoxylomannan test and GeneXpert.TB, respectively, which were missed by both pre-depletion and post-depletion mNGS. When all traditional methods were negative, diagnosis was adjudicated by at least two physicians based on mNGS and clinical manifestations such as patients' symptoms, chest CT and responses to empirical antibiotics. Among 205 patients, 15 (7.32%) were diagnosed with non-infectious diseases. The remaining 190 (92.68%) were diagnosed with infectious pneumonia. The microbiological etiology was confirmed in 115 (60.53%) patients (Fig. 2**E**). In these cases, mNGS achieved a detection rate of 93.04% (107/115), compared with 49.57% (57/115) by traditional methods.

Of all patients enrolled, mNGS provided useful diagnostic clue in 123 patients (60.00%), 64 of which were primarily diagnosed by mNGS (all conventional methods were negative). On the other hand, conventional tests were useful in 67 patients (32.68%), 8 of which were solely diagnosed by conventional methods (Fig. 2**F**). We found mNGS to be beneficial in diagnosing lung abscess (especially caused by anaerobic bacteria), tuberculosis, aspergillosis, bacterial pneumonia (mainly *Streptococcus pneumoniae*), *Pneumocystis jiroveci* pneumonia (PJP), Chlamydia/Mycoplasma pneumonia, non-tuberculous Mycobacteria (NTM) pulmonary infections etc., for which conventional methods were either not readily available in all hospitals (*i.e. Pneumocystis jiroveci*, chlamydia/mycoplasma), difficult to perform (*i.e.* culture of anaerobes) or time-consuming (*i.e.* culture of Mycobacteria).

On the other hand, the lack of pathogens in mNGS were useful for clinicians to rule out microbial infections, including 6 cases of interstitial pneumonia, 4 cases of lung carcinoma, one case of heart failure, one case of hypersensitivity pneumonitis, one case of chronic obstructive pulmonary disease (COPD) and one case of radiation pneumonitis. In these patients, mNGS detected com-

https://doi.org/10.1016/j.jinf.2021.09.018
0163-4453/© 2021 The British Infection Association. Published by Elsevier Ltd. All rights reserved.

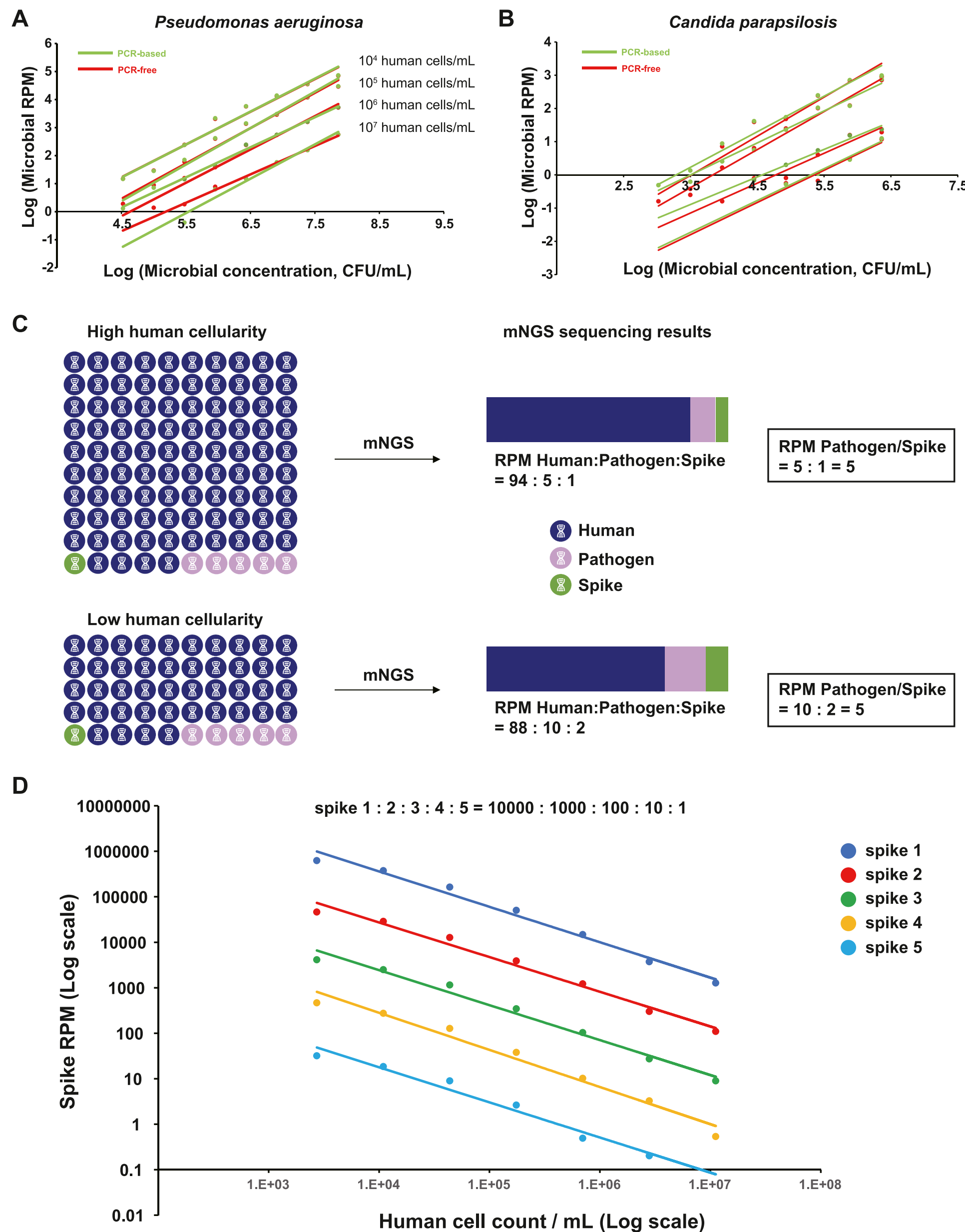

Fig. 1. Sensitivity of mNGS in is negatively affected by host cells. Different concentrations (CFU/mL) of inactivated *Pseudomonas aeruginosa* and *Candida parapsilosis* were mixed with different amount of human cells (10^4 – 10^7 copies/mL Jurkat cells) and mNGS was performed. The microbial RPM was plotted against microbial concentrations (**A, B**). Each dot represented an individual sequencing library. PCR-based library preparation was shown as green while PCR-free library preparation was shown as red. A spike-in nucleic acid could be used as an internal control to calculate the relative quantity of both host and microbial DNA to spike DNA (**C**). Five different spike molecules were prepared and tested in artificial samples containing varying concentrations of human cells (Jurkat) and spike RPM was plotted against the cell count (**D**).

mon respiratory colonizers such as *Haemophilus, Streptococcus, Prevotella, Actinomyces* species, common skin colonizers such as *Propionibacterium, Malassezia* species, and human herpesviruses, *Candida* species, Torque teno virus (**Supplementary Table 2**).

In summary, prior to the decision of using mNGS for diagnosis, one should be aware of both the utility and drawbacks of this technique. It is suitable for detecting a wide array of pathogens, especially rare and polymicrobial infections. We showed in this study that mNGS is effective for detecting pathogens for which conventional methods are lacking or inaccessible. However, due to the “pan-pathogen” aspect of mNGS, a comprehensive understanding of respiratory infections and sufficient knowledge in microbiology are important for interpretation of results. Moreover, host depletion causes both increases and decreases of microbial reads and should be carefully evaluated and performed with caution.

Supplementary Fig. 1. The heat map of microbial RPM before and after host depletion. The RPM of each detected microbial species in pre-depletion and post-depletion samples was plotted. The sample IDs were shown in Y-axis while the microbial species were shown in X-axis.

Supplementary Fig. 2. The conventional microbiological tests used for pathogen detection. The traditional diagnostic tests were performed in parallel with mNGS.

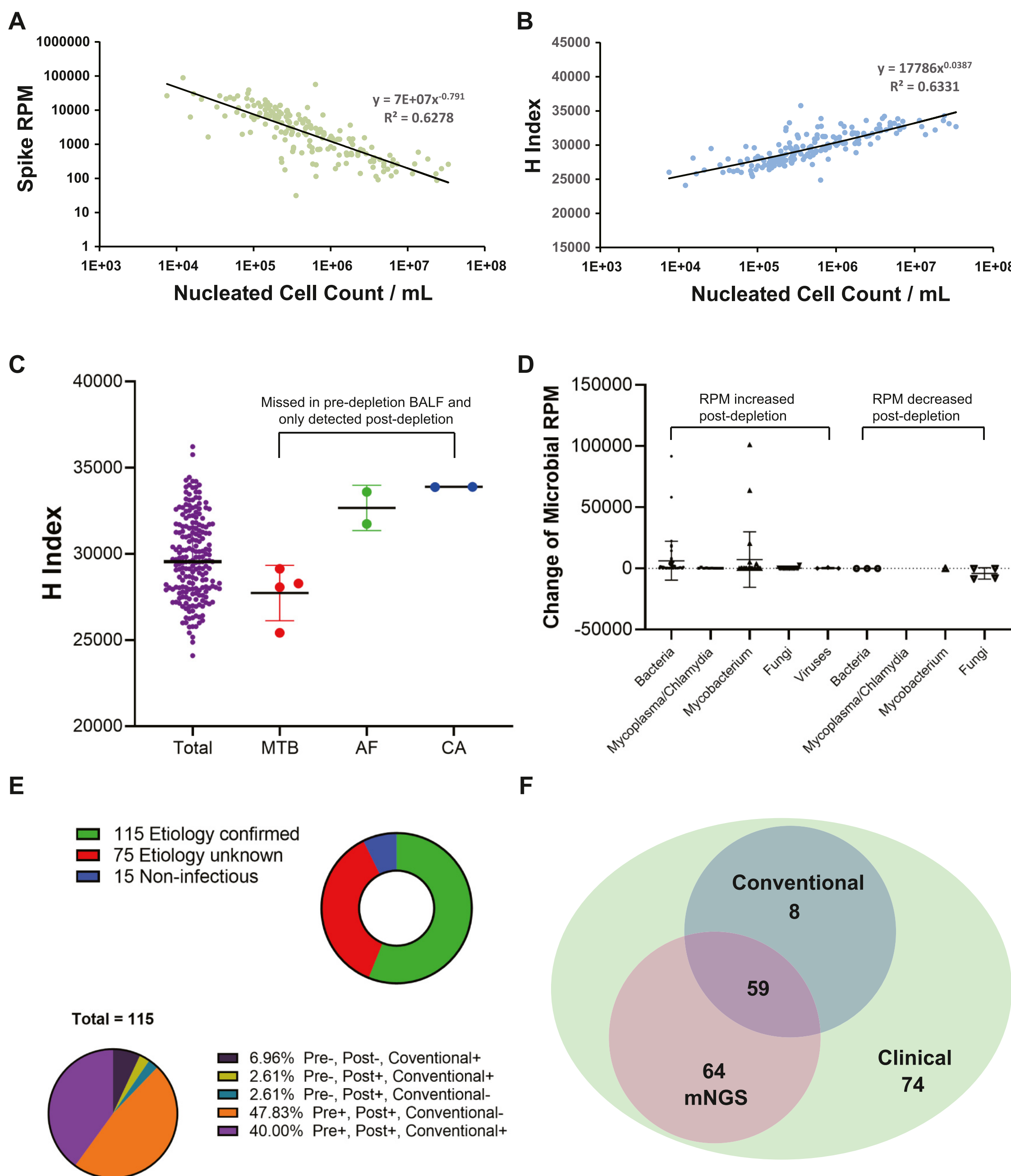

Fig. 2. Spike-in internal control for host DNA quantification. The nucleated cell count was measured by AO/PI dual fluorescence staining and plotted for 205 BALF samples. A linear regression was plotted using spike RPM (**A**) or host index (**B**) against the nucleated cell count. For each sample, mNGS was carried out twice: one for the original sample, the other for the sample with host cells removed. The host indices for all original samples were shown as box plot with Interquartile range. In addition, the host indices were shown for pre-depletion samples that missed microorganisms in the original mNGS but were detected after host depletion and confirmed by reference methods (**C**). The change of microbial RPM before and after host depletion was plotted in (**D**). The percentage of diagnosed patients and patients with known etiology were shown in (**E**). The number of patients diagnosed via mNGS, conventional methods and/or clinical manifestations were shown in (**F**).

Declaration of Competing Interest

OC, XH, JW and CL are employees of Hangzhou Matridx Biotechnology CO., Ltd. The rest of the authors declared no conflict of interest.

CRediT authorship contribution statement

Hua Zhou: Data curation, Formal analysis, Writing – review & editing. **Chuan Ouyang:** Data curation, Formal analysis, Writing – review & editing. **Xu Han:** Supervision, Data curation, Writing – review & editing. **Lisha Shen:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Jian Ye:** Supervision, Data curation, Writing – review & editing. **Zhixian Fang:** Supervision, Data curation, Writing – review & editing. **Wenyu Chen:** Supervision, Data curation, Writing – review & editing. **Aifeng Chen:** Supervision, Data curation, Writing – review & editing. **Qiang Ma:** Supervision, Data curation, Writing – review & editing. **Maohong Hua:** Supervision, Data curation, Writing – review & editing. **Junfei Zhu:** Supervision, Data curation, Writing – review & editing. **Xiaomai Wu:** Supervision, Data curation, Writing – review & editing. **Xuemei Lin:** Supervision, Data curation, Writing – review & edit-

ing. **Yuexiang Shui:** Supervision, Data curation, Writing – review & editing. **Changsheng Zhou:** Supervision, Data curation, Writing – review & editing. **Kai Fang:** Supervision, Data curation, Writing – review & editing. **Junwei Du:** Supervision, Data curation, Writing – review & editing. **Zhihui Huang:** Supervision, Data curation, Writing – review & editing. **Gang Wang:** Supervision, Data curation, Writing – review & editing. **Qun Lv:** Supervision, Data curation, Writing – review & editing. **Weina Huang:** Supervision, Data curation, Writing – review & editing. **Jun Wang:** Supervision, Data curation, Writing – review & editing. **Xiaoting Hua:** Supervision, Data curation, Writing – review & editing. **Jianying Zhou:** Supervision, Data curation, Writing – review & editing. **Chao Liu:** Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Yunsong Yu:** Conceptualization, Visualization, Writing – review & editing.

Ethics

This study was approved by the Ethics committee of the First Affiliated Hospital, Zhejiang University School of Medicine (Approval ID: IIT20200032A). Informed consent was obtained from all participants.

Funding

This work was supported by a research grant from National Natural Science Foundation of China (81971897) and a grant of The National Key Research and Development Program of China (2018YFE0102100). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability

The raw data of mNGS and traditional tests can be found in **Supplementary Table 1**. Raw sequencing files (fastq) can be accessed via accession number PRJNA734131 in SRA of NCBI (https://www.ncbi.nlm.nih.gov/sra).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jinf.2021.09.018.

Reference

1. Peng JM, Du B, Qin HY, Wang Q, Shi Y. Metagenomic next-generation sequencing for the diagnosis of suspected pneumoni in immunocompromised patients. *J Infect* 2021;**82**(4):22–7 AprPubMed PMID: 33609588.
2. Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock GProfessional practice C Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch Pathol Lab Med* 2017;**141**(6):776–86 JunPubMed PMID: 28169558.
3. Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, Jeanes C, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat Biotechnol* 2019;**37**(7):783–92 JulPubMed PMID: 31235920.
4. Luan Y, Hu H, Liu C, Chen B, Liu X, Xu Y, et al. A proof-of-concept study of an automated solution for clinical metagenomic next-generation sequencing. *J Appl Microbiol* 2021;**131**(2):1007–16 Jan 13PubMed PMID: 33440055.
5. Gu W, Deng X, Lee M, Sucu YD, Arevalo S, Stryke D, et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat Med* 2021;**27**(1):115–24 JanPubMed PMID: 33169017.

Hua Zhou[1]
Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China

Chuan Ouyang[1], Xu Han
Medical Department, Hangzhou Matridx Biotechnology Co., Ltd, Bd 2-4, 2073 Jinchang Rd, Hangzhou 310030, China

Lisha Shen
Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China

Jian Ye
Department of Respiratory and Critical Care Medicine, Hangzhou First People's Hospital, School of Medicine, Zhejiang University. Hangzhou 310001, China

Zhixian Fang, Wenyu Chen
Department of Respiratory and Critical Care Medicine, Jiaxing First Hospital, Jiaxing 314001, China

Aifeng Chen
Department of Respiratory and Critical Care Medicine, Hangzhou Red Cross Hospital Affiliated to Zhejiang University, Hangzhou 310003, China

Qiang Ma
Department of Respiratory and Critical Care Medicine, Yuhang District Second People's Hospital, Hangzhou 311121, China

Maohong Hua
Department of Respiratory and Critical Care Medicine, The Third People's Hospital of Xiaoshan, Hangzhou 311203, China

Junfei Zhu
Department of Respiratory and Critical Care Medicine, Taizhou Central Hospital, Taizhou 318000, China

Xiaomai Wu
Department of Respiratory and Critical Care Medicine, Taizhou Hospital of Zhejiang Province, Taizhou 317000, China

Xuemei Lin
Department of Respiratory and Critical Care Medicine, The People's Hospital of Jiangshan, Quzhou 324100, China

Yuexiang Shui
Department of Respiratory and Critical Care Medicine, The People's Hospital of Lanxi, Jinhua 321100, China

Changsheng Zhou
Department of Respiratory and Critical Care Medicine, People's Hospital of Cangnan, Wenzhou Medical University, Wenzhou 325800, China

Kai Fang
Department of Respiratory and Critical Care Medicine, Quhua Hospital, Quzhou 324004, China

Junwei Du
Department of Respiratory and Critical Care Medicine, Fuyang people's Hospital, Hangzhou 311400, China

Zhihui Huang, Gang Wang
Department of Respiratory and Critical Care Medicine, Anji people's Hospital, Huzhou 313300, China

Qun Lv
Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Hangzhou Normal University, Hangzhou 324000, China

Weina Huang
Department of Respiratory and Critical Care Medicine, Ningbo First Hospital, Ningbo 315010, China

Jun Wang
Medical Department, Hangzhou Matridx Biotechnology Co., Ltd, Bd 2-4, 2073 Jinchang Rd, Hangzhou 310030, China

Xiaoting Hua
Department of Infectious Diseases, Sir Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310020, China

Jianying Zhou
Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China

Chao Liu*
Medical Department, Hangzhou Matridx Biotechnology Co., Ltd, Bd 2-4, 2073 Jinchang Rd, Hangzhou 310030, China

Yunsong Yu**
Department of Infectious Diseases, Sir Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310020, China
Key Laboratory of Microbial Technology and Bioinformatics of Zhejiang Province, Hangzhou 310016, China
Regional Medical Center for National Institute of Respiratory Diseases, Sir Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310020, China

*Corresponding author. **Corresponding author at: Department of Infectious Diseases, Sir Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310020, China.
E-mail addresses: liuchao@matridx.com (C. Liu), yvys119@zju.edu.cn (Y. Yu)

[1] These authors contributed equally to this work.