



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Application of next generation sequencing technology on contamination monitoring in microbiology laboratory

Yan Xiao ^a, Li Zhang ^b, Bin Yang ^a, Mingkun Li ^{b,c,*}, Lili Ren ^{a,*}, Jianwei Wang ^a

^a NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100730, China

^b Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

^c Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

ARTICLE INFO

Article history:

Received 12 November 2018

Received in revised form 7 January 2019

Accepted 21 February 2019

Available online 1 March 2019

Keywords:

Laboratory contamination

Biosafety

Next generation sequencing

Deep sequencing

ABSTRACT

The surveillance and prevention of pathogenic microbiological contamination are the most important tasks of biosafety management in the lab. There is an urgent need to establish an effective and unbiased method to evaluate and monitor such contamination. This study aims to investigate the utility of next generation sequencing (NGS) method to detect possible contamination in the microbiology laboratory. Environmental samples were taken at multiple sites at the lab including the inner site of centrifuge rotor, the bench used for molecular biological tests, the benches of biosafety cabinets used for viral culture, clinical sample pre-treatment and nucleic acids extraction, by scrubbing the sites using sterile flocked swabs. The extracted total nucleic acids were used to construct the libraries for deep sequencing according to the protocol of Ion Torrent platform. At least 1G raw data was obtained for each sample. The reads of viruses and bacteria accounted for $0.01 \pm 0.02\%$, and $77.76 \pm 12.53\%$ of total reads respectively. The viral sequences were likely to be derived from gene amplification products, the nucleic acids contaminated in fetal bovine serum. Reads from environmental microorganisms were also identified. Our results suggested that NGS method was capable of monitoring the nucleic acids contaminations from different sources in the lab, demonstrating its promising utility in monitoring and assessing the risk of potential laboratory contamination. The risk of contamination from reagents, remnant DNA and environment should be considered in data analysis and results interpretation.

© 2019 Chinese Medical Association Publishing House. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The prevention and control of contamination caused by cultured viruses, bacteria and other infectious materials are the most important part in biosafety management in the laboratory [1–3]. Microbiological contaminations in laboratories have been frequently reported in instruments [4], rooms [5,6], and the operators [7,8]. These contaminations are mainly from the performance of experiments, for example, bacteria/virus isolation and culture, gene amplification, and clinical samples preparation. As such, it is important to establish an effective and unbiased method to evaluate and monitor the risk of lab contaminations.

The potential contamination in the environment of a microbiology lab is very complicated, as usually the source and background of contaminants is unknown. However, it is hard to conduct effective contamination surveillance in a common laboratory due to the methodology limitation of contaminant detections. The isolation of microbial species is difficult because of its tedious work and poor sensitivity. Molecular tests such as PCR is sensitive, however it can only detect the gene of

known pathogens. It is desired to develop a comprehensive technology which can detect the potential contaminants of unknown resources, characterized by culture-independency and high sensitivity.

The development of next-generation sequencing (NGS) technology provided a culture-independent method to obtain microorganisms genome information with high sensitivity [9,10], and has been widely used to detect the pathogenic microorganisms [11–13]. In this study, we evaluated the performance of NGS as a promising tool on the surveillance and risk assessment of potential lab contaminations. Frequently used microbiology laboratories which included biosafety laboratory level I (BSL-I) and II (BSL-II) were investigated in our study. Various sites were inspected in different rooms. We aimed to establish an unbiased and effective evaluating technique involving NGS to monitor the laboratory contamination.

2. Materials and methods

2.1. Sample collection

Four rooms were investigated in the study. Room 1 is a BSL-II used for viral isolation and culture. Room 2 is a BSL-II, used for pre-treatment of respiratory samples from patients with respiratory tract infections and also used for nucleic acid extraction. Room 3 is a BSL-I, where a negative-

* Corresponding authors: NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100101, China.

E-mail addresses: renlili@ipbcams.ac.cn (Lili Ren), limk@big.ac.cn (Mingkun Li).

HIGHLIGHTS

Scientific question

This study investigated the performance of next generation sequencing (NGS) technology in monitoring laboratory microbiological contamination.

Evidence before this study

Contamination is a big issue in microbiology laboratories, while it is hard to conduct effective contamination surveillance in a common laboratory due to the lack of proper methodology. The isolation of microbial species is difficult which needs tedious lab work and has low sensitivity. Although molecular tests such as PCR are sensitive, only limited number of known microbes could be detected simultaneously.

New findings

A holistic assessment of the biological background in the environment of a microbiology laboratory was conducted through NGS technology in this study. In general, the richness of viral and bacterial sequences is positively associated with the frequency of laboratory activities in the room and on the bench. The data indicated that viral reads were mainly from experimental activities, materials and environment. All types of bacteria seem to be derived from the external environment. The gene fragments from exogenous microorganisms have a great impact on the deep sequencing results, especially for the clinical samples with low microbial biomass. Application of the NGS technique on biosafety monitoring would enable us to disentangle the background noise presenting in the sequencing data, which in turn help to distinguish true contamination caused by microorganisms from false positives.

Significance of the study

The study demonstrated that NGS could be a promising approach in monitoring the contamination from different sources in the lab, suggesting its promising usage in monitoring and assessing the risk of potential laboratory contamination.

pressure PCR hood used for nucleic acid template adding was placed. Room 4 is a BSL-I, used for molecular tests and immunological experiments. The sampling sites included the biosafety cabinet, the bench, and the inner sites of centrifuge rotor (Table 1).

For each site, five different locations were selected, sterile flocked swabs (Copan Diagnostics, Inc., Murrieta, CA) soaked in sterile saline were used to collect samples by scrubbing repeatedly on each sampling point for 10 times. All sampling swabs from one site were rinsed into 1 ml sterile saline in one tube. The swabs were discarded after repeat pressing. Swab rinsed into 1 ml saline without sampling was used as negative control, with code number of H0. A positive control was included, which was sputum collected from a pneumonia patient infected

with human coronavirus (HCoV) 229E determined previously by RT-PCR method [14]. All procedures were performed according to the biosafety regulations and requirements.

2.2. Sequencing library construction

Total 400 µl of each sample was used for nucleic acids extraction by using QIAamp® viral RNA mini Kit (Qiagen, Hilden, Germany) and the elution volume was 80 µl. Since both the DNA and RNA molecules are of interest to us, no DNase was used in our protocol. All nucleic acids were used for library construction. The OD260/280 value of total nucleic acids was measured by NanoDrop Spectrophotometer (ND-1000, Thermo Fisher, Wilmington, DE, USA). The RNA was specifically quantified by Qubit® 2.0 Fluorometer (Thermo Fisher, Wilmington, DE, USA). The libraries were constructed by using Ion Total RNA-Seq Kit v2 (Thermo Fisher, Wilmington, DE, USA). Total 30–50 ng of RNA was fragmented by RNase III (Thermo Fisher, Wilmington, DE, USA), then hybridized and ligated with the Ion Adaptor Mix v2 at 65 °C for 10 min, 30 °C for 30 min. The reverse transcribed single-stranded cDNA was amplified with Ion 5' and 3' PCR Primer v2. The sequencing libraries were qualified and quantified by using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and Qubit® 2.0 Fluorometer.

2.3. Deep sequencing and data analysis

Each sequencing library was diluted to 15 pM. Templated Ion Sphere Particles were prepared and enriched by using Ion OneTouch 200 Template Kit. Deep sequencing was performed by using Ion Personal Genome Machine™ system with Ion PGM 200 Sequencing Kit in Ion 318 Chips (Thermo Fisher, Wilmington, DE, USA). The sequencing output from each sample was no less than 1 G. The raw data reads were filtered by a minimum length of 36 bp, and the remaining reads were aligned to the NCBI nt database (Feb 2016 version) by MegaBLAST [15] (-evalue 1e-10 -max_target_seqs 10 -max_hsp 1 -qcov_hsp_perc 60). Taxonomy was assigned by MEGAN (-ms 100 -sup 1 -me 0.01 -top 10) [16]. The community composition was analyzed at both the genus and species levels. Only reads that could be mapped to specific species with high confidence were considered for genome mapping. Alpha diversity (observed genera/species and Shannon index) calculation and principal coordinate analysis (PCoA) (based on abundance weighted Jaccard distance) were performed with Qiime [17] (v1.9.1). Mann-Whitney test was performed with GraphPad prism 7.

3. Results

3.1. Data overview

The quality of nucleic acids measured by OD 260/280 was 1.8–2.2 for all samples. Similar quantity of RNA was yielded from H1–H9 samples (88.8–127.2 ng, median of 108.80 ng) (Figure 1A), which indicated similar biomass in all sampled locations. As expected, the amount of nucleic acids from H0 was ultralow (below the detection limit) and all nucleic acids were used to construct the sequencing library. After sequencing and data filtering, a range from 3,046,082 to 6,000,548 reads from H1 to H8 was used for taxonomy classification and more than 66.69% of

Table 1
Sampling sites and samples code.

Room no.	Laboratory function	Biosafety level	Biosafety cabinet	Bench	Centrifuge rotor
1	Viral isolation and culture	II	H1	H2	H3
2	Clinical samples pretreatment and nucleic acids extraction	II	H4	H5	H6
3	Nucleic acids templates amplification	I	H7 ^a	\	\
4	Molecular biological and immunological experiment	I	\	H8	H9

^a H7 is a negative-pressure hood used for nucleic acid template adding for PCR.

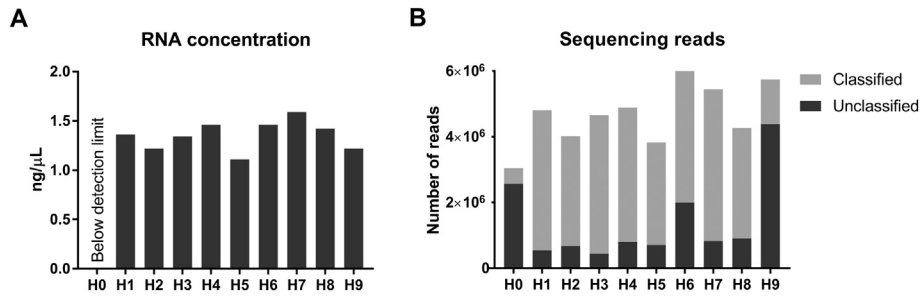


Figure 1. Quantities of RNA and sequencing reads used for taxonomy classification. (A) Quantities of RNA extracted from samples of H0–H9. (B) The total number of sequencing reads and the proportions of taxonomic classified (dark gray) and unclassified reads (light gray).

reads from these samples could be classified. The samples from H0 (480,074 reads) and H9 (1,354,249 reads) had dramatically low amount of reads, and only 15.76% and 23.58% of reads were successfully aligned and classified to a certain taxonomic rank (Figure 1B). The data obtained from the positive control was consistent with the real-time PCR results. The reads of known pathogen, HCoV 229E accounted to 72.4% of all microbial reads, thus verified the reliability of NGS procedure.

3.2. Detections of viral sequences

As the major activities in the laboratory involved respiratory viruses, we first analyzed the viral reads in the data. The results demonstrated that viral sequences were in a relatively low abundance ($0.01 \pm 0.02\%$), with the viral reads ranged from 64 to 733 (median of 148). The number of viral species ranged from 8 to 34 (median of 12) in different samples including H0. The distribution of virus species in each sampling site was shown in Figure 2. The samples collected from Room 4, which was used for molecular biological and immunological experiments, had the highest number of viral species (15 species). The three most abundant viruses were cercopithecine herpesvirus 5 (CeHV-5) (23% reads, observed at 7 sites), influenza A viruses (IFVA) (16%, 8 sites) and lactococcus phage 936 sensu lato (8%, 2 sites).

Due to the high sensitivity of the NGS method, viral reads that derived from routine laboratory work were identified, such as IFVA, human enterovirus 71 viruses, HCoV HKUI and NL63 (seen in Figure S1). These reads were mainly from Room 2 and Room 4. Among these,

reads of IFVA accounted for the highest proportion. They were detected in 6 sites except the biosafety cabinet. However, the obtained gene sequences were all close to the primer positions for viral genes amplification (Figure 3), as PCR tests were performed on the viral positive clinical samples within a month before sampling.

The sequences of bovine diarrhea virus were only found on the biosafety cabinet in Room 1, which was used for viral isolation and culture experiments (Figure 2). The existence of bovine diarrhea viral nucleic acids in fetal bovine serum (FBS) was further confirmed by PCR methods (data not shown) [18], which may represent contamination of bovine serum used for cell culture as reported also in previous studies [18–21].

The reads of CeHV-5 were identified at all sampling sites except H2, H3 and H9. This virus was previously isolated from the tissue of rhesus macaque [22], and had been detected in the cell line obtained from the kidney of rhesus macaque [23]. There are two rhesus macaque kidney cell lines, LLC-MK2 and Vero cells, used in the lab for virus isolation during the study period. However, we did not find CeHV-5 nucleic acids in these cultured cells by PCR. There were also viral sequences aligned to human mastadenovirus, polyomavirus, hepatitis virus, human metapneumovirus and rotavirus, but their read numbers and covered lengths were very limited, thus were likely to be false positives (Figures 2 and S1).

3.3. Diversity and composition of the sequences of biological substances

The alpha diversity of genera measured by observed genera and Shannon index did not differ significantly among H1–H9 (Figures 4A,

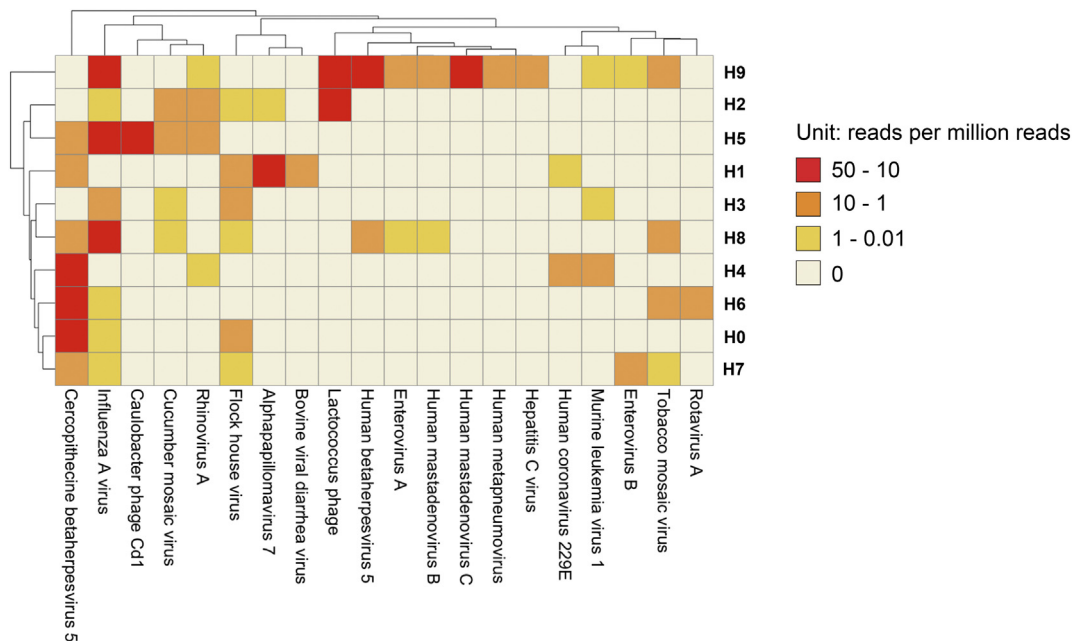


Figure 2. Heatmaps of the top 20 most abundant virus. The abundances were normalized by the number of reads assigned to the species level for each sample.

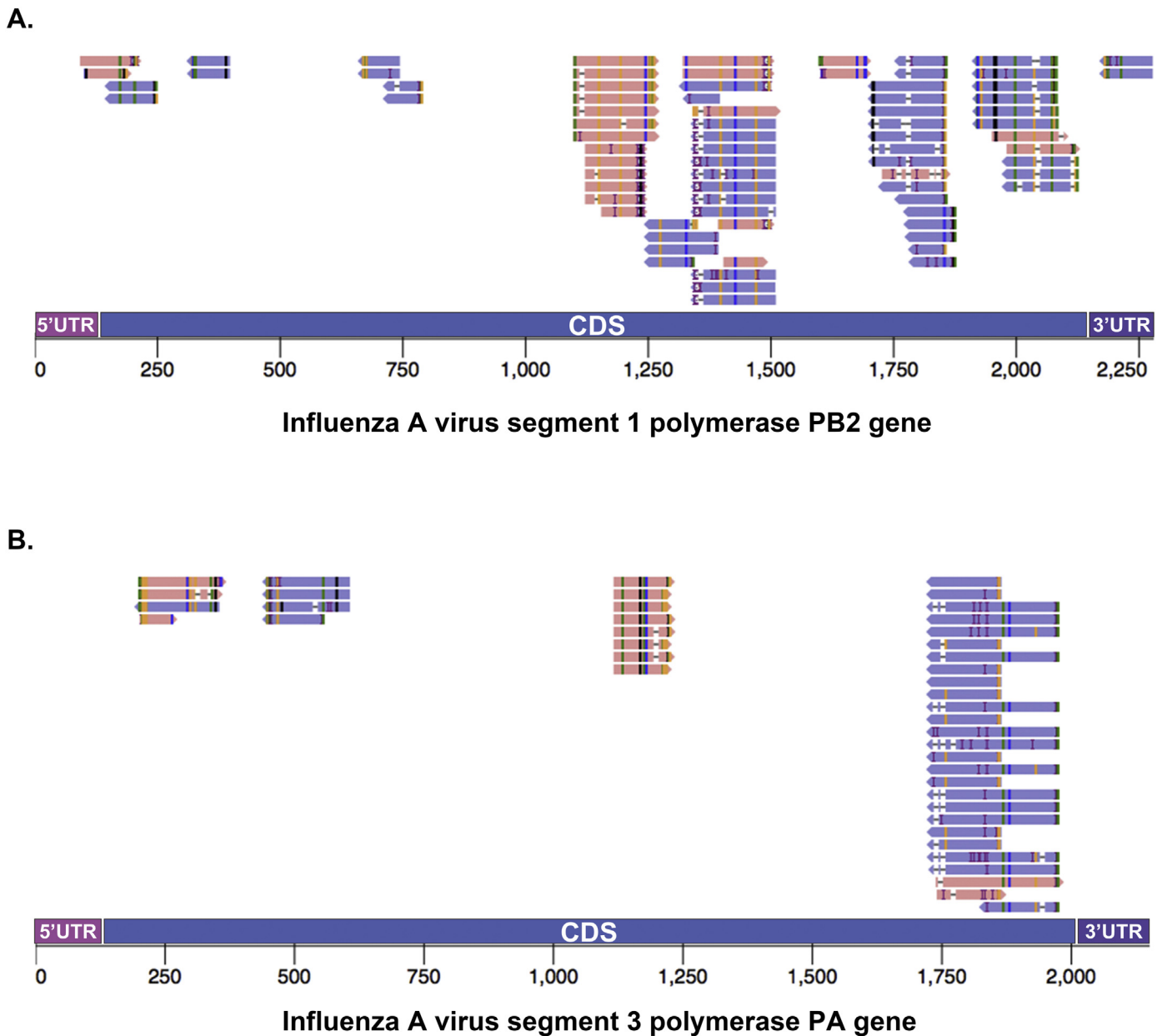


Figure 3. Schematic diagram of the influenza A viral reads coverage on genes. (A) Reads coverage on segment 1 gene. (B) Reads coverage on segment 3 gene.

2B). However, when going down to the species level, H2 and H8 showed much higher diversity than other samples except for the negative control (Figures 4C, 2D). Besides H2 and H8, another sample collected from the operating bench, H5, also had considerable higher alpha diversity, indicating a generally high diversity of biological substances on the bench. It is noteworthy that H0 had high evenness of biological substances background, as shown by the high Shannon index at both the genus and species levels (Figure 4B, D), which might reflect that the contamination from reagents and/or environment is more severe for low biomass samples.

A tight clustering of H1, H2 and H3, which were all taken from the viral culture room (Room 1), was highlighted on PCoA plot that based on weighted Jaccard distance at the genus level (Figure 4E). This suggests a considerable impact of the physical distance/boundary on the spread of biological substances.

Of all the biomass sequences classified to the domain level, Bacteria is the most abundant in all samples ($77.76 \pm 12.53\%$), followed by fungi and all other Eukaryota, accounting for $5.89 \pm 4.61\%$ and $16.32 \pm$

12.21% of the classified reads, respectively. The amount of Archaea ($0.01 \pm 0.01\%$) and Viruses ($0.01 \pm 0.02\%$) is relatively low.

Zooming into the species level, the top 20 abundant ones include four species belonging to the phylum *Actinobacteria* of Bacteria, twelve species belonging to the phylum *Proteobacteria*, two species belonging to the phylum *Chordata* of Eukaryota, and two species belonging to the phylum *Streptophyta* of Eukaryota. For the Bacteria domain, *Micrococcus luteus* (*M. luteus*) is outstandingly abundant, accounting for $58.88 \pm 31.72\%$ of the reads classified to the species level. This result is reasonable because *M. luteus* distributes widely in soil, water, dust and air, and it is part of the normal flora of human skin, mouth, mucosae, oropharynx and upper respiratory tract [24]. The rest of most abundant bacteria includes both gram-positive (*Actinobacteria*) and gram-negative (*Proteobacteria*) bacteria, which originate from a diverse range of environments: *Escherichia coli*, *Sphingomonas*, *Pseudomonas putida* and *Variovorax paradoxus* are frequently observed in soil and water; *Cutibacterium acnes* is a human skin-associated organism; *Moraxella catarrhalis* and *Acinetobacter baumannii* are common

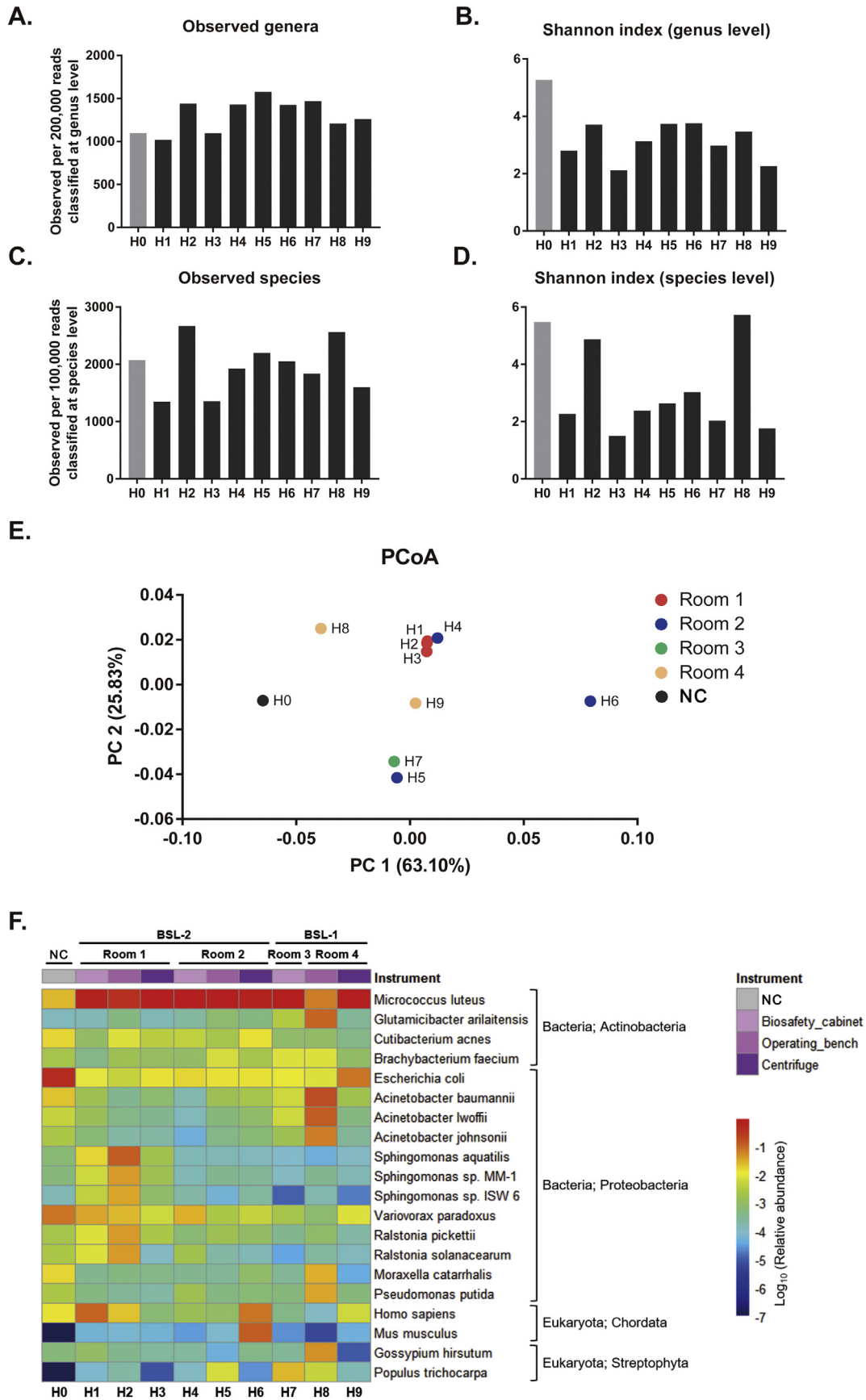


Figure 4. Diversity and compositions of identified bacteria. (A–D) Alpha diversity calculated by (A) number of species, (B) Shannon index at the genus level, (C) number of species and (D) Shannon index at the species level. (E) Principal coordinate analysis (PCoA) plot based on the weighted Jaccard distance between genera. In A–E, the numbers of reads were rarefied to 200,000 and 100,000 per sample at the genus level and species level, respectively. (F) Heatmap showing the top 20 most abundant species. The abundances were normalized by the number of reads assigned to the species level for each sample. NC, negative control.

opportunistic pathogens. *Ralstonia* had been identified as a common contaminant of DNA extraction kits or PCR reagents [25].

For the Eukaryota domain, the abundances of human-originated reads in H1 (10.26%) and H6 (7.93%) were higher than that in other samples (below 2.15%). They were possibly derived from personnel entered the rooms for equipment maintenance at that time. Additionally, H6 had a dramatically high percentage of mouse-originated reads (12.9%) compared to that in other samples (below 0.01%). This could be associated with the isolation of mouse lymphocytes at that location. Regarding plant-originated biological substances, *Gossypium hirsutum*, known as upland cotton and often present in clothes, had a much higher percentage in H8 (4.89%) than in other samples (below 0.05%, Figure 4F), suggesting a greater impact of human activity on the operating bench in the BSL-1 room. The relative abundance of *Populus trichocarpa*, namely *California poplar*, was generally higher in the BSL-1 rooms with ordinary pressure system comparing with that in the BSL-2 rooms with negative pressure system ($1.01 \pm 1.32\%$ vs $0.12 \pm 0.27\%$, $p = 0.17$, Mann-Whitney test, Figure 4F). This indicates that the environmental contamination could be affected by the season, considering that the samples were taken in April, when poplar blossomed in Beijing.

4. Discussion

NGS technology is an effective approach for laboratory management and microbe monitoring [26]. In this study, we carried out a holistic assessment of the biological background in the environment of a microbiology laboratory. In general, the diversity of viral and bacterial sequences is positively associated with the frequency of laboratory activities in the room and on the bench.

We are more concerned about the distribution of viruses in this microbiology laboratory, where respiratory viruses-related experiments were performed regularly. Firstly, all the viral reads obtained from the environment samples were compared to the positive control. It showed the HCoV-229E reads dominant in the positive control was only detected in H1 and H4 samples with less than 10 reads. The two sampling sites were biosafety cabinet bench in Room 1 and 2, used for virus culture and clinical samples treating, indicating an occasional contamination from samples or nucleic acids. No other samples showed positive on HCoV 229E, which helped to ensure us the quality control of NGS procedure in this study. There are viral reads in the negative control (H0) mapping to CeHV-5, IFVA and flock house virus, which similar to the distribution in H7, the PCR hood used for adding nucleic acids templates, indicating a residual nucleic acids contamination. Viral reads related to the experiments performed in recent one month were also identified in these environment samples. In a summary, we considered these viral sequences were mainly from three parts. (1) Experimental activities. The obtained short viral reads are mainly mapped to the terminal gene region or the PCR primers targeting regions (see Figure 3). No complete viral genome was identified even in samples collected from viral isolation and culture sites. Moreover, the reads of human mastadenovirus, polyomavirus, enterovirus, and human betaherpesvirus 5, were found only in samples collected from bench, centrifuge rotor (H8, H9) in BSL-I and PCR hood (H7), correlated to the experiments performed in these sites. Of note, since RNA were preferentially amplified in our protocol, DNA contamination (e.g., PCR products) level should be underestimated, thus we are presenting a lower boundary of DNA contamination. The results indicated these viral reads were mainly from the gene amplification products or residual degraded nucleic acids. As the viral reads related to the viruses used in recent one month could be detected, the residual DNA fragments could help to trace experimental activities within a month. (2) Experimental materials. For example, the fetal bovine serum was confirmed to be contaminated by the bovine viral diarrhea virus [18–21]. Besides that, the reverse transcriptase which derived from murine leukemia virus [27–29], and the recombinant protein expression system which involves the use of viral vectors are potential sources of contamination [30–32]. (3) Environment. Reads of the tobacco mosaic virus and

cucumber mosaic virus were found in most of the samples except negative control and the samples collected from biosafety cabinets in BSL-II (H1 and H4). As reported, the two kinds of viruses are widely distributed in water and soil [33,34]. The hepatitis C virus and rotavirus, not used in the lab, were found in samples from the centrifuge rotor (H6) and the bench in BSL-I with less than ten reads. Whether these viral reads were from environment or not need to be investigated further.

The largest part of biomass in our data is contributed by bacteria, of which *M. luteus* is the most prevalent. All types of bacteria seem to be derived from the external environment [24]. Room 4 has the most abundant microbial species, as frequent molecular biological experiments and immunology related experiments occurred in this room. A large number of microbial sequences in all centrifuge rotors were found, which emphasized the necessity of routine clean and disinfection for this equipment. The gene fragments from exogenous microorganisms have a great impact on the deep sequencing results, especially for the clinical samples with low microbial biomass, such as cerebrospinal fluid, blood, and bronchoalveolar lavage fluid [25,35–38]. Application of the NGS technique on biosafety monitoring would enable us to disentangle the background noise presenting in the sequencing data, which in turn help to distinguish true microorganisms from false positives. Meanwhile, laboratory contamination should be suspected when obtaining the complete genome sequence of cultured viruses, or abnormal enrichment of specific microbes.

5. Conclusions

NGS method was capable of monitoring the nucleic acids contaminations from different sources in the lab, demonstrating its promising utility in monitoring and assessing the risk of potential laboratory contamination. The risk of contamination from reagents, remnant DNA and environment should be considered in data analysis and results interpretation.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bsheal.2019.02.003>.

Acknowledgements

This study was supported by CAMS Innovation Fund for Medical Sciences (2016-I2M-1-014, 2017-I2M-3-017), the National Major Science & Technology Project for Control and Prevention of Major Infectious Diseases in China (2017ZX10103004, 2018ZX10305409, 2018ZX10301401, 2018ZX10732401), and Fondation Mérieux. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

References

- [1] D. Soolingen, H.J. Wisselink, R. Lumb, R. Anthony, A. van der Zanden, C. Gilpin, Practical biosafety in the tuberculosis laboratory: containment at the source is what truly counts, *Int. J. Tuberc. Lung Dis.* 18 (2014) 885–889.
- [2] J. Alimonti, A. Leung, S. Jones, J. Gren, X. Qiu, L. Fernando, et al., Evaluation of transmission risks associated with in vivo replication of several high containment pathogens in a biosafety level 4 laboratory, *Sci. Rep.* 25 (2014) 5824.
- [3] M. Magnuson, R. Campisano, J. Griggs, S. Fitz-James, K. Hall, L. Mapp, M. Mullins, et al., Analysis of environmental contamination resulting from catastrophic incidents in a biosafety level 4 laboratory, *Sci. Rep.* 25 (2014) 5824.
- [4] M.L. Yarbrough, J.H. Kwon, M.A. Wallace, T. Hink, A. Shupe, V.J. Fraser, et al., Frequency of instrument, environment, and laboratory technologist contamination during routine diagnostic testing of infectious specimens, *J. Clin. Microbiol.* 56 (2018), e00225-18.
- [5] X.L. Wang, J. Song, Q.Q. Song, J. Yu, X.N. Luo, G.Z. Wu, J. Han, Viral contamination source in clinical microbiology laboratory, *Biomed. Environ. Sci.* 29 (2016) 609–611.
- [6] E. Tamburini, V. Donegà, M.G. Marchetti, P. Pedrini, C. Monticelli, A. Balbo, Study on microbial deposition and contamination onto six surfaces commonly used in

- chemical and microbiological laboratories, *Int. J. Environ. Res. Public Health* 12 (2015) 8295–8311.
- [7] R.J. Hay, Operator-induced contamination in cell culture systems, *Dev. Biol. Stand.* 75 (1991) 193–204.
- [8] L.S. Ng, W.T. Teh, S.K. Ng, L.C. Eng, T.Y. Tan, Bacterial contamination of hands and the environment in a microbiology laboratory, *J. Hosp. Infect.* 78 (2011) 231–233.
- [9] A.H. Németh, A.C. Kwasniewska, S. Lise, R.S. Parolin, E.B. Becker, K.D. Bera, et al., Next generation sequencing for molecular diagnosis of neurological disorders using ataxias as a model, *Brain* 136 (2013) 3106–3118.
- [10] H. Al-Mousa, M. Abouelhoda, D.M. Monies, N. Al-Tassan, A. Al-Ghoniaim, B. Al-Saud, et al., Unbiased targeted next-generation sequencing molecular approach for primary immunodeficiency diseases, *J. Allergy Clin. Immunol.* 137 (2016) 1780–1787.
- [11] F. Yan, Y. Xiao, M. Li, H. Zhang, R. Zhang, H. Zhou, et al., Metagenomic analysis identified human rhinovirus B91 infection in an adult suffering from severe pneumonia, *Am. J. Respir. Crit. Care Med.* 195 (2017) 1535–1536.
- [12] L. Ren, D. Yang, X. Ren, M. Li, X. Mu, Q. Wang, et al., Genotyping of human rhinovirus in adult patients with acute respiratory virus infections identified predominant infections of genotype A21, *Sci. Rep.* 7 (2017) 41601.
- [13] J. Yang, F. Yang, L. Ren, Z. Xiong, Z. Wu, J. Dong, et al., Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach, *J. Clin. Microbiol.* 49 (2011) 3463–3469.
- [14] L. Ren, R. Gonzalez, Z. Wang, Z. Xiang, Y. Wang, H. Zhou, et al., Prevalence of human respiratory viruses in adults with acute respiratory tract infections in Beijing, 2005–2007, *Clin. Microbiol. Infect.* 15 (2009) 1146–1153.
- [15] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410.
- [16] D.H. Huson, S. Beier, I. Flade, A. Górski, M. El-Hadidi, S. Mitra, H.J. Ruscheweyh, R. Tappu, MEGAN Community Edition - interactive exploration and analysis of large-scale microbiome sequencing data, *PLoS Comput. Biol.* 12 (2016), e1004957.
- [17] J.G. Caporaso, J. Kuczynski, J. Stombaugh, et al., QIIME allows analysis of high-throughput community sequencing data, *Nat. Methods* 7 (2010) 335–336.
- [18] L.V. Uryvaev, A.V. Dedova, L.V. Dedova, K.S. Ionova, N.A. Parasjuk, T.K. Selivanova, et al., Contamination of cell cultures with bovine viral diarrhoea virus (BVDV), *Bull. Exp. Biol. Med.* 153 (2012) 77–81.
- [19] K. Nagayama, K. Oguma, H. Sentsui, Survey on vertical infection of bovine viral diarrhoea virus from fetal bovine sera in the field, *J. Vet. Med. Sci.* 77 (2015) 1531–1534.
- [20] M. Giammaroli, J.F. Ridpath, E. Rossi, M. Bazzucchi, C. Casciari, G.M. De Mia, Genetic detection and characterization of emerging HoBi-like viruses in archival foetal bovine serum batches, *Biologicals* 43 (2015) 220–224.
- [21] L. Gagnieur, J. Cheval, M. Gratigny, C. Hébert, E. Muth, M. Dumarest, et al., Unbiased analysis by high throughput sequencing of the viral diversity in fetal bovine serum and trypsin used in cell culture, *Biologicals* 42 (2014) 145–152.
- [22] L. Perelygina, L. Zhu, H. Zurkühlen, R. Mills, M. Borodovsky, et al., Complete sequence and comparative analysis of the genome of herpes B virus (Cercopithecine herpesvirus 1) from a rhesus monkey, *J. Virol.* 77 (2003) 6167–6177.
- [23] S.A. Baylis, N. Shah, A. Jenkins, N.J. Berry, P.D. Minor, Simian cytomegalovirus and contamination of oral poliovirus vaccines, *Biologicals* 31 (2003) 63–73.
- [24] C.L. Greenblatt, J. Baum, B.Y. Klein, S. Nachshon, V. Koltunov, R.J. Cano, *Micrococcus luteus* - survival in Amber, *Microb. Ecol.* 48 (2004) 120–127.
- [25] S.J. Salter, M.J. Cox, E.M. Turek, S.T. Calus, W.O. Cookson, M.F. Moffatt, et al., Reagent and laboratory contamination can critically impact sequence-based microbiome analyses, *BMC Biol.* 12 (2014) 87.
- [26] J. Besser, H.A. Carleton, P. Gerner-Smidt, R.L. Lindsey, E. Trees, Next-generation sequencing technologies and their application to the study and control of bacterial infections, *Clin. Microbiol. Infect.* 24 (2018) 335–341.
- [27] O. Malik, H. Khamis, S. Rudnizky, A. Kaplan, The mechano-chemistry of a monomeric reverse transcriptase, *Nucleic Acids Res.* 45 (2017) 12954–12962.
- [28] M. Baba, R. Kakue, C. Leucht, P. Rasor, H. Walch, D. Ladiges, et al., Further increase in thermostability of Moloney murine leukemia virus reverse transcriptase by mutational combination, *Protein Eng. Des. Sel.* 30 (2017) 551–557.
- [29] Y. Katano, T. Hisayoshi, I. Kuze, H. Okano, M. Ito, K. Nishigaki, et al., Expression of moloney murine leukemia virus reverse transcriptase in a cell-free protein expression system, *Biotechnol. Lett.* 38 (2016) 1203–1211.
- [30] T.Z. Salem, C.P. Seaborn, C.M. Turney, J. Xue, H. Shang, X.W. Cheng, The influence of SV40 polyA on gene expression of baculovirus expression vector systems, *PLoS One* 10 (2015), e0145019.
- [31] K. Tsai, D.G. Courtney, B.R. Cullen, Addition of m6A to SV40 late mRNAs enhances viral structural gene expression and replication, *PLoS Pathog.* 14 (2018), e1006919.
- [32] D.H. Xu, X.Y. Wang, Y.L. Jia, T.Y. Wang, Z.W. Tian, X. Feng, et al., SV40 intron, a potent strong intron element that effectively increases transgene expression in transfected Chinese hamster ovary cells, *J. Cell. Mol. Med.* 22 (2018) 2231–2239.
- [33] S. Shrestha, S. Shrestha, J. Shindo, J.B. Sherchand, E. Haramoto, Virological quality of irrigation water sources and pepper mild mottle virus and tobacco mosaic virus as index of pathogenic virus contamination level, *Food. Environ. Virol.* 10 (2018) 107–120.
- [34] W. Chen, W. Liu, H. Jiao, H. Zhang, J. Cheng, Y. Wu, Development of a concentration method for detection of tobacco mosaic virus in irrigation water, *Virol. Sin.* 29 (2014) 155–161.
- [35] T. Newsome, B.J. Li, N. Zou, S.C. Lo, Presence of bacterial phage-like DNA sequences in commercial Taq DNA polymerase reagents, *J. Clin. Microbiol.* 42 (2004) 2264–2267.
- [36] T. Mohammadi, H.W. Reesink, C.M. Vandenbroucke-Grauls, P.H. Savelkoul, Removal of contaminating DNA from commercial nucleic acid extraction kit reagents, *J. Microbiol. Methods* 61 (2005) 285–288.
- [37] H.A. Barton, N.M. Taylor, B.R. Lubbers, A.C. Pemberton, DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database, *J. Microbiol. Methods* 66 (2006) 21–31.
- [38] M. Laurence, C. Hatzis, D.E. Brash, Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes, *PLoS One* 9 (2014), e97876.