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.
Original Article

VSITA, an Improved Approach of Target Amplification in the Identification of Viral Pathogens

ZHANG Yi1,6, ZHANG Chen1,6, LI Bo2, LI Yang1, HE Xiao Zhou1, LI Acher1, WU Wei1, DUAN Su Xia1,3, QIU Fang Zhou1,3, WANG Ji1, SHEN Xin Xin1, YANG Meng Jie1, LI De Xin1,6, and MA Xue Jun1,6

1. Key Laboratory for Medical Virology, National Health and Family Planning Commission, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2. State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100730, China; 3. Pediatric Research Institute, Children's Hospital of Hebei Province, Shijiazhuang 050031, Hebei, China

Abstract

Objective Unbiased next generation sequencing (NGS) is susceptible to interference from host or environmental sequences. Consequently, background depletion and virome enrichment techniques are usually needed for clinical samples where viral load is much lower than background sequences.

Methods A viral Sequence Independent Targeted Amplification (VSITA) approach using a set of non-ribosomal and virus-enriched octamers (V8) was developed and compared with traditionally used random hexamers (N6). Forty-five archived clinical samples of different types were used in parallel to compare the V8 and N6 enrichment performance of viral sequences and removal performance of ribosomal sequences in the step of reverse transcription followed by quantitative PCR (qPCR). Ten sera samples from patients with fever of unknown origin and 10 feces samples from patients with diarrhea of unknown origin were used in comparison of V8 and N6 enrichment performance following NGS analysis.

Results A minimum 30 hexamers matching to viral reference sequences (sense and antisense) were selected from a dataset of random 4,096 (4^6) hexamers (N6). Two random nucleotides were added to the 5’ end of the selected hexamers, and 480 (30 × 4^2) octamers (V8) were obtained. In general, VSITA approach showed higher enrichment of virus-targeted cDNA and enhanced ability to remove unwanted ribosomal sequences in the majorities of 45 predefined clinical samples. Moreover, VSITA combined with NGS enabled to detect not only more viruses but also achieve more viral reads hit and higher viral genome coverage in 20 clinical samples with diarrhea or fever of unknown origin.

Conclusion The VSITA approach designed in this study is demonstrated to possess higher sensitivity and broader genome coverage than traditionally used random hexamers in the NGS-based identification of viral pathogens directly from clinical samples.

Key words: Virus; Next generation sequencing; Non-ribosomal; Virus targeted

Biomed Environ Sci, 2018; 31(4): 272-279 doi: 10.3967/bes2018.035 ISSN: 0895-3988 www.besjournal.com (full text) CN: 11-2816/Q Copyright ©2018 by China CDC
INTRODUCTION

With the development of urbanization and globalization, viruses previously confined to small, remote geographic areas is keeping spreading all over the world through global business[1]. Classical viral diseases are invading novel populations. For example, the transmission of Middle East respiratory syndrome (MERS) coronavirus from camel to human in Arabic countries, and to remote areas, as far as Asia[2-7], the transmission of yellow fever virus from Africa to Asia[8-12], and the transmission of Zika virus from South America to the whole world[13,14-17]. Even in the area with inconvenient transportation, well-defined viruses can cause new headaches for public health, like the outbreak of Ebola in 2014-2016[18-25].

The identification of epidemic pathogens is therefore challenging. Traditionally, serological tests and PCR-based methods are used as first-line identification methods in outbreak of epidemic diseases[26]. However, these methods rely on pre-defined specific antibody, primer, probes, and do not work with new pathogens. The development of next generation sequencing (NGS), with unprecedented sequencing throughput, has promised to address these problems[27-31]. With the use of random oligo nucleic acids instead of predefined primers, unbiased sequencing is able to identify any virus sequence, known or unknown, local or imported, or even the pathogens that have never been reported[32].

However, unbiased sequencing or shotgun sequencing is susceptible to interference from environmental sequences. According to reported results as well as our own experience in processing traditional shotgun sequences, over 95% of success reads are commonly found to be host or environmental ribosomal RNA[33]. Consequently, background depletion techniques are particularly useful for samples where viral load is much lower than environment sequences. Some studies reported different methods for background depleton, but virome enrichment techniques are usually needed to access a broader set of viruses[34].

In the present study, we aimed to develop a Viral Sequence Independent Targeted Amplification (VSITA) approach using a set of non-ribosomal and virus-enriched octamers (V8) and compare with traditionally used random hexamers (N6) in the reverse transcription step followed by library construction and NGS analysis, attempting to reduce the interference from unwanted RNA sequences and improve NGS-based virus identification in clinical samples.

MATERIAL AND METHODS

Design of the V8 for VSITA

First, we removed a total of 3,985 hexamers that primed reverse-transcription of human or bacterial ribosomal RNA (Accession Number: NR_046235.3 of human 45S pre-ribosomal sequence, J01859.1 for Escherichia coli 16S ribosomal RNA, and NR_037007.2 for Staphylococcus aureus 16S ribosomal RNA) from a candidate dataset of 4,096 (4^6) random hexamers. Second, we constructed a single FASTA sequence file with all the virus-related sequences available from NCBI (8,584 viral genome sequences) (release data 8 July 2014) and corresponding complementary sequences (17,168 virus-related sequences in total). Third, we aligned the remaining 111 hexamers with both virus-related sequences, and scored each of them according to the numbers of perfect matching to either sequence. Forth, we then obtained the first 30 hexamers with the highest scores in matching to either sequence. Thus these 6+2N octamers (V8) (listed in Table 1) were used in reverse-transcription step for VSITA.

Table 1. Octamer (V8) Sequences Used for VSITA (5'-3')

| Motif     | Motif     | Motif     | Motif     | Motif     | Motif     | Motif     |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| NNNGCAAT  | NNGATATC  | NNGCATTG  | NNGTCTAG  | NNAGTATA  | NNAAGTAT  | NNAATGTG  |
| NNAATTGT  | NNAATCTA  | NNAATTTA  | NNCGTATG  | NNCGTGTT  | NNCAATGC  | NNCAATGC  |
| NNATACTT  | NNATGCGC  | NNATTTTA  | NNCTAGAC  | NNCTCGAG  | NNTAGATT  | NNTAGATT  |
| NNCAATTG  | NNCTACAG  | NNCTTAGA  | NNCTTAGG  | NNTTATA   | NNTTATA   | NNTTATA   |
| NNTAAAAT  | NNTATAAA  | NNTTATAG  | NNTTATA   | NNTTATA   | NNTTATA   | NNTTATA   |
|           |           |           |           |           |           |           |
Clinical Samples and Pre-preparation

Forty-five archived clinical samples of different types (serum, feces, and throat swabs) were included in this study. These samples were previously tested by quantitative RT-PCR assay (qRT-PCR) (Supplementary Table 1 available in www.besjournal.com) and found to be positive for either of 14 different virus types/subtypes [Japanese encephalitis virus, Rotavirus, Norovirus, Human Adenovirus (hADV), Dengue virus, Influenzavirus A H1N1, Influenzavirus A H3N2, Influenzavirus A swirl H1N1 (swiH1N1), Enterovirus 71 (EV71), Coxsackievirus A16 (CA16), Cytomegalovirus (CMV), Human immunodeficiency virus (HIV), Hepatitis B virus (HBV), Hepatitis C Virus (HCV)]. These 45 samples were used in parallel to compare the V8 and N6 enrichment performance of viral sequences and removal performance of ribosomal sequences in the step of reverse transcription followed by quantitative PCR (qPCR). Additional two separate groups, one group containing 10 serum samples from patients with fever of unknown origin (F1-F10) and one containing 10 feces samples from patients with diarrhea of unknown origin (D1-D10), were also included in this study. These 20 samples were used in comparison of V8 and N6 enrichment performance following NGS analysis.

For all the 65 samples, total RNA was extracted with RNease Mini Plus Kit (QIAGEN) following manufacturer’s instruction. To evaluate the performance of VSITA, both 4,096 random hexamers (N6) and VSITA primer set (V8) were used in parallel to synthesize first strand of cDNA with Superscript III (Invitrogen) for each sample following manufacturer’s instruction. For second strand synthesis of each sample, Klenow fragment was added to first strand synthesis system for 2-hour incubation at 37 °C.

Pre-evaluation—Quantitative Real-time PCR Assay for 18S rRNA, 28S rRNA and Viral Genome

For the 45 previously validated positive samples, both N6-primed and V8 primed cDNAs were evaluated by qPCR assays of ribosomal RNA and viral target sequence in parallel according to the protocols described in previous reports (Supplementary Table 1). Previously reported sequences of primers and probes were used in this study to ensure reliability of real-time PCR results (as listed in Supplementary Table 1). For comparison, relative quantity of target sequence using V8-primed cDNA vs. N6 counterpart was calculated as follows:

$$R = 2^{\left(C_{N6} - C_{V8}\right)}$$ (1)

R: relative quantity of V8-primed target cDNA vs. N6-primed target cDNA; $C_{N6}$: Average Ct value of quantitative real-time PCR using N6-synthesized cDNA of four replicates; $C_{V8}$: Average Ct value of quantitative real-time PCR using V8-synthesized cDNA of four replicates.

Evaluation by NGS

For the 20 unidentified clinical samples, both V8 and N6 primed cDNAs were pre-enriched by multiple displacement amplification (MDA) using phi29 DNA polymerase (QIAGEN) according to manufacturer’s instruction prior to library construction to guarantee enough amounts of templates for sequencing. Enriched cDNA products were evaluated with Agilent high sensitivity DNA assay on Agilent 2100 bioanalyzer. Fragmentation, end repair, adaptor ligation, size selection and purification were performed successively following instruction of Ion Torrent Hi-Q kit (Thermo Fisher Scientific). Adaptors with sequence barcodes were used to identify different libraries. After library evaluation with Agilent 2100 and real-time PCR, equal amount of each of the 40 libraries (20 samples for either V8 or N6 treatment) were grouped into four libraries and added to four 316v2 chips for sequencing. Sequencing was performed by 316 Chip Kit v2 on Ion Torrent PGM systems (Thermo Fisher Scientific) following instruction of Ion Torrent.

Data Analysis

Raw data from sequencing was generated by software of Ion Torrent PGM system and downloaded in BAM format. The raw data was analyzed with our in-house software virus identification pipeline (VIP)\(^{(95)}\).

RESULTS

Pre-evaluation by Real-time PCR

Relative quantity (the ratio) of V8-synthesized cDNA of target sequences vs. those N6-synthesized cDNA of the same sample was calculated to evaluate the ability of V8 primers in enriching target viral sequences and eliminating unwanted ribosomal sequences. Compared to N6, most V8-synthesized cDNA (39/45 for 18S, 31/45 for 28S) showed obvious decrease in the quantities of human 18S and 28S
A novel approach of target amplification of viral pathogen. 275

ribsosomal RNA sequences (Figure 1A and 1B). The discrepancy occurred in partial samples infected with Rotavirus, hADV, swH1N1, CMV, and HIV and all the samples infected with HBV and HCV. On the other hand, viral sequences were obviously enriched by V8 in most (38/45) samples (Figure 1C). The inconsistency occurred in partial samples infected with Norovirus, hADV, H3N2, coxsackievirus 16, HIV, HBV and HCV. In general, the relative quantities of 18S and 28S sequences were under 1 (Figure 1D), indicating the better performance of VSITA to inhibit the amplification of sequences from 18S and 28S. Meanwhile, the relative quantities of viral sequences were over 1 (Figure 1D), representing the better enrichment of viral genome sequences by VSITA.

**Fragment Evaluation of Enriched cDNA**

Fragment distribution of the enriched cDNA was evaluated by Agilent 2100 Bioanalyzer system. The obvious difference in the fragment distribution was observed between V8 and N6 priming. Taking the first-strand synthesis of 2 samples as examples for comparison (Figure 2A vs. 2B, 2C vs. 2D), the N6 priming showed higher diversity of the fragments, the fragments were more evenly distributed along the length axis (Figure 2A, 2C) while V8 priming tended to generate larger fragments more concentrated in some certain regions (Figure 2B, 2D). Results from other samples were showed in Supplementary Figure 1 (available in www.besjournal.com).

**NGS Evaluation**

Of all the 10 samples from patients with fever of unknown origin, N6 primers approach succeeded to identify 185 reads of dengue virus (DENV III) in sample F10 and covered 25.1% of dengue virus genome, while V8 approach identified 2,257 reads and covered 94.88% of the same dengue virus genome (Table 2). The presence of dengue virus genome in sample F10 was later validated by Real-time PCR (Supplementary Table 1). Of 4 samples

![Figure 1. V8 versus N6 ratio of different RNA levels quantified by qPCR. (A) Human 18S ribosomal RNA level; (B) Human 28S ribosomal RNA level; (C) Viral RNA level; (D) Overview of relative quantity. The calculation of relative quantity of V8 versus N6 was showed in material and methods.](image-url)
Figure 2. Fragment length analysis with Agilent 2100 bioanalyzer with: (A) N6 library from sample F1; (B) V8 library from sample F1; (C) N6 library from sample D1; (D) V8 library from sample D1.
(F1, F2, F3, and F7, Table 2), a few reads were generated by V8 but not by N6 approach. However, very few reads from F1, F2, and F3 were found to match with Hantavirus and Phlebovirus and real-time PCR assay failed to amplify any sequence from these samples. For sample F7, V8 approach was only able to identify 46 reads (22.31% of genome) due to low quantity of dengue virus genome (real-time PCR Ct value of 35.14). Both methods failed in the detection of the remaining 5 samples.

Virus sequences were identified from 5 out of 10 diarrhea samples of unknown origin (D1, D5, D6, D8, and D10) by both V8 and N6 approaches. Both methods found co-infection of Echovirus 6 and Norovirus in sample D1 with high possibility (with genome coverage of 92.33% and 90.13% for Echovirus 6, 86.53% and 86.93% for Norovirus, respectively). Of the other four samples, infection of Coxsackievirus A6, Poliovirus, Coxsackievirus A4, Coxsackievirus B2 was found by both methods and validated by real-time PCR. For virus identification with NGS, V8 showed more viral reads hit and higher genome coverage in all the samples tested (Table 2).

**DISCUSSION**

Shotgun NGS sequencing with random primer-based library construction has become the preferred choice for investigation of unknown infections in many laboratories. However, clinical sample is always a combination of different genomes from the host, the pathogen, and environmental species. Under such circumstance, viral sequence is usually at disadvantage in random library construction, regardless of actual viral load. Ribosomal RNA, above all, is predominantly amplified in random sequencing, and therefore became major interference in shotgun NGS-based pathogen identification, especially in viral infection cases. Unlike prokaryotes and eukaryotes, there is little similarity between virus genomes. Accordingly, efforts have been made to enrich viral sequences in NGS investigation of complex clinical samples. Endoh et al. has reported a set of 96 non-ribosomal hexanucleotides, which preferentially primes the virus genome RNA, however, the viral enrichment abilities of both random hexamers and 96 non-ribosomal hexanucleotides primer sets were nearly comparative in model experiments as described in our recent report. We speculate that the reduced diversity in non-ribosomal hexamers (from $4^6 = 4,096$ down to 96) might explain the limited coverage or sensitivity of viral pathogens found by 96 non-ribosomal hexamers.

To address this situation, we developed a novel VSITA approach by designing V8 primer set for virus-targeted selection in this study. After the removal of ribosomal RNA-targeted hexamers, we further narrowed virus-targeted hexamer sets down to an even smaller group containing only 30 virus-targeted hexamers. To expand the diversity of 30 primer set and improve detection sensitivity of viral pathogen, we intentionally added two random nucleotides (NN) to the 5’ end of all the hexamers. As a result, the final V8 primer set was composed of

| Sample Number | Identified Virus | Ct Value* | V8 | N6 |
|---------------|-----------------|-----------|----|----|
|               |                 |           | Reads Hit | Coverage (%) | Viral Reads Percentage (%) | Reads Hit | Coverage (%) | Viral Reads Percentage (%) |
| F1            | HFRS            | neg       | 3  | 2.76 | 0           | 0 | 0 | 0 |
| F2            | HFRS            | neg       | 1  | 0.38 | 0           | 0 | 0 | 0 |
| F3            | SFTS            | neg       | 3  | 0.94 | 0           | 0 | 0 | 0 |
| F7            | DENV III        | 35.14     | 46 | 22.31 | 0.08       | 0 | 0 | 0 |
| F10           | DENV III        | 25.97     | 2,257 | 94.88 | 0.61       | 185 | 25.10 | 0.19 |
| D1            | Echovirus 6     | 24.49     | 1,134 | 92.33 | 8.11       | 817 | 90.13 | 4.42 |
| D1            | Norovirus       | 28.75     | 48,344 | 86.53 | 8.11       | 25,276 | 86.93 | 4.42 |
| D5            | Coxsackievirus A6| 24.09  | 233 | 79.54 | 0.50       | 22 | 39.27 | 0.08 |
| D6            | Poliovirus      | 26.70     | 60  | 71.94 | 0.40       | 20 | 32.59 | 0.14 |
| D8            | Coxsackievirus A4| 26.50  | 171 | 78.26 | 0.80       | 51 | 47.55 | 0.30 |
| D10           | Coxsackievirus B2| 23.47  | 346 | 96.09 | 1.90       | 116 | 88.58 | 0.50 |

*Note.* Tests with reported real-time PCR methods to validate the identification of viruses.
480 (30 × 4² = 480) non-ribosomal virus-targeted octamers. We infer that the virus-targeted specificity of V8 will not be obviously affected as the specificity of a primer is predominantly determined by its 3’ end sequence.

Comparison of the virus-enrichment in cDNA synthesis between the V8 and random hexamers (N6) approaches was carried out by qPCR assays of 18S rRNA, 28S rRNA, and viral sequences. Our results showed V8 approach revealed improved virus enrichment and removal of human ribosomal RNA sequences in the majority of 45 archived clinical samples. Among a few samples (7/45), V8 approach showed inconsistent results (lower than N6) in virus (Norovirus, hADV, H3N2, coxsackievirus, HIV, HBV and HCV) enrichment. This phenomenon might be due to the difference in matrix constitution and targeted viral concentrations in extracted RNA of these clinical samples. Also, among some samples (6/45 for 18S, 14/45 for 28S), the performance of random hexamers was better in depletion of human ribosomal RNA than V8. Especially in the case of 28S sequences, N6 showed better performance than V8 in all of the 3 HBV-positive and 4 HCV-positive samples. The reason might be the nonspecific 28S amplification using cyber green-based qPCR analysis because subsequent melting curve analysis indicated they were nonspecific products. Though dsDNA viruses such as hADV and HBV were not primed in the step of reverse transcription by VSITA, they were involved in the virus-targeted enrichment in the PCR step due to the specificity of V8. More efforts will be needed to fine-tune the balance between virus-targeted sensitivity and specificity of V8 primer set in the future work. Nevertheless, given a large variety of virus-positive clinical samples of different types (serum, feces, and throat swabs) were tested, our data suggested VSITA is a worthwhile alternative to improve the enrichment of virus targets in the clinical samples.

Furthermore, 20 clinical samples with fever or diarrhea of unknown causes were used to evaluate the performance of V8 and N6 in parallel, the fragment length analysis of generated library was conducted with Agilent 2100 Bioanalyzer. The results showed that V8 produced longer fragments distribution and lower length diversity as compared to N6. This phenomenon might be attributed to the difference of V8 and N6 primer sets in the specificity and diversity. V8 has stronger specificity inclining to amplify limited viral genome regions, while N6 has higher diversity tending to generate fragments with a broad range. The sequencing results with ion torrent PGM system revealed an obviously superior performance using V8. By the use of V8 primers, one more case of dengue virus infection was found and validated using real-time PCR in serum sample tests. In the testing of fecal samples, both methods were able to report the causative viruses in several cases, including one case of multiple infections. However, V8 showed apparently higher genome coverage and more reads hit for all virus-positive samples in comparison with N6, except for one case with Norovirus infection, where N6 covered slightly larger region on the Norovirus genome (86.93% vs. 86.53%), although V8 achieved more reads hit. Overall, NGS data implied that V8 approach was superior to N6 in the identification of virus from clinical samples.

Other than virus-targeted amplification, efforts have also been made in pre-treatment of samples to improve virus identification with NGS. It is reported that ultracentrifugation of clinical samples improved virus discovery by removal of free RNA and DNA as well as bacteria[39]. Pre-degradation of free nucleic acid with cocktail enzyme system has also been reported[33]. The enrichment strategy in this study can be combined with those reported methods to provide an integrated and comprehensive enrichment solution.

In this study, only serum, swab, and feces samples were used to test the reliability of the VSITA approach followed by NGS. The performance of VSITA in other sample types, such as cerebrospinal fluid (CSF), urine, and other viral infection needs to be further validated in future study.

In conclusion, the VSITA approach designed in this study is demonstrated to possess higher sensitivity and broader genome coverage than traditionally used hexamers in NGS-based identification of viral pathogens directly from clinical samples.

CONFLICT OF INTERESTS

All the authors approved the final manuscript and they have no conflict of interest to declare.

ETHICAL APPROVAL

All aspects of the study were performed in accordance with national ethics regulations and approved by the Institutional Review Boards of National Institute for Viral Disease Control and Prevention, Center for Disease Control and Prevention of China.
ACKNOWLEDGEMENTS

We acknowledge the Pediatric Research Institute, Children’s Hospital of Hebei Province, Shijiazhuang, Hebei, China for providing fecal clinical samples, the National Laboratory for Hemorrhagic Fever, National Institute for Viral Disease Control and Prevention (IVDC), Chinese Center for Disease Control and Prevention (CDC) for providing serum clinical samples.

Received: January 24, 2018; Accepted: April 3, 2018

REFERENCES

1. Haug CJ, Kienny MP, Murgue B. The Zika Challenge. N Engl J Med, 2016; 374, 1801-3.
2. Su S, Wong G, Liu Y, et al. MERS in South Korea and China a potential outbreak threat? Lancet, 2015; 385, 2349-50.
3. MERS-the latest threat to global health security. Lancet, 2015; 385, 2324.
4. Drosten C, Meyer B, Muller MA, et al. Transmission of MERS-coronavirus in household contacts. N Engl J Med, 2014; 371, 828-35.
5. Azhar EI, El-Kafrawy SA, Farraj SA, et al. Evidence for camel-to-human transmission of MERS coronavirus. N Engl J Med, 2014; 370, 2499-505.
6. Perlman S, McCray PB Jr. Person-to-person spread of the MERS coronavirus-an evolving picture. N Engl J Med, 2013; 369, 466-7.
7. Assiri A, McGeer A, Perl TM, et al. Hospital outbreak of Middle East respiratory syndrome coronavirus. N Engl J Med, 2013; 369, 407-16.
8. Markoff L. Yellow fever outbreak in Sudan. N Engl J Med, 2013; 368, 869-91.
9. Chen Z, Liu L, Lv Y, et al. A fatal yellow fever virus infection in China: description and lessons. Emerg Microbes Infect, 2016; 5, e69.
10. The L Yellow fever a global reckoning. Lancet, 2016; 387, 1348.
11. Ahmed QA, Memish ZA. Yellow fever from Angola and Congo: a storm gathers. Trop Doct, 2017; 47, 92-6.
12. Wilder-Smith A, Leong WY. Importation of yellow fever into Singapore. Nature, 2013; 501, 263-4.
13. Ferguson NM, Cucunuba ZM, Dorigatti I, et al. EPIDEMIOLOGY. The Yellow fever a global reckoning. Lancet, 2016; 387, 1348.
14. Jouannic JM, Frisher S, Leparc-Goffart I, et al. Zika virus infection in French Polynesia. Lancet, 2016; 387, 1051-2.
15. Lessler J, Chaisson LH, Kucirka LM, et al. Assessing the global threat from Zika virus. Science, 2016; 353, 83-4.
16. Petersen LR, Jamieson DJ, Powers AM, et al. Zika Virus. N Engl J Med, 2016; 374, 1552-63.
17. The L Zika’s emerging threat for the Asia-Pacific region. Lancet, 2016; 388, 1026.
18. Gire SK, Goba A, Andersen KG, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science, 2014; 345, 1369-72.
19. Whitby CJ, Farrar J, Ferguson N, et al. Infectious disease: tough choices to reduce Ebola transmission. Nature, 2014; 515, 192-4.
20. Althaus CL. Ebola superspreading. Lancet Infect Dis, 2015; 15, 507-8.
21. Tong YG, Shi WF, Liu D, et al. Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone. Nature, 2015; 524, 93-6.
22. Zinsser K, Morrison K, Anema A, et al. The velocity of Ebola spread in parts of west Africa. Lancet Infect Dis, 2015; 15, 1005-7.
23. Currie J, Grenfell B, Farrar J. Infectious diseases. Beyond Ebola. Science, 2016; 351, 815-6.
24. Rico A, Brody D, Coronado F, et al. Epidemiology of Epidemic Ebola Virus Disease in Conakry and Surrounding Prefectures, Guinea, 2014-2015. Emerg Infect Dis, 2016; 22, 178-83.
25. Wang Q, Zhang Y, Wang HY, et al. Detection and Analysis of Ebola Virus in Sierra Leone-China Friendship Biosafety Laboratory from March 11 to April 20, 2015. Biomed Environ Sci, 2016; 29, 443-7.
26. Boonham N, Kreuze J, Winter S, et al. Methods in virus diagnostics: from ELISA to next generation sequencing. Virus Res, 2014; 186, 20-31.
27. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet, 2016; 17, 333-51.
28. Han Y, Gong L, Sheng J, et al. Prediction of virological response by pretreatment hepatitis B virus reverse transcriptase quasispecies heterogeneity: the advantage of using next-generation sequencing. Clin Microbiol Infect, 2015; 21, 797.e1-8.
29. Harrison A, Binder H, Buhot A, et al. Physico-chemical foundations underpinning microarray and next-generation sequencing experiments. Nucleic Acids Res, 2013; 41, 779-96.
30. Marx V. Next-generation sequencing. The genome jigsaw. Nature, 2013; 501, 263-8.
31. Saliba AE, Westermann AJ, Gorski SA, et al. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res, 2014; 42, 8845-60.
32. Schlaberg R, Chiu CY, Miller S, et al. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. Arch Pathol Lab Med, 2017; 141, 776-86.
33. Neill JD, Bayles DO, Ridpath JF. Simultaneous rapid sequencing of multiple RNA virus genomes. J Virol Methods, 2014; 201, 68-72.
34. Quince C, Walker AW, Simpson JT, et al. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol, 2013; 31, 833-44.
35. Li Y, Wang H, Nie K, et al. VIP: an integrated pipeline for metagenomics of virus identification and discovery. Sci Rep, 2016; 6, 23774.
36. Endoh D, Mizutani T, Kirisawa R, et al. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. Nucleic Acids Res, 2005; 33, e65.
37. Armour CD, Castle JC, Chen R, et al. Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. Nat Methods, 2014; 1005-7.
38. Armstrong CD, Castle JC, Chen R, et al. Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. Nat Methods, 2014; 201, 647-9.
39. Quince C, Walker AW, Simpson JT, et al. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol, 2013; 31, 833-44.
40. Li Y, Wang H, Nie K, et al. VIP: an integrated pipeline for metagenomics of virus identification and discovery. Sci Rep, 2016; 6, 23774.
41. Endoh D, Mizutani T, Kirisawa R, et al. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. Nucleic Acids Res, 2005; 33, e65.
42. Neill JD, Bayles DO, Ridpath JF. Simultaneous rapid sequencing of multiple RNA virus genomes. J Virol Methods, 2014; 201, 68-72.
43. Quince C, Walker AW, Simpson JT, et al. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol, 2013; 31, 833-44.
44. Li Y, Wang H, Nie K, et al. VIP: an integrated pipeline for metagenomics of virus identification and discovery. Sci Rep, 2016; 6, 23774.
45. Endoh D, Mizutani T, Kirisawa R, et al. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. Nucleic Acids Res, 2005; 33, e65.