

Chinese Expert Consensuses on the Application of Nanopore Sequencing in the Detection of Pathogenic Microorganisms

Grassroots Committee Division of Therapeutic Drug Monitoring, Chinese Pharmacological Society; Expert Committee of Precision Medicine for Clinical Treatment of Guangdong Pharmaceutical Association; the Writing Group of the *Consensuses of Experts on the Application of Nanopore Sequencing in the Detection of Pathogenic Microorganisms*

ABSTRACT To improve the diagnosis and treatment of critical infectious diseases, standardize the clinical application of nanopore sequencing technology, and promote the benign development of this technology.

METHODS Using the nominal group method, a writing group of multidisciplinary experts (pharmacy, laboratory, clinical) was formed. After joint discussion, the outline of "Expert consensus on the application of nanopore sequencing technology in the detection of pathogenic microorganisms" was determined. The Delphi method was used to collect the external review opinions of experts, and the opinions were sorted out, summarized, analyzed, fed back and revised. Finally, the medical experts of the leading unit held a finalized meeting to discuss and establish a consensus. Using the nominal group method, a writing group of multidisciplinary experts (pharmacy, laboratory, clinical) was formed. After joint discussion, the outline of "Expert consensus on the application of nanopore sequencing technology in the detection of pathogenic microorganisms" was determined. The Delphi method was used to collect the external review opinions of experts, and the opinions were sorted out, summarized, analyzed, fed back and revised. Finally, the medical experts of the leading unit held a finalized meeting to discuss and establish a consensus.

RESULTS & CONCLUSIONS Although the current sequencing technology has undergone multiple generations of changes, its performance tends to be optimized and stable, but improving accuracy and reducing sequencing cost are still its main challenges. After the standardized application of nanopore sequencing technology, Nanopore sequencing technology will have a broad application prospect in the field of pathogenic microorganism detection, clinical infection microorganism diagnosis and real-time pathogenic microorganism monitoring.

KEYWORDS Third-generation sequencing; Nanopore sequencing; Pathogenic microorganisms

Infectious disease is a general term for diseases caused by pathogenic microorganisms, and it is still a major threat to global public health. Rapid and accurate detection of pathogenic microorganisms is of great clinical value for their diagnosis and treatment. There are three types of detection for pathogenic microorganisms in clinical laboratories: (1) Pathogen isolation and culture is the gold standard for the diagnosis of infectious diseases. However, many pathogenic microorganisms are difficult or even impossible to culture due to the long culture period and low positive rate, so the effective detection of pathogens can be hard, resulting in delayed or missed diagnosis; Besides, the delayed etiological diagnosis can cause improper use or overuse of broad-

spectrum antibiotics, leading to the increase of drug resistance of pathogenic bacteria and the prevalence of highly drug-resistant pathogenic microorganisms in clinic. (2) Immunology-based detection, such as antigen and antibody detection, features low cost and ease of operation, but the results can be easily disturbed, leading to false positives/negatives; (3) Detection based on molecular biology, such as fluorescence PCR and hybridization chip, excels in rapidity, sensitivity and specificity, but it belongs to targeted testing, which cannot detect unknown pathogens, and the number of targets detected at a time is limited.

With the rise of genomics, high-throughput sequencing provides a new technical means for the identification of pathogenic microorganisms. Rapid and accurate identification of pathogenic microorganisms of infectious diseases can be achieved through the determination of nucleic acid sequences of samples with higher throughput and shorter duration. As a result, it has been increasingly used in the diagnosis and epidemic prevention of clinical infectious diseases. High-throughput sequencing, also known as the new generation sequencing, generally refers to the next-generation sequencing technique (NGS) and the third-generation sequencing technique (TGS). NGS has a high sequencing throughput, but its sequencing reading length is short, which leads to the difficulty of genome splicing in subsequent bioinformatics analysis and the high GC preference in sequencing. These shortcomings limit its application and promote the development of TGS^[1]. TGS has been gradually applied in various clinical fields because of its long sequencing reading length, which partly makes up for the deficiency of NGS. TGS is mainly divided into targeted next-generation sequencing (tNGS), metagenomics next-generation sequencing (MNGS) and whole genome sequencing (WGS) according to their different detection purposes and methods. TGS can be divided into (1) nanohole electrical signal sequencing, represented by the nanopore single molecule sequencing technology of Oxford Nanopore Technologies (ONT), and (2) single molecule fluorescence signal sequencing, represented by the single molecule realtime (SMRT) technology of Pacific Biosciences (PacBio) according to their different sequencing principles.

The third-generation sequencing based on nanopore single molecule sequencing (hereinafter referred to as "third-generation nanopore sequencing") has the advantages of real-time, portability, long sequencing reading length, and simple operation, which can be used for sample detection at bedside and in the field, meet the needs of clinical real-time detection, shorten the diagnosis time of clinical pathogenic microorganisms and reduce antibiotic abuse. Therefore, nanopore sequencing has played an important role in the fields of pathogenic microorganism detection, diagnosis and prevention of infectious diseases in recent years. During the COVID-19 pandemic, China's CDC and several provincial and municipal designated hospitals used nanopore sequencing technology to carry out full-length sequencing and typing, which provided support for immediate epidemic prevention and control^{[2]-[3]}. Currently, there have been many multi-center studies and articles published in China based on nanopore sequencing pathogen detection. For example, a multi-center study led by the First Affiliated Hospital of Zhejiang University School of Medicine and involving 8 hospitals applied nanopore targeted sequencing technology (NTS) to the pathogen diagnosis of alveolar lavage fluid in patients with pneumonia. According to the research results, NTS has the advantages of long reading length, short detection period, low cost and covering the core pathogens, and has excellent performance in the etiological diagnosis of patients with respiratory pneumonia, so it can be used as a reliable means for the etiological

diagnosis of patients with pulmonary infection^[4]. Another example is a multi-center study led by the First Affiliated Hospital of Suzhou University and participated by ten 3A hospitals in Jiangsu Province. The results show that the overall sensitivity and specificity of nanopore mNGS detection are significantly higher than that of culture method for pathogen identification and drug resistance detection of many different types of samples^[5]. In conclusion, nanopore sequencing has unique advantages and broad application prospects in the field of clinical pathogenic microorganism detection.

However, the existing consensus on high-throughput sequencing specification of pathogenic microorganisms is mainly based on the second-generation sequencing platform, such as the *China Expert Consensus on the Application of Metagenome High-throughput Sequencing Technique in Pathogen Detection of Infectious Diseases*, the *Expert Consensus on Clinical Application of China Metagenomics Second-generation Sequencing Technology to Detect Infectious Pathogens*, and the Expert Consensus on the Application of High-throughput Sequencing Technique in the Diagnosis of Mycobacterial Diseases^[6]-Error! Reference source not found. Because of its characteristics, the third-generation sequencing based on nanopore single molecule sequencing is quite different from the second-generation sequencing in terms of database building process and data analysis, but there is still a lack of technical specifications for nanopore sequencing. This paper standardizes the whole process of nanopore sequencing, including sample collection and preservation, detection, bioanalysis and report interpretation, and gives recommendations on key issues, and establishes the *Consensuses of Experts on the Application of Nanopore Sequencing in the Detection of Pathogenic Microorganism*, which is applicable to three directions of nanopore sequencing (tNGS, mNGS and WGS).

1. Consensus-making method

This consensus was jointly initiated by the Grassroots Committee Division of Therapeutic Drug Monitoring of Chinese Pharmacological Society and the Expert Committee of Precision Medicine for Clinical Treatment of Guangdong Pharmaceutical Association, and co-lead by Guangzhou Red Cross Hospital Affiliated to Jinan University, Huazhong University of Science and Technology Union Shenzhen Hospital (Nanshan Hospital), Sixth Affiliated Hospital of Sun Yat-sen University, Guangdong Provincial People's Hospital, First Affiliated Hospital of Sun Yat-sen University and Memorial Hospital of Sun Yat-sen University. Multidisciplinary medical and laboratory experts in China were organized for the compilation and validation of the consensus.

Nominal group technique is adopted in this consensus, that is, a writing group of experts from many disciplines (pharmacy, laboratory and clinical) discussed and determined the consensus outline through online and offline meetings. The outline mainly includes the principle, characteristics, basic process and requirements of nanopore sequencing, bioanalysis and report issuance and interpretation. The writing group systematically searched, analyzed and summarized the contents involved in the outline, and made a consensus according to the current situation, clinical needs and research evidence in China. Foreign revisers are multidisciplinary experts with rich clinical experience. Delphi method was used to collect the opinions of foreign reviewers, and their opinions were sorted, summarized, analyzed, fed back and revised. Finally, the medical experts of the lead unit held a final meeting to discuss and establish this consensus.

2. Principle and characteristics of nanopore sequencing

2.1 Principle of nanopore sequencing

Nanopore sequencing refers to a sequencing practice integrating single molecule detection and current signal conduction. In principle, it gets rid of the elution of NGS and PCR amplification (sequencing while synthesizing). Double-stranded nucleic acid is untied by motor protein connected to the connector, and single-stranded nucleic acid molecules pass through the nanopore driven by field stress. As different bases generate current signals with different blocking degrees when they pass through nanopores, the base information can be deduced by recording the change of current signal, so as to complete the sequencing of nucleic acid molecules^{[1][9]}. This technology marks a giant leap in detection from short to ultra-long reads, and from optical to electron conduction detection.

2.2 Characteristics of nanopore sequencing

(1) The longest sequence that can be detected by nanopore sequencing reaches Mb level^[9], and long reading length and long sequence result in higher specificity in the identification of pathogenic species at the same level with strong homology. Nanopore single molecule sequencing can span some complex regions of the genome to complete the assembly of complex genomes and the discovery of unknown structural variations^{[10]-[11]};

(2) The average speed of nanopore single-hole sequencing is 450bp/s^[9], and multi-channel nanopore sequencing can be performed in parallel to quickly obtain high throughput to be sequenced;

(3) Library building needs no fragmentation, and sequencing can be performed by directly connecting the splice and sequencing the nucleic acid sequence, which is easy to operate^[12];

(4) Sequencing independent of PCR amplification avoids the preference brought by amplification to sequencing. The original DNA or RNA can be directly sequenced without bias, and the original base modification information is retained^{[13]-[14]};

(5) Real-time base identification and analysis can terminate sequencing on demand, significantly shortening the sequencing and analysis process to ensure dynamic detection^[12];

(6) Nanopore chips can be used repeatedly after cleaning until the nanopore protein is inactivated, reducing the demand for sample collection and improving the utilization of chips;

(7) Nanopore sequencers are small and portable, and is not limited by external resources such as power, laboratory space and servers, making them suitable for real-time sequencing near beds, in the field or in outbreak areas^{[15]-[16]}.

Recommendation 1: Due to the potential bio-safety hazard of pathogenic samples, it is recommended to carry out pathogenic microorganism detection projects locally to avoid improper transportation; According to the flexibility of nanopore sequencing, sample pretreatment and computer sequencing should be completed immediately after sampling.

Recommendation 2: In the case that infectious diseases are highly suspected according to clinical manifestations with unknown pathogens, and it is necessary to quickly identify the infectious pathogenic microorganisms, but the sample size is small and the startup cost is high, it is recommended to use the third-generation nanopore sequencing, which is more flexible and can identify and analyze the bases in real time, so as to achieve the balance between timeliness and cost of

detection of infectious pathogenic microorganisms.

Recommendation 3: For the prevention and control of emerging infectious diseases, when it is necessary to analyze the evolution and variation of pathogens based on the full-length genome, it is recommended to use the third-generation nanopore sequencing with long sequencing length for the full-length genome detection and molecular traceability of pathogenic microorganisms.

Recommendation 4: It is recommended to use the third-generation nanopore sequencing with long sequencing reading length to distinguish the subtypes of pathogenic microorganisms with high homology (such as Mycobacterium tuberculosis and non-Mycobacterium tuberculosis) to increase its specificity.

3. Basic process and requirements of nanopore sequencing

3.1 Sample collection, transportation and preservation

3.1.1 Sample collection

For patients with clinically suspected infectious diseases, samples are collected for nanopore sequencing. The sensitivity of this technology mainly depends on the proportion of human nucleic acids in the sample, the contents of microorganisms and the extraction efficiency of microbial nucleic acids. The following principles should be followed in the collection of samples for infectious diseases to ensure the effectiveness of detection^{Error! Reference source not found.-Error! Reference source not found.}: (1) Try to collect samples on the day when patients are admitted to hospital or before using antibacterial drugs; (2) Try to collect samples of primary lesions; (3) Sterile operation should be strictly carried out in sample collection; (4) Aseptic operation should be strictly carried out when collecting test samples from aseptic parts to avoid pollution; (5) Necessary measures should be taken to minimize the pollution of the samples collected from the infected or contaminated parts. The sample collection methods are as follows^{Error! Reference source not found.-Error! Reference source not found.}.

(1) Blood, bone marrow and other hypercoagulable samples: Use EDTA anticoagulant tubes or free DNA preservation tubes to take 3-5 ml of blood from elbow vein. Thoroughly disinfect the skin before collection, and avoid collection from the vein where antibiotics are administered unless catheter-related bloodstream infection is suspected. Gently mix the sample upside down for 8-10 times, so as to fully mix the blood sample and the protective agent to avoid coagulation or hemolysis. Tight the tube cover immediately, and send it for test after barcode labeling.

(2) Respiratory tract samples: Throat swab: Throat swab samples are usually only used to diagnose upper respiratory tract infections. Ask the patient to open his/her mouth and pronounce "ah" to expose the throat, and use a tongue depressor if necessary; Take out the long sterile cotton swab from the tube, and quickly wipe the secretions of the palatal arch, pharynx and tonsil on both sides. Try to squeeze and collect some pus on the tonsil. Place 2-3 swabs in a tube containing preservation solution, tight the tube cover immediately, and send it for test after barcode labeling.

Bronchoalveolar lavage fluid (BALF): BALF is the best sample for suspected pneumonia or lower respiratory tract infection. Introduce a fiberoptic bronchoscope after the patient's throat is locally anesthetized, and lavage the lung or sub-lung segment below the bronchus where the lesion is located for several times through the

bronchoscope. Inject 20-60 ml of sterile saline each time at 37°C or room temperature, usually repeat for 4-5 times, and collect the liquid with full negative pressure suction. The sample contains about 10 ml of secretions from bronchial endings and alveoli; Discard the samples that may be contaminated in the first tube, collect the second one, and send it for test immediately after barcode labeling.

Sputum: Sputum samples are mainly used for auxiliary diagnosis of lower respiratory tract infection. Patients should rinse their mouths with clear water or normal saline for 2-3 times under the guidance of medical staff (those with dentures should take off their dentures first), and then cough up 3-5 ml of deep sputum in a sterile cup. Do not take saliva or nasopharyngeal cavity secretions; Collect the sample from the airway through a sputum aspirator if the patients cannot expectorate themselves. Tight the tube cover immediately, and send it for test after barcode labeling.

(3) Cerebrospinal fluids: Cerebrospinal fluids are the most important sample for diagnosis of central nervous system infection. After the skin is disinfected, insert an empty needle with a core needle into the space between the 3rd and 4th or the 4th and 5th lumbar vertebrae. Insert the needle into the arachnoid space, and remove the core needle to collect cerebrospinal fluids. Take about 3-5 ml of cerebrospinal fluids after the second tube to prevent pollution, and send it for test immediately after barcode labeling.

(4) Other sterile body fluids: Pleural/ascites/joint effusion: Obtained by clinicians through puncture or surgical methods. Locate the puncture site by imaging or ultrasound, and perform anesthesia after disinfecting the skin of the puncture position. Puncture the target site with a vent needle, and extract 5 ml of pleural/ascites/joint effusion samples in a sterile tube, and send it for test immediately after barcode labeling.

Urine: A midstream clear urine sample usually collected by the patients themselves. Collect 5 ml of midstream urine into a sterile sampling tube early in the morning, and send it for test immediately after barcode labeling.

(5) Abscess samples: Open abscess: Clean the wound surface with saline solution, collect the secretion at the base of deep wound or ulcer with a swab, or cut off the tissue at the edge of deep lesion.

Closed abscess: Thoroughly disinfect the local skin or mucosal surface of the lesions, extract 3 ml of pus with a syringe, or take part of the abscess wall for test after incision and drainage.

Fistula or sinus abscess: Preferably collect the deepest tissue during surgical exploration. Place it into a sterile sampling tube, and send it for test immediately after barcode labeling.

(6) Tissue samples: Deep tissue samples can only be obtained during surgical exploration. Collect enough tissue samples, generally $\geq 5\text{mm}^3$, at least about the size of mung beans. Small pieces of tissue should be moisturized with 1-2 ml of sterile water or saline solution. Large pieces should be cut and packed into a sterile sampling tube, and sent for test immediately after barcode labeling.

(7) Aqueous humor/vitreous humor samples: Aqueous humor samples are collected by trained ophthalmologists in the operating room. After anesthesia and routine cleaning of conjunctival, puncture into the anterior chamber parallel to the iris plane at the limbus corneoscleralis with 1 ml of sterile syringe, and extract about 0.1 ml of aqueous humor samples.

Puncture the vitreous cavity 10 mm perpendicular to the sclera in the posterior flat part of the corneoscleral margin to the center of the eyeball, and extract vitreous humor samples as many as possible (≥ 0.2 ml). Place it into a sterile sampling tube, and send it for test immediately after barcode labeling.

(8) Fecal samples: Fecal samples are generally collected by the patients themselves, and natural defecation or rectal swab are usually selected according to the testing requirements. For natural defecation, the parts that are easy to be polluted in the front section should be discarded, and the feces in the middle section should be about 0.5-1 g, the size of soybean grains; For rectal swab, wash around the anus with soapy water, and insert a cotton swab stained with sterile saline into the anus 4-5 cm (2-3 cm for children). The cotton swab is in contact with the surface of rectal mucosa. Gently rotate the swab until feces can be clearly seen on the swab. Preferably collect ≥ 2 swabs. Place it into a sterile sampling tube, and send it for test immediately after barcode labeling.

(9) Bacterial liquid samples: Bacterial liquid samples for tNGS or WGS are usually obtained by culture, which requires suitable separation, purification and culture conditions to obtain pure cultures from clinical samples. For bacterial liquid samples for non-targeted PCR-Free WGS, the nucleic acids extracted from the collected samples should be greater than 1 μg , and the bacterial liquids be placed into a 5 ml sterile sampling tube for test.

3.1.2 Sample transportation and preservation

(1) Sample transportation: The type of delivery depends on actual situations. If the samples can be delivered to the laboratory and the test can be started within 24 hours, they can be transported with ice bags at low temperature; If the samples can be delivered within 24-72 hours, they should be transported with dry ice, and sample pretreatment and nucleic acid extraction be carried out immediately after they arrive at the laboratory to prevent the degradation of nucleic acids from affecting the test results^{Error! Reference source not found.}. During transportation, violent bumps should be avoided as much as possible to prevent pollution risks caused by liquid leakage. If highly pathogenic or new infectious diseases are suspected, they should be packaged and transported in strict accordance with the requirements of China's infectious diseases law and other relevant laws and regulations^{Error! Reference source not found.}. If they are sent to a third-party company for sequencing, the nucleic acids should be extracted preferably under biosafety protection conditions, and then transported with dry ice.

(2) Sample preservation: The samples should be sent for test immediately after collection. If not, they can be temporarily stored at 4°C for no more than 7 hours, and they should be sent for test on the same day. If the delivery cannot be completed in a short period of time, they should be immediately stored in the refrigerator below -20°C for no more than 7 days. For long-term preservation, they should be immediately stored

in the refrigerator at -80°C , and the RNA sequencing samples be directly stored in the refrigerator at -80°C . Repeated freezing and thawing should be preferably avoided during preservation, generally ≤ 3 times.

4. Detection Process

4.1 Sample pretreatment

Ensure that the samples are collected and preserved properly to avoid pollution or nucleic acid degradation, and timely record its initial state.

Each laboratory should establish standard operating procedures for sample pretreatment, including sample liquefaction, sample concentration and host removal. For different types of clinical samples, targeted pretreatment procedures should be established. For example, sputum, a highly viscous sample, needs to be liquefied, and a standardized procedure should be established for the classification of sample viscosity, the volume of liquefied liquid and the liquefaction time during the liquefaction process to ensure the integrity of microbial nucleic acid; For blood samples, it is necessary to decide whether to centrifuge for plasma separation according to the actual testing requirements; For tissue samples, in order to improve the efficiency of nucleic acid extraction, it is necessary to chop them in advance before operation^[7]. Whether it is necessary to remove host cells or nucleic acids before sample extraction should be considered according to the sample type and application direction. Targeted sequencing is targeted to capture specific pathogenic microorganism nucleic acids, and human nucleic acids have little interference with the detection results. However, for macro-gene sequencing, if the content of host cells in the sample is high, the sensitivity of detection of pathogenic microorganisms will be reduced under the condition of constant sequencing data. Therefore, a verified method can be considered to remove host cells or nucleic acids before extracting high-host background samples^[6].

4.2 Nucleic acid extraction

Establish a complete nucleic acid extraction method, which needs to be verified by repeatability and tolerance to ensure the integrity and purity of the extracted nucleic acid. See "4.3 Verification of nucleic acid quality" for specific verification indicators. Standardized extraction procedures should be established according to the types of clinical specimens and microbial species to ensure the extraction effect of different types of pathogenic microorganisms. For example, the breaking effect is guaranteed for special microorganisms such as fungi and mycobacteria, and the extraction effect is guaranteed for RNA viruses and free nucleic acids. Each test samples should include internal reference, negative control and positive control to evaluate whether there are any abnormalities such as pollution caused by operation or environment in each batch of samples^[7]. If an automatic extractor is used, it should be ensured that cross-contamination is avoided and the time is within a suitable range.

When extracting nucleic acid, whether DNA and RNA need to be extracted separately should be considered according to the application direction and detection target. If the detection scope includes new or unknown pathogens (including RNA virus) and needs to be detected by macro-gene, DNA and RNA should be extracted separately in nucleic acid extraction to improve the undifferentiated detection rate of various pathogens.

4.3 Verification of nucleic acid quality

Each laboratory should establish standards for qualified nucleic acid samples, including purity, integrity and nucleic acid content. The A_{260}/A_{280} of high-quality DNA should be 1.7~1.9, and A_{260}/A_{230} be greater than 2. The A_{260}/A_{280} of high-quality RNA should be 1.8~2.0, and $A_{260}/A_{230} > 2$ ^{Error! Reference source not found.}. Appropriate methods should be used to detect the integrity of nucleic acid. If the degradation of nucleic acid is serious, it needs to be re-extracted. The nucleic acid samples extracted each time should be quantitatively determined by using Qubit fluorescent dye to ensure that the nucleic acid quantity meets the requirements of subsequent tests.

4.4 Library preparation

Library preparation for TGS refers to the process of repairing the extracted nucleic acids and connecting the sequencing splice. If multiple samples are sequenced in parallel, sequencing splice can be connected after connecting the tag splice. Different types of connection have different reagents, time and initial investment, so it is necessary to clarify their standardized operation procedures. Library preparation varies according to the different applications of the sequencing technology. For example, tNGS should ensure the specific amplification or capture efficiency of pathogenic microorganisms; mNGS focuses more on unbiased amplification; For WGS, the sequencing can be performed without interrupting nucleic acids, and only 1 μ g of pure bacterial liquid samples are needed for WGS. Samples with low initial amount can be sequenced after amplification. For example, WGS can be performed for 2019-nCoV after amplification with a Ct value of only 35.

PCR amplification and fragmentation can be carried out according to the initial amount of nucleic acids and the final application, so as to improve the data output. Amplification may lead to the disproportionate expansion of flora or contaminated sequence, placing greater demands on analysis. Upon the establishment of the library, it should be tested timely on the computer. Since the sequencing splice contains protein components, it is not suitable for ethanol treatment or long-term preservation.

4.5 Verification of library quality

To determine the nucleic acid quality and the library outcomes, different quality standards should be adopted according to how the library is constructed. Qubit fluorochrome is used to identify library concentrations, and a certain concentration needs to be reached for a high-quality library. The concentration and library throughput are related to what the application is for and the length of nucleic acid sequence in the library. It is necessary to determine the length of nucleic acid sequence in the library, and adjust the library size accordingly. For example, to detect a sequence of about 5,000 bp with WGS requires 170 ng of nucleic acids for a sample size of 50 fmol^[21].

4.6 Computer sequencing

Nanopore chips contains biological protein pores, and its preservation and use should conform to the standard process to ensure the sequencing performance. The quality control procedures and standards should be clear before use. Nanopore sequencing chips can be used several times, and there should be clear requirements for the standards of available chips. The number of available holes in the chip used for the first time should be >800. As the number of use increases, the available holes in the chip are reduced, so do the data production speed and hole status to varying degrees. Whether the chip can continue to be used should be evaluated according to the amount of data

needed and the test requirements. The minimum application standard is 50 holes according to the quality control standard and test report of nanopore chip made by ONT manufacturer.

The amount of sequencing data required depends on the expected use (pathogen identification, drug resistance detection, WGS, etc.), the human nucleic acid ratio in the sample, and the quality of the chip. Nanopore sequencers output sequence information of nucleic acids in real time, including length, quality, splice effect, the status of holes, the stability of the chip, and the speed through the holes, and determine the sequencing standards according to different applications. Generally, the splice rate should be $\geq 90\%$, the use ratio of holes $\geq 75\%$, and the through-hole speed of the sequence should be above 300 bp/s, and the sequencing quality value should be greater than 9. The sequencing should be stopped immediately and remedied if the above requirements are not met.

Recommendation 5: There should be standardized procedures for cleaning and using nanopore chips, determine the quality control standards before use, select appropriate parameters such as data volume and sequencing time according to sample type and application direction, and monitor sequencing indicators during sequencing, such as sequencing speed and quality, and sequencing status of holes.

5. Bioanalysis and report issuance

5.1 Bioanalysis

Bioanalysis includes the following steps: sequence splitting, splice and low-quality sequence filtering, data statistics, comparison with reference database and species identification, filtering background microorganisms, identifying cross-contamination, and species annotation (species taxonomy, pathogenicity, clinical significance, etc.).

5.1.1 Sequence splitting

When constructing a sequencing library, each sample is added with a tag, and multiple samples are mixed in batches for computing (the multiple samples mixed are added with a chip). Samples are distinguished by identifying the tags after sequencing (also known as barcode, generally 24 bp^[22] and belongs to synthetic sequence fragments).

5.1.2 Splice and low-quality sequence filtering

The sequencing data of each sample obtained after the raw data is split needs to be filtered, including sequencing splice and low-quality sequences (sequences with shorter length or lower average quality), and the obtained high-quality sequences are used as the input data for microbial identification. For TGS, porechop^[23] can be used to filter the splice and cut it off from the sequence. Tools such as NanoFilt^[26] and Filtrong^[28] can be used to filter length and low-quality sequences, and the threshold value of minimum length can be selected according to the fragment length for different detection products. Generally, the whole sequence should be filtered if it has an average sequencing quality value $< Q7$ ^{[24]-[25]}.

5.1.3 Statistics

NanoStats^[26] can be used for the statistics on sequencing data, including sequence number (reads), base number (bases), sequence-length distribution, average sequencing quality values, etc. There are different requirements for statistical indicators of different sample types and product testing performance, and relevant tests are needed to obtain

statistical indicators.

The principles to be referenced in determining the index of sequencing data amount include the following: (1) Application direction: Compared with tNGS, mNGS detects more microbial targets and is more influenced by human host nucleic acid, so mNGS generally needs more data than tNGS. (2) Product detection limit: The lower the detection limit, the higher the data required. (3) Sample type: For alveolar lavage fluid samples, the abundance of pathogenic microorganisms is high, which may be detected under the condition of low data volume; For blood samples, the abundance of pathogenic microorganisms is relatively low, and the content of human cells is high, which often requires a higher amount of data to detect.

As for the index of sequence length distribution, the sequence fragment of mNGS is randomly broken from the microbial genome, and the length distribution is generally random, while the fragment length of tNGS depends on the primer amplicon length of the microorganism to be detected. For example, for microorganisms with smaller genomes such as viruses, it is easier to break into short fragments (as short as 100bp) during the test^[27], so the primer amplicon needs to be designed shorter, which can also affect the length of sequencing fragments. Therefore, the difference of amplicon length can directly affect the requirement of sequence length, which is usually not less than the shortest amplicon length.

5.1.4 Reference database comparison and species identification

Minimap2^[29] and other software are used to compare the long reads of TGS with the reference database to obtain the information on species identification and taxonomy. The reference database should be constructed with high-quality reference genome database, such as FDA-ARGOS^[30] and NCBI RefSeq^[31]. Samples of pathogenic microorganisms may have host sequences. If they are human samples, the comparison sequence database should include human reference genome databases, such as Human GRCh37/hg19 and Human GRCh38/hg38 genome databases (<http://genome.ucsc.edu/>). They are commonly used human gene databases published in February 2009 and December 2013 respectively. The comparison database should be updated regularly to filter and delete redundant, incorrect or poorly assembled reference sequences, and include the genome information of newly discovered pathogenic microorganisms. The laboratory should construct the pathogen spectrum according to the sample type, and each microorganism should contain enough sequence characteristics to represent the species and genus level.

5.1.5 Filtering background microorganisms

There are reagent engineering bacteria, environmental microorganisms and laboratory residual microorganisms in pathogen detection, which can cause sequencing pollution and false positive results. Therefore, the laboratory needs to build a background database to filter contaminated sequences, and set up a no template control (NTC) for certain testing samples to monitor the background microorganisms and filter them^[32].

5.1.6 Determining cross contamination

The cross-contamination between samples in the same batch can be caused by tag synthesis errors, pollution between samples, aerosol pollution caused by opening the lid during the test, etc., especially the strong positive sample is easy to pollute other samples and lead to false positive^[33]. The cross-contamination ratio of samples can be

estimated by designing tests or adding internal reference sequences, so as to provide reference for the interpretation of subsequent results.

5.1.7 Species annotation

It is necessary to establish an annotation database of clinical pathogenic microorganisms, including relevant information of each microorganism, such as species taxonomy (bacteria, fungi, viruses, mycoplasma/chlamydia/spirochete/rickettsia, parasites, etc.), pathogenicity (pathogenic and conditional pathogenicity), planting information, clinical significance, etc. Species annotation can be of help to the subsequent report interpretation.

5.2 Report interpretation

Report interpretation refers to a process of eliminating interference and determining the infected pathogenic microorganisms. There are many types of clinical samples, and they differ in terms of colonization bacteria, microbial proportion and spatial heterogeneity, which leads to confusion in clinical interpretation. Clinicians need to determine whether the detected pathogenic microorganisms meet the clinical diagnosis according to the patient's medical history, infection indicators, imaging, clinical manifestations, complications, epidemiological history, contact history, etc., and combined with the quality control requirements of test data^{Error! Reference source not found.}. If conditions permit, cross-validation can be further carried out by other technologies. It should be noted that due to the different sequencing platforms, processes and depths, this part of the interpretation is of reference value to most situations, except those that do not meet special circumstances.

Currently, the identification of pathogenic microorganism follows Koch's postulates^{Error! Reference source not found.}: (1)The microorganism exists in patients with similar diseases, but not in healthy individuals; (2) It must be able to be separated, cultured and purified; (3) It can cause the same disease when inoculated in susceptible animals, and can be isolated from the inoculated animals; (4) It can cause disease in every individual^{Error! Reference source not found.}. However, TGS cannot completely meet the traditional Koch rule as a clinical test method. As a supplement to this rule, TGS's detection of pathogenic microorganisms has the following requirements for the interpretation of the results.

5.2.1 Data quality control

The detection range varies markedly for different application scenarios, that is, the detection target may contain different numbers of pathogenic microorganisms, drug resistance genes and virulence genes, so the indicators for evaluating sequencing quality automatically differ. For example, mNGS requires more data because of the large number of pathogenic microorganisms detected and the influence of host nucleic acids, while tNGS is limited to a small number of specific pathogenic microorganisms, and a small amount of data can meet the requirements. Before interpretation, it should be determined whether it meets the requirements of sample quality control according to the technical requirements of different products, such as data size, fragment length, sequencing quality and internal reference content. If any sample does not meet the quality control conditions, the testing steps and sample status should be traced back as soon as possible to determine the reasons and take further measures.

5.2.2 Positive thresholds and interpretation standards

Theoretically, all microorganisms in samples can be detected, but due to the differences in genome length and sample types of different microorganisms, it is impossible to establish an invariable standard for negative and positive interpretation of all microorganisms. The laboratory should establish and verify the positive thresholds and interpretation standards according to the expected use, sample types, detection targets and technical characteristics^{Error! Reference source not found.}. ROC curve analysis is a useful tool to determine the optimal threshold of the training set of clinical samples with known results, and an independent verification set is used to verify the preset threshold^[35]. Viruses rarely survive in the environment, so a small number of specific sequences can be detected as positive (such as <3 specific sequences)^[35]. The background bacteria can be monitored by NTC control samples to avoid reporting environmental bacteria, symbiotic bacteria and conditional pathogens that are not related to clinical practice. Generally, the higher the number of sequences, the greater the possibility of pathogenic microorganisms (dozens of specific sequences). Independent interpretation criteria can be adopted for pathogenic bacteria that are highly concerned in clinic and difficult to detect, such as *Mycobacterium tuberculosis*, *Yersinia pestis* and *Brucella*, that is, a specific sequence can be detected to be positive^[36]. Because the parasite genome is complex and similar to the human genome, it should be interpreted after the sequence specificity is strictly confirmed^{Error! Reference source not found.}. If the detected sequence marks a new species, it is not limited by the threshold value, but the homology comparison results need to be given.

In the analysis and interpretation of drug-resistant genes, the database of drug-resistant genes such as CARD^[38] can be considered, which shows the distribution of drug-resistant genes in different species. In addition, species inference should be made combining the abundance and comparison of pathogenic microorganisms, and further comparison verification or drug sensitivity test should be made when necessary. mNGS is greatly influenced by human nucleic acids with more pathogenic targets to be detected, so it often requires a higher amount of data to detect drug-resistant genes. However, the probability of detecting the corresponding drug-resistant genes increases when pathogenic microorganisms have higher abundance. tNGS is less affected by human sources, and specific primers or probes are designed to capture drug-resistant genes, so the amount of data needed is lower with higher probability of detecting drug-resistant genes. Because of the long reading length of the third generation nanopore sequencing, the alignment and localization of drug-resistant genes are more accurate than that of the second generation with short reading length, so it has higher accuracy in drug resistance detection.

5.2.3 Verification of test results

If the detected microorganisms conform to the characteristics of the disease, they may be pathogenic microorganisms that cause infection, but it cannot be confirmed just based on the number of sequences. The inspector should consider the coverage and specificity of sequences on the genome^{Error! Reference source not found.}, and investigate the cross-contamination of the same batch of samples with cross-contamination coefficient. In addition, PCR and other methods can be used to verify pathogenic microorganisms, such as influenza virus, adenovirus and Covid-19 in respiratory samples; *Shigella*, *Salmonella* and *Norovirus* in stool samples; *Enterovirus*, *herpes simplex virus* and *west Nile virus* in CSF; *Brucella*, *Bartonella* and *human immunodeficiency virus* in blood. If an unnamed microorganism with high sequence number is detected, we should be highly alert to the emergence of new species^{Error! Reference source not found.}.

5.2.4 Interpretation of microbial population detected from sterile site samples

A variety of pathogenic microorganisms can be co-detected in aseptic abscess samples, including brain abscess, cervical abscess, parapharyngeal abscess, oral abscess, etc. The number of microbial sequences detected may reach the positive thresholds, and most of them are strict anaerobic bacteria and facultative anaerobic bacteria. This phenomenon caused by microbial population is called that a population or a microbial ecosystem causes a disease, rather than a pathogen of Koch's postulates causing a disease. Although the theory of population pathogenicity needs further discussion, this phenomenon will be more verified in clinic with the wide application of deep sequencing technology^[39]. The detection of microbial populations (including different categories or species of the same kind) in a clinical sample (especially abscess) from a sterile site should not be easily regarded as pollution. If the existence of these microorganisms meets the clinical diagnosis, they should be fully covered with antibacterial drugs.

5.2.5 Verification of results of non-culturable or difficult-to-culture microorganisms

mNGS is used to detect the nucleic acid sequences of all pathogenic microorganisms in the sample, including non-culturable or difficult-to-culture microorganisms, which can not be reproduced by traditional culture methods, and there is also a lack of serological or antigen detection methods. Microorganisms besides viruses, spirochetes, rickettsia and parasites can be verified by nucleic acid detection methods, such as first-generation sequencing technology^{Error! Reference source not found.}.

5.2.6 Interpretation of pathogenic, colonized and contaminant microorganisms

The patient's immune status, basic diseases and sample sources should be considered upon the determination of the conditional pathogen as the pathogen. If there are a large number of background bacteria or miscellaneous bacteria sequences without dominant microorganisms, pollution should be considered first, and conditional pathogens second^{Error! Reference source not found.}. After surgery or other invasive operations, if the samples from sterile sites show a single bacteria and the sequence number may not be high, hospital infection should be considered based on clinical practice, and it should be distinguished from background bacteria, and pollution should not be blindly considered. In different sample types, the same sequence number may be interpreted in different ways^{Error! Reference source not found.}. Negative mNGS test is significant for eliminating infection, but it should also be interpreted correctly based on clinical manifestations^{Error! Reference source not found.}.

5.2.7 Treatment requirements of highly infectious microorganisms

For special pathogenic microorganisms with high infectivity, the laboratory should formulate special reporting procedures according to the regulations of the administrative department of health, and report the test results to the Center for Disease Control and Prevention (CDC). If there are suspected *Vibrio cholerae* O1 or O139, *Yersinia pestis*, Ebola virus or new pathogenic microorganisms, other methods, such as PCR and serology, should be adopted as soon as possible. It should be reported to the clinical and CDC systems immediately if the test results are consistent with the sequencing results^{Error! Reference source not found.}.

Recommendation 6: TGS improves the fidelity of the original sample sequence due

to its long reading length and PCR-free sequencing process. The accuracy of pathogen detection reflected by a single sequence is higher than that of the second generation sequencing with short reading length. Therefore, a small number of pathogenic microorganisms detected by TGS should also be collectively evaluated based on the patient's clinical manifestations, characteristics of pathogenic microorganisms, sample types and laboratory environment.

In conclusion, this consensus standardizes the application of nanopore sequencing in the detection of pathogenic microorganisms of infectious diseases, in the hope of promoting its clinical application and development. The current sequencing technology has undergone generations of updates with increasingly optimized and stable performance, but improving accuracy and reducing costs are still the main challenges. With the improvement of nanopore sequencing platforms and the reduction of sequencing costs, combined with the unique advantages of the platform itself, such as flexibility, portability and suitability for in-hospital detection, this technology is expected to be widely applied in the field of on-site pathogenic microorganism detection, clinical infection microorganism diagnosis, and real-time pathogenic microorganism monitoring.

6. Supplementary provisions

6.1 Declaration of interest

All consensus experts and drafting members have completed the declaration of interest form as required. All authors declare no conflict of interest.

6.2 Consensus updates

This consensus will be updated in 2-3 years.

6.3 Acknowledgement

Thanks to Mr. ZHANG Yu (Zhejiang Digena Diagnosis Technology Co., Ltd.) for verifying the technical routes and methods in this consensus and giving suggestions for revision.

The Writing Group of the *Consensuses of Experts on the Application of Nanopore Sequencing in the Detection of Pathogenic Microorganisms*

Writing Group Leaders:

ZHANG Shuyao, Chief Pharmacist, Guangzhou Red Cross Hospital Affiliated to Jinan University

HUANG Yihui, Deputy Chief Physician, Guangzhou Red Cross Hospital Affiliated to Jinan University

LI Xiaoyan, Chief Pharmacist, Department of Pharmacy, Sixth Affiliated Hospital of Sun Yat-sen University

ZHONG Shilong, Researcher, Department of Pharmacy, Guangdong Provincial People's Hospital

LI Aiguo, Chief Physician, Guangzhou Red Cross Hospital Affiliated to Jinan University

ZHENG Zhihua, Chief Pharmacist, Guangdong Pharmaceutical Association

HOU Tiewing, Chief Physician, 1.Hospital Office, Shenzhen Nanshan People's Hospital, Shenzhen, Guangdong 518052, China 2.Shenzhen University Medical School, Shenzhen, Guangdong 518073, China E-mail: sz_houtiewing@yeah.net

Expert Group Members (in alphabetical order):

CHEN Jie, Chief Pharmacist, Department of Pharmacy, the First Affiliated Hospital of Sun Yat-sen University

CHEN Yanfang, Chief Pharmacist, Guangzhou Eighth People's Hospital, Guangzhou Medical University

GAO Fenfei, Professor, Pharmacology Laboratory of Shantou University Medical College

GENG Yan, Chief Physician, Clinical Laboratory, The Second Affiliated Hospital of Xi'an Jiaotong University

HE Yanling, Chief Pharmacist, Department of Pharmacy, Jinshazhou Hospital of Guangzhou University of Chinese Medicine

HONG Wenxin, Chief Physician, Department of Infectious Diseases, Guangzhou Red Cross Hospital Affiliated to Jinan University

HOU Tiejing, 1.Hospital Office, Shenzhen Nanshan People's Hospital, Shenzhen, Guangdong 518052, China 2.Shenzhen University Medical School, Shenzhen, Guangdong 518073, China E-mail: sz_houtiejing@yeah.net

HUANG Bin, Chief Physician, Clinical Laboratory, The First Affiliated Hospital of Sun Yat-sen University.

HUANG Zuojun, Deputy Chief Pharmacist, Guangzhou Red Cross Hospital Affiliated to Jinan University

LI Wei, Chief Physician, Chief of Department of Neurology, Henan Provincial People's Hospital

LI Huiling, Chief Physician, Department of Tropical Medicine, Hainan Hospital of Chinese PLA General Hospital

LI Yanli, Chief Physician, Department of Respiratory Intervention, The First People's Hospital of Shangqiu

LI Yilei, Chief Pharmacist, Department of Pharmacy, Southern Hospital, Southern Medical University

LIU Jiayun, Chief Technician, Clinical Laboratory, The First Affiliated Hospital of Air Force Medical University of PLA

MA Yan, Chief Physician, Department of Pediatrics, Shandong Provincial Third Hospital

MEI Qinghua, Chief Pharmacist, Department of Pharmacy, Guangdong Second Provincial General Hospital

MEI Shiyue, Chief Physician, Institute of Pediatric Medicine, Henan Children's Hospital

PENG Liang, Chief Technician, Clinical Laboratory, The Fifth Affiliated Hospital of Guangzhou Medical University

PENG Jie, Chief Physician, Department of Infectious Diseases, Southern Hospital, Southern Medical University

QIU Kaifeng, Chief Pharmacist, Department of Pharmacy, Memorial Hospital of Sun Yat-sen University

WANG Xiaoqin, Chief Technician, Clinical Laboratory, The First Affiliated Hospital of Xi'an Jiaotong University

WU Shibo, Chief Physician, Department of Respiratory Medicine, Ningbo Medical Center Lihuili Hospital

WU Junyan, Chief Pharmacist, Department of Pharmacy, Memorial Hospital of Sun

Yat-sen University

YAN Haiyan, Chief Technician, Clinical Laboratory, Shenshan Medical Center, Memorial Hospital of Sun Yat-sen University

YAN Pengke, Chief Pharmacist, The Third Affiliated Hospital of Guangzhou Medical University

Yu Zhiqiang, Chief Technician, Clinical Laboratory, Dongguan People's Hospital, The Tenth Affiliated Hospital of Southern Medical University

ZHENG Zhang, Professor, Jinan University College of Pharmacy

ZHENG Lei, Chief Technician, Clinical Laboratory, Nanfang Hospital, Southern Medical University

ZHONG Honglan, Chief Pharmacist, Department of Pharmacy, Guangzhou Chest Hospital

ZHOU Hua, Chief Physician, Department of Respiratory and Critical Care, The First Affiliated Hospital of Zhejiang University School of Medicine

ZOU Shangrong, Chief Pharmacist, Guangzhou Eighth People's Hospital, Guangzhou Medical University

ZHU Youfeng, Chief Physician, The Affiliated Hospital, Ji'nan University

ZHANG Rui, Chief Physician, The Affiliated Hospital, Ji'nan University

References

[1] Liu Y, Xu XM. Progress in clinical detection based on nanopore sequencing technology[J]. *Chin J Lab Med*.2022, 45(3):4.DOI:10.3760/cma.j.cn114452-20210618-00383.

[2] Li J, Wang H, Mao L, et al. Rapid genomic characterization of SARS-CoV-2 viruses from clinical specimens using nanopore sequencing. *Sci Rep*. 2020;10(1):17492.

[3] Zhao N, Zhou N, Fan H, et al. Mutations and phylogenetic analyses of SARS-CoV-2 among imported COVID-19 from abroad in Nanjing, China. *Front Microbiol*. 2022;13:851323.

[4] Lin Q, Yao Y, Li X, et al. The application of nanopore targeted sequencing for pathogen diagnosis in bronchoalveolar lavage fluid of patients with pneumonia: a prospective multicenter study. *Infect Dis (Lond)*. 2024;56(2):128-137.

[5] Zhao M, Zhang Y, Chen L, et al. Nanopore sequencing of infectious fluid is a promising supplement for gold-standard culture in real-world clinical scenario. *Front Cell Infect Microbiol*. 2024;14:1330788.

[6] Clinical Microbiology Group of Laboratory Medicine Branch of Chinese Medical Association, Clinical Microbiology Group of Microbiology and Immunology Branch of Chinese Medical Association, Clinical Microbiology and Infection Branch of China Association for the Promotion of International Exchange of Health Care. China expert consensus on the application of metagenome high-throughput sequencing technique in pathogen detection of infectious diseases[J]. *Chin J Lab Med*, 2021,44(2):107-120. DOI:10.3760/cma.j.cn114452-20210206-00794.

[7] Editorial Committee of Chinese Journal of Infectious Diseases. Expert consensus on clinical application of China metagenomics second-generation sequencing technology to detect infectious pathogens[J]. *Chinese Journal of Infectious Diseases*,2020,38(11):681-689.

[8] High-throughput Sequencing Consensus Expert Group. Expert consensus on the application of high-throughput sequencing technique in the diagnosis of mycobacterial diseases. *Chin J Infect*, 2023,41(03):175-182.

[9] Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*. 2021 Nov;39(11):1348-1365.

[10] Moss EL, Maghini DG, Bhatt AS. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. *Nat Biotechnol*. 2020;38(6):701-707.

[11] Cretu Stancu M, van Roosmalen MJ, Renkens I, et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat Commun*. 2017;8(1):1326.

[12] <https://nanoporetech.com/platform/technology/advantages-of-nanopore-sequencing>

[13] Leger A, Amaral PP, Pandolfini L, et al. RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat Commun*. 2021;12(1):7198. Tourancheau A, Mead EA,

- [14] Zhang XS, Fang G. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat Methods*. 2021;18(5):491-498.
- [15] Hansen S, Dill V, Shalaby MA, et al. Serotyping of foot-and-mouth disease virus using oxford nanopore sequencing[J]. *J Virol Methods*, 2019, 263: 50-53.
- [16] Hoenen T. Sequencing of ebola virus genomes using nanopore technology. *Bio Protoc*. 2016;6(21):e1998.
- [17] WS/T 640-2018, Collection and transfer of clinical microbiological test samples[S].
- [18] Hospital Infection Control Branch of Chinese Preventive Medicine Association. Guidelines for collection and inspection of clinical microbiological specimens[J]. *Chinese Journal of Nosocomiology*, 2018,28(20):9.
- [19] Ma H, Yang XJ, Sui DW, et al. Biosafety management of pathogenic microorganism laboratory[J]. *Chinese Journal of Public Health Engineering*, 2010(2):2.DOI:CNKI:SUN:ZGWX.0.2010-02-032.
- [20] Laboratory Medicine Branch of Chinese Medical Association. Expert consensus on standardization of clinical application of quantitative macro genome sequencing technique to detect pathogenic microorganisms in qualcomm [J]. *Chin J Lab Med*, 2020,43(12):1181-1195.
- [21] <https://nanoporetech.com/support>
- [22] https://community.nanoporetech.com/technical_documents/chemistry-technical-document/v/chtd_500_v1_revae_07jul2016
- [23] Wick RR, Judd LM, Gorrie CL, et al. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom*. 2017;3(10):e000132. Published 2017 Sep 14.
- [24] Roberts LW, Enoch DA, Khokhar F, et al. Long-read sequencing reveals genomic diversity and associated plasmid movement of carbapenemase-producing bacteria in a UK hospital over 6 years. *Microb Genom*. 2023;9(7):mgen001048. doi:10.1099/mgen.0.001048
- [25] Petersen C, Sørensen T, Westphal KR, et al. High molecular weight DNA extraction methods lead to high quality filamentous ascomycete fungal genome assemblies using Oxford Nanopore sequencing. *Microb Genom*. 2022;8(4):000816. doi:10.1099/mgen.0.000816IF: 3.9 Q2
- [26] Wouter De Coster, Sven D'Hert, Darrin T Schultz, et al. visualizing and processing long-read sequencing data, *Bioinformatics*, Volume 34, Issue 15, August 2018, Pages 2666–2669, <https://doi.org/10.1093/bioinformatics/bty149>
- [27] Guo Y, Li Z, Li L, et al. A dual-process of targeted and unbiased Nanopore sequencing enables accurate and rapid diagnosis of lower respiratory infections. *EBioMedicine*. 2023;98:104858. doi:10.1016/j.ebiom.2023.104858
- [28] <https://github.com/rrwick/Filtlong>
- [29] Li, H. (2018). *Minimap2: pairwise alignment for nucleotide sequences*. *Bioinformatics*, 34:3094-3100.
- [30] Sichtig H, Minogue T, Yan Y, et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. *Nat Commun*. 2019;10(1):3313. Published 2019 Jul 25.
- [31] Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005;33(Database issue):D501-D504.
- [32] Lao HY, Wong LL, Hui Y, et al. The clinical utility of Nanopore 16S rRNA gene sequencing for direct bacterial identification in normally sterile body fluids. *Front Microbiol*. 2024;14:1324494. Published 2024 Jan 9. doi:10.3389/fmicb.2023.1324494
- [33] Xu Y, Lewandowski K, Lumley S, et al. Detection of viral pathogens with multiplex nanopore MinION sequencing: Be careful with cross-talk. *Front Microbiol*. 2018;9:2225. Published 2018 Sep 19. doi:10.3389/fmicb.2018.02225
- [34] Cao MK, Wang MX. New thinking of Koch's rule in modern medicine[J]. *Medicine and Philosophy*, 1991,(12): 8-10.
- [35] Miller, Steve et al. “Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid.” *Genome research* vol. 29,5 (2019): 831-842.
- [36] Fan S, Wang X, Hu Y, et al. Metagenomic next-generation sequencing of cerebrospinal fluid for the diagnosis of central nervous system infections: A Multicentre Prospective Study[J]. *bioRxiv*, 2019: 658047.
- [37] Expert Group on Clinical Application of Metagenomics Sequencing technique in Moderate and Severe Infections, Sepsis and Shock Committee of China Research Hospital Association, Chinese

Society for Microbiology Microbial Toxin Committee, et al. Expert consensus on clinical application of metagenomics sequencing technique in moderate and severe infections (1st Edition)[J]. Chin Crit Care Med, 2020,32(5):531-536.

[38] Alcock BP, Huynh W, Chalil R, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2023;51(D1):D690-D699. doi:10.1093/nar/gkac920

[39] Salzberg SL, Breitwieser FP, Kumar A, et al. Next-generation sequencing in neuropathologic diagnosis of infections of the nervous system. Neurol Neuroimmunol Neuroinflamm. 2016;3(4):e251. Published 2016 Jun 13.

[40] Miao Q, Hu BJ, Chen X, et al. An interpretation method based on the second-generation sequencing of pathogenic macro-genes, CN110335656A[P].2019.

[41] China Preventive Medicine Association. T/CPMA 010-2020 General criteria for pathogen screening based on high-throughput sequencing[Z].2020-07-01.