EBioMedicine 74 (2021) 103694

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Commentary

The multidimensional nature of metagenomics drives interdisciplinary diagnostics

Jutte J.C. de Vries

Department of Medical Microbiology, Leiden University Medical Center

ARTICLE INFO

Article History:
Received 28 October 2021
Accepted 28 October 2021

Introduction

Clinical metagenomics is slowly becoming part of today's clinician's armamentarium to identify infectious diseases [1–3]. The clinical application of metagenomics landmarks the revolutionary introduction of this pathogen-agnostic approach, enabling the detection of all pathogens including uncultivable, variant, rare, resistant, and previously undiscovered ones. The current most frequent application is for difficult to diagnose patients suspected of having an infectious syndrome but negative by routine testing [4, 5]. In the current issue of this journal, Guo and colleagues describe the application of metagenomics for combined microbial and oncogenic diagnostics [6]. Simultaneous detection of infectious and non-infectious diseases may be the next level of diagnostics to aim for. This commentary aims to describe the highlights of the manuscript, the context of current clinal practice, and to discuss the significance for the diagnostic field.

Efficient use of sequencing data drives evolution of the diagnostic landscape

Common grounds drive collaboration. The wet lab protocols for sequencing of pathogens and the human genome share common grounds: the mutual target format of double-stranded DNA genome is the subject of investigation of both human genetic and bacterial/viral DNA analyses. One of the previous studies providing evidence for the use of clinical genetic sequence data for the detection of viral pathogens is reported by Chesnais and colleagues [7]. The authors retrospectively analyze sequence data that was originally produced for non-invasive prenatal testing (NIPT), using a viral identification pipeline. The authors were able to detect viral sequences, including human cytomegalovirus, which is potentially pathogenic in utero. This proof-of-principle study accentuated the potential use in retrospective settings, and is food for thought of potential applications in prospective settings. The same principle of joint interest in sequence data applies for human exome sequencing and viral transcriptomics, both targeting poly-A tailed RNA sequences. Westermann and colleagues have combined profiling of both host and pathogen expression using one set of RNASeq data, referred to as "dual RNA-seq" or "dual transcriptomics" [8]. The design of so-called pan-genomic pipelines has been encouraged, to study for example virus-host interactions [9]. In short, pioneering studies as described above have provided evidence for the potential of combined microbial and human gene (expression) analysis, for research and potentially clinical applications. This combined approach may be referred to as "multidimensional metagenomics".

The clinical syndrome directs interdisciplinary diagnostics

Syndromic testing is referred to as the simultaneous testing of a broad panel of pathogens, typically by means of targeted molecular methods. Clinical metagenomics may be seen as the most comprehensive syndromic approach currently available. However, in clinical practice, the differential diagnosis can comprise of a combination of both infectious and non-infectious causes. For these patients, syndromic testing does not suffice and tailer-made diagnostics in the context of personalized medicine is required. For example, the differential diagnosis of encephalitis usually includes viral and inflammatory causes. Interdisciplinary collaboration by sharing of sequence data for multiple, discipline-specific analyses would be most beneficial in patients with overlapping clinical symptomatology or underlying etiological factors, such as microbiology and immunology, prenatal screening (NIPT) and microbiology, and microbiology and oncology.

Guo and colleagues report in the current issue the innovative and advanced use of metagenomic sequence data for human oncogenic analysis of lung tissue samples [6]. Both Illumina and Nanopore metagenomic sequencing were used and sequences were analyzed using an experimental pipeline for combined detection of DNA pathogens and cancer. Artificial intelligence (AI) assisted the design of the cancer pipeline and it is anticipated that machine learning will be increasingly used to optimize the performance of bioinformatic analyses. Illumina sequencing resulted in 83% sensitivity of diagnosing cancer based on experimental chromosomal instability

DOI of original article: http://dx.doi.org/10.1016/j.ebiom.2021.103639.
E-mail address: jjcdevries@lumc.nl

https://doi.org/10.1016/j.ebiom.2021.103694

(duplication/deletion) analysis when compared to histology results. Using the same sequence dataset, infections as defined by culture and other microbiological tests were detected with 78% sensitivity. The lowest sensitivity was for detection of aspergillus which may be largely explained by the non-molecular nature of the galactomannan antigen reference test. With regard to pipeline specificity, while commonly human sequences are removed from microbial pipelines to increase specificity, co-analysis of microbial and host sequences resulted in false positive microbial hits as complicating factor. Nanopore sequencing analysis of human sequences in this setting was constrained due to the low throughput format intrinsic to this technique, and resulted in low sensitivity (35%) for pathogen detection. Other publications using Nanopore sequencing for pathogen detection based on other sample types with lower host background have reported higher sensitivity, exemplifying the variety of dependencies and their impact on test performance [10]. Potential advantages of combined diagnostics would be reduced sampling, shorter turn-around-time when compared to culture, and lower sequencing costs.

While practical steps will need to be taken in order to improve the technical performance required for clinical applications, this progressive study further builds towards the scenario of interdisciplinary diagnostics based on joint use of metagenomic sequencing data.

Declaration of Competing Interest

None

Acknowledgements

None

References

[1] Wilson MR, Sample HA, Zorn KC, Arevalo S, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. N Engl J Med 2019;380:2327–40 10.1056/NEJMoa1803396.
[2] López-Labrador FX, Brown JR, Fischer N, Harvala H, et al. Recommendations for the introduction of metagenomic high-throughput sequencing in clinical virology, part I: Wet lab procedure. J Clin Virol 2021;134:104691 Jan10.1016/j.jcv.2020.104691. Epub 2020 Nov 18. PMID: 33278791.
[3] de Vries JJC, Brown JR, Couto N, Beer M, et al. Recommendations for the introduction of metagenomic next-generation sequencing in clinical virology, part II: bioinformatic analysis and reporting. J Clin Virol 2021;138:104812 May10.1016/j.jcv.2021.104812. Epub 2021 Mar 26. PMID: 33819811.
[4] Carbo EC, Blankenspoor I, Goeman JJ, Kroes ACM, et al. Viral metagenomic sequencing in the diagnosis of meningoencephalitis: a review of technical advances and diagnostic yield. Expert Rev Mol Diagn 2021;5 Oct10.1080/14737159.2021.1985467. Epub ahead of print. PMID: 34607520.
[5] Reyes A, Carbo EC, Harinxma Thoe Slooten JSV, Kraakman MEM, et al. Viral metagenomic sequencing in a cohort of international travellers returning with febrile illness. J Clin Virol 2021;143:104940 Oct10.1016/j.jcv.2021.104940. Epub 2021 Aug 3. PMID: 34416523.
[6] Guo Y, Li H, Chen H, et al. Metagenomics Next-Generation Sequencing to Identify Pathogens and Cancer from in Lung Biopsy Tissues. EBioMedicine 2021 10.1016/j.ebiom103639.
[7] Chesnais V, Ott A, Chaplais E, et al. Using massively parallel shotgun sequencing of maternal plasmatic cell-free DNA for cytomegalovirus DNA detection during pregnancy: a proof of concept study. Sci Rep 2018;8(4321) https://doi.org/10.1038/s41598-018-22414-6.
[8] Westermann A, Gorski S, Vogel J. Dual RNA-seq of pathogen and host. Nat Rev Microbiol 2012;10:618–30 https://doi.org/10.1038/nrmicro2852.
[9] Marschall. The Computational Pan-Genomics Consortium, Computational pan-genomics: status, promises and challenges. Brief Bioinform 2018;19(1) VolumeIssueJanuary118−1310.1093/bib/bbw089.
[10] Gu W, Deng X, Lee M, et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. Nat Med 2021;27(1):115–24 Jan10.1038/s41591-020-1105-z. Epub 2020 Nov 9. PMID: 33169017.