BMC Genomics

RESEARCH

Open Access

Optimizing next-generation sequencing efficiency in clinical settings: analysis of read length impact on cost and performance

Pedro Milet Meirelles[1,2*], Pablo Alessandro B Viana[1,7], Diogo Antonio Tschoeke[3,4], Laise de Moraes[5], Luciane Amorim Santos[5], Manoel Barral-Netto[5,6,7], Ricardo Khouri[5,6] and Pablo Ivan P Ramos[7]

Abstract

Background The expansion of sequencing technologies as a result of the response to the COVID-19 pandemic enabled pathogen (meta)genomics to be deployed as a routine component of surveillance in many countries. Scaling genomic surveillance, however, comes with associated costs in both equipment and sequencing reagents, which should be optimized. Here, we evaluate the cost efficiency and performance of different read lengths in identifying pathogens in metagenomic samples. We carefully evaluated performance metrics, costs, and time requirements relative to choices of 75, 150 and 300 base pairs (bp) read lengths in pathogen identification.

Results Our findings revealed that moving from 75 bp to 150 bp read length approximately doubles both the cost and sequencing time. Opting for 300 bp reads leads to approximately two- and three-fold increases, respectively, in cost and sequencing time compared to 75 bp reads. For viral pathogen detection, the sensitivity median ranged from 99% with 75 bp reads to 100% with 150–300 bp reads. However, bacterial pathogens detection was less effective with shorter reads: 87% with 75 bp, 95% with 150 bp, and 97% with 300 bp reads. These findings were consistent across different levels of taxa abundance. The precision of pathogen detection using shorter reads was comparable to that of longer reads across most viral and bacterial taxa.

Conclusions During disease outbreak situations, when swift responses are required for pathogen identification, we suggest prioritizing 75 bp read lengths, especially if detection of viral pathogens is aimed. This practical approach allows better use of resources, enabling the sequencing of more samples using streamlined workflows, while maintaining a reliable response capability.

Keywords Metagenomics, Health surveillance, Cost efficiency, Pathogen detection

*Correspondence:
Pedro Milet Meirelles
pmeirelles@ufba.br
[1] Institute of Biology, Federal University of Bahia (UFBA), Salvador, Bahia 41745-715, Brazil
[2] National Institute for Interdisciplinary Transdisciplinary Studies in Ecology and Evolution (IN-TREE), Salvador, Brazil
[3] Health Systems Engineering Laboratory, Alberto Luiz Coimbra Institute of Graduate Studies and Engineering Research (COPPE), Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
[4] Institute of Biology, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil
[5] Laboratory of Precision Medicine and Public Health (MESP 2), Gonçalo Moniz Institute, Oswaldo Cruz Foundation (Fiocruz), Salvador, Bahia, Brazil
[6] Federal University of Bahia School of Medicine, Salvador, Brazil
[7] Center for Data and Knowledge Integration for Health (CIDACS), Gonçalo Moniz Institute, Oswaldo Cruz Foundation (Fiocruz), Salvador, Bahia, Brazil

Background

COVID-19 served as a reminder of the limitations inherent in current disease surveillance systems, underscoring the need to harness data for informed decision-making [1, 2]. Genomics approaches have played key roles by identifying the emergence of new SARS-CoV-2 variants, including variants of interest (VOIs) and variants of concern (VOCs) that have displayed intricate dynamics of emergence and replacement after the identification of the initial Wuhan-Hu-1 strain in late 2019 [3, 4]. The expansion of sequencing technologies in many countries was accelerated by the need to track VOIs/VOCs, which represents a positive outcome of the pandemic. However, low/middle-income countries (LMICs) are limited in their ability to support sequencing capacity due to several factors, including the high cost of sequencing kits and consumables, the need for specialized laboratory personnel, and challenges ensuring comprehensive coverage (particularly for countries with large territories) [5].

One key aspect to evaluate is the trade-off between the sequencing depth and overall costs. One way to achieve increased depth is by sequencing longer read sizes. Longer reads can provide a more detailed view of genomic structures and facilitate more accurate pathogen identification and characterization. Nevertheless, it is imperative to benchmark the added costs of sequencing longer reads, which also includes extended sequencing durations and heightened computational resource demands, considering the potential advantages they offer in terms of improved precision and enhanced accuracy for pathogen identification. Reaching a commensurate balance between read length and sequencing cost is critical for the sustainable implementation of advanced genomic surveillance systems.

In this study, we investigated the cost efficiency and test performance of different Illumina read lengths, considering their impact on pathogen detection performance, cost, and sequencing time. Our analysis comprises the examination of accuracy, specificity, sensitivity, and precision (i.e., positive predictive value), across various simulated metagenomic datasets. Our focus was primarily on assessing the identification performance of specific pathogens, particularly those frequently associated with human outbreaks.

Methods

Mock metagenomes

To assess the performance of pathogen detection in metagenomes with varying read lengths, we analyzed several simulated metagenomes (i.e., mock metagenomes). Metagenomes were created using InSilicoSeq (version 2.0.1) [6]. Each composition was randomly generated based on predefined throat taxonomic profiles obtained from the Metagenomic Sequence Simulator (MeSS (https://github.com/metagenlab/MeSS), enriched with metadata information from all taxonomic levels using TaxonKit (version 0.17.0) [7]. Information on pathogenic taxa was included from CZID (https://czid.org/pathogen_list), Illumina Respiratory Pathogen ID/AMR Enrichment Panel (RPIP) kit (https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/respiratory-pathogen-id-panel.html) and Viral Surveillance Pathogen (VSP) targets (https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/viral-surveillance-panel.html). The mock metagenomes were generated with sequencing errors that mimic those typically introduced by DNA sequencing platforms, with reads of 75, 150, and 300 base pairs (bp) in length. A total of 48 distinct mock compositions were created, resulting in 144 synthetic metagenomes (Table S1). These synthetic datasets included 489 unique taxa, comprising 34 viral pathogens and 183 bacterial pathogens.

We generated mock metagenomes with and without related lineages to test the hypothesis that closely related species would influence the ability to classify/identify taxa. When defining the composition of closely related samples, we randomly selected species from the same genus or from different genera within the same family. To be more specific in the problem we are looking at, we also focus on mocks with and without pathogens following the same logic (see Supplementary Information for more details).

Mock metagenomes annotation

Our initial analysis started with quality control aimed at evaluating sequencing quality and filtering out individual reads. We applied a Phred quality score threshold of 20, a minimum read length requirement of 50, and a maximum allowable number of N's set at 2, all performed with the fastp software (version 0.20.1) [8].

Kraken2 (version 2.1.2) was used for taxonomic identification [9]. This tool relies on k-mer profiles and employs the Lowest Common Ancestor (LCA) algorithm for precise classification. At this step, we used the standard plus PFP Kraken2 database provided by the developers of this tool (available at https://benlangmead.github.io/aws-indexes/k2, updated on Jun 5th, 2024), which contains a diverse range of genomes from the NCBI RefSeq, including archaea, bacteria, viruses, plasmids, humans, protozoa, fungi, and plants.

Performance metrics

In our assessment of pathogen detection performance at various read lengths, we examined the degree of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) of the taxonomic annotation.

These metrics were analyzed at the read-level separately for each taxon in the metagenome composition based on their identification accuracy. Reads correctly identified as belonging to the target taxon were deemed 'true positives', while those that failed to be recognized as such, despite originating from the specified taxon, were categorized as 'false negatives'. Conversely, 'false positives' refer to reads erroneously assigned to this taxon when they should not have been. 'True negatives' were defined as reads accurately not classified as belonging to the taxon in question.

This classification schema was uniformly applied across all taxa present in the metagenome. Using these classifications, we constructed confusion matrices for each taxon, enabling the calculation of key performance metrics, including Sensitivity, Specificity, Accuracy, and Precision (Table S2) [10].

Statistical analysis

To examine whether there were variations in the overall pathogen detection performance across different read sizes and mock sample compositions, we employed the Friedman test, followed by pairwise comparisons using the Nemenyi-Wilcoxon-Wilcox all-pairs test for a two-way balanced complete block design. To test if the sensitivity was correlated with taxa abundance, we performed the Spearman correlation test. For all analyses, we considered significant differences when p-values were < 0.05. We conducted all statistical analysis and data visualization using the R software (version 4.3.0) with the package *rstatix* (version 0.7.2) and package *PMCMRplus* (version 1.9.8).

Results

We focused on evaluating the performance of different read lengths (75, 150, and 300 bp) in terms of their ability to correctly identify metagenomic samples containing both viral and bacterial pathogenic reads (simulating infected individuals), while also evaluating the robustness of the compared strategies to correctly identify samples in which no pathogenic reads were present.

For viral pathogens, when using 75 bp read length, the sensitivity median was 99%. This increased to 100% when using 150 bp and 300 bp reads (Fig. 1A). For bacterial pathogens, the sensitivity medians per read length (75 bp, 150 bp and 300 bp) were as follows: 87%, 95%, and 97% (Fig. 1C). When considering both bacterial and viral taxa, regardless of pathogenicity, sensitivity medians observed were as follows: 94% for a 75 bp read length, 97% for a 150 bp read length, and 98% for a 300 bp read length (Figure S1). While longer read lengths demonstrated improved sensitivity, the findings suggest that even with reduced read lengths (75 bp), sensitivity resulted in statistically similar performance metrics (compared to 150 bp and 300 bp reads) for many taxa (Figure S1), particularly for viral pathogens. We also found that sensitivity was not correlated with specific taxa abundance (Spearman correlation; p-value < 0.001 and $\rho = 0.043$; Figure S2).

Precision, which measures the accuracy of true positive predictions, remained consistently high across all read lengths, even with the shorter 75 bp reads. For viral pathogens, precision medians were measured at 100% for all read lengths (Fig. 1B). For bacterial pathogens, precision medians were as follows: 75 bp: 99.7%, 150 bp: 99.8%, and 300 bp: 99.7% (Fig. 1D). Considering all taxa, precision medians were measured at approximately 100% for all read lengths (Figure S2). These findings demonstrate that positive predictions remain highly accurate even when shorter read lengths are used.

Our analysis also assessed specificity and accuracy, which are critical metrics in pathogen identification. Specificity medians were 100% for all taxa and read lengths, indicating the ability to correctly identify samples in which pathogenic taxa were absent, i.e., true negatives (Figure S2). Similarly, accuracy medians exceeded 99.8% for all taxa and read lengths.

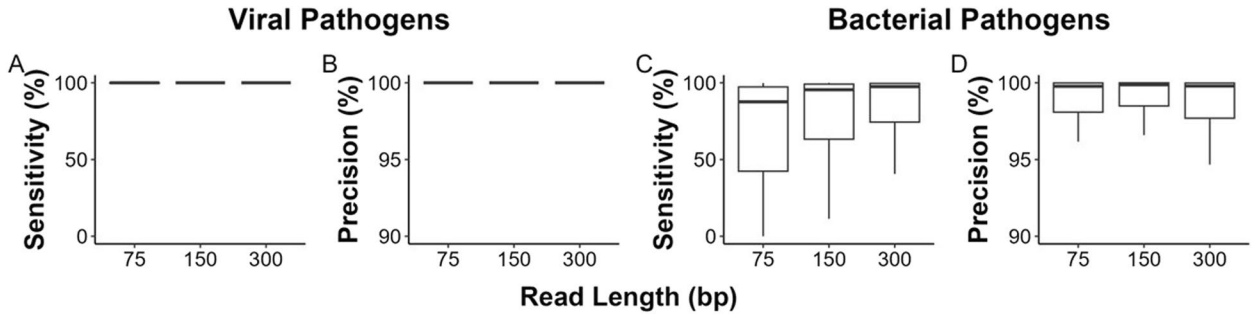

Fig. 1 – Sensitivity and precision for viral (**A** and **B**) and bacterial (**C** and **D**) pathogens species identification for 75, 150, and 300 bp read lengths. Each boxplot represents the distribution of these metrics, and the black horizontal line within each box indicates the median value for the respective metric at each read length

We also tested whether the metric values significantly varied among different read lengths. For viral pathogens, significant differences in accuracy and sensitivity were observed (Table S3), with the predominant differences observed between 75 bp read lengths compared to the other read lengths (Table S4). In the case of bacterial pathogens and the overall analysis of all taxa, our investigation revealed statistically significant differences across all metrics (accuracy, specificity, sensitivity, and precision) (Table S5 and S6). Pairwise comparisons between read lengths also demonstrated significant differences, except for specificity and precision on bacterial pathogens when comparing 75 bp to 150 bp (Table S7), as well as precision bacterial pathogens and all taxa when contrasting 75 bp to 300 bp (Tables S7 and S8).

Among the taxa examined, a subset demonstrated exceptional identification performance across all metrics. Notably, these taxa consistently achieved minimum values of accuracy exceeding 99%, specificity surpassing 99%, sensitivity exceeding 98%, and precision exceeding 92% across all read lengths. Among these taxa are several relevant respiratory viruses, such as rhinovirus subtypes A, B, and C, SARS-CoV-2, and human mastadenovirus subtypes B, C, and E (Fig. 2). These findings are particularly encouraging, highlighting the robustness of pathogen identification for these specific viral pathogens across varying read lengths.

Certain viruses presented challenges in achieving satisfactory classification metrics (>50% performance) irrespective of read lengths. Human rhinovirus sp., Parechovirus A, and Enterovirus subtypes A and B were among the taxa that did not fare well in the performance metrics analysis. Of these, Parechovirus A and Enterovirus A showed some improvements in sensitivity when larger 300 bp reads were applied. The sensitivity for Parechovirus A increased from 5.4% at 75 bp to 25.4% at 300 bp. Similarly, the classification of Enterovirus A was improved from 37.5% at 75 bp to 46.6% when 300 bp reads were employed. However, despite these improvements, sensitivity rates for all these viruses remained consistently below 50% across all read lengths. We highlight that sensitivity remained constant to the same taxon, for both easy and hard-to-find taxa, among different types of samples, demonstrating robustness in pathogen detection regardless of the underlying sample taxonomic profile (Fig. 3 and S3).

In our pursuit of assessing the costs of various Illumina read lengths within the context of pathogen identification, we have chosen to focus on sequencing costs, which constitute ~ 23 to 70% of the total wet-lab-related project cost (Table S9). This decision is driven by the immediate and substantial economic impact that optimizing read lengths can have on large-scale projects, particularly in the context of public health surveillance initiatives. Of note, the fixed costs, including sample

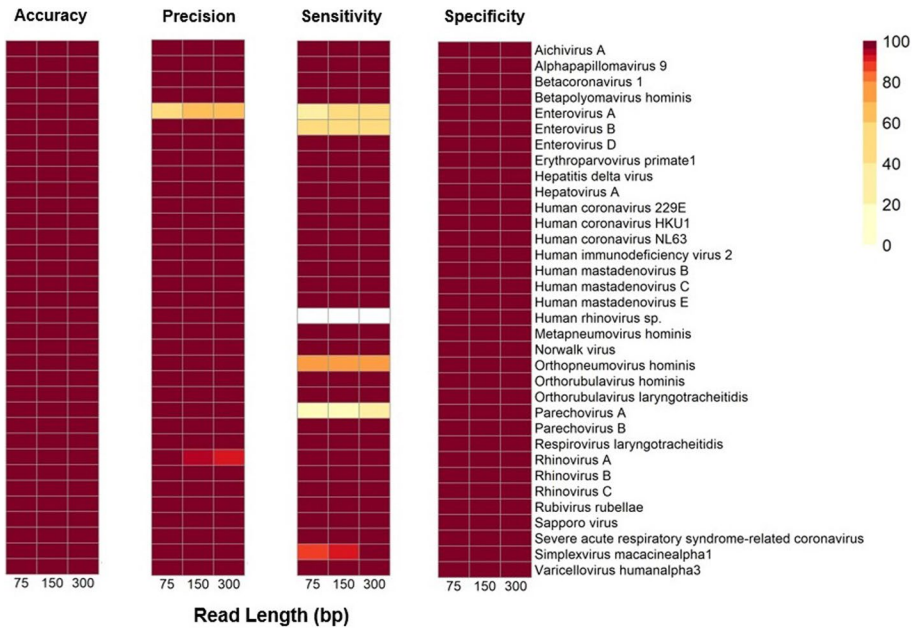

Fig. 2 – Accuracy, precision, sensitivity, and specificity for viral pathogen species identification for 75, 150, and 300 bp read lengths. Each row in the heatmap corresponds to a specific viral pathogen species. At the same time, the color gradient indicates the mean percentage value of each metric, calculated for the taxon considering the samples in which it was present. Dark red color represents higher values and white represents lower values

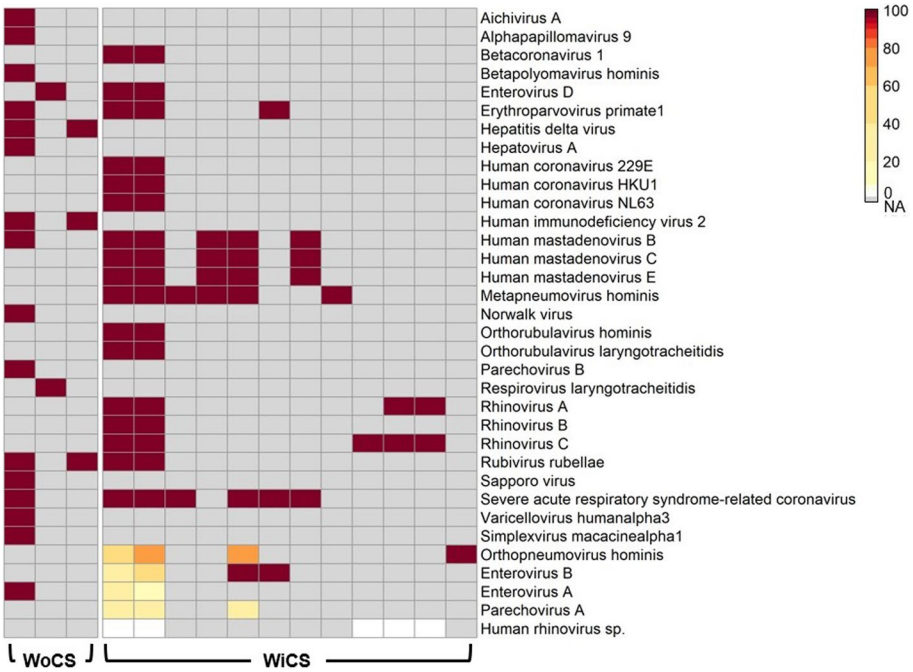

Fig. 3 – Sensitivity for viral pathogens. We categorized samples as WoCS (Without Close Strain) and WiCS (With Close Strain). In this plot we report the results of 300 bp read length. Each column represents a sample, the gray color indicates the absence of the taxon, the color gradient indicates the sensitivity for the taxon, dark red represents higher values and white represents lower values. The WoCS and WiCS groups were composed of the mock metagenomes from PNPT-WoCS plus NPT-WoCS and PNPT-WiCS plus NPT-WiCS, respectively. We show only the samples whose composition has at least one virus from the pathogenic list, so the direct sum of the number of samples of these groups may not be the same as the WoCS and WiCS samples in the plot

extraction and library construction for shotgun protocol (Table S10), were held constant across different read lengths. As a model, we used the Illumina MiSeq instrument, a benchtop sequencer widely adopted in many molecular biology and clinical diagnostic laboratories because of its flexibility of use and cost accessibility. Table S9 details the sequencing features and scenario characteristics of the experiment associated with each read length. Financial costs were a primary consideration, revealing distinctive expense levels for each read length. We performed all cost estimations in Brazilian Real (BRL), applying the closing exchange rate from the Central Bank of Brazil on October 5, 2023, where R$ 1.00 equals US$ 0.19. Although we recognize that these costs may vary significantly across different countries, the proportional relationships would remain consistent among the different MiSeq sequencing run kits used. At a multiplexing level of 25 samples per run (i.e., 2 million reads per sample), the cost of paired 75 bp sequencing using a V3-150 cycle kit is $67.27 per sample, while the cost of paired 150–300 bp sequencing using a V3-600 cycle kit is $113.56 per sample. These costs are considerably lower than using the V2-300 cycle kit for paired 75–150 bp sequencing, which would cost $129.33 per sample.

Additionally, our approach considered sequencing run time, which increases with increasing read length, and can impact the turn-around time of outbreak investigations and the efficiency of large-scale projects. A 75 bp read length required a time investment up to 21 h (exclusively considering the sequencing time), while the 150 bp read length extended this up to 36 h. The 300 bp read length, chosen for its greater depth of sequence coverage, required the most significant time investment, with sequencing spanning up to 66 h.

To gain a comprehensive perspective, we conducted a comparison relative to the 75 bp read length, serving as the analysis baseline (1x). Transitioning to a 150 bp read length resulted in a time requirement approximately 1.7 times greater than that of the baseline, while the cost also increased approximately 1.7 times in comparison. The move to 300 bp read lengths further escalated both the time expenditure and monetary costs. Sequencing with a 300 bp read length required approximately 3.1 times the time investment of the 75 bp baseline, with approximately 1.7 times increase in costs (Figure S4).

Discussion

In this study, we evaluated the impact of different Illumina read lengths on metagenomic pathogen identification using in silico mock metagenomes. We discovered that while longer read lengths enhanced sensitivity, short 75 bp reads maintained similar levels of precision, specificity, and accuracy. Notably, this approach demonstrates the feasibility of metagenomic sequencing for pathogen identification, even when resource constraints limit the use of NGS to layouts with shorter read lengths.

Our findings echo prior research highlighting the trade-offs between read lengths and sensitivity in metagenomic pathogen identification, particularly for viral pathogen detection [11–14]. An intriguing result from our analysis is the lack of a clear correlation between the abundance of reads from a specific taxon and its identification sensitivity. This suggests that the identification sensitivity of certain taxa goes beyond read abundance, underscoring the intricate dynamics of pathogen identification within metagenomic datasets depending on its genome composition or availability of genomes in reference databases [9]. The reasons may rise from the variety of strains for a given species, the similarity of the genome with close species, the complexity of the genome, and the genome size, among others. In addition, while some taxa exhibit outstanding accuracy and sensitivity, others pose unique challenges, emphasizing the need for ongoing optimization in metagenomic sequencing techniques and analysis [15, 16]. However, it's important to acknowledge limitations, such as our reliance on in silico mock shotgun metagenomes, which may not fully replicate real-world complexities. Additionally, our findings are contingent on reference database completeness and accuracy, and limitations in these databases can impact results. It is, also, worth noting that the application of targeted sequencing approaches (tNGS), i.e., RPIP, may further enhance performance metrics, potentially impacting the costs of read lengths in relation to pathogen identification performance [14, 17].

The implications of our findings are particularly relevant for large-scale projects, enabling optimization in resource allocation while maintaining a rapid response capacity to infectious disease outbreaks, crucial for public health policy. Our results suggest the judicious use of shorter read lengths, such as 75 bp, for metagenomic sequencing. This strategy ensures similar performance metrics, significantly reduces costs, and expedites results acquisition. This strategic use of resources allows large-scale projects to optimize their workflows, effectively allocate their budgets, and maintain a swift response capacity in the face of infectious disease outbreaks.

Conclusion

In conclusion, our study highlights the advantages of prioritizing 75 bp read lengths during disease outbreaks, where rapid pathogen identification is critical. Shorter reads significantly reduce sequencing time by approximately threefold and lower costs by approximately twofold compared to longer read lengths. Despite their brevity, 75 bp reads demonstrate comparable precision across most viral and bacterial taxa, although sensitivity may vary, particularly in bacterial identification. This pragmatic approach optimizes resource allocation, allowing for increased sample throughput and streamlined workflows without compromising response reliability.

Moving forward, future research should continue to explore cost-effective strategies, including targeted next-generation sequencing approaches, to enhance pathogen detection capabilities. Efforts to overcome proprietary constraints on innovative technologies, such as Illumina's RPIP panel, will be essential for maximizing the broader scientific community's access to advanced diagnostic tools and ensuring rapid and effective pathogen surveillance and response.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10778-1.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

We thank the fruitful scientific discussions held within the Rede Genômica Fiocruz group. This work is part of the Alert-Early System of Outbreaks with Pandemic Potential (ÆSOP; http://aesop.health), an initiative under development by Brazil's Oswaldo Cruz Foundation (Fiocruz) and the Federal University of Rio de Janeiro (UFRJ). The initial concept development of ÆSOP was supported by Fundação Oswaldo Cruz (Fiocruz), and implementation of the system is financially supported by the Rockefeller Foundation (grant 2023-PPI-007 awarded to MB-N). MBN, PIPR, RK are research fellows for Brazil's National Council for Scientific and Technological Development (CNPq). The authors disclose that the manuscript was entirely written by humans. OpenAI's GPT-3.5 was used to perform grammar/language corrections to improve clarity and readability of the initial drafts, followed by revisions/adjustments/corrections performed by the authors.

Authors' contributions

Conceptualization: PMM, RK and PIPR. Data curation, bioinformatics and statistical analysis: PMM, DAT, PABV and LM. Resources: MBN. Writing—original draft: PMM and PIPR. Writing—review & editing: MBN, RK, LA, DAT, PABV and LM. All authors have read and agreed to the published version of the manuscript.

Authors' information

Not applicable.

Funding

Grant 2023-PPI-007 awarded to MB-N by Rockefeller Foundation.

Availability of data and materials
Data is provided within the manuscript or supplementary information files. All codes used are available on GitHub at https://github.com/cidacslab/aesop-metegenomics-read-length.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 16 February 2024 Accepted: 4 September 2024
Published online: 12 September 2024

References

1. Morgan OW, Aguilera X, Ammon A, Amuasi J, Fall IS, Frieden T, et al. Disease surveillance for the COVID-19 era: time for bold changes. Lancet. 2021;397:2317–9.
2. Al Knawy B, Adil M, Crooks G, Rhee K, Bates D, Jokhdar H, et al. The Riyadh Declaration: the role of digital health in fighting pandemics. Lancet. 2020;396:1537–9.
3. Biswas N, Mallick P, Maity SK, Bhowmik D, Mitra AG, Saha S, et al. Genomic surveillance and phylodynamic analyses reveal the emergence of novel mutations and co-mutation patterns within SARS-CoV-2 variants prevalent in India. Front Microbiol. 2021;12:703933.
4. Ghosh N, Nandi S, Saha I. A review on evolution of emerging SARS-CoV-2 variants based on spike glycoprotein. Int Immunopharmacol. 2022;105:108565.
5. Ladner JT, Sahl JW. Towards a post-pandemic future for global pathogen genome sequencing. PLoS Biol. 2023;21:e3002225.
6. Gourlé H, Karlsson-Lindsjö O, Hayer J, Bongcam-Rudloff E. Simulating Illumina metagenomic data with InSilicoSeq. Bioinformatics. 2019;35:521–2.
7. Shen W, Ren H. TaxonKit: a practical and efficient NCBI taxonomy toolkit. J Genet Genomics. 2021;48:844–50.
8. Chen S, Zhou Y, Chen Y, Gu J, fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20:1–13.
9. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20:1–13.
10. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K et al. Foundational statistical principles in Medical Research: sensitivity, specificity, positive predictive value, and negative predictive value. Med (Kaunas). 2021;57(5):503.
11. Yang S, Johnson MA, Hansen MA, Bush E, Li S, Vinatzer BA. Metagenomic sequencing for detection and identification of the boxwood blight pathogen Calonectria pseudonaviculata. Sci Rep. 2022;12:1–14.
12. Takeuchi S, Kawada J, ichi, Horiba K, Okuno Y, Okumura T, Suzuki T, et al. Metagenomic analysis using next-generation sequencing of pathogens in bronchoalveolar lavage fluid from pediatric patients with respiratory failure. Sci Rep 2019. 2019;9(1):9:1–11.
13. Pichler I, Schmutz S, Ziltener G, Zaheri M, Kufner V, Trkola A, et al. Rapid and sensitive single-sample viral metagenomics using Nanopore Flongle sequencing. J Virol Methods. 2023;320:114784.
14. Zhang D, Zhang J, Du J, Zhou Y, Wu P, Liu Z, et al. Optimized sequencing adaptors enable Rapid and Real-Time metagenomic identification of pathogens during runtime of sequencing. Clin Chem. 2022;68:826–36.
15. de Vries JJC, Brown JR, Fischer N, Sidorov IA, Morfopoulou S, Huang J et al. Benchmark of thirteen bioinformatic pipelines for metagenomic virus diagnostics using datasets from clinical samples. J Clin Virol. 2021;141:104908.
16. Waite DW, Liefting L, Delmiglio C, Chernyavtseva A, Ha HJ, Thompson JR. Development and validation of a bioinformatic workflow for the Rapid Detection of Viruses in Biosecurity. Viruses. 2022;14(10):2163.
17. Lin R, Xing Z, Liu X, Chai Q, Xin Z, Huang M, et al. Performance of targeted next-generation sequencing in the detection of respiratory pathogens and antimicrobial resistance genes for children. J Med Microbiol. 2023;72:001771.

Publisher's Note