Check for updates

Pathogen Detection Using Metagenomic Next-Generation Sequencing of Plasma Samples from Patients with Sepsis in Uganda

Brian S. Grundy,[a] Hardik Parikh,[a] Shevin Jacob,[b] Patrick Banura,[c] Chris C. Moore,[a] Jie Liu,[d] Eric R. Houpt[a]

aDivision of Infectious Diseases and International Health, University of Virginia, Charlottesville, Virginia, USA
bLiverpool School of Tropical Medicine, Liverpool, United Kingdom
cMinistry of Health, National Disease Control Department, Kampala, Uganda
dSchool of Public Health, Qingdao University, Qingdao, China

ABSTRACT Metagenomic sequencing is a promising new method for pathogen detection. We aimed to detect pathogens from archived plasma using metagenomic sequencing in a previously well-characterized cohort of 254 predominantly HIV-infected patients with sepsis in Uganda. We used Illumina sequencing and the Chan Zuckerberg ID metagenomics platform to sequence and identify pathogens. On average, each plasma sample yielded $3,404,737 \pm 2,201,997$ reads (mean $\pm$ standard deviation), of which $220,032 \pm 416,691$ ($6.3\% \pm 8.6\%$) were identified as nonhuman reads. Using a background model filter, 414 genus-specific pathogen identifications were found in the 254 samples. Nineteen pathogens were previously detected positive by quantitative PCR (qPCR), compared to sequencing, which demonstrated 30.2% sensitivity and 99.5% specificity. Sensitivity was higher for viral pathogens than nonviral pathogens (37% versus 5%). For example, HIV viremia was detected in 69% of samples using qPCR, and sequencing revealed 70% sensitivity and 92% specificity. There were 75 genus-specific potential pathogens identified by sequencing in this cohort, including hepatitis B and Epstein-Barr virus (EBV), among several others. qPCR showed a prevalence of hepatitis B and EBV viremia of 17% and 45%, respectively. In-hospital mortality was associated with a lower qPCR threshold cycle value for EBV (adjusted odds ratio, 0.85; $P < .001$) but not for hepatitis B or HIV. In conclusion, a broad range of potential pathogens were identified by metagenomic sequencing in patients with sepsis in Uganda. Unexpectedly high rates of hepatitis B and EBV viremia were found. Whether these viral infections in HIV patients with sepsis are clinically important requires further study.

IMPORTANCE The use of next-generation sequencing (NGS) in blood samples is an emerging technology for clinical microbiology labs. In this work, we performed NGS on plasma samples from a well-characterized cohort, where all samples had been previously tested by PCR for 43 pathogens. Therefore, we could compare sequencing performance against that of PCR and identify clinical correlates. A broad range of potential pathogens were identified by metagenomic sequencing in patients with sepsis in Uganda, particularly viruses, which we confirmed by PCR. In addition to HIV viremia, unexpectedly high rates of hepatitis B and EBV viremia were found, which may have important clinical implications.

KEYWORDS human immunodeficiency virus, metagenomics, sepsis, sub-Saharan Africa

The etiology of sepsis in HIV patients in resource-limited settings is often unknown (1–4). A better understanding of the causative organisms of sepsis can help guide clinicians to select appropriate antimicrobial therapy and avoid the use of unnecessary antibiotics.

Many potential bloodstream infections are often not identified in routine blood cultures (5). For example, there is a high rate of *Mycobacterium tuberculosis* bloodstream

infections in HIV and sepsis (6, 7), but mycobacterial cultures are required for diagnosis. In addition, the diagnosis of several pathogens, including *Leptospira*, *Coxiella*, *Rickettsia*, and *Brucella*, typically requires serologic testing (8, 9). Many viruses can also cause a sepsis-like picture, and there are several latent viruses that can have additional management implications (10–12). *Plasmodium* and other parasites can also present with severe illness.

A comprehensive diagnostic evaluation for these pathogens, including bacterial cultures, serologies, and antigen testing, is costly and complex. Our previous work using quantitative PCR (qPCR) to detect bacteria, viruses, fungi, and protozoa showed good sensitivity and specificity for pathogen detection from human blood; however, there still remained a large number of specimens with no potential pathogen identified (13, 14). Using a qPCR TaqMan Array card (ThermoFisher), we previously detected a potential pathogen in 72% of a cohort of mostly HIV-infected inpatients with sepsis from Uganda (14). Therefore, many patients still had no pathogens detected, plus not all pathogens detected were causal of sepsis.

Therefore, in this study we aimed to use metagenomic next-generation sequencing (mNGS) for unbiased detection of potential pathogens, including bacteria, fungi, viruses, and parasites, from the available paired plasma samples from this well-characterized cohort. We compared the results with prior qPCR results for whole blood and clinical data.

RESULTS

There were 671 participants in the full cohort, 336 of whom were previously tested by qPCR and 254 for which we had available plasma. Of the 254 participants, the median (interquartile range [IQR]) age was 35 (28 to 40) years, and 110 (43%) were female. The median (IQR) CD4$^+$ count was 60 (13 to 169) cells/μL, the median (IQR) lactate was 3.9 (3.0 to 4.9) mmol/liter, and 65 (27%) died in hospital. The clinical and laboratory data from the 254 plasma samples were not significantly different from the previous qPCR study and the total enrolled cohort (14, 15).

The plasma samples from Uganda sepsis patients, 4 healthy American plasma samples, and 5 water controls underwent extraction, indexing, pooling, and sequencing. Excluding controls, there was a total of 878,422,177 150-bp paired-end sequencing reads. On average (mean $\pm$ SD), each plasma sample had 3,404,737 $\pm$ 2,201,997 reads, of which 220,032 $\pm$ 416,691 reads (6.3% $\pm$ 8.6%) were identified as nonhuman reads. Figure S1 in the supplemental material shows the steps in sequencing and analysis. In total, 414 genus-specific pathogen (295 viral and 119 nonviral) identifications were made in the 254 samples, of which 75 unique genus-specific pathogens (16 viral and 59 nonviral) met threshold criteria (i.e., Z-score of >1 above background and >10 reads per million, for both nucleic acids and protein alignment). A total of 207 (81%) samples had at least one pathogen detected, and 139 (55%) samples had at least 2 pathogens detected. The median (IQR) number of pathogens per sample was 2 (1 to 3). Figure 1 highlights the viral pathogens identified in each sample and their relative abundances in number of reads per million. Figure S2A and B show bacterial and eukaryotic pathogens identified, respectively.

For sequencing detections, we evaluated the breadth of sequencing coverage against each pathogen reference NCBI genome. For different pathogens, the percentages of genome coverage were highly variable. Viral detections covered an average of 37% $\pm$ 33% of the reference genomes, compared to 16% $\pm$ 27% for nonviral pathogens (Fig. 1 and Fig. S2A and B).

Of these genus-level pathogen detections by sequencing, qPCR had been previously performed for 16 pathogens on paired whole-blood samples. For this work, we also performed qPCR for HIV, Epstein-Barr virus (EBV), and hepatitis B virus on the same plasma samples tested by mNGS. Sequencing demonstrated 30.2% sensitivity and 99.5% specificity compared with targeted qPCR (Table 1). Sensitivity of sequencing was higher for viral pathogens than nonviral pathogens (average 37% versus 5%; $P < 0.001$). Detection of HIV using sequencing had a sensitivity and specificity of 70% and 92%, respectively. We
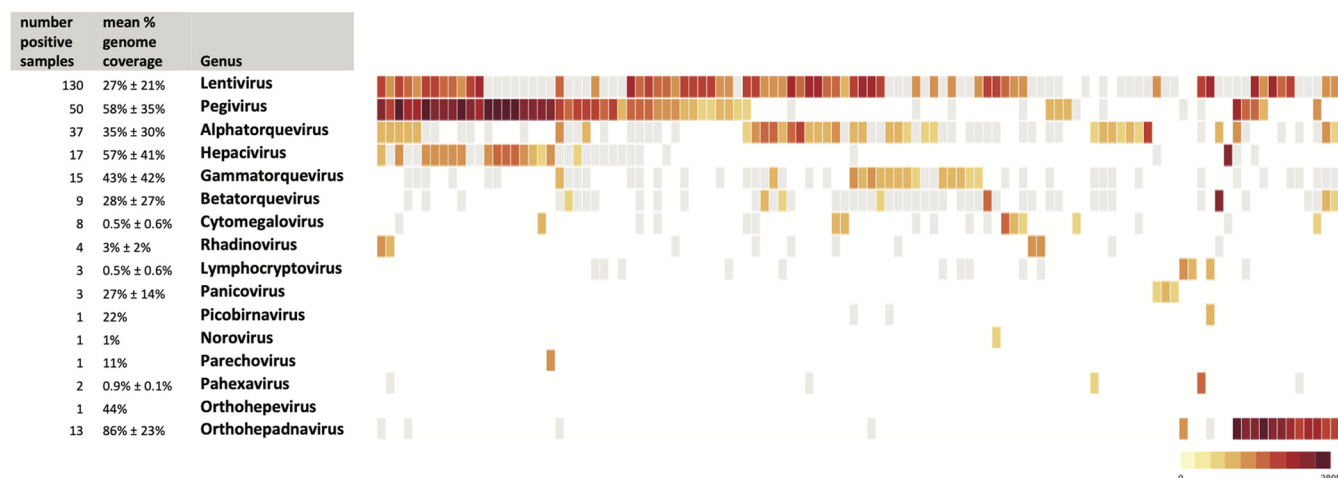
FIG 1 Heatmap of virus genus-specific pathogens identified in plasma samples from patients with sepsis in Uganda. Each column represents a plasma sample from a patient with sepsis in Uganda. Each row represents a virus genus-specific pathogen identified. Colored logarithmic scale represents reads per million (rPM), with darker red representing the highest rPM. Gray boxes represent reads identified for that genus in the sample but not meeting filter threshold criteria. Samples with no viral detection were excluded (n = 77), and samples with *Lentivirus* (HIV) as the only virus detected were excluded (n = 68).

performed receiver operating curve (ROC) analysis comparing mNGS reads per million normalized to the genome size for all pathogens tested by qPCR. For viral pathogens, the optimal cutoff was 2.4 reads per million per genome size (megabases). This would yield an overall sensitivity and specificity of 57% and 94% and an area under the curve (AUC) of 0.763. For nonviral pathogens, an optimal cutoff could not be identified, given the low AUC of 0.48 (Fig. S3).

Sequencing detected *Lentivirus* (HIV) in 130 samples, *Orthohepadnavirus* (hepatitis B virus) in 14 samples, *Lymphocryptovirus* (Epstein-Barr virus) in 3 samples, and *Cytomegalovirus* in

TABLE 1 Comparison of qPCR-positive whole-blood samples to sequencing results from plasma samples

Pathogen[a]	No. of samples (mean C_T, range) for samples that were positive by qPCR and:		No. of samples negative by qPCR and:		Sequencing vs qPCR sensitivity/specificity
	mNGS positive	mNGS negative	mNGS positive	mNGS negative	
Viral pathogens					
CMV	7 (26.8, 21–31.4)	59 (32.3, 20.9–37)	1	187	11%/99%
Dengue virus	0	5	0	249	0%/100%
EBV*	3 (21, 17.9–23.3)	107 (35.4, 28.6–39.8)	0	144	3%/100%
Hepatitis B virus*	14 (15.2, 10.7–22)	28 (32.4, 22.2–35.3)	0	212	33%/100%
Hepatitis E virus	0	1	1	252	0%/100%
HIV*	124 (28.3, 21.4–37.3)	52 (32.7, 23–39.2)	6	72	70%/92%
Bacterial pathogens					
S. pneumoniae	1	17	1	235	6%/100%
E. coli	0	4	0	250	0%/100%
Rickettsia spp.	1	5	0	248	17%/100%
P. aeruginosa	0	1	9	244	0%/96%
K. pneumoniae	0	2	0	252	0%/100%
S. aureus	0	2	2	250	0%/99%
Salmonella spp.	0	2	0	252	0%/100%
Leptospira spp.	0	2	0	252	0%/100%
Coxiella spp.	0	1	0	253	0%/100%
M. tuberculosis	1	43	0	207	2%/100%
Other pathogens					
Plasmodium	2	15	0	234	12%/100%
Cryptococcus	0	6	0	248	0%/100%
Toxoplasma	0	2	0	252	0%/100%

[a]Asterisks denote pathogens detected by qPCR with plasma sample following sequencing.

8 samples. *Hepacivirus* was detected in 17 samples, but only 1 sample aligned to hepatitis C virus species (and was hepatitis C qPCR positive), with the remaining samples aligning to GB virus B species. By qPCR, HIV viremia was detected in 71% of samples, EBV was detected in 45%, and hepatitis B virus was detected in 17% of samples. The quantity of virus detected by qPCR (i.e., threshold cycle [C_T] value) correlated with the number of sequencing reads (Fig. 2).

Mortality was associated with a lower C_T value for EBV (adjusted odds ratio of 0.85, $P < 0.001$) but not for hepatitis B virus or HIV (Fig. 3). There was no significant correlation between total sequencing reads, pathogen reads, or pathogen-specific reads with regard to mortality or lactate concentration.

DISCUSSION

In this study, we demonstrated the ability of metagenomic next-generation sequencing from plasma samples to detect potential etiologic pathogens contributing to sepsis in Uganda. We found that the mNGS method was relatively insensitive for many bacterial pathogens but was moderately successful for many viruses. We successfully identified HIV in the majority of samples and detected new pathogens, including EBV and hepatitis B virus, among several others that may be important.

On average, each plasma sample had 3,404,737 ± 2,201,997 reads, of which 220,032 ± 416,691 reads (6.3% ± 8.6%) were identified as nonhuman potential pathogen reads. This depth was similar to other studies using mNGS sequencing from plasma or serum for pathogen detection (16), and the percentage of nonhost reads from plasma samples was higher than that in some other studies (16, 17). This could have been contributed by the use of human rRNA depletion in the cDNA library preparation, the extreme illness of our cohort, and/or the high HIV rate, since HIV was the leading pathogen identified.

Background filtering allowed for an unbiased way to remove potential contaminant reads from the data, as highlighted in Fig. 1, without which nearly every sample had some level of pathogen reads including in healthy control plasma samples and water samples (for instance, these samples revealed bacterial reads such as *Escherichia* and *Klebsiella*, among others). If a pathogen was detected in our water and healthy controls, such as several bacterial species, they were filtered out, even if they were potentially a true pathogen in the samples. While the filtering thresholds were somewhat arbitrary, they could be adjusted. We aimed to keep high filter thresholds to maintain high specificity at the expense of sensitivity. The appropriate filter threshold was difficult to determine and can differ for different pathogens based on several factors, including size of pathogen genome, level of pathogen burden in serum, and depth of sequencing (18–20).

The sensitivity of mNGS compared to targeted qPCR was low, at 30% overall, and even lower for bacteria at 5%. This was not surprising to us and can likely explained by several factors, including the read depth, limited sample volume, age of the samples (10+ years), different sample types (sequencing on plasma versus qPCR on whole blood), stringency of the background filtering thresholds used, and the potential presence of inhibitory substances to sequencing in human plasma, which have all been previously described (21–23). Several studies have shown similarly variable sensitivity and specificity of mNGS versus PCR (24–27). Additionally, our extraction and library preparation methods aimed to sequence both RNA and DNA viruses, bacteria, and fungi by using RNA sequencing, which may have preferentially favored some pathogens over others, as previously described (27–29).

We evaluated the percent reference genome coverage of pathogens in each sample to better characterize identified pathogens and increase specificity, by requiring that the reads mapped to diverse sites and were less likely to have been a single contaminant region or genome database anomaly. This approach, however, would favor pathogens with smaller reference genomes or reference accessions that use incomplete genomes. This likely contributed to the higher viral genome coverage.

A total of 75 genus-specific pathogens were identified that passed the filter threshold criteria. The most prevalent pathogen, not surprisingly, was HIV (*Lentivirus*), which was
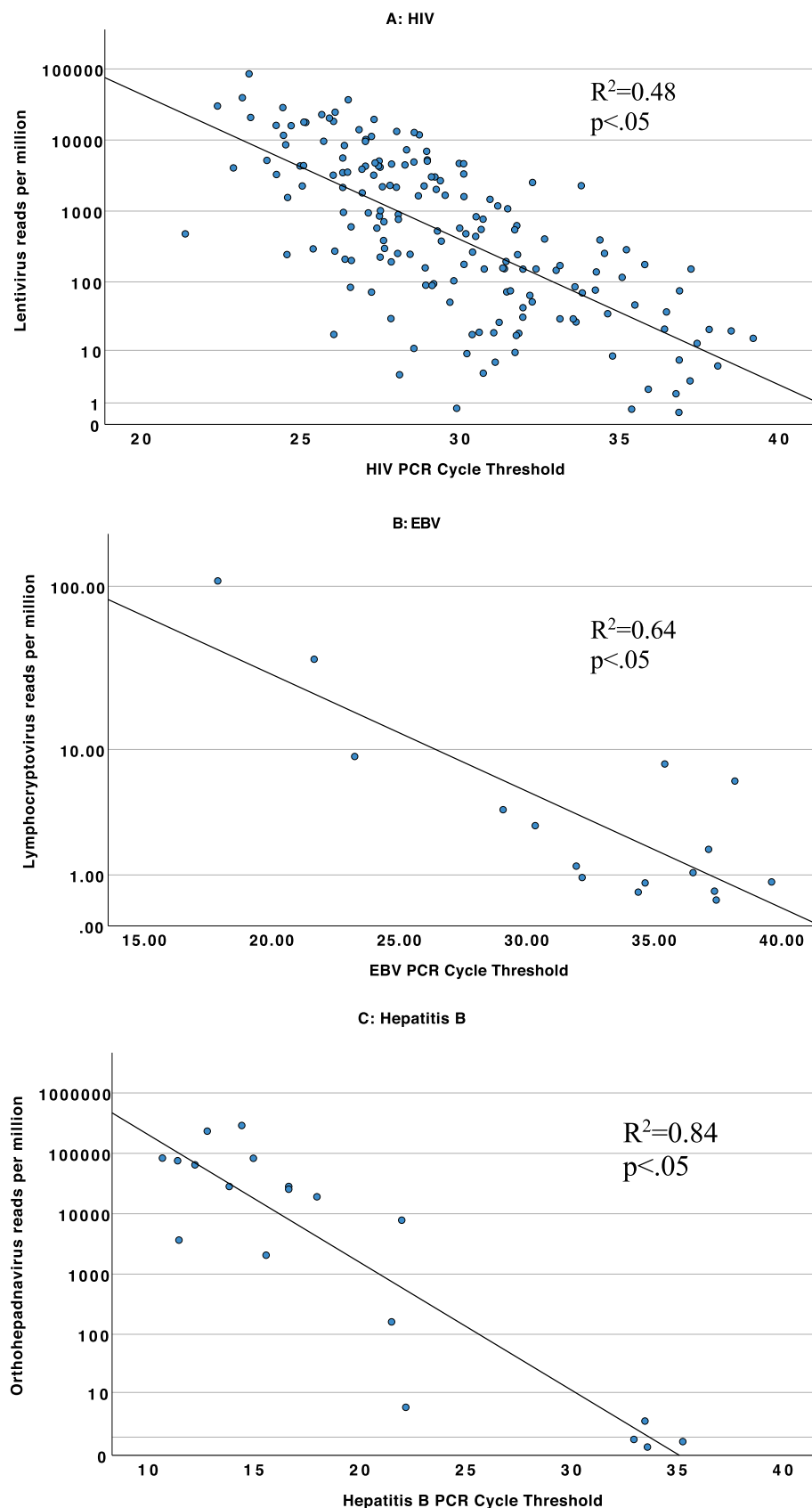
FIG 2 Comparison of qPCR cycle threshold values to sequencing reads per million for HIV, EBV, and hepatitis B virus. (A) Comparison for HIV; (B) comparison for EBV; (C) comparison for hepatitis B virus. Samples that were positive by both PCR and sequencing are shown. All sequencing reads without threshold filters are shown. Note that the y axis of each scatterplot is on a logarithmic scale.
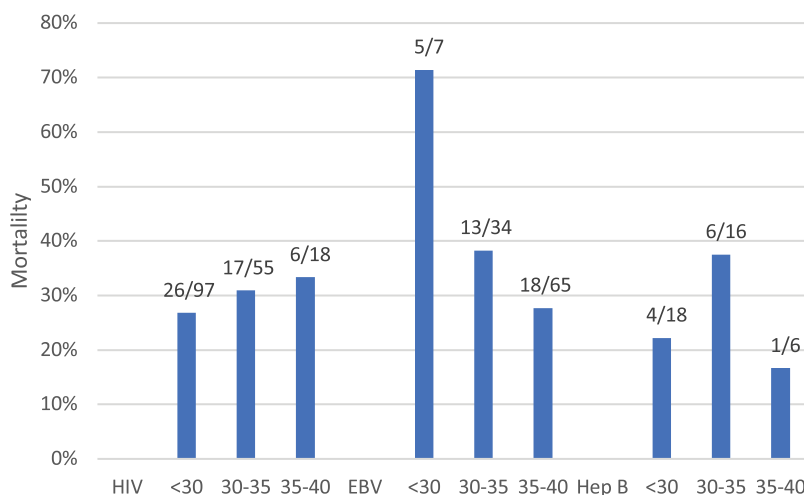
FIG 3 Mortality association with HIV, EBV, and hepatitis B virus by C_T value. The x axis shows C_T ranges separated for each pathogen, and the y axis shows the in-hospital mortality of each pathogen by C_T value. The fractions above each bar represent the numbers of deaths within each C_T range as the numerator and the number of samples in that C_T range as the denominator.

detected by this method in 130 samples. HIV was detected by qPCR in 176 samples. Other prevalent pathogens included pegivirus, seen in 50 samples, which is highly prevalent globally but not known to be of clinical importance (30). The pegivirus prevalence here was similar to that noted in an HIV-infected cohort in Ghana (20% versus 16%, respectively) (10). Several of the torque viruses were highly prevalent in this cohort, which was not surprising as these viruses are ubiquitous human viruses approaching possibly 100% prevalence in the human population, with no known clinical significance (31). Additionally, hepatitis B (*Orthohepadnavirus*) was seen in 13 samples, which can have significant implications if acute infection contributing to sepsis or the chronic management of this virus. *Hepacivirus* (the genus of hepatitis C virus) was detected by sequencing in 17 samples, but on analysis of species-level alignments in only one sample was hepatitis C detected (and confirmed by qPCR), with the remaining samples aligning to the closely related GB virus B, of which viremia is known to be highly prevalent but not known to cause human disease (32). *Rhadinovirus* (human herpesvirus 8 or Kaposi's sarcoma-associated herpesvirus) was detected in 4 samples; this genus is often associated with HIV and has important clinical implications of potential malignancy (33). Several of the other genera identified, including many bacteria and fungi (as highlighted in Fig. S2A and B), are atypical for human bloodstream infections and may have represented environmental contamination, commensal organisms, or etiologic pathogens in this highly immunocompromised HIV-infected population.

Sequencing of HIV was evaluated in this cohort, given its high prevalence and importance in contributing to the risk of opportunistic and atypical infections in sepsis. The sensitivity and specificity of mNGS for the detection of HIV compared to qPCR were 70% and 92%, respectively. We suspect some reduced sensitivity of molecular diagnostics for HIV in this cohort, given 23% of HIV-positive patients were known to be taking antiretroviral therapy and a portion may have been virologically suppressed (15). Despite this possibility, the cohort had a low median CD4 level, suggestive of poor virological control, but qPCR and mNGS results did not correlate with CD4 count. Mortality was not associated with qPCR for HIV, which highlighted other potential pathogens as the primary cause of this cohort's acute illness, sepsis, and mortality.

EBV (*Lymphocryptovirus*) was only detected in 3 samples using this mNGS method, but when tested by qPCR, 45% of samples were found to be positive. These findings presumably reflected the insensitivity of mNGS compared to targeted qPCR. The high prevalence of EBV viremia by qPCR in this cohort was higher than we anticipated and higher than reported in a similar study of febrile illness in Kenya (29%) (34). This high-level viremia likely reflected the severity of illness in this cohort and the activation of

EBV-infected immune cells (35). It is possible that the association between high EBV viremia and mortality may be clinically relevant or related to potential EBV-related malignancy; however, we were unable to evaluate this explicitly in this cohort.

Hepatitis B (*Orthohepadnavirus*) viremia was also unexpectedly found in 17% of samples by qPCR. This was higher than expected, given the WHO estimate of HBsAg seroprevalence of 6.1% in sub-Saharan Africa and a reported 10.9% viremia in symptomatic HIV-infected patients in Ghana (10, 36). While again not likely to be significantly contributing to sepsis or associated with mortality in this cohort, the high prevalence of viremia has implications for the acute and chronic management of hepatitis B and the frequent coinfection with HIV.

This study had several limitations. The plasma samples had been cryopreserved for about 12 years prior to sequencing, while the comparison whole-blood samples had been cryopreserved for about 9 years prior to PCR (14, 37); therefore, the prolonged storage could have affected detection yield. Similarly, the extraction and sequencing methods used and the background model from healthy human and water control samples could have variably impacted the pathogens detected. The turnaround time and cost of both qPCR and mNGS are a limitation for the routine use of these molecular diagnostics in clinical care. The 48-plex qPCR using the TaqMan Array card had an approximate turnaround time of a few hours at about $100 per sample, while mNGS required a few days at a cost of about $500 per sample (37). While there are several available bioinformatics methods, we used the Chan Zuckerberg ID bioinformatics pipeline in this study; however, when we examined the results using an in-house method, we found similar results. Specifically, we evaluated our results using a different bioinformatics pipeline with K-mer-based species identification and found similar results (Table S1). Notably, all such analyses are dependent on the quality of the NCBI reference library used. The bioinformatics methods used may also likely identify pathogens with smaller genomes and are limited by the depth of sequencing. Analyses of clinical correlations, such as lactate concentrations, and mortality with individual pathogens were limited by the relatively small number of specific pathogen detections in this cohort.

In summary, in a predominantly HIV-infected population with sepsis, we were able to use mNGS to detect pathogens in plasma and compare these results to prior qPCR, culture, and clinical data. While relatively insensitive compared to qPCR, our methods were able to detect several new infections, especially viral infections, that could be further assessed by qPCR. mNGS may serve as a useful adjunctive tool to identify new pathogens not detected by other methods.

MATERIALS AND METHODS

Sample collection. Patient enrollment and specimen collection have been described previously (15). Individuals were ≥18 years of age with sepsis and were enrolled in a fluid resuscitation study from the medical casualty units of Mulago National Referral Hospital and Masaka Regional Referral Hospital in Uganda from May 2008 to May 2009. Baseline demographic, clinical, and routine lab data were obtained. Briefly, among the 336 patients tested previously by PCR, we tested all available cryopreserved plasma specimens (*n* = 254). We also collected blood from 4 healthy American volunteers as negative controls. All plasma samples were cryopreserved at −80°C until they were tested in 2021. qPCR detection for 43 targets from the paired whole blood (collected at the same time as the plasma samples) has been previously described (14).

Ethics statement. Written informed consent was obtained from each patient or a surrogate if the patient could not provide consent. All work was approved by the University of Virginia Institutional Review Board.

Library preparation and sequencing. Using 1 mL of plasma from each patient, total nucleic acid was extracted using the QIAamp circulating nucleic acid kit (Qiagen) following the kit protocol into a 30-μL elution volume. cDNA was created from 6 μL of elution mixture using the REPLI-g WTA single-cell kit (Qiagen) according to the manufacturer's instructions, with the addition of human rRNA removal using the QIAseq FastSelect −rRNA HMR kit (Qiagen). In total, 263 samples (254 plasma samples, 4 healthy human plasma samples, and 5 water controls) were indexed and pooled using the Nextera DNA Flex library prep kit and run on an Illumina Hiseq 2500 system to obtain 150-bp paired-end reads. Confirmatory qPCRs were performed on the extracted plasma for HIV, EBV, and hepatitis B virus using previously described primers and probes (13, 38–40).

Data processing. Microbial pathogens were identified from raw sequencing reads using Chan Zuckerberg ID (v6.6), a cloud-based, open source bioinformatics platform. In brief, initial host read

filtering was performed using spliced transcripts alignment to a reference (STAR) algorithm, followed by removal of duplicate or low-quality and low-complexity sequences. Next, reads were aligned to the host genome of interest using bowtie2 to remove any remaining host reads. The nonhuman reads were then aligned to the NCBI nucleotide and protein database (NCBI index date, 20 April 2020), using GSNAPL and RAPSearch, respectively.

All Chan Zuckerberg ID scripts and user instructions are available at https://github.com/chanzuckerberg/idseq-dag, and the graphical user interface web application for sample upload is available at https://github.com/chanzuckerberg/idseq-web. Further user guides and explanations can be found at https://chanzuckerberg.zendesk.com/hc/en-us. Using suggested threshold filters from CZidseq, a Z-score was calculated for both nucleic acid and protein alignments for each genus relative to a background of nontemplate (water only) and healthy human plasma controls, whereby a genus with a Z-score of >1 and with >10 reads per million and for both nucleic acids and protein alignment in a sample were considered positive. For further specificity in the detection of new pathogens, the percent coverage for each pathogen to the reference NCBI accession was evaluated. Based on the sequencing results, we selected HIV, HBV, and EBV to further evaluate the results by targeted qPCR. qPCR was considered positive for detection if the C_T was <40.

Statistics. We calculated sensitivity and specificity for sequencing results compared to previous qPCR results. We also performed ROC analysis by comparing pathogen reads per million without any background thresholds, divided by each reference NCBI genome size against PCR. We evaluated relationships between sequencing and qPCR results with in-hospital mortality, CD4, HIV status, and lactate concentration with logistic regression and adjusted for age and sex. P values were 2-tailed, and values of <0.05 were considered significant (SPSS, version 28, IBM Corp, Armonk, NY).

Data availability. Sequencing data from this study are publicly available under the accession number PRJNA898830. Data that were processed through the Chan Zuckerberg ID pipeline are publicly available at https://czid.org/ under the project title "Uganda Sepsis Metagenomics Nextseq."

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.7 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Thwaites CL, Lundeg G, Dondorp AM, Adhikari NKJ, Nakibuuka J, Jawa R, Mer M, Murthy S, Schultz MJ, Thien BN, Kwizera A. 2019. Infection management in patients with sepsis and septic shock in resource-limited settings, p 163–184. In Dondorp AM, Dünser MW, Schultz MJ (ed), Sepsis management in resource-limited settings. Springer International Publishing, Cham, Switzerland.

2. Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. 2010. Critical care and the global burden of critical illness in adults. Lancet 376:1339–1346. https://doi.org/10.1016/S0140-6736(10)60446-1.

3. Crump JA, Ramadhani HO, Morrissey AB, Saganda W, Mwako MS, Yang LY, Chow SC, Morpeth SC, Reyburn H, Njau BN, Shaw AV, Diefenthal HC, Shao JF, Bartlett JA, Maro VP. 2011. Invasive bacterial and fungal infections among hospitalized HIV-infected and HIV-uninfected adults and adolescents in northern Tanzania. Clin Infect Dis 52:341–348. https://doi.org/10.1093/cid/ciq103.

4. Jacob ST, Moore CC, Banura P, Pinkerton R, Meya D, Opendi P, Reynolds SJ, Kenya-Mugisha N, Mayanja-Kizza H, Scheld WM, Promoting Resource-Limited Interentions for Sepsis Management in Uganda (PRISM-U) Study Group. 2009. Severe sepsis in two Ugandan hospitals: a prospective observational study of management and outcomes in a predominantly HIV-1 infected population. PLoS One 4:e7782. https://doi.org/10.1371/journal.pone.0007782.

5. Prasad N, Murdoch DR, Reyburn H, Crump JA. 2015. Etiology of severe febrile illness in low- and middle-income countries: a systematic review. PLoS One 10:e0127962. https://doi.org/10.1371/journal.pone.0127962.

6. Crump JA, Ramadhani HO, Morrissey AB, Saganda W, Mwako MS, Yang LY, Chow SC, Njau BN, Mushi GS, Maro VP, Reller LB, Bartlett JA. 2012. Bacteremic disseminated tuberculosis in sub-saharan Africa: a prospective cohort study. Clin Infect Dis 55:242–250. https://doi.org/10.1093/cid/cis409.

7. Jacob ST, Pavlinac PB, Nakiyingi L, Banura P, Baeten JM, Morgan K, Magaret A, Manabe Y, Reynolds SJ, Liles WC, Wald A, Joloba ML, Mayanja-Kizza H, Scheld WM. 2013. Mycobacterium tuberculosis bacteremia in a cohort of HIV-infected patients hospitalized with severe sepsis in Uganda: high frequency, low clinical suspicion [corrected] and derivation of a clinical prediction score. PLoS One 8:e70305. https://doi.org/10.1371/journal.pone.0070305.

8. Halliday JEB, Carugati M, Snavely ME, Allan KJ, Beamesderfer J, Ladbury GAF, Hoyle DV, Holland P, Crump JA, Cleaveland S, Rubach MP. 2020. Zoonotic causes of febrile illness in malaria endemic countries: a systematic review. Lancet Infect Dis 20:e27–e37. https://doi.org/10.1016/S1473-3099(19)30629-2.

9. Allan KJ, Biggs HM, Halliday JE, Kazwala RR, Maro VP, Cleaveland S, Crump JA. 2015. Epidemiology of leptospirosis in Africa: a systematic review of a neglected zoonosis and a paradigm for 'One Health' in Africa. PLoS Negl Trop Dis 9:e0003899. https://doi.org/10.1371/journal.pntd.0003899.

10. Compston LI, Li C, Sarkodie F, Owusu-Ofori S, Opare-Sem O, Allain JP. 2009. Prevalence of persistent and latent viruses in untreated patients infected with HIV-1 from Ghana, West Africa. J Med Virol 81:1860–1868. https://doi.org/10.1002/jmv.21614.

11. Compston LI, Sarkobie F, Li C, Candotti D, Opare-Sem O, Allain JP. 2008. Multiplex real-time PCR for the detection and quantification of latent and persistent viral genomes in cellular or plasma blood fractions. J Virol Methods 151: 47–54. https://doi.org/10.1016/j.jviromet.2008.03.023.

12. Collenberg E, Ouedraogo T, Ganamé J, Fickenscher H, Kynast-Wolf G, Becher H, Kouyaté B, Kräusslich HG, Sangaré L, Tebit DM. 2006. Seroprevalence of six different viruses among pregnant women and blood donors in rural and urban Burkina Faso: a comparative analysis. J Med Virol 78: 683–692. https://doi.org/10.1002/jmv.20593.

13. Liu J, Ochieng C, Wiersma S, Ströher U, Towner JS, Whitmer S, Nichol ST, Moore CC, Kersh GJ, Kato C, Sexton C, Petersen J, Massung R, Hercik C, Crump JA, Kibiki G, Maro A, Mujaga B, Gratz J, Jacob ST, Banura P, Scheld WM, Juma B, Onyango CO, Montgomery JM, Houpt E, Fields B. 2016. Development of a TaqMan array card for acute-febrile-illness outbreak investigation and surveillance of emerging pathogens, including Ebola virus. J Clin Microbiol 54:49–58. https://doi.org/10.1128/JCM.02257-15.

14. Moore CC, Jacob ST, Banura P, Zhang J, Stroup S, Boulware DR, Scheld WM, Houpt ER, Liu J. 2019. Etiology of sepsis in Uganda using a quantitative polymerase chain reaction-based TaqMan array card. Clin Infect Dis 68:266–272. https://doi.org/10.1093/cid/ciy472.

15. Jacob ST, Banura P, Baeten JM, Moore CC, Meya D, Nakiyingi L, Burke R, Horton CL, Iga B, Wald A, Reynolds SJ, Mayanja-Kizza H, Scheld WM, Promoting Resource-Limited Interventions for Sepsis Management in Uganda (PRISM-U) Study Group. 2012. The impact of early monitored management on survival in hospitalized adult Ugandan patients with severe sepsis: a prospective intervention study. Crit Care Med 40:2050–2058. https://doi.org/10.1097/CCM.0b013e31824e65d7.

16. Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, Egger R, Dimitrov B, Juan YF, Sheu J, Wang J, Kalantar K, Langelier C, Ruel T, Mpimbaza A, Wilson MR, Rosenthal PJ, DeRisi JL. 2019. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. PLoS One 14:e0218318. https://doi.org/10.1371/journal.pone.0218318.

17. Kisat MT, Odenheimer-Bergman A, Markus H, Joseph B, Srivatsan SN, Contente-Cuomo T, Khalpey Z, Keim P, O'Keeffe T, Askari R, Salim A, Rhee P, Murtaza M. 2021. Plasma metagenomic sequencing to detect and quantify bacterial DNA in ICU patients suspected of sepsis: a proof-of-principle study. J Trauma Acute Care Surg 91:988–994. https://doi.org/10.1097/TA.0000000000003396.

18. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12:87. https://doi.org/10.1186/s12915-014-0087-z.

19. Zinter MS, Mayday MY, Ryckman KK, Jelliffe-Pawlowski LL, DeRisi JL. 2019. Towards precision quantification of contamination in metagenomic sequencing experiments. Microbiome 7:62. https://doi.org/10.1186/s40168-019-0678-6.

20. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6:226. https://doi.org/10.1186/s40168-018-0605-2.

21. Dark P, Blackwood B, Gates S, McAuley D, Perkins GD, McMullan R, Wilson C, Graham D, Timms K, Warhurst G. 2015. Accuracy of LightCycler SeptiFast for the detection and identification of pathogens in the blood of patients with suspected sepsis: a systematic review and meta-analysis. Intensive Care Med 41:21–33. https://doi.org/10.1007/s00134-014-3553-8.

22. Nga TV, Karkey A, Dongol S, Thuy HN, Dunstan S, Holt K, Tu I, Campbell JI, Chau TT, Chau NV, Arjyal A, Koirala S, Basnyat B, Dolecek C, Farrar J, Baker S. 2010. The sensitivity of real-time PCR amplification targeting invasive Salmonella serovars in biological specimens. BMC Infect Dis 10:125. https://doi.org/10.1186/1471-2334-10-125.

23. Hosokawa-Muto J, Fujinami Y, Mizuno N. 2013. Assessment of viable bacteria and bacterial DNA in blood and bloodstain specimens stored under various conditions. J Forensic Leg Med 20:1035–1040. https://doi.org/10.1016/j.jflm.2013.09.024.

24. Castro JC, Rodriguez-R LM, Harvey WT, Weigand MR, Hatt JK, Carter MQ, Konstantinidis KT. 2018. imGLAD: accurate detection and quantification of target organisms in metagenomes. PeerJ 6:e5882. https://doi.org/10.7717/peerj.5882.

25. Zhang B, Zhou J, Gui R, Li Z, Zu Y, Wang J, Yu F, Zhang Y, Zhao H, Ji Z, Song Y. 2021. Metagenomic next generation sequencing in the detection of pathogens in cerebrospinal fluid of patients after alternative donor transplantation: a feasibility analysis. Front Cell Infect Microbiol 11:720132. https://doi.org/10.3389/fcimb.2021.720132.

26. Ebinger A, Fischer S, Höper D. 2021. A theoretical and generalized approach for the assessment of the sample-specific limit of detection for clinical metagenomics. Comput Struct Biotechnol J 19:732–742. https://doi.org/10.1016/j.csbj.2020.12.040.

27. Gu W, Deng X, Lee M, Sucu YD, Arevalo S, Stryke D, Federman S, Gopez A, Reyes K, Zorn K, Sample H, Yu G, Ishpuniani G, Briggs B, Chow ED, Berger A, Wilson MR, Wang C, Hsu E, Miller S, DeRisi JL, Chiu CY. 2021. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. Nat Med 27:115–124. https://doi.org/10.1038/s41591-020-1105-z.

28. Gu W, Miller S, Chiu CY. 2019. Clinical metagenomic next-generation sequencing for pathogen detection. Annu Rev Pathol 14:319–338. https://doi.org/10.1146/annurev-pathmechdis-012418-012751.

29. Miller S, Chiu C. 2021. The role of metagenomics and next-generation sequencing in infectious disease diagnosis. Clin Chem 68:115–124. https://doi.org/10.1093/clinchem/hvab173.

30. Yu Y, Wan Z, Wang JH, Yang X, Zhang C. 2022. Review of human pegivirus: prevalence, transmission, pathogenesis, and clinical implication. Virulence 13:324–341. https://doi.org/10.1080/21505594.2022.2029328.

31. Lolomadze EA, Rebrikov DV. 2020. Constant companion: clinical and developmental aspects of torque teno virus infections. Arch Virol 165:2749–2757. https://doi.org/10.1007/s00705-020-04841-x.

32. Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P. 2011. The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. J Gen Virol 92:233–246. https://doi.org/10.1099/vir.0.027490-0.

33. Etta EM, Alayande DP, Mavhandu-Ramarumo LG, Gachara G, Bessong PO. 2018. HHV-8 seroprevalence and genotype distribution in Africa, 1998–2017: a systematic review. Viruses 10:458. https://doi.org/10.3390/v10090458.

34. Masakhwe C, Ochanda H, Nyakoe N, Ochiel D, Waitumbi J. 2016. Frequency of Epstein - Barr virus in patients presenting with acute febrile illness in Kenya. PLoS One 11:e0155308. https://doi.org/10.1371/journal.pone.0155308.

35. Nowalk A, Green M. 2016. Epstein-Barr virus. Microbiol Spectr 4. https://doi.org/10.1128/microbiolspec.DMIH2-0011-2015.

36. Sonderup MW, Dusheiko G, Desalegn H, Lemoine M, Tzeuton C, Taylor-Robinson SD, Spearman CW. 2020. Hepatitis B in sub-Saharan Africa: how many patients need therapy? J Viral Hepat 27:560–567. https://doi.org/10.1111/jvh.13247.

37. Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY. 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. N Engl J Med 370:2408–2417. https://doi.org/10.1056/NEJMoa1401268.

38. Biswas S, Rust LN, Wettengel JM, Yusova S, Fischer M, Carson JN, Johnson J, Wei L, Thode T, Kaadige MR, Sharma S, Agbaria M, Bimber BN, Tu T, Protzer U, Ploss A, Smedley JV, Golomb G, Sacha JB, Burwitz BJ. 2022. Long-term hepatitis B virus infection of rhesus macaques requires suppression of host immunity. Nat Commun 13:2995. https://doi.org/10.1038/s41467-022-30593-0.

39. Liu J, Gratz J, Amour C, Nshama R, Walongo T, Maro A, Mduma E, Platts-Mills J, Boisen N, Nataro J, Haverstick DM, Kabir F, Lertsethtakarn P, Silapong S, Jeamwattanalert P, Bodhidatta L, Mason C, Begum S, Haque R, Praharaj I, Kang G, Houpt ER. 2016. Optimization of quantitative PCR methods for enteropathogen detection. PLoS One 11:e0158199. https://doi.org/10.1371/journal.pone.0158199.

40. Morris T, Robertson B, Gallagher M. 1996. Rapid reverse transcription-PCR detection of hepatitis C virus RNA in serum by using the TaqMan fluorogenic detection system. J Clin Microbiol 34:2933–2936. https://doi.org/10.1128/jcm.34.12.2933-2936.1996.