Next-generation sequencing of microbial cell-free DNA for rapid noninvasive diagnosis of infectious diseases in immunocompromised hosts [version 1; peer review: 1 approved]

Jose F. Camargo, Asim A. Ahmed, Martin S. Lindner, Michele I. Morris, Shweta Anjan, Anthony D. Anderson, Clara E. Prado, Sudeb C. Dalai, Octavio V. Martinez, Krishna V. Komanduri

1Infectious Diseases, University of Miami Miller School of Medicine, Miami, FL, 33136, USA
2Karius, Inc., Redwood City, CA, USA
3Department of Pharmacy, University of Miami Sylvester Comprehensive Cancer Center, Miami, FL, 33136, USA
4Department of Microbiology, University of Miami Miller School of Medicine, Miami, FL, 33136, USA
5Division of Hematology Oncology, University of Miami Miller School of Medicine, Miami, FL, 33136, USA

Abstract

Background: Cell-free DNA (cfDNA) sequencing has emerged as an effective laboratory method for rapid and noninvasive diagnosis in prenatal screening testing, organ transplant rejection screening, and oncology liquid biopsies.

Methods: Here we report our experience using next-generation sequencing (NGS) for detection of microbial cfDNA in a cohort of ten immunocompromised patients with febrile neutropenia or deep-seated infection.

Results: Among five hematological malignancy patients, for whom a microbiological diagnosis was established, pathogen identification by cfDNA NGS demonstrated 100% positive agreement with conventional diagnostic laboratory methods. Further, cfDNA identified the etiological agent in two patients with culture negative sepsis who had undergone hematopoietic stem cell transplant.

Conclusion: These data support the clinical utility of measurement of microbial cfDNA sequencing from peripheral blood for rapid noninvasive diagnosis of infections in immunocompromised hosts. Larger studies are needed.

Keywords

Cell-free microbial DNA, next generation sequencing, infection, immunocompromised host, hematopoietic stem cell transplant
**Introduction**

Infections are a leading cause of morbidity and mortality among immunocompromised individuals. Bacteremia occurs in up to 25% of all patients with neutropenia and fever. Infection is a leading cause of non-relapse mortality among hematopoietic cell transplantation (HCT) recipients. The incidence of bacteremia is higher than 40% and 90%, respectively, within the first 100 days post-transplant. The cumulative incidence rates of proven/probable invasive fungal infections during the first year after allogeneic HCT with non-myeloablative conditioning is 19%. Infection is also a common complication of chimeric antigen receptor-modified T (CAR-T)-cell immunotherapy with 28-day cumulative incidence of 23% after CAR-T-cell infusion.

Establishing a microbiological diagnosis of infectious diseases in this vulnerable population is often challenging for a number of reasons. i) Prior exposure to antibiotics and antifungals which confounds the yield of blood cultures; indeed, most patients with neutropenia and fever will have no infectious etiology documented. ii) Low sensitivity of mycobacterial and fungal cultures; some microorganisms, such as fastidious bacteria, mycobacteria and dimorphic fungi require longer incubation periods; and blood cultures in almost half of patients with candidemia are negative. iii) Tissue biopsies are often precluded due to the risk of bleeding in the setting of thrombocytopenia, coagulopathy in those with liver disease or hemodynamic instability in critically ill patients. A delay in diagnosis in patients with invasive fungal infection results in higher mortality.

Thus, there is an unmet need for novel, rapid, cost-effective, noninvasive diagnostic methods in the field.

Cell-free DNA (cfDNA) technology has been used successfully in noninvasive prenatal testing, organ transplant rejection screening, and oncology liquid biopsies. In recent years, this technology has been developed for use in infectious disease diagnostics. Detection of microbial cfDNA by next generation sequencing (NGS) is an accurate and precise way of identifying and quantifying pathogens. The Karius® Test relies on sequencing of microbial cfDNA circulating in plasma to identify over 1,000 pathogens, including bacteria, viruses and fungi, from a 5 ml blood sample. This novel diagnostic tool has recently validated in a study showing that microbial cfDNA NGS identified 94% of microbes identified by conventional blood culture in patients with sepsis and has excellent correlation with quantitative PCR testing in patients with cytomegalovirus (CMV).

Recent reports indicate that NGS measuring microbial cfDNA is useful in the diagnosis of cases of Streptococcus pneumoniae-related hemolytic uremic syndrome, Coxiella burnetii endocarditis, invasive Mycobacterium chimaera infection, Nocardia cyriaciigeorgica pneumonia, Capnocytophaga canimorsus sepsis, M. tuberculosis complex and M. haemophilum infections, M. bovis aortitis; Candida spp., Aspergillus spp., non-Aspergillus molds invasive infections; Pneumocystis jirovecii pneumonia (PJP), Toxoplasma gondii infection and chorioamnionitis, among others. Among 21 patients with culture-positive infective endocarditis, cfDNA NGS identified the same organism as blood cultures in 20 patients (95% sensitivity) and additionally identified Enterococcus faecalis in one out of the three patients with definitive culture-negative endocarditis.

Of note, in this study the cfDNA NGS test identified pathogens causing endocarditis in patients pre-treated with antibiotics up to 30 days prior to initial sample collection.

Here we evaluated the clinical utility of NGS for detection of microbial cfDNA in plasma in a cohort of ten patients receiving chemotherapy or transplants with episodes of febrile neutropenia, sepsis or documented infection.

**Methods**

**Study subjects**

Adult patients followed at the Sylvester Comprehensive Cancer Center were enrolled between July 31 and October 2, 2018. Inclusion criteria were: i) age >18 years old; ii) patients must have received chemotherapy or transplant; and iii) must have had a febrile illness or documented infection (e.g., positive blood cultures, clinical/radiographic evidence of pneumonia). There were no exclusion criteria. In this proof-of-concept study, most of the patients had an established diagnosis of infection prior to NGS testing. The study was approved by the University of Miami Institutional Review Board (IRB approval #20080899), consistent with principles in the Declaration of Helsinki. Each participant provided written informed consent for their inclusion in the study. No sample size calculation was done; instead the number of patients enrolled was entirely dependent on the number of cfDNA kits made available for the pilot study.

**Sample collection and processing**

Blood samples (5 mL) were collected in BD vacutainer plasma preparation tubes. Samples were collected at the time of suspected or confirmed infection diagnosis. Within 1 hour of sample collection, tubes were spun down at 1,100 RCF for 10 min at room temperature. Samples were shipped overnight to Karius, Inc. (Redwood City, CA).

**Measurement of cfDNA using NGS**

Cell-free DNA was extracted from plasma, NGS libraries were prepared, and sequencing was performed on an Illumina NextSeq®500. Sequencing reads identified as human were removed, and remaining sequences were aligned to a curated pathogen database. Any of over 1,000 organisms in the Karius clinical reportable range found to be present above a predefined statistical threshold were reported as previously described. The quantity for each organism identified was expressed in Molecules Per Microliter (MPM), the number of DNA sequencing reads from the reported organism present per microliter of plasma.

**The Karius® Test**

**Reference database and QC.** Reference genomes for Homo sapiens and microorganisms (bacteria, viruses, fungi/molds, and other eukaryotic pathogens) were retrieved from the National Center for Biotechnology Information (NCBI) ftp site (NCBI, U.S. National Library of Medicine (NLM), Human Genome, release GRCh38.p7, and NCBI, U.S. NLM,.
Microbial Genomes, respectively). Sequence similarities between microorganism references were inspected to identify taxonomic mislabeling and sequence contamination. From the reference genomes passing these quality controls, a subset was selected to maximize sequence diversity. As part of the selection process, NCBI BioSample data were used to ensure the inclusion of reference genomes from both clinical and non-clinical isolates. The final reference genome dataset included over 21,000 reference genomes, containing over 2.7 million sequences. Selected sequences were collected into a single FASTA file and used to generate our microorganism reference BLAST database. A subset of these taxa, including 1251 clinically significant microorganisms, was used as the clinical reportable range.

Clinical reportable range (CRR). The selection of organisms in our clinical reportable range (CRR) was performed as follows. A candidate list was generated by two board-certified infectious disease physicians by including (a) DNA viruses, culturable bacteria, additional fastidious and unculturable bacteria, mycobacteria, and eukaryotic pathogens from the standard text18 and a number of infectious disease references, (b) organisms in the pathogen database referenced in published case reports, and (c) reference genomes sequenced from human clinical isolates (as indicated by the NCBI BioSample resource) with publications supporting pathogenicity. Organisms from the above list that were associated with high-quality reference genomes, as determined by our reference database QC process (see above), were used to further narrow the range. Finally, organisms at risk of generating common false-positive calls because of sporadic environmental contamination were removed. The sequence database is continuously curated to minimize human cross-reactivity as well as cross-reactivity between pathogens and is screened to mitigate contamination with sequences from human or other organisms.

Sequencing. Plasma samples were thawed, centrifuged at 16,000 RCF for 10 min, and spiked with a known concentration of synthetic DNA molecules for quality control purposes. Cell-free DNA was extracted from 0.5 mL plasma using a magnetic bead-based method (Omega Bio-tek Mag-Bind® cfDNA kit; catalog number M3298-01, Norcross, GA). DNA libraries for sequencing are constructed using a modified Ovation® Ultralow System V2 library preparation kit (NuGEN, San Carlos, CA). Negative controls (buffer only instead of plasma) and positive controls (healthy plasma spiked with a known mixture of microbial DNA fragments) were processed alongside patient samples in every batch. Samples were multiplexed with other samples and sequenced on an Illumina NextSeq® 500.

Analysis pipeline. Primary sequencing output files were processed using bcl2fastq (v2.17.1.14) to generate the demultiplexed sequencing reads files. Reads were filtered based on sequencing quality and trimmed based on partial or full adapter sequence. The bowtie2 (version 2.2.4) tool was used to align the remaining reads against Karius’ human and synthetic-molecules references. Sequencing reads exhibiting strong alignment against the human references or the synthetic molecule references were collected and excluded from further analysis. Remaining reads were aligned against Karius’ proprietary microorganism reference database using NCBI-blast (version 2.2.30+). A mixture model was used to assign a likelihood to the complete collection of sequencing reads that included the read sequence probabilities and the (unknown) abundances of each taxon in the sample. An expectation-maximization algorithm was applied to compute the maximum likelihood estimate of each taxon abundance. Only taxa whose abundances rejected the null hypothesis of originating from environmental contamination (as calculated from the negative controls) at high significance levels were reported. The quantity for each organism identified was expressed in molecules per microliter (MPM), the number of DNA sequencing reads from the reported organism present per microliter of plasma. The entire process from DNA extraction through analysis was typically completed within 28 hours.

Results

Background patient information
The characteristics of the patients studied are presented in Table 1. The median age was 56 years (range, 20–65) with 60% of participants males. Except for a kidney transplant recipient, all other patients had underlying hematological malignancy and/or had received an HCT. All but one (patient #2) were admitted in the hospital at the time of clinical evaluation. All the patients were receiving antimicrobials at the time of plasma sample collection. Three patients had neutropenia (absolute neutrophil count <500/µL) at the time of febrile illness. All febrile patients had blood cultures collected within 24 hours of plasma sample collection for NGS.

Results of NGS for detecting microbial cfDNA
The kidney transplant recipient had an Aspergillus fumigatus deep-seated abdominal abscess, and Aspergillus cfDNA levels, although detected in plasma, were below the positive reporting threshold. However, among patients with hematological malignancy in whom a microbiological diagnosis was established (n=5), cfDNA NGS testing correlated with other methods in all cases (100% sensitivity). This included patients with proven/probable invasive aspergillosis, PJP, Stenotrophomonas maltophilia bacteremia, CMV and adenovirus viremia. Among four patients with hematological malignancy with negative standard laboratory testing, the NGS test identified causes of bacterial sepsis in two patients (Table 1), both of whom had a compatible clinical scenario and experienced good clinical response to antibiotic therapy with resolution of fever and hypotension.

Discussion
Here we report our experience using cfDNA NGS in the evaluation of immunocompromised patients—predominantly those with hematological malignancy—with febrile illness or documented invasive infections. The study cohort included a heterogeneous group of clinical scenarios, including deep-seated pyogenic abdominal infection, pulmonary nodules/pneumonia, neutropenic fever, and septic shock. The results of this proof-of-concept study, where most of the patients had an established diagnosis of infection prior to NGS testing, complement recent reports studying the use of cfDNA NGS in immunocompromised hosts.
Table 1. Clinical characteristics of study subjects and results of next-generation sequencing of cell-free DNA.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age, gender</th>
<th>Underlying disease</th>
<th>Clinical scenario</th>
<th>Sample from CVC</th>
<th>Days of antibiotics/antifungals prior to blood draw</th>
<th>Conventional diagnostic method results</th>
<th>Microbial cfDNA pathogen results</th>
<th>MPM</th>
<th>Reference values</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65F</td>
<td>Kidney transplant</td>
<td>Pyogenic intra-abdominal infection</td>
<td>No</td>
<td>18/182</td>
<td>Aspergillus fumigatus detected by PCR and culture in abdominal fluid</td>
<td>Negative (Aspergillus fumigatus)</td>
<td>15&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt;10</td>
<td>No&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>21M</td>
<td>NHL, HCT day +342</td>
<td>Mediastinal lymphadenopathy</td>
<td>No</td>
<td>0/8</td>
<td>Negative fungal serologies and antigens BAL and lymph node tissue cultures negative</td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>20M</td>
<td>AML, HCT day +9</td>
<td>Neutropenic fever, diarrhea</td>
<td>Yes</td>
<td>8/2</td>
<td>CMV detected &lt;137 IU/mL (subsequently peaked at 2,621 IU/mL) Blood cultures and C. difficile PCR negative</td>
<td>Cytomegalovirus</td>
<td>108</td>
<td>&lt;10</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>64F</td>
<td>B-ALL MMUD day +291</td>
<td>Fever, cough, lung mass</td>
<td>Yes</td>
<td>6/5</td>
<td>Pneumocystis jirovecii BAL PCR+</td>
<td>Pneumocystis jirovecii</td>
<td>263</td>
<td>&lt;10</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>37M</td>
<td>Relapsed DLBCL after CAR-T</td>
<td>Neutropenic fever, weakness, diarrhea, cough</td>
<td>Yes</td>
<td>21/5</td>
<td>Adenovirus 480 copies/mL (subsequently peaked at 2,600 copies/mL)</td>
<td>Adenovirus</td>
<td>845</td>
<td>&lt;10</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>56M</td>
<td>AML, MMUD day +290</td>
<td>Pulmonary nodules (recently diagnosed IA) admitted with SOB</td>
<td>Yes</td>
<td>6/21</td>
<td>CMV detected &lt;137 IU/mL (subsequently peaked at 440 IU/mL) Repeat BAL negative</td>
<td>Cytomegalovirus</td>
<td>93</td>
<td>&lt;10</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>44M</td>
<td>DLBCL</td>
<td>Fevers, pulmonary nodules</td>
<td>Yes</td>
<td>3/3</td>
<td>Blood cultures negative</td>
<td>Rothia mucilaginosa</td>
<td>20</td>
<td>&lt;10</td>
<td>No</td>
</tr>
<tr>
<td>Patient</td>
<td>Age, gender</td>
<td>Underlying disease</td>
<td>Clinical scenario</td>
<td>Sample from CVC</td>
<td>Days of antibiotics/antifungals prior to blood draw</td>
<td>Conventional diagnostic method results</td>
<td>Microbial cfDNA pathogen results</td>
<td>MPM</td>
<td>Reference values&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Correlation&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------------------</td>
<td>-----</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td>8</td>
<td>60F</td>
<td>MDS, HCT day+160, GI-GVHD</td>
<td>Septic shock, multi-organ failure</td>
<td>Yes</td>
<td>15/10</td>
<td>Blood cultures negative</td>
<td><em>Escherichia coli</em></td>
<td>2,492</td>
<td>&lt;17</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Lactobacillus rhamnosus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Torque teno virus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>55F</td>
<td>Multiple myeloma</td>
<td>Pneumonia</td>
<td>Yes</td>
<td>2/0</td>
<td>Negative BAL studies</td>
<td>Negative</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>58M</td>
<td>AML</td>
<td>Neutropenic fever, pulmonary nodules, sepsis</td>
<td>Yes</td>
<td>120/129</td>
<td><em>S. maltophilia</em> in blood cultures</td>
<td><em>Stenotrophomonas maltophilia</em></td>
<td>236,594</td>
<td>&lt;83</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Aspergillus oryzae</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Staphylococcus epidermidis</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Refers to empiric or targeted therapy only. It does not include days of antimicrobial prophylaxis.

<sup>b</sup> Blood cultures were obtained within 24h of plasma sample for NGS in all patients and resulted as negative unless specified otherwise in the table.

<sup>c</sup> Reference value is the 97.5th percentile in self-reported healthy adults for whom the Karius<sup>®</sup> Test was performed.

<sup>d</sup> Correlation between Karius<sup>®</sup> Test and standard laboratory methods.

<sup>e</sup> Aspergillus fumigatus reads were present in the raw data but below the threshold for a positive test result.

<sup>f</sup> Initial cfDNA testing performed 7 weeks prior had only identified *S. epidermidis* and EBV. At that time, BAL and transbronchial biopsy results were unrevealing.

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BAL, bronchoalveolar lavage; CAR-T, chimeric antigen receptor-modified T-cell immunotherapy; cfDNA, cell-free DNA; CMV, cytomegalovirus; CVC, central venous catheter; DLBCL, diffuse large B cell lymphoma; GI-GVHD, gastrointestinal graft-versus-host disease; HCT, hematopoietic cell transplantation; F, female; M, male; MPM, molecules per microliter; NGS, next-generation sequencing; NLH, Non-Hodgkin lymphoma; SOB, shortness of breath.
In a recent study of 55 patients with neutropenic fever, cfDNA testing had positive agreement with conventional blood cultures in 9 of 10 patients in whom blood cultures identified a causative organism of sepsis. Using clinical adjudication by three infectious diseases specialists, cfDNA NGS had a sensitivity of 85.4% (41/48) and specificity of 100% (7/7)\(^6\). Thus, this test is a promising diagnostic tool in neutropenic fever, a clinical scenario where conventional work up fails to identify an etiological agent in a majority of cases\(^4\). Another study evaluated 40 patients with prolonged neutropenia and fever (>96h) despite administration of antibiotics for suspected fungal infection (the authors excluded patients who had received antifungal therapy for >3 days); in this study cfDNA NGS identified fungal pathogens including \textit{Aspergillus fumigatus}, \textit{Rhizopus} spp., \textit{Candida albicans}, \textit{Candida glabrata} and \textit{Pneumocystis jirovecii}\(^5\).

In a recent report by Hong \textit{et al.}\(^2\), in seven out of nine subjects (including seven immunocompromised hosts) with proven deep-seated invasive fungal infection, plasma NGS testing detected the same fungus identified from the biopsy tissue at the genus level. The fungi identified by plasma NGS included \textit{Aspergillus} spp. and non-\textit{Aspergillus} molds such as \textit{Scedosporium}, \textit{Rhizopus}, and \textit{Cunninghamella}\(^2\). In that report, there was one case where the plasma sample was obtained after at least 15 days of anti-\textit{Aspergillus} therapy, and NGS testing did not identify the causal organism of invasive fungal infection. Similarly, for the kidney transplant patient reported here with invasive aspergillosis, in whom \textit{Aspergillus fumigatus} cfDNA levels in plasma were detected below the reporting threshold, several months of antifungal therapy had been administered prior to the time of plasma collection. Thus, prolonged antifungal therapy prior to sample collection (e.g., >7-14 days) might interfere with detection of fungal DNA.

Although NGS has been used for screening of allograft rejection in solid organ transplant recipients\(^17\)–\(^19,21\), there is limited data with the use of NGS for diagnosis of infections in this population. A recent study demonstrated strong correlation between clinical test results and cfDNA derived from CMV in a cohort of lung transplant recipients\(^23\). In addition, cfDNA revealed undiagnosed cases of infection with microsporidia and pathogenic viruses, including adenovirus and human herpesvirus 6 among lung transplant patients\(^2\).

Recently, Fung \textit{et al.} reported three patients who received allogeneic HCT transplant in whom NGS cfDNA facilitated the diagnosis of an uncommon presentation of \textit{Chlamydia trachomatis} and recurrent and metastatic complications of \textit{Staphylococcus aureus} bacteremia before standard microbiology\(^9\).

The fact that in our cohort cfDNA NGS testing identified the cause of febrile illness in two patients with culture-negative sepsis who had a compatible clinical syndrome and responded well to antibiotic therapy supports the notion that NGS testing can be a useful diagnostic tool, particularly when conventional blood cultures are negative. The Karius\(^6\) Test pathogen-specific reference ranges have been established using cfDNA levels from healthy donors. Patient #8 had detectable levels of Torque teno virus, which belongs to \textit{Anelloviridae} family and is considered to lack pathogenic potential; this suggests the possibility that cfDNA NGS might on occasion yield detection of members of the commensal microbiota or viroma. To our surprise, however, even though many of the patients tested had mucosal barrier damage (e.g., mucositis) allowing for bacterial translocation from the gut, the Karius\(^6\) Test did not show a non-specific gut flora signal. The test was negative in patients in whom we failed to establish a microbiological diagnosis for their febrile illness, and when positive, typically correlated with conventional laboratory testing. Whether the currently defined cfDNA thresholds are optimal for identifying and quantifying pathogens of clinical relevance in highly vulnerable immunocompromised hosts will require further study. Importantly, the turnaround time for results was consistently within 48 hours, which is quite rapid considering that samples were shipped overnight from our institution located in Florida to the Karius Inc. laboratory in California.

Lack of control group, small number of patients and the heterogeneity of the cohort in terms of underlying diseases and causes of immunosuppression represent major limitations of this report. Larger clinical trials evaluating plasma NGS in patients with cancer and undergoing transplant are ongoing (NCT03226158, NCT03262584, NCT02912117, NCT02804464). Until larger cohort data becomes available, our observations suggest that detection of microbial cfDNA using NGS is valuable for the rapid noninvasive diagnosis of infectious complications following chemotherapy or transplantation.

**Conclusion**

Our data, along with a number of recent reports, support the clinical utility of the measurement of microbial cfDNA in peripheral blood using NGS for rapid noninvasive diagnosis of infections in immunocompromised hosts. As with other novel laboratory diagnostics used in clinical practice, the results of cfDNA NGS technology need to be interpreted with caution and in conjunction with other laboratory, radiological and clinical findings. Larger studies are needed to validate these findings.

**Data availability**

**Underlying data**

Microbial cfDNA NGS for Rapid Noninvasive Diagnosis of Infectious Diseases in Immunocompromised Hosts, BioProject accession number PRJNA554271.

---

**Grant information**

The author(s) declared that no grants were involved in supporting this work.

**Acknowledgments**

This work was supported by Karius, Inc., Redwood City, CA.
References


Open Peer Review

Current Peer Review Status: ✔

Version 1

Reviewer Report 04 September 2019

https://doi.org/10.5256/f1000research.21682.r52481

© 2019 Muller W. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

William Muller

1 Division of Infectious Diseases, Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, IL, USA
2 Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Camargo et al. report on results of plasma next-generation sequencing (NGS) for infectious diagnosis in a series of ten patients at risk for infection due to underlying immunocompromise. In addition to demonstrating concordance of NGS testing with conventional microbiologic diagnostic testing methods in seven of the ten patients, NGS testing identified a possibly causative organism in two cases in which conventional test results were negative. In one of the ten patients, PCR and culture detected *Aspergillus*; while NGS was able to detect this organism, it was below the assay detection limit. The authors conclude that this pilot study supports the conclusion that there may be clinical benefit for using this test in this population of patients, warranting more rigorous studies of test performance.

Overall, the manuscript does not contain significant flaws that would preclude indexing. The authors appropriately acknowledge the limitations of this small pilot study, which by design is not able to determine detailed test performance characteristics such as sensitivity or specificity. Accordingly, my comments are fairly minor and include:

- The methods and results sections report the NGS data in “molecules per microliter” (MPM), which is defined as “the number of DNA sequencing reads from the reported organism present per microliter of plasma.” Although this is a fairly straightforward concept, in the context of sequencing it is not totally clear how the more commonly used concepts of sequencing “breadth” and depth (e.g., as discussed by Sims et al., 2014) apply to MPM. Is it possible to explain MPM in a little more detail? E.g., is there a minimum depth of sequencing that needs to be satisfied for there to be "one" MPM? Does every base in a sequence need to have a certain number of reads?

- The methods also refer to the removal of possible “false-positive calls” from common environmental contaminants – can any examples of organisms that would fit these criteria be provided? Since immunocompromised patients may be at risk of infection from uncommon organisms associated with their environment (particularly uncommon moulds), it is important to understand what may not be reported.
• It’s not completely clear from the text whether results of NGS testing were used for clinical care. Although in most cases it is stated that an established diagnosis of infection was made prior to NGS testing, at least two patients had negative conventional testing and were reported to respond to therapy directed at the organisms identified by NGS. Were these patients responding to empiric treatment, or did NGS results direct the treatment?

• Several places in the manuscript refer to the abdominal infection in the kidney transplant patient as “deep-seated,” but only once is the more medically precise term “abscess” used to describe this infection. Also, can more details about this infection be provided? This is important as the patient appears to have had negative testing (unless the detection limit was lowered below that which the assay typically uses), providing additional clues to situations in which NGS testing may provide false negative results (more on this in the next comment).

• All of the patients had received some antibiotic (antibacterial and/or antifungal) treatment at the time of NGS testing, and one of the more impressive and useful aspects of the test is that it may be able to detect organisms in the setting of effective treatment. However, although a duration of treatment is reported in Table 1, whether this treatment would have been effective against the organism identified is not clear, complicating the interpretation of the data. For example, the Discussion notes that “several months” of treatment had been given to the patient with the *Aspergillus* abscess prior to NGS testing, but it is not totally clear if it was all directed against *Aspergillus* – if so, the negative result is less concerning (and the ability to detect the organism below the reportable threshold is still impressive). Similarly, did patient 10 receive 129 days of antifungal treatment with anti-*Aspergillus* activity prior to his positive test? Although it is understandable why prophylactic treatment would not generally be reported for all patients, was patient 4 receiving any prophylaxis against Pneumocystis?

There are a handful of fairly minor editorial corrections also, including:

• On page 4 the “clinically reportable range” is referred to as “our” clinically reportable range – although it is acknowledged that three of the authors are from the company which performed the NGS testing, given that certain details of the testing are ultimately proprietary perhaps a more generic statement (as simple as “the clinically reportable range”) would be preferred.

• In that same paragraph, although the Mandell textbook is considered by many to be the “go-to” reference for clinical infectious diseases, perhaps it could be referred to as “a clinical infectious diseases reference textbook” or some less subjective phrase? Also, it appears the 8th edition is being referenced, which I believe is actually from 2015 (not 2014)?

• The third paragraph in the left column on page 7 contains the phrase “there is limited data…” – as data is plural this should be corrected to “there are limited data…”.

References

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**  
Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**  
Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**  
Yes

**Are the conclusions drawn adequately supported by the results?**  
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Clinical pediatric infectious diseases with an emphasis on transplant ID; antimicrobial clinical trials in pediatric patients

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

**Author Response 11 Sep 2019**

**Jose Camargo**, University of Miami Miller School of Medicine, Miami, USA

**Reviewer’s comment:** Camargo *et al.* report on results of plasma next-generation sequencing (NGS) for infectious diagnosis in a series of ten patients at risk for infection due to underlying immunocompromise. In addition to demonstrating concordance of NGS testing with conventional microbiologic diagnostic testing methods in seven of the ten patients, NGS testing identified a possibly causative organism in two cases in which conventional test results were negative. In one of the ten patients, PCR and culture detected *Aspergillus*; while NGS was able to detect this organism, it was below the assay detection limit. The authors conclude that this pilot study supports the conclusion that there may be clinical benefit for using this test in this population of patients, warranting more rigorous studies of test performance.

Overall, the manuscript does not contain significant flaws that would preclude indexing. The authors appropriately acknowledge the limitations of this small pilot study, which by design is not able to determine detailed test performance characteristics such as sensitivity or specificity. Accordingly, my comments are fairly minor and include:

**Author’s response:** Thank you for this thorough review of the manuscript and your insightful comments that have improved the clarity and quality of the paper. Below there is a point-by-point response your inquiries.

**Reviewer’s comment:**

- The methods and results sections report the NGS data in “molecules per microliter” (MPM), which is defined as “the number of DNA sequencing reads from the reported organism present per microliter of plasma.” Although this is a fairly straightforward concept, in the
context of sequencing it is not totally clear how the more commonly used concepts of sequencing “breadth” and depth (e.g., as discussed by Sims et al., 2014) apply to MPM. Is it possible to explain MPM in a little more detail? E.g., is there a minimum depth of sequencing that needs to be satisfied for there to be “one” MPM? Does every base in a sequence need to have a certain number of reads?

**Author’s response:** We revised the text to clarify the connection between MPM and sequencing depth. In the Karius Test, the total number of reads observed per organism is often orders of magnitude lower than in applications where the concepts of sequencing breadth and depth are typically used (such as genome assembly). There is indeed a minimum amount of sequencing information that must be obtained in order for a sample to pass quality control criteria which is proportional to the concentration of cell-free DNA in that patient’s sample. Importantly, the MPM value is not affected by sequencing depth or human cell-free DNA concentration in the sample.

**Reviewer’s comment:**

- The methods also refer to the removal of possible “false-positive calls” from common environmental contaminants – can any examples of organisms that would fit these criteria be provided? Since immunocompromised patients may be at risk of infection from uncommon organisms associated with their environment (particularly uncommon moulds), it is important to understand what may not be reported.

**Author’s response:** The manuscript was revised to provide further clarity on process to guarantee that the Karius Test does not include frequently observed environmental contaminants. Here, it is important to mention that those taxa are not removed on a case-by-case basis but are completely out of scope for the whole test.

**Reviewer’s comment:**

- It’s not completely clear from the text whether results of NGS testing were used for clinical care. Although in most cases it is stated that an established diagnosis of infection was made prior to NGS testing, at least two patients had negative conventional testing and were reported to respond to therapy directed at the organisms identified by NGS. Were these patients responding to empiric treatment, or did NGS results direct the treatment?

**Author’s response:** Two aspects of the study should be considered here. i) This was a pilot, proof-of-concept, study in which the majority of patients had an established diagnosis prior to Karius test. ii) We found that NGS correlated well with results from standard diagnostic evaluation. Consequently, the NGS results did not largely influence clinical decision making in this cohort except perhaps in patient #5 where adenovirus PCR in blood was ordered (to confirm DNAemia and monitor viral kinetics which often guide initiation of antiviral therapy) after NGS yielded adenovirus; and in patient #8 in whom clinical presentation was very suggestive of gastrointestinal sepsis due to bacterial translocation in the setting of GVHD, and although blood cultures were negative decision was made to continue empiric antibiotic therapy to complete a course in view of clinical improvement and results of NGS testing.

**Reviewer’s comment:**

- Several places in the manuscript refer to the abdominal infection in the kidney transplant patient as “deep-seated,” but only once is the more medically precise term “abscess” used to describe this infection. Also, can more details about this infection be provided? This is important as the patient appears to have had negative testing (unless the detection limit was lowered below that which the assay typically uses), providing additional clues to situations in which NGS testing may provide false negative results (more on this in the next comment).
Author’s response: We edited the abstract and manuscript to avoid the term “deep-seated”. This was a 65-year-old patient who presented with a 1-month history of fever, generalized body aches, malaise and abdominal pain. Ten months prior, the patient received a kidney transplant complicated with perinephric abscess due to *Aspergillus fumigatus* that required multiple abdominal washouts and several months of antifungal therapy. Namely, at the time of admission when NGS was sent, the patient had received >6 months of voriconazole (serum levels: 0.5-6.1 mcg/ml; target: 1.5-5 mcg/ml), and micafungin was added for persistent fungal infection. We have added some of this clinical information to the footnote of Table 1. At the time of blood sample for NGS the patient was receiving dual antifungal therapy. We suspect this is the reason why the Karius test was negative. We describe the possibility of false negative results in patients receiving antifungal therapy in the discussion: “prolonged antifungal therapy prior to sample collection (e.g., >7-14 days) might interfere with detection of fungal DNA.”

Reviewer’s comment:
- All of the patients had received some antibiotic (antibacterial and/or antifungal) treatment at the time of NGS testing, and one of the more impressive and useful aspects of the test is that it may be able to detect organisms in the setting of effective treatment. However, although a duration of treatment is reported in Table 1, whether this treatment would have been effective against the organism identified is not clear, complicating the interpretation of the data. For example, the Discussion notes that “several months” of treatment had been given to the patient with the *Aspergillus* abscess prior to NGS testing, but it is not totally clear if it was all directed against *Aspergillus* – if so, the negative result is less concerning (and the ability to detect the organism below the reportable threshold is still impressive). Similarly, did patient 10 receive 129 days of antifungal treatment with anti-*Aspergillus* activity prior to his positive test? Although it is understandable why prophylactic treatment would not generally be reported for all patients, was patient 4 receiving any prophylaxis against Pneumocystis?

Author’s response: Your point is well taken. The kidney transplant patient received >6 months of anti-*Aspergillus* therapy and was receiving combination of voriconazole plus micafungin at the time of sample collection for NGS testing. Patient 10 had indeed received 129 days of anti-mold prophylaxis/empiric treatment with various agents (including posaconazole, isavuconazole, liposomal amphotericin B [L-AmB]), and was receiving combination of isavuconazole, L-AmB and micafungin at the time of NGS testing. We believe significant fungal burden in the setting of refractory leukemia and prolonged neutropenia facilitated the detection of fungal DNA despite triple antifungal therapy. Patient 4 was not receiving prophylaxis against *Pneumocystis* (atovaquone had been discontinued 3 months prior presentation since CD4>400) but had received 3 days of TMP-SMX treatment dose at the time of NGS testing. We have edited the discussion to clarify that all patients (except those with viral infections) were on active antimicrobial therapy at the time of NGS testing.

Reviewer’s comment:
There are a handful of fairly minor editorial corrections also, including:
- On page 4 the “clinically reportable range” is referred to as “our” clinically reportable range – although it is acknowledged that three of the authors are from the company which performed the NGS testing, given that certain details of the testing are ultimately proprietary perhaps a more generic statement (as simple as “the clinically reportable range”) would be preferred.
Author’s response: This has been modified.

Reviewer’s comment:
- In that same paragraph, although the Mandell textbook is considered by many to be the “go-to” reference for clinical infectious diseases, perhaps it could be referred to as “a clinical infectious diseases reference textbook” or some less subjective phrase? Also, it appears the 8th edition is being referenced, which I believe is actually from 2015 (not 2014)?

Author’s response: Text has been modified. 2014 is correct.

Reviewer’s comment:
- The third paragraph in the left column on page 7 contains the phrase “there is limited data…” – as data is plural this should be corrected to “there are limited data….”

Author’s response: Thank you. This grammar mistake has been corrected in the revised version.

Competing Interests: No competing interests were disclosed.