ARTICLE
https://doi.org/10.1038/s41564-018-0349-6

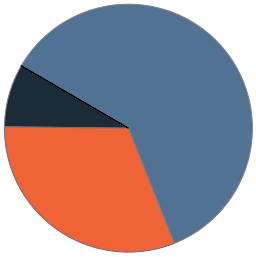
M

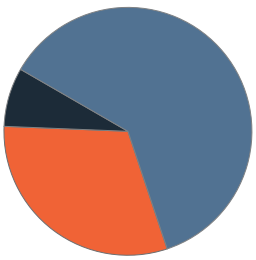
[1]Karius, Inc., Redwood City, CA, USA. [2]Department of Emergency Medicine, Stanford University, Palo Alto, CA, USA. [3]These authors contributed equally: Timothy A. Blauwkamp, Simone Thair, Judith C. Wilber, Samuel Yang. *e-mail: tim.blauwkamp@kariusdx.com

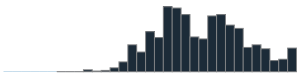
that ranged from kilobases to megabases in length. It also included both exclusive pathogens and frequent colonizers, as well as microorganisms that are commonly found as high-level environmental contaminants. Two pairs of closely related organisms, *Escherichia coli*/*Shigella flexneri* and *Staphylococcus aureus*/*Staphylococcus epidermidis*, were included to ensure fidelity of species discrimination during coinfections (Supplementary Table 2). Genomic DNA (gDNA) from each reference microbe was sheared to a typical microbial cfDNA fragment length and spiked into human plasma obtained from asymptomatic donors. As the concentration of human cfDNA in plasma can range over 1,000-fold, potentially affecting the sensitivity of sequencing-based assays[40,41], test performance was characterized in three different human plasma matrices representing 'low human', 'medium human' and 'high human' cfDNA samples (Fig. 2e).

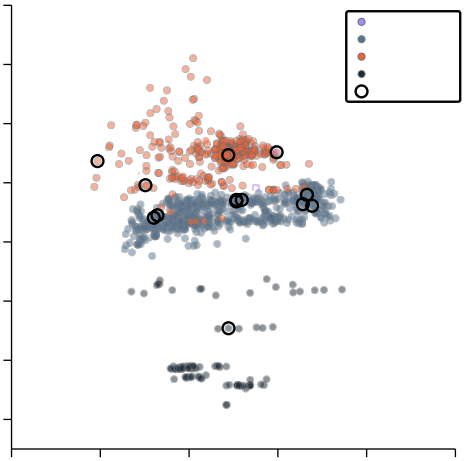
In addition to the experiments with reference microorganisms in the laboratory, performance metrics were also assessed in silico across a broader range of microorganisms than is possible in the clinical laboratory. The design of these in silico experiments tested performance in the face of genetic divergence between clinical isolates and reference genomes. This was accomplished primarily by blinding the analysis pipeline to the strain from which the simulated reads were drawn, thereby forcing the analysis to proceed in the absence of the exact strain match in the database. The conclusions drawn from the contrived and in silico samples were further verified using clinical samples where appropriate.

Analytical performance characterization. *Limit of detection.* The limit of detection (LoD) was determined for each of the thirteen reference microorganisms in each of the low-, medium- and high-human-DNA-plasma matrices (Fig. 3). At a typical sequencing depth, the LoD ranged from 33 to 74 molecules of microbe-specific cfDNA per microlitre of plasma (MPM) for most reference organisms in the low human plasma (Supplementary Table 3). The LoD range increased slightly to 39–103 MPM in the high human plasma. The performance of *Pseudomonas aeruginosa* differed from the other reference microorganisms in two ways. First, the LoD in the low human plasma was approximately tenfold higher than most other reference microorganisms, at 415 MPM. Second, the calculated LoD improved as the human cfDNA increased, decreasing to 132 MPM in the high-human-plasma matrix. Both of these aberrant properties of *P. aeruginosa* can be explained by the high levels of contaminating DNA from this species in the environment (Fig. 2d). The sensitivity of sequencing-based tests depends on the number of sequencing reads obtained, so we also sub-sampled every sample down to the minimum number of reads required to pass the quality controls (Supplementary Table 3). Sub-sampling to this level removed approximately 90% of the reads from each sample of typical depth and the LoD showed a corresponding approximately tenfold increase, ranging from 326 to 596 MPM for most organisms and from 4,159 to 1,341 MPM for *P. aeruginosa*, across all plasma backgrounds. The LoD at a variety of intermediate sequencing depths between full and minimum depth is shown in Fig. 3b

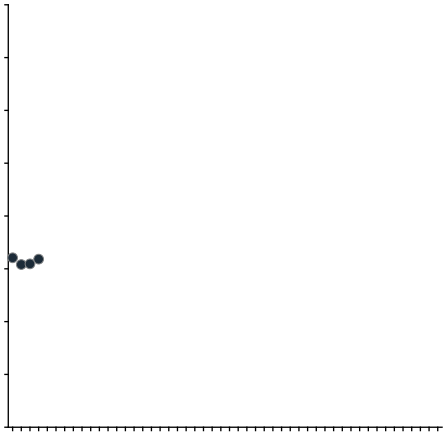








coefficient of variation lower than 50% while maintaining linearity with higher concentrations. The LoQ was calculated for each of the 13 reference microorganisms in each of the three human plasma backgrounds. Strong linearity was observed for all microorganisms in all human plasma matrices across the entire measured concentration range of 10 to 316,000



results available and were included in the analysis. The demographic and clinical characteristics of the enrolled patients are summarized in Supplementary Table 4. More than one-quarter (27.7%; 97 of 350) of the patients received antimicrobial treatment within two weeks preceding presentation. The mean length of patient hospital stay was 4.7 days, 6% required care in the intensive care unit during their stay and four patients died during hospitalization.

Compared with initial blood culture, cfDNA sequencing had a sensitivity of 93.7% (59 of 63; confidence interval (CI) of 84.5–98.2%; Table 2 and Fig. 5a). Discordant positive results included unculturable bacteria, bacteria from patients that were pre-treated with antimicrobials, viruses and eukaryotic pathogens. Additional microbiological testing over the first seven days of admission, including tissue and fluid cultures, serology, nucleic acid testing and subsequent blood cultures, identified 69 additional microbiological causes of the sepsis alert, of which microbial cfDNA sequencing identified 53, yielding an overall sensitivity of 84.9% ($n = 112$ of 132) compared with all microbiological testing. The discordant positive results from this comparison were adjudicated according to the criteria outlined in Methods and Supplementary Fig. 8 to generate a composite reference standard for the aetiology of the sepsis alert. In comparison to this composite reference standard, cfDNA sequencing demonstrated

Diversity robustness

	Bioinformatic cross-reactivity	Diversity robustness near the LoD	Unconstrained diversity robustness
Simulations	1,250	125	125
Diversity robustness	Exact match	No exact match; ≥1 assembly; <3% divergence	No exact match; ≥1 assembly; unconstrained divergence
Simulated infection level	High	Near LoD	High
PPV (%)	99.4 (*n*=1,250 of 1,257)	99.2 (*n*=121 of 122)	92.1 (*n*=117 of 127)
Specificity per analyte (%)	99.9995 (1,561,243 true negative calls of 1,561,250 total negative calls)	99.9994 (156,124 true negative calls of 156,125 total negative calls)	99.994 (156,115 true negative calls of 156,125 total negative calls)

Cross-reactivity was assessed using in silico simulations to measure the rate of false positive calls at a variety of simulated infection levels and with varying degrees of genetic distance between the infecting microbe and the genomes present in the database to simulate genetic diversity of clinical isolates. For bioinformatic cross-reactivity experiments, the infecting microbe exactly matched an assembly in the database. For diversity rubustness near the LoD, the infecting microbe did not have an exact assembly match in the database, but did have at least one assembly in the database less than 3% diverged from the infecting microbe. For unconstrained diversity robustness, the infecting microbe did not have an exact assembly match in the database, but did have at least one assembly in the database with the same species label (no limit on divergence).

Positive and negative agreement of microbial cfDNA sequencing versus initial blood culture, all microbiological testing and composite reference standard

Patient characteristics (*n*=348)	NGS positive	NGS negative	Agreement (%)	95% CI (%)
Positive by initial blood culture	59	4	93.7	84.5–98.2
Negative by initial blood culture	171	114	40.0	34.3–45.9
Positive by all microbiological testing	112	20	84.8	77.6–90.5
Negative by all microbiological testing	112	104	48.2	44.3–55.0
Positive by composite reference standard	169	13	92.9	88.1–96.1
Negative by composite reference standard	62	104	62.7	54.8–70.0

The composite reference standard includes the results from all microbiological tests (including the initial blood culture) performed within seven days of presentation and clinical adjudication. The NGS false negatives compared to initial blood culture included *Listeria monocytogenes*, coagulase-negative *S. aureus*, *Streptococcus agalactiae* and *Stenotrophomonas maltophilia* (this organism was not included in the NGS-test reportable range). NGS agreement with other methods was calculated as described in Supplementary Figures 7 and 8.

a sensitivity of 92.9% (169 of 182; CI of 88.1–96.1%; Table 2 and Supplementary Table 5). Overall, the identification of sepsis alert aetiology was higher for cfDNA sequencing (169 of 348) than for both blood culture (63 of 348) and all microbiological testing combined over seven days (132 of 348). Of the 96 subjects that received antimicrobial treatment within two weeks preceding presentation, cfDNA sequencing identified pathogens that are classified as definite or probable causes of the sepsis alert in 46 (47.9%), whereas blood culture identified pathogens in only 19 (19.6%).

Among the 166 samples for which no definite or probable microbiological cause of the sepsis alert was identified, no significant microbial cfDNA was detected in 104 by cfDNA sequencing, resulting in a specificity for sepsis alert aetiology of 62.7% (104 of 166; CI of 55.2–70.4%; Table 2). The organisms identified by cfDNA adjudicated as possible or unlikely causes of the sepsis alert included a number of reactivated herpesviruses, chronic infections such as *H. pylori* and human papillomavirus, microorganisms likely to be commensals and possible causes of non-sepsis-related acute infection (Supplementary Tables 6 and 7).

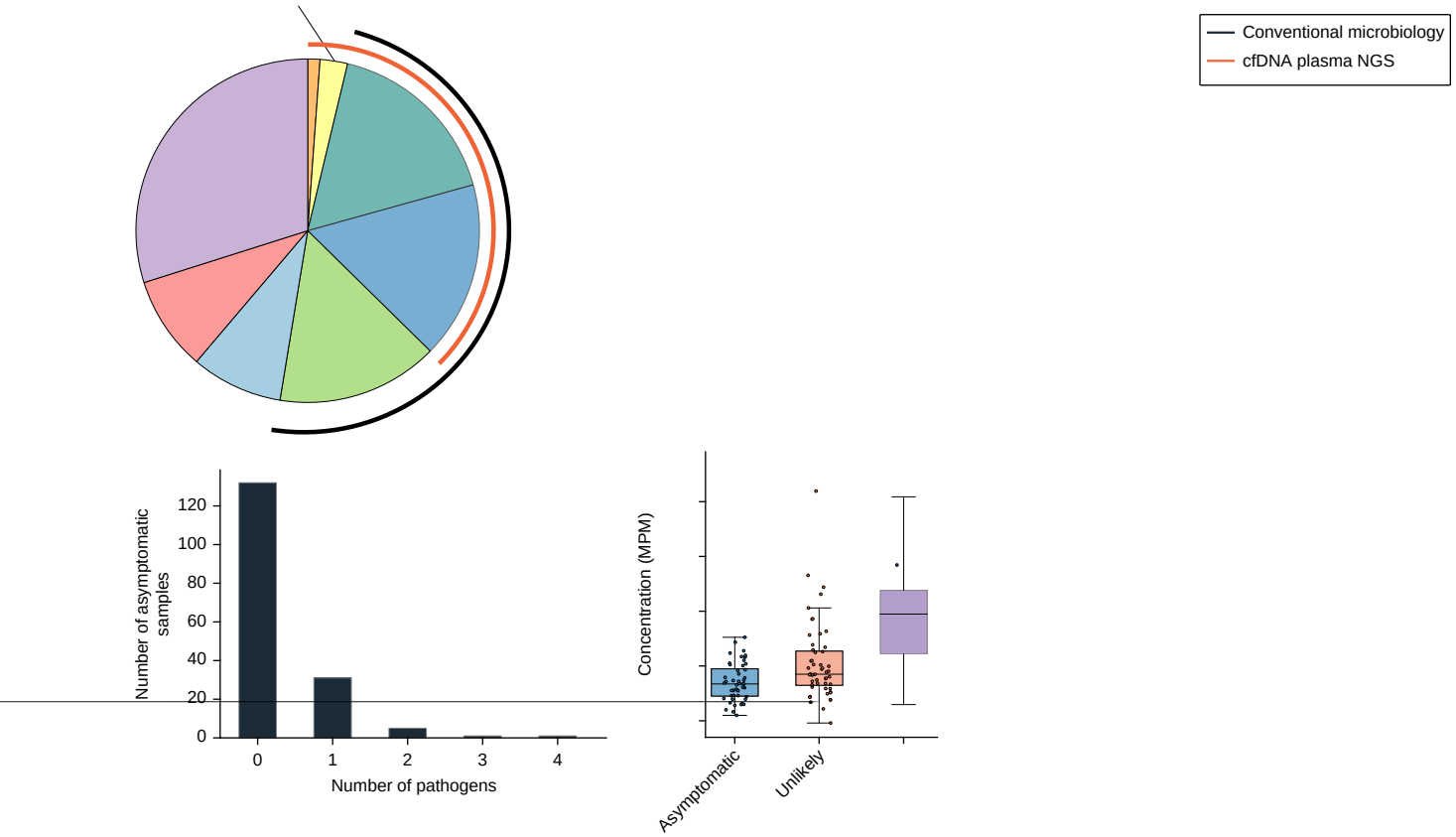
For samples with a composite reference positive result, the estimated time to result for cfDNA sequencing was compared with conventional testing (Fig. 5b). For cfDNA sequencing, the time to result was represented using the median shipping and testing times for the last 400 samples run in the Clinical Laboratory Improvement Amendments of 1988 (CLIA)-certified lab. This time to result (53.0 h) was significantly shorter than the median time to positive result for conventional testing, based on the electronic medical records (EMR) (92.4 h; $P=0.0004$; Fig. 5b). Although cfDNA sequencing did not provide antimicrobial susceptibility testing, this analysis demonstrates that the sensitivity and speed with which species-level identification is provided by cfDNA sequencing may offer significant benefit to patients.

As microbial cfDNA is a new biomarker, additional context was provided by assessing test performance on 167 asymptomatic donors. Of the asymptomatic samples, 77.2% (129 of 167) had no microorganisms reported (Fig. 5c). Among the 22.8% of samples in which microbial cfDNA was reported, a single species was detected in most cases. In general, the concentrations of microbial cfDNA detected from asymptomatic donors were lower than the concentrations of microbial cfDNA detected from patients with confirmed infecting pathogens and similar to the concentration of microorganisms classified as unlikely causes of the sepsis alert (Fig. 5d). The most frequently detected microorganisms in these patients included *H. pylori*, *Klebsiella pneumoniae* and *Haemophilus influenzae*; the latter two are commonly present as human commensals (Supplementary Table 8).

Clinical laboratory experience with the first 2,000 samples. We characterized test performance across the first 2,000 patient plasma samples from across the United States submitted to our CLIA-certified, College of American Pathologists (CAP)-accredited laboratory for microbial cfDNA sequencing (Table 3). Due to collection-site handling errors or excessive shipping delays, 46 samples were not tested. Of the samples tested, 98.1% were reported, with 87.6% of the reports delivered the operating day after on-time sample receipt. Microorganisms were reported in 53.7% of all of the tested samples, covering 318 different species of bacteria, fungi, parasites and viruses. Among the positive reports, 49.6% reported a single organism, with the remainder reporting two or more organisms (Supplementary Fig. 9). Although no patient-specific information was used in determining which microorganisms to report, the analysis of passively collected ICD-10 codes associated with approximately half of the samples suggested that the most common use for the test was in immunocompromised patients, followed by sepsis, endocarditis and complicated pneumonia.

Discussion

The results presented here show that microbial cfDNA sequencing offers the potential to reliably identify a wide variety of infections



throughout the body from a plasma sample in a clinically useful time frame. A number of challenges to the provision of high-quality diagnostic testing for clinical metagenomics applications have been previously identified[17,33,34] and we present additional considerations. Here, we attempt to address all of these challenges through a combination of traditional and metagenomic-specific validation strategies, including the use of 13 representative microorganisms that were chosen to probe the boundaries of analytical performance in 348 contrived samples, thousands of in silico simulations that tested the integrity of the bioinformatics pipeline in the face of clinical-isolate divergence, 580 clinical samples that assessed performance in patient samples and 2,000 samples that were run through our CLIA-certified laboratory and reported in real-time.

Having developed the test for low bias, we did not observe significant differences in analytical performance as a result of differing GC-content, genome size, superkingdom, genetic similarity among coinfecting organisms or differences of up to 3% between the detected strains and reference genome. Elevated levels of human cfDNA background in the sample had only minor effects on sensitivity and precision. The level of environmental contamination did influence test sensitivity, but only for the 5–8 microorganisms with the highest environmental backgrounds among 1,250 microorganisms probed. The sequencing depth also influenced LoD, but the processing methods for this test are designed to provide similar sequencing depth for all samples, such that 95% of the samples tested fell into a range of sequencing depths where sensitivity was consistent.

The positive agreement between this test and blood culture (93.7%) of patients with a sepsis alert is equal to or better than other direct molecular diagnostic methods, including real-time PCR panels and PCR combined with electrospray ionization[25,26]. Adjudication of results from the SEP-SEQ study also showed that this test identified a greater number of aetiological causes of the sepsis alert than standard-of-care testing. However, the sensitivity and breadth of microorganisms detected, combined with the diversity of the microbiome compositions across patients[17,34], makes it challenging to achieve high diagnostic specificity. The analytical validation experiments demonstrated very low levels of falsely reporting of microbial cfDNA that was not in the original plasma sample (Fig. 4), consistent with high reproducibility of cfDNA detection across independent

runs (Supplementary Fig. 5). Therefore, the microbial cfDNA identified in 22.8% of the asymptomatic samples is probably derived from the original donor specimen, representing cfDNA primarily from commensal organisms or subclinical colonization (Supplementary Table 8). Similarly, the cfDNA species reported in the possible or unlikely SEP-SEQ samples (37.3% of negative composite reference samples) is presumed to originate from these sources in addition to incidental findings unrelated to the sepsis alert.

Although it would be convenient if the cfDNA concentration alone were indicative of true infection, both the microbe identity and the location of infection are likely to influence the concentration

NCBI) with publications supporting pathogenicity. Organisms from the above list that were associated with high-quality reference genomes, as determined by our reference database quality control 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 final list was defined as the CRR of this test (Supplementary Table 1; complete list available at www.kariusdx.com/pathogen-list, v3.1.1).

Analytical validation. *Reference materials.* Genomic DNA from 14 microorganisms was obtained from either the ATCC or NIST. Because the human mastadenovirus B genome was available only in small quantities, larger amounts were produced by seven non-overlapping PCR amplicons of approximately 5 kb each. Enzymatic shearing of each reference microbe genome was accomplished with DNaseI or Fragmentase (New England Biolabs) to create semi-randomly fragmented gDNA. Sheared gDNA was puri ed using Oligo Clean and Concentrator (Zymo Research) and quanti ed by uorometery (Qubit, ermoFisher Scienti c). e fragment length distributions of the sheared gDNAs were evaluated by electrophoresis (TapeStation 2200, Agilent) and optimized to obtain consistent ranges across the 14 genomes (60–90 bp modal length), corresponding to the distribution of pathogen cfDNA found in clinical samples[48]. Quanti ed sheared gDNA was spiked into plasma pooled from 8–10 healthy donors (ZenBio) to create the low human contrived samples. Two additional healthy human plasma pools were generated by spiking the low-human-plasma pool with puri ed human mononucleosomes (EpiCypher) to simulate samples with di erent amounts of human cfDNA, for example, the medium- and high-human-plasma matrices. As measured with QuantIT PicoGreen (ermoFisher Scienti c), the extracted cfDNA levels averaged 0.23 ng μl^{-1} for low- (no added mononucleosomes), 0.81 ng μl^{-1} for medium- and 1.86 ng μl^{-1} for high-human-plasma matrices. e human cfDNA concentration was 100 to 1,000,000 times more abundant than sheared pathogen gDNA in the contrived samples used for this study. One of the microorganisms used for validation, *Clostridium sporogenes*, is not a member of our reportable range at present due to its confounded phylogenetic relationship with *Clostridium botulinum*[49] and was therefore excluded from all analyses. e puri ed human mononucleosomal DNA (EpiCypher) was found to contain signi cant levels of contaminating *E. coli* DNA and therefore analyses of *E. coli* in the medium- and high-human plasma were excluded from all analyses.

LoD. An LoD value was estimated for each of the 13 representative pathogens in high-, medium- and low-human-background levels at a variety of sequencing depths. Sheared gDNA from each of the representative pathogens were mixed at nominally equivalent molar concentrations, spiked at 10,000 MPM into healthy human plasma and diluted with healthy human plasma over 7 0.5-log serial dilutions ranging from 10,000 to 10

Clinical validation. The SEP-SEQ study (NCT02730468) was a prospective observational study including patients presenting to the Emergency Department at the Stanford University Medical Center identified by a sepsis alert triage model designed for the early identification and intervention in patients with sepsis[51]. The study protocol and patient informed consent were reviewed and approved by the University's IRB (Stanford) and the study was conducted in compliance

13. Kothari, A., Morgan, M. & Haake, D. A. Emerging technologies for rapid identi cation of bloodstream pathogens. *Clin. Infect. Dis.*

nature research

Corresponding author(s): Timothy A. Blauwkamp

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☐	☒ The exact sample size ( ) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *O* *M*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

Data

Policy information about