nature microbiology

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

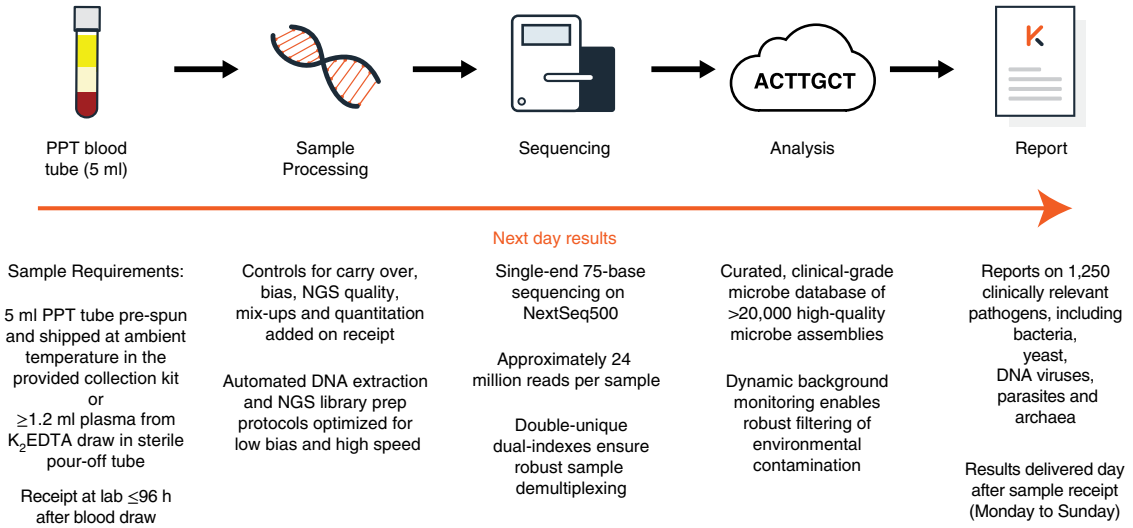


Fig. 1 | The Karius test workflow. Blood plasma from a routine draw is isolated and shipped overnight at ambient temperature to the Karius CLIA/CAP laboratory. Sample-specific controls are added on receipt and an automated liquid-handling platform performs cfDNA extraction and NGS-library preparation. The NGS libraries are multiplexed, inspected for quality and sequenced. A custom-built analysis pipeline uses a clinical-grade database to identify microbial DNA fragments found in plasma. Pathogens with plasma DNA levels that are significantly higher than real-time background thresholds are listed on the patient report, along with the concentration of the microbial cfDNA in plasma.

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, [illegible] and [illegible], 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).

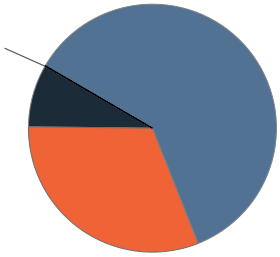
In addition to the experiments with reference microorganisms in the laboratory, performance metrics were also assessed in silico across a broader range of microorganisms than is possible in the clinical laboratory. The design of these in silico experiments tested performance in the face of genetic divergence 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. [illegible]. 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 [illegible] 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 [illegible] 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 [illegible], across all plasma backgrounds. The LoD at a variety of intermediate sequencing depths between full and minimum depth is shown in Fig. 3b. Of all of the samples run in the analytical validation, 95% had sequencing coverage of at least 181,000 unique whole-assay internal normalization control (WINC) molecules, above which the median LoD of all reference microorganisms in all human plasmas varied from 34 to 74 MPM.

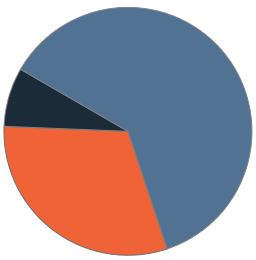
Although it was feasible to measure the LoD directly for the 13 reference microorganisms, our reportable range spans 1,250 taxa with a potentially broader range of sensitivities, due, for example, to variations in the level of background DNA or divergence between clinical strains and our reference database. We designed an in silico experiment to test that the LoDs estimated above were representative of the assay sensitivity across the broader clinical reportable range (CRR). Briefly, we selected 125 microorganisms from the CRR at random and simulated a number of reads that corresponded to the LoD. These reads were added to reads from a sequenced library of human cfDNA and run through our analytical pipeline while blinded to the assembly from which the simulated reads were generated. The simulated organism was correctly identified in 121 of 125 simulations for a sensitivity of 97% and the positive predictive value (PPV) was 99% (121 of 122), consistent with the expected 95% sensitivity at the LoD (Fig. 3c).

[illegible]. The limit of quantitation (LoQ) for this test was defined as the lowest nominal input concentration of microbial cfDNA that had a precision corresponding to a

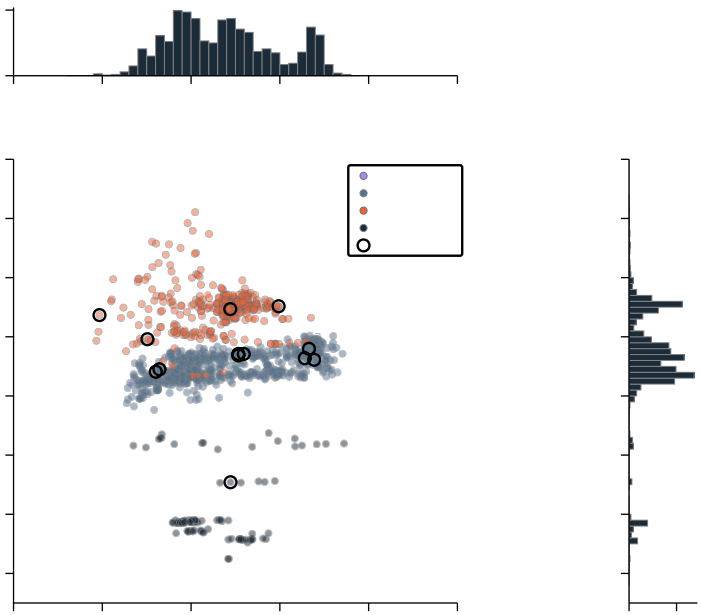
a



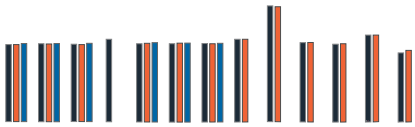
b



c



d



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

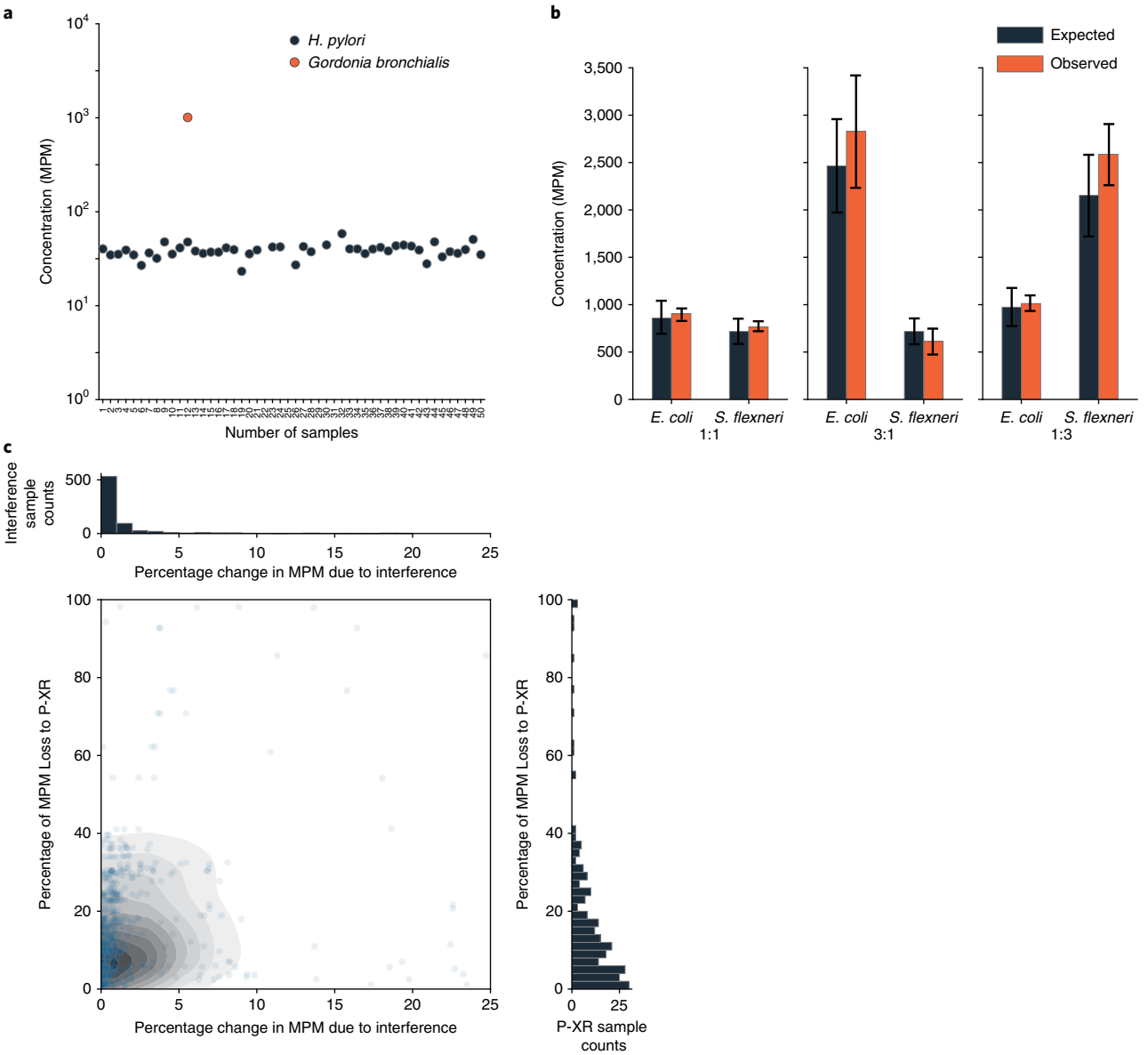


Fig. 4 | Analytical specificity. a, Analytical specificity as a function of environmental contamination was investigated using repeated measurements of a single, thoroughly characterized healthy plasma sample composed of eight different donors. The results from 50 replicate measurements taken across 9 days are shown. *H. pylori* cfDNA was known to be present in this sample, whereas other microorganisms were not. **b**, Potential interference during coinfection was assessed using healthy human plasma samples spiked with various ratios of genetically similar microbial cfDNA reference materials. The reported plasma cfDNA concentrations are shown for *E. coli* and *S. flexneri* when coinfection was contrived at ratios of 1:1, 3:1 and 1:3. The expected plasma cfDNA concentration was calculated from samples where each organism was spiked into a sample alone. The mean ± s.d. from three replicates is shown. $P > 0.05$ for all comparisons ($P = 0.58$, 0.56 and 0.13 for *E. coli* and 0.41, 0.13 and 0.32 for *S. flexneri*, at 1:1, 3:1 and 1:3 ratios, respectively), two-tailed *t*-tests. **c**, In silico assessment of accuracy in coinfected samples. The *y* axis shows the bias in concentration due to pathogen cross-reactivity (P-XR) in mono-infection and the *x* axis shows the additional bias due to interference during coinfection. Histograms show the number of in silico simulated samples with various levels of MPM bias due to pathogen cross reactivity (right) and interference from co-infecting organisms (above).

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% (= 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

Table 1 | 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).

Table 2 | 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 [illegible] and human papillomavirus, microorganisms likely to be commensals and possible causes of non-sepsis-related acute infection (Supplementary Tables 6 and 7).

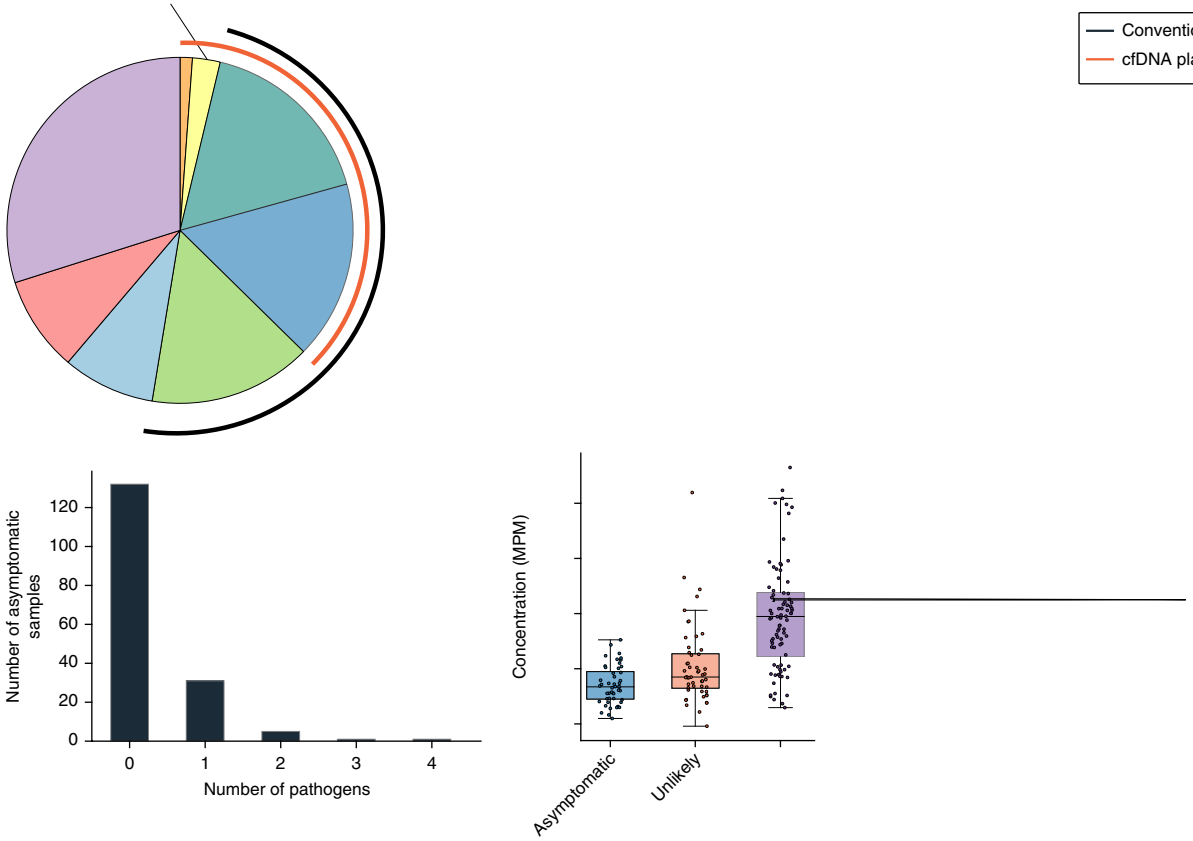
For samples with a composite reference positive result, the estimated time to result for cfDNA sequencing was compared with conventional testing (Fig. 5b). For cfDNA sequencing, the time to result was represented using the median shipping and testing times for the last 400 samples run in the Clinical 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; [illegible] =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 [illegible] and [illegible]; 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 of microbial cfDNA observed in plasma. The cfDNA concentrations associated with definite calls spanned four orders of magnitude (20–450,000 MPM). The cfDNA concentrations of microorganisms that were detected in asymptomatic donors, however, did not exceed 300 MPM and investigation of the unlikely causes of sepsis alert with concentrations greater than 300 MPM showed that these were primarily reactivated herpesviruses, chronic infections with human papillomavirus and probable causes of acute infection. Therefore, high concentrations of microbial cfDNA were typically associated with true infections, whereas low concentrations were associated with both true infections and commensal/colonizer/contaminant microorganisms of unknown clinical significance. Regardless of the cfDNA concentrations, the entire clinical picture must be considered when determining the clinical significance of a microbe detected by cfDNA sequencing.

The performance characteristics reported here point towards clinical applications in which the benefits of sensitivity, non-invasive sampling and broad testing outweigh the limitations of sequencing cost, turn-around time and identification of microorganisms beyond those related to a specific indication. In light of these trade-offs, use cases that could potentially benefit from this approach include rule-out testing for sepsis when the standard of care was not sensitive enough or too invasive, testing immunocompromised patients presenting with non-specific symptoms of infection or testing before invasive procedures with high cost or morbidity. Future studies that directly examine clinical utility will be important to establish the specific indications for which microbial cfDNA sequencing should be used.

Methods

Clinical-grade microbial cfDNA sequencing for infectious disease. . All samples were processed with locked and version-controlled protocols, analytical pipelines and reference databases in the Karius CLIA-certified and CAP-accredited laboratory. Following receipt, plasma samples were thawed, if frozen, spiked with a known concentration of synthetic normalization molecule controls (see Methods below) and centrifuged at 16,000g for 10 min to remove residual cells. Cell-free DNA was extracted from 0.25 ml plasma using a modified Mag-Bind cfDNA Kit (Omega Biotek) in Hamilton STAR liquid handling workstations. DNA libraries for sequencing were constructed using customized dual-indexed Ovation Ultralow System V2 library preparation kits (NuGEN) in Hamilton STAR liquid handling workstations. Sequencing libraries were pooled with environmental and assay control samples that were processed alongside the test samples in every batch. Pooled sample libraries were purified and up to 24 libraries per batch were multiplexed and sequenced on Illumina NextSeq500 sequencers using a 75-cycle single-end, dual index sequencing kit. On average, approximately 24 million reads were obtained for each sample.

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. . 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 purified using Oligo Clean and Concentrator (Zymo Research) and quantified by fluorometery (Qubit, ThermoFisher Scientific). The 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]. Quantified 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 purified human mononucleosomes (EpiCypher) to simulate samples with different amounts of human cfDNA, for example, the medium- and high-human-plasma matrices. As measured with QuantIT PicoGreen (ThermoFisher Scientific), 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. The 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, , is not a member of our reportable range at present due to its confounded phylogenetic relationship with [49] and was therefore excluded from all analyses. The purified human mononucleosomal DNA (EpiCypher) was found to contain significant levels of contaminating DNA and therefore analyses of in the medium- and high-human plasma were excluded from all analyses.

. 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 MPM per microbe. Twelve replicates of each dilution in each of the high-, medium- and low-human-plasma matrices were analysed over 12 different days according to the standard workflow. Probit analysis of 12 replicates at each concentration in each healthy plasma matrix was used to establish the LoD for each reference microbe in each plasma matrix as follows. We assumed that sensitivity (the probability of detection of each replicate) varies with log concentration according to a cumulative normal distribution, that is, a Probit model. This model implied that the number of positives at each concentration followed a binomial distribution and the whole dilution series had a likelihood given by the products of the binomial distributions for each concentration. The parameters of the cumulative normal distribution associating concentration with sensitivity were fitted to maximize the likelihood of the dilution series and the LoD determined by inverting the fitted distribution to determine the log concentration corresponding to a sensitivity of 95%. Note that for this experiment, we did not apply our cross-reactivity call filters due to the special nature of these contrived samples (highly multiplexed with all organisms at nominally equal concentrations).

The variation in LoD as a function of sequencing coverage was determined by sub-sampling to a level where a single unique pathogen read would correspond to MPMs of 12, 6, 4, 3, 2, 1.5, 1.2, 1 and 0.6 (or unique WINC counts of 25,000, 50,000, 75,000, 100,000, 150,000, 200,000, 250,000, 300,000 and 500,000 unique reads). We used the complete fastq in each replicate while performing microbial abundance estimation and then multiplied each taxon abundance by the ratio of the targeted WINC count to the observed WINC count before the determination of statistical significance over background.

. After computing the LoD for each combination of microorganism and human background level, we performed an in silico experiment to estimate the LoD across 125 additional microorganisms. We began by determining which taxa on the CRR corresponded to at least one pair of assemblies that were at most 3% diverged (as estimated from overlaps in -mer content). Then, for each of a random selection of 125 of these taxa, we selected an assembly at random from those that have at least one similar assembly. To determine the number of reads of this assembly to simulate, we selected one of the 13 representative pathogens and a human background level at random to represent the wet-laboratory efficiency of capturing and sequencing reads of the simulated assembly. We found the lowest concentration at which the test had at least 95% sensitivity for this pathogen–background combination at the minimum sequencing depth. We then randomly selected one of the (sub-sampled) technical replicates at this concentration and determined the number of reads of the pathogen. Next, we simulated the same number of reads from the assembly chosen for simulation, selecting for each read a random start location and orientation, and a length drawn from the distribution observed in the wet laboratory representative. These reads were then added to those of a sequenced library of the low human cfDNA matrix without any spiked pathogens and run through our analytical pipeline blinded to the simulated assembly by discarding any alignments to it before abundance estimation. Finally, we determined whether the CRR taxon of the simulated assembly was reported by the analytical pipeline.

. Precision samples were contrived in the low-human-plasma matrix with sheared microbial DNA from all 13 reference microorganisms spiked at a nominal concentration of 1,000 MPM. One-hundred replicate samples were generated and stored frozen at −80 °C until the day of testing. Two aliquots were thawed and processed according to the standard workflow on each day of testing, including independent addition of internal control mixtures on each day. Precision was calculated according to the Clinical and Laboratory Standards Institute EP05-A2, Appendix C guidelines for establishing within-run and within-laboratory precision using one batch a day[50].

. The best fits for our analysis of linearity were determined as follows. We sought to fit the data on a log scale (appropriate to dilution series) and to ensure that, on a linear scale, the fit was linear and intersected the origin. To achieve both of these objectives, we log-log transformed these data (consisting of nominal and observed concentrations) and then performed a best linear fit with a fixed slope of unity.

. The extent of cross-reactivity between closely related species during coinfection was assessed first using pre-analytical mixtures. Contrived coinfections in healthy human plasma were generated using two pairs of genetically similar and clinically relevant microorganisms, with or with . Each pair of microorganisms was tested at concentration ratios of 1:3, 1:1 and 3:1. Triplicates of each contrived coinfection sample were tested and the MPM values averaged. The change in MPM of each organism, relative to the MPM observed when each organism was present alone in the sample, was calculated.

In silico experiments testing interference were divided into two sections, the first tested for inaccuracy of the observed concentration due to cross-reactivity of mono-infecting microorganisms with a similar taxon in the database and the second evaluated the accuracy of the observed concentration in the presence of closely related coinfecting microorganisms due to interference.

First, 125 pairs of assemblies from different CRR taxa that belonged to the same family were selected (250 assemblies total), where each assembly used for simulation was less than 3% divergent from another assembly in the same taxon (as estimated by -mer distances). Then, we spiked 10,000 simulated reads from each of the 250 assemblies into different healthy human plasma samples down-sampled to the minimum sequencing coverage allowed by the quality control criteria. We then ran these simulated mono-infection samples through our analytical pipeline, blinded to the spiked assemblies to better simulate the diversity expected from clinical isolates.

Second, we assessed the fraction of reads that were lost to interference from a coinfection with a related organism. Reads were drawn randomly for each assembly in three ratios (1:3, 1:1 and 3:1) and then spiked into a healthy human plasma sample that was sub-sampled to the unique WINC count minimum. These read sets were run through the analytical pipeline blinded to both assemblies. Assemblies (125 pairs) were run, each with 1:1, 3:1 and 1:3 ratios. The fraction of reads lost due to interference during coinfection is the difference between the concentration observed during coinfection versus mono-infection, divided by the concentration observed during mono-infection.

. To determine whether MPM is dependent on sequencing depth, we sub-sampled sequencing reads from a sample containing all 13 reference microorganisms in a medium human background and a unique WINC molecule count of 600,581. We then sampled (without replacement) the fraction of reads that resulted, on average, in unique WINC counts of 25,000, 50,000, 100,000, 250,000 and 500,000. We performed this sub-sampling ten times for each of the unique WINC target counts and re-executed the analytical portion of the test on these 50 collections of sampled reads. The results are shown in Supplementary Fig. 3b where we average the MPMs of the representative pathogens across the ten replicates at each sampling level.

. Samples that were utilized in the quantitative evaluation of the assay included plasma from 25 CMV positive individuals obtained from a biorepository of de-identified remnant blood samples (Discovery Life Sciences) annotated with quantitative PCR test results (Roche cobas CMV, Roche Molecular Systems). All samples were obtained through Institutional Review Board (IRB)-approved protocols with appropriate informed consent provided for sample use (Schulman IRB).

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 identification of bloodstream pathogens. **59**, 272–278 (2014).
14. Simner, P. J., Miller, S. & Carroll, K. C. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. **66**, 778–788 (2018).
15. Greninger, A. L. et al. Clinical metagenomic identification of encephalitis and assembly of the draft genome: the continuing case for reference genome sequencing. **7**, 113 (2015).
16. Wilson, M. R. et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. **370**, 2408–2417 (2014).
17. Schlaberg, R. et al. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. **141**, 776–786 (2017).
18. Naccache, S. N. et al. Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. **60**, 919–923 (2015).
19. Stokowski, R. et al. Clinical performance of non-invasive prenatal testing (NIPT) using targeted cell-free DNA analysis in maternal plasma with microarrays or next generation sequencing (NGS) is consistent across multiple controlled clinical studies. **35**, 1243–1246 (2015).
20. Song, K., Musci, T. J. & Caughey, A. B. Clinical utility and cost of non-invasive prenatal testing with cfDNA analysis in high-risk women based on a US population. **26**, 1180–1185 (2013).
21. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. **105**, 16266–16271 (2008).
22. Schutz, E. et al. Graft-derived cell-free DNA, a noninvasive early rejection and graft damage marker in liver transplantation: a prospective, observational, multicenter cohort study. **14**, e1002286 (2017).
23. Bloom, R. D. et al. Cell-free DNA and active rejection in kidney allografts. **28**, 2221–2232 (2017).
24. De Vlaminck, I. et al. Noninvasive monitoring of infection and rejection after lung transplantation. **112**, 13336–13341 (2015).
25. De Vlaminck, I. et al. Circulating cell-free DNA enables noninvasive diagnosis of heart transplant rejection. **6**, 241ra277 (2014).
26. Snyder, T. M., Khush, K. K., Valantine, H. A. & Quake, S. R. Universal noninvasive detection of solid organ transplant rejection. **108**, 6229–6234 (2011).
27. Aravanis, A. M., Lee, M. & Klausner, R. D. Next-generation sequencing of circulating tumor DNA for early cancer detection. **168**, 571–574 (2017).
28. Lanman, R. B. et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. **10**, e0140712 (2015).
29. Bettegowda, C. et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. **6**, 224ra224 (2014).
30. Dawson, S. J. et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. **368**, 1199–1209 (2013).
31. Abril, M. K. et al. Diagnosis of sepsis by whole-genome next-generation sequencing. **3**, ofw144 (2016).
32. Hong, D. K. et al. Liquid biopsy for infectious diseases: sequencing of cell-free plasma to detect pathogen DNA in patients with invasive fungal disease. **92**, 210–213 (2018).
33. Lefterova, M. I., Suarez, C. J., Banaei, N. & Pinsky, B. A. Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. **17**, 623–634 (2015).
34. Dunne, W. M. Jr, Westblade, L. F. & Ford, B. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. **31**, 1719–1726 (2012).
35. Kim, D. et al. Optimizing methods and dodging pitfalls in microbiome research. **5**, 52 (2017).
36. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. **12**, 87 (2014).
37. Weiss, S. et al. Tracking down the sources of experimental contamination in microbiome studies. **15**, 564 (2014).
38. Naccache, S. N. et al. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. **87**, 11966–11977 (2013).
39. (Food and Drug Administration, 2016).
40. Chang, C. P. et al. Elevated cell-free serum DNA detected in patients with myocardial infarction. **327**, 95–101 (2003).
41. Lo, Y. M., Rainer, T. H., Chan, L. Y., Hjelm, N. M. & Cocks, R. A. Plasma DNA as a prognostic marker in trauma patients. **46**, 319–323 (2000).
42. Vincent, J. L., Martinez, E. O. & Silva, E. Evolving concepts in sepsis definitions. **23**, 29–39 (2011).
43. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **30**, 2114–2120 (2014).
44. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. **9**, 357–359 (2012).
45. Camacho, C. et al. BLAST+: architecture and applications. **10**, 421 (2009).
46. Xia, L. C., Cram, J. A., Chen, T., Fuhrman, J. A. & Sun, F. Accurate genome relative abundance estimation based on shotgun metagenomic reads. **6**, e27992 (2011).
47. Bennett J. E., Dolin, R. & Blaser, M. J. 8th edn (Saunders, Philadelphia, 2015).
48. Burnham, P. et al. Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. **6**, 27859 (2016).
49. Kalia, V. C. et al. Analysis of the unexplored features of (16S rDNA) of the Genus **12**, 18 (2011).
50. 2 (NCCLS, 2004).
51. Dellinger, R. P. et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock, 2012. **39**, 165–228 (2013).
52. Bagdasarian, N., Rao, K. & Malani, P. N. Diagnosis and treatment of in adults: a systematic review. **313**, 398–408 (2015).
53. Focosi, D., Antonelli, G., Pistello, M. & Maggi, F. Torquetenovirus: the human virome from bench to bedside. **22**, 589–593 (2016).
54. Karius, Inc. https://www.kariusdx.com/pathogen-list/3.1.1 (2018).

Acknowledgements

The authors would like to thank S. Sinha for assistance with the preparation of the manuscript, as well as H. Quach, R. Davila, S. Madan, V. Baichwal, C. Ho, H. Seng, R. Aquino, A. Parham, R. Mann, I. Brown, P. Callagy, A. Visweswaran, C. Keller, C. Bucsit and A. Araya for their contributions to these validation studies.

Author contributions

S.T., M.J.R., L.B., M.S.L., I.D.V., T.K., F.C.C., S.V., G.D.W., A.C., Z.N.R., G.M.-S., L.H., S.Balakrishnan, J.V.Q., D.H., D.K.H. and M.L.V. designed and carried out experiments, analysed data and summarized the results. T.A.B., S.T., M.L.V., M.K., S.Bercovici, J.C.W. and S.Y. analysed data and supervised the work. T.A.B., S.Bercovici, S.T., D.H. and D.K.H. wrote the paper.

Competing interests

This study was funded by Karius, Inc. and describes the validation of a product developed by Karius, Inc. All authors (excepting S.T., J.V.Q. and S.Y.) are current or former employees and/or share -holders of Karius, Inc. This does not alter our adherence to Nature Microbiology policies on sharing data and materials.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-018-0349-6.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.A.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

n

Onl common tests should be described solel b name describe more comple techniques in the Methods section

F t r P

Gi e P alues as e act alues hene er suitable

d r

State e plicitl hat error bars represent e g SD SE CI

Our eb collection on statistics for biologists ma be useful