

FedUni ResearchOnline

<https://researchonline.federation.edu.au>

Copyright Notice

This is the published version of:

Flaherty, B. R., Barratt, J., Lane, M., Talundzic, E., & Bradbury, R. S. (2021). Sensitive universal detection of blood parasites by selective pathogen-DNA enrichment and deep amplicon sequencing. *Microbiome*, 9(1), 1.

Available online at: <https://doi.org/10.1186/s40168-020-00939-1>

Copyright © The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

RESEARCH

Open Access



Sensitive universal detection of blood parasites by selective pathogen-DNA enrichment and deep amplicon sequencing

Briana R. Flaherty^{1,2†}, Joel Barratt^{1,2*†} , Meredith Lane^{1,3}, Eldin Talundzic⁴ and Richard S. Bradbury^{1,5*}

Abstract

Background: Targeted amplicon deep sequencing (TADS) has enabled characterization of diverse bacterial communities, yet the application of TADS to communities of parasites has been relatively slow to advance. The greatest obstacle to this has been the genetic diversity of parasitic agents, which include helminths, protozoa, arthropods, and some acanthocephalans. Meanwhile, universal amplification of conserved loci from all parasites without amplifying host DNA has proven challenging. Pan-eukaryotic PCRs preferentially amplify the more abundant host DNA, obscuring parasite-derived reads following TADS. Flaherty et al. (2018) described a pan-parasitic TADS method involving amplification of eukaryotic 18S rDNA regions possessing restriction sites only in vertebrates. Using this method, host DNA in total DNA extracts could be selectively digested prior to PCR using restriction enzymes, thereby increasing the number of parasite-derived reads obtained following NGS. This approach showed promise though was only as sensitive as conventional PCR.

Results: Here, we expand on this work by designing a second set of pan-eukaryotic primers flanking the priming sites already described, enabling nested PCR amplification of the established 18S rDNA target. This nested approach facilitated introduction of a second restriction digestion between the first and second PCR, reducing the proportional mass of amplifiable host-derived DNA while increasing the number of PCR amplification cycles. We applied this method to blood specimens containing *Babesia*, *Plasmodium*, various kinetoplastids, and filarial nematodes and confirmed its limit of detection (LOD) to be approximately 10-fold lower than previously described, falling within the range of most qPCR methods.

Conclusions: The assay detects and differentiates the major malaria parasites of humans, along with several other clinically important blood parasites. This represents an important step towards a TADS-based universal parasite diagnostic (UPDx) test with a sufficient LOD for routine applications.

Keywords: Molecular parasitology, Amplicon sequencing, Blood microbiota, Parasite biodiversity, Parasite detection, Molecular diagnosis

* Correspondence: nsk9@cdc.gov; jbarratt@cdc.gov; r.bradbury@federation.edu.au

†Briana R. Flaherty and Joel Barratt contributed equally to this work.

¹Parasitic Diseases Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The versatility of targeted amplicon deep sequencing (TADS) enables detection of new pathogens as well as known pathogens that are difficult-to-culture and has facilitated the detailed characterization of bacterial, viral, and fungal communities from diverse biological and environmental specimens. These technologies are often used to characterize communities of microbial pathogens in human clinical samples and have been applied to the detection of antimicrobial resistance as part of routine surveillance networks [1–9]. However, the application of TADS and other next-generation sequencing (NGS)-based technologies, such as shotgun metagenomics, to the characterization of eukaryotic communities comprised of protozoa and helminths has been slow to advance, and the potential use of these technologies for parasite detection in routine diagnostic settings has not been rigorously explored.

Due to its non-specific nature, shotgun metagenomic approaches may identify unknown and/or unanticipated infectious agents from complex clinical specimens wherein few clues are provided as to the etiological agent of disease. In these circumstances, without metagenomic data, clinicians are limited in their ability to recommend the most appropriate tests for a differential diagnosis, making patient management challenging. Shotgun metagenomic sequencing of cerebrospinal fluid (CSF) has recently been applied to such complex clinical cases with great success, confirming multiple protozoal and helminthic infections, including four cases of *Balamuthia mandrillaris*-induced granulocytic amoebic encephalitis [10–12], one of *Taenia solium* neurocysticercosis [6], and four cases of meningitis caused by *Angiostrongylus cantonensis* [7, 13]. Only moderate success has been reported when applying metagenomic testing of plasma for residual pathogen DNA, even for viral and bacterial pathogens [14–16]. In a recent study of febrile illness in Uganda, this approach detected a large number of malaria infections not identified by microscopy; however, it also failed to detect some microscopy-confirmed infections [17]. Serum metagenomics has not yet been applied to rarer parasitic infections of the blood, such as trypanosomiasis, babesiosis, and leishmaniasis, or for the detection of helminthic microfilariae.

The success of metagenomic sequencing in detecting pathogens in CSF is potentially owed to the low complexity of CSF in terms of its limited extraneous DNA content compared to other biological matrices, such as tissue and stool. Accordingly, in cases of infectious meningitis or encephalitis involving host white cell inclusions in CSF greater than 200 cell/mm³, the sensitivity of CSF analysis is reduced by interference from background sequencing reads derived from host DNA [7]. Development of parasite-

specific diagnostic primer sets is an obvious solution for overcoming the impact of host DNA interference, but this has been hampered by the conserved nature of eukaryotic housekeeping loci (e.g., genes encoding the rRNA subunits), causing pan-parasitic primers to co-amplify host DNA and obscuring parasitic infections (Flaherty 2018). This mechanism of interference has also hampered the application of direct metagenomic sequencing of parasite DNA from specimen matrices containing high numbers of nucleated cell types, such as blood and tissue [18]. Using primers not broadly specific for all eukaryotes, protozoan infections have been detected incidentally via TADS. For example, following amplification of a fragment of 28S rDNA using primers designed to be broadly specific for fungi, Gomez et al. [19] identified two *Toxoplasma gondii* infections and one *Trypanosoma cruzi* infection from brain tissue, as well as one *Leishmania* spp. infection from a skin biopsy [19]. That assay was described as having “partial protozoal coverage” although it was not designed to detect parasites specifically [19].

Targeted amplicon deep sequencing of loci amplified using pathogen-specific primers has facilitated successful detection and identification of parasitic infections in blood and feces, both from human and animal samples [20–26]. These assays have generally been restricted to a single genus or genetically similar group, such as trypanosomes, rhabditid and strongylid helminths, or protozoa [22–25]. While this strategy effectively mitigates the impact of interfering host DNA on the sensitivity of TADS-based diagnostics, it still requires prescient knowledge of the parasite the diagnostician expects to detect. Ultimately, when utilizing genus or species-specific primers, the value of TADS as pan-specific diagnostic approach is lost.

A pan-parasitic TADS method with potential diagnostic utility was recently described by Flaherty et al. [18]. That method used a broadly specific pan-eukaryotic primer pair targeting a region of the 18S rDNA gene that possessed restriction enzyme cut sites only in vertebrates. This facilitated restriction enzyme digestion of human 18S rDNA specifically, as the human 18S amplicon possessed these restriction sites that are absent in blood protozoa and filarial nematodes. This digestion step, when performed on total DNA extracts prior to PCR, significantly reduced the number of host-derived reads following TADS. This concomitantly increased the number of reads recovered from blood parasites and enabled detection of parasites in some specimens that appeared negative when undigested specimens were tested. This method detected all protozoa and helminths investigated, which included the most commonly observed parasites in human blood [18]. A drawback of

this method was its limit of detection (LOD), which was comparable to most conventional PCR assays. The method was therefore less amenable to routine parasite diagnosis when compared to the more sensitive, cheaper, and less time-consuming qPCR methods already available [18]. Irrespective of this limitation, the study confirmed that taking advantage of the restriction endonuclease cut sites present in the human 18S rDNA is an effective approach for selective amplification of homologous parasite DNA; it reduced host-derived reads by more than 50%. Additionally, the number of parasite-derived reads increased by a factor of 5 to 10 times compared to their paired undigested samples using this procedure [18].

In light of the promise of this method, we sought to expand on the observations of Flaherty et al. [18], towards the development of a TADS-based, pan-parasite diagnostic test for blood with a sufficient LOD to be viable for routine use. We designed a second set of pan-eukaryotic primers that flank the priming sites previously described [18], enabling nested PCR amplification of the original 18S rDNA target. This two-step nested PCR approach facilitated introduction of a secondary restriction enzyme digestion step performed between the first and second PCR, reducing the proportional mass of amplifiable host-derived DNA while increasing the number of PCR amplification cycles (Fig. 1). We applied this assay to human blood specimens confirmed positive for various parasites using routine diagnostic methods (e.g., qPCR and microscopy) including several specimens containing multiple parasite species. Our assay is approximately 10-fold more sensitive than that described by Flaherty et al. [18] and has a LOD comparable to (and exceeding) some published real-time PCRs. This marked improvement in our LOD makes our nested universal parasite diagnostic (UPDx) test more amenable for routine parasite detection in a diagnostic laboratory setting than any previously described TADS-based method for parasites.

Methods

Assay design and study rationale

The TADS-based method previously described by Flaherty et al. [18] utilized a pair of pan-eukaryote PCR primers targeting an ~ 200-bp region of 18S rDNA with BamHI-HF and XmaI restriction cut sites present in all vertebrates assessed but absent in all blood protozoa and helminths, to the best of our knowledge. These sites could be exploited to digest human DNA within whole DNA extracts in a targeted fashion, prior to PCR, to encourage preferential amplification of parasite DNA. Given the demonstrated utility of this locus [18], we sought to utilize the same locus but introduce several modifications that increase its amenability to routine

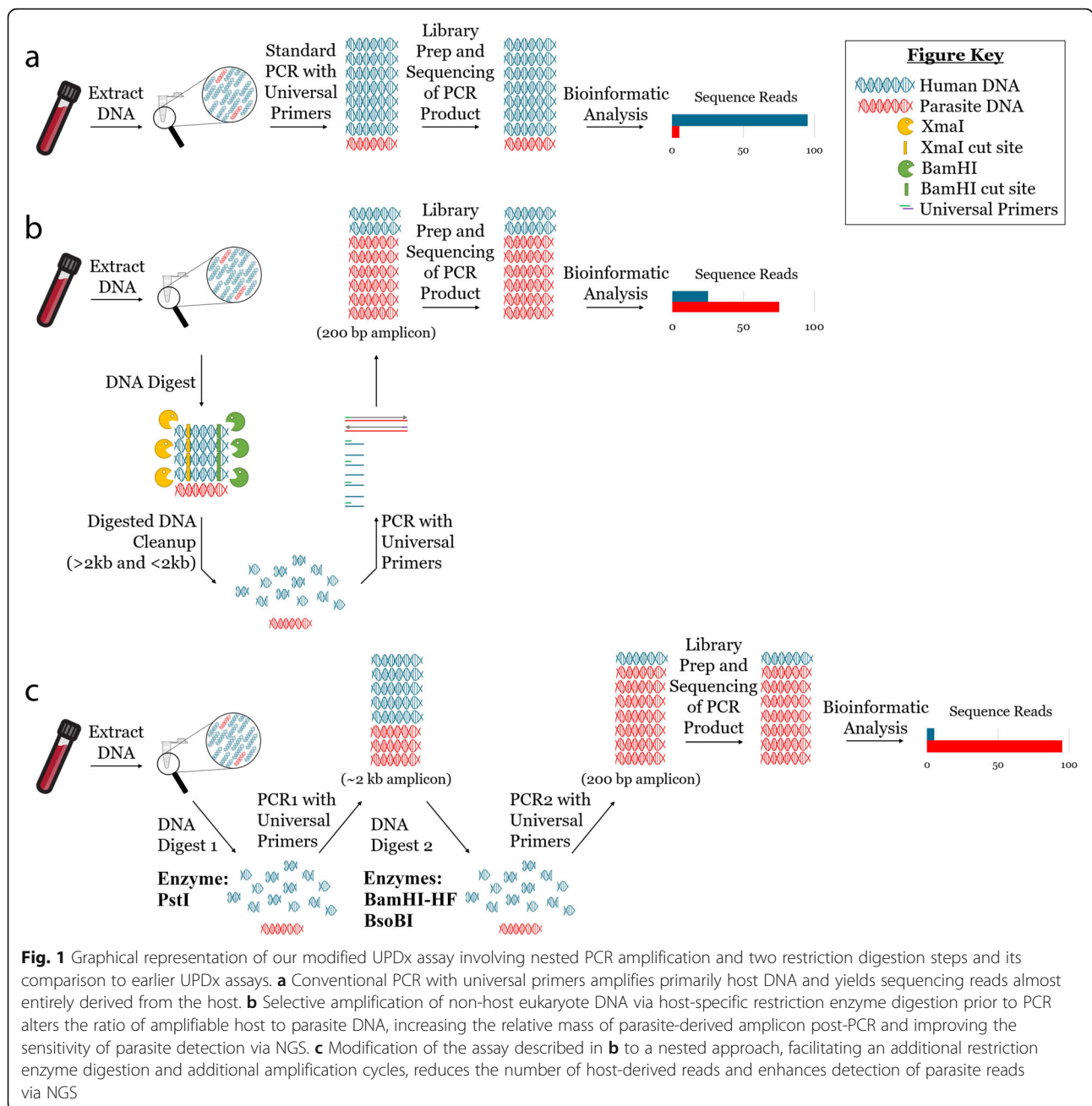
use. To improve the assays' LOD, we designed an additional set of pan-eukaryotic primers with priming sites flanking the original ~ 200-bp target described by Flaherty et al. [18], facilitating a nested PCR amplification of the same amplicon, which increased the number of amplification cycles.

This nested approach also introduced an opportunity for two restriction enzyme digestions to be performed: one on the total DNA extract prior to the first PCR (Digest 1: D1) and a second performed on the product of the first PCR but preceding the second PCR (Digest 2: D2). The larger ribosomal DNA region captured by the new "outer" nested primers contains a PstI restriction enzyme cut site within the human 18S rDNA target sequence, which was utilized as the restriction digestion target for D1. The second digestion is performed between the first and second PCR, taking advantage of restriction sites originally described by Flaherty et al. [18]. However, a modification of the original assay was implemented, by replacing XmaI with the BsoBI restriction enzyme during D2, noting that BsoBI also has restriction sites within our target amplicon that do not exist for the blood protozoa and filarial nematodes examined here, and confirming an additional choice for restriction enzyme digestion during D2. While it may be of no consequence for this assay, we note that BsoBI is not sensitive to CpG methylation. CpG methylation is an epigenetic mechanism utilized by all mammals for gene silencing as well as during embryonic development [27], and XmaI activity is impaired by CpG methylation.

To assess the performance of our modified nested TADS-based approach, we compare it directly to the original method described by Flaherty et al. [28] and to a qPCR used for routinely detecting *Plasmodium falciparum* in the Parasitic Diseases Branch at CDC. We apply our modified assay to an assortment of blood specimens positive for a range of parasites comprising either blood from healthy donors spiked with parasites or their DNA, or positive clinical specimens confirmed as part of the routine diagnostic activities performed in the Parasitic Diseases Branch at CDC. The precise methodologies are detailed in later sections.

Source of samples

The majority of human clinical blood samples used in this study were originally submitted to the CDC Parasitic Diseases Branch for confirmatory diagnosis of parasitic infections. Following diagnosis, samples were de-identified and frozen in 200 µL aliquots at - 80 °C for later use. De-identification involved the collector of the specimens providing an aliquot to the authors of this study in an ambiguously marked container (e.g., "*P. falciparum* specimen 1") so that its linkage back to the patients was not possible. Samples acquired in this



manner include the following: *P. falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale*, *Babesia microti*, *Babesia divergens*-like variant MO1, *T. cruzi* (HIV-positive and HIV-negative clinical samples), *Loa loa*, and NPF (no parasite found)—used as negative controls. A blood spot on filter paper collected in the Democratic Republic of Congo from a person with *Mansonella perstans* microfilaremia was provided by Dr. Vita Cama of the CDC Parasitic Diseases Branch. Acute *T. cruzi* infection EDTA blood DNA extracts were generously provided by Dr. Stella Chenet and Dr. Maria Isabel Jercic of the Institute of Public Health, Chile (Santiago).

B. divergens was obtained from the infected blood of laboratory-raised gerbils, and *Babesia duncani* was obtained from the infected blood of laboratory-raised guinea pigs, routinely maintained in CDC's animal facility (Table 1). Bioinformatic analysis confirmed that the restriction enzyme cut sites were present in all the respective animal species for all relevant animal samples utilized here. For some rare blood-borne parasites (*Plasmodium knowlesi* and *Brugia malayi*), either animal blood or human blood collected during previous studies and stored at -80°C were used. Rare blood parasites that could not be acquired as true clinical samples at the

Table 1 Host, source, and original identification method of samples used in this study

Parasite	Specimen number	Sample type	Host	Species identification diagnostic method/s
<i>Plasmodium falciparum</i>	Specimen 1	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [29]
<i>Plasmodium vivax</i>	Specimen 2	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [29]
<i>Plasmodium ovale</i>	Specimen 3	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [29]
<i>Plasmodium malariae</i>	Specimen 4	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [29]
<i>Plasmodium knowlesi</i>	Specimen 5	EDTA blood	<i>Macaca mulatta</i>	Microscopy
<i>Babesia microti</i>	Specimen 6	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [30]
<i>Babesia divergens</i>	Specimen 7	EDTA blood	<i>Meriones unguiculatus</i>	Microscopy and PCR w/s [30]
<i>Babesia duncani</i>	Specimen 8	EDTA blood	<i>Meriones unguiculatus</i>	Microscopy and PCR w/s [30]
<i>Babesia divergens</i> -like variant MO1	Specimen 9	EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [30]
<i>Trypanosoma cruzi</i> (infections)	Specimens 10 and 11	EDTA blood	<i>Homo sapiens</i>	Real-time PCR (Qvarnstrom 2012)
<i>Trypanosoma cruzi</i> (culture)	Specimen 12	RPMI culture in EDTA blood	<i>Homo sapiens</i>	Microscopy
<i>Trypanosoma brucei</i>	Specimen 13	HMI-9 culture in EDTA blood	<i>Homo sapiens</i>	Microscopy
<i>Leishmania infantum</i>	Specimen 14	RPMI culture in EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [31]
<i>Leishmania donovani</i>	Specimen 15	RPMI culture in EDTA blood	<i>Homo sapiens</i>	Microscopy and PCR w/s [31]
<i>Brugia malayi</i>	Specimen 16	EDTA blood	<i>Felis catus</i>	Microscopy
<i>Loa loa</i>	Specimen 17	EDTA blood	<i>Homo sapiens</i>	Real-time PCR [32]
<i>Mansonella perstans</i>	Specimen 18	Blood spot on filter paper	<i>Homo sapiens</i>	Real-time PCR [33, 34]
NPF (no parasite found)	N/A	EDTA blood	<i>Homo sapiens</i>	Microscopy, PCR [29–31]

w/s with Sanger sequencing of PCR product, CDC Centers for Disease Control and Prevention, FR3 Filariasis Research Reagent Resource Center, PDB Parasitic Diseases Branch, UGA University of Georgia, NIH National Institutes of Health

time of the study were recreated by spiking uninfected human blood with cultured parasites—*Leishmania infantum*, *Leishmania donovani*, *T. cruzi*, and *Trypanosoma brucei* subsp. *rhodesiense* cultures were added to whole human blood at a ratio of 1:100. All blood samples were collected into EDTA anticoagulant. Full details regarding specimen source, matrix, original parasite identification method, and DNA extraction method are provided in Table 1. Simulated mixed parasite infections were prepared according to the descriptions in Table S1, which can be found in Supplementary File S1 on page 4.

3D7 parasite culture

Plasmodium falciparum 3D7 parasites were grown according to routine methods [28]. Parasites were cultured in human O⁺ red blood cells at 4% hematocrit under a gas mixture of 90% nitrogen, 5% oxygen, and 5% CO₂. Cultures were maintained in 25 or 75 cm² tissue culture flasks at 37 °C and in complete culture medium made up of RPMI containing 25 mM HEPES, 0.05 mg/mL hypoxanthine, 2.2 mg/mL NaHCO₃, 0.5% inactivated

O⁺ human serum, 2 g/L glucose, and 0.01 mg/mL gentamicin. Culture growth was assessed daily via Giemsa stain and passaged or supplemented as necessary.

Primer design

To design pan-eukaryote “outer” nested primers, Geneious software (Biomatters Inc, Newark, NJ, USA) was used to align the 18S rRNA genes from publicly available sequences of 29 species of protozoa, 30 species of helminths, and *Homo sapiens*. A list of these sequences is provided in Supplementary File S1, Appendix A. The PCR1 (outer) primer sequences are as follows: TTGATCCTGCCAGTAG TCATATGC (outer forward) and GGTGTGTACAAAGG GCAGG GAC (outer reverse), generating a PCR product of approximately 2 kb. Inner primers for the universal nested PCR were identical to those previously described by Flaherty et al. [18], with primer sequences as follows: CCGGAG AGGGAGCCTGAGA (inner forward) and GAGCTGGA ATTACCGCGG (inner reverse), generating a PCR product of approximately 200 bp. Primers were synthesized at the CDC Biotechnology Core Facility.

Universal Blood Parasite Detection

DNA from 200 μ L of parasite-free or parasite-infected whole blood was extracted using a QIAamp DNA Blood Mini QIAcube Kit (Qiagen) and a Qiagen QIAcube for automated sample preparation. Samples and negative extraction controls were processed according to the kit protocol and eluted into 50 μ L of Qiagen Buffer EB. Following extraction, 8.5 μ L of DNA extract was digested for 2 h at 37 °C with 10 units of PstI (0.50 μ L) and 1 μ L of 10X CutSmart Buffer (digest D1). Digested DNA (2 μ L of that digest) was subsequently amplified (PCR1) in a volume of 25 μ L per reaction. For PCR1 cycling, samples were denatured at 98.0 °C for 30 s followed by 15 cycles of 98.0 °C for 10 s, primer annealing at 67.0 °C for 30 s and extension at 72.0 °C for 2 min, and a final extension at 72.0 °C for 2 min. Thereafter, the entire 25 μ L product of PCR1 was again digested for 1 h at 37 °C by directly adding 10 units of BamHI-HF (0.5 μ L), 10 units of BsoBI (1 μ L), and 2.5 μ L of 10X CutSmart Buffer (digest D2) to the original PCR tube. Two microliters of that digested product were then transferred into PCR2 (20 μ L per reaction). All restriction enzymes were purchased from NEB (Ipswich, MA, USA). For PCR2, samples were denatured at 98.0 °C for 30 s followed by 30 cycles of 98.0 °C for 10 s, primer annealing at 67.0 °C for 30 s and extension at 72.0 °C for 45 s, and a final extension at 72.0 °C for 2 min. All PCR reactions (both PCR1 and PCR2) contained Q5 Buffer, dNTPs, forward and reverse primers (1.25 μ L of 10 μ M stock), Q5 High GC Enhancer, and Q5 High-Fidelity DNA Polymerase (NEB, Ipswich, MA, USA), in the concentrations specified in the manufacturer's instructions. Only the reaction volumes varied between PCR1 (25 μ L) and PCR2 (20 μ L). Each PCR run was accompanied by at least three negative controls—DNA extracts of parasite-negative blood. To establish the conditions described here, assay parameters were optimized systematically using DNA from cultures of *Leishmania infantum* spiked into human blood and *P. knowlesi*-containing *Macaca mulatta* blood; optimal conditions were considered to be those that most reduced the number of vertebrate-derived reads in the final sequence dataset while maximizing the number of parasite-derived reads. An overview of these optimization experiments is provided in Supplementary File S1. Specimens were processed using the method described here in addition to comparing different restriction digestion procedures: specimens prepared using only Digest 1, only Digest 2, or both digestion steps. Each of these three reaction conditions was applied to the same specimens in triplicate and sequenced on three separate MiSeq sequencing runs. These results were compared to a single replicate prepared for each specimen analyzed using the method described by Flaherty et al. [18]. The significance of

differences between all conditions were assessed using a 2-way ANOVA with Tukey's multiple comparisons posttest.

Amplicon sequencing

Following PCR2 samples were analyzed on a 1.5% agarose gel, cleaned with the Monarch PCR & DNA Cleanup Kit (< 2 kb), and eluted with 20 μ L elution buffer, and final amplicon concentration was determined using a Qubit 2.0 Fluorometer with the Qubit dsDNA High Sensitivity Assay Kit (Life Technologies, Grand Island, NY, USA). These final concentrations were used downstream to normalize the DNA during library preparation. Samples were then diluted 1:5 in NEB Elution Buffer and transferred to the CDC Biotechnology Core Facility's Genome Sequencing Lab for library preparation and DNA amplicon sequencing using the NEBNext Ultra DNA Library Prep Kit for Illumina (NEB), the MiSeq Reagent Nano Kit v2 (Illumina), and sequencing on an Illumina MiSeq Sequencing platform (Illumina). Alternatively, the same library preparation and sequencing procedure as previously described was performed within the Parasitic Diseases Branch at CDC and sequenced on the CDC Division of Parasitic Diseases and Malaria Sequencing Facility's Illumina MiSeq. Each specimen was analyzed as a single replicate using this workflow, except for the specimens included in the limit of detection experiments (described below) which were performed in triplicate. All raw reads have been made publicly available on the NCBI Sequence Read Archive under BioProject accession number PRJNA437674. Specimens relevant to the present study under this BioProject include the text "nUPDx method validation" in their title.

Bioinformatic analysis

Analysis of sequence data was undertaken using a custom workflow designed in Geneious (Geneious Prime, version 11: www.geneious.com). This workflow first removed primer sequence from either end of the Illumina reads (250 bases, paired end) using the Trim Ends plugin allowing 3 mismatches and a minimum match length of 5 bases. Illumina adapter trimming was next performed using BBDuk. Low-quality ends were trimmed at either end (minimum Phred score of 20), and reads shorter than 50 bases in length were discarded. Paired reads were then merged using default parameters. Next, a BLASTN search was performed against a database constructed from human 18S rDNA sequences and some human 18S pseudogenes (GenBank Accession numbers: X03205.1, XR_003508809.1, NG_055289.1, NG_054751.1, AC129664.7) with the percent identity set to 99%, a word size of 11, and the qcov_hsp_perc flag set to 60. Reads obtaining BLASTN hits to this

database using these parameters were discarded. The remaining reads were assembled using the Geneious de novo assembler applying the following custom parameters: a minimum overlap of 50 bases and a minimum overlap identity of 100%. The resulting haplotypes were next subjected to a BLASTN search (default parameters) against a reference database containing a set of 18S sequences compiled from a range of parasites (GenBank Accession numbers are provided in Tables 2 and 3), and the BLASTN hit/s obtaining the nearest match to the haplotype/s detected were exported to text. Finally, the proportion of merged reads used to construct each haplotype was calculated as a percentage of the total number of merged reads generated for that specimen. To distinguish between positive and negative results, we employed the same cutoff system described by Flaherty et al. [18], where the number of merged reads used to construct a parasite-derived haplotype was calculated as a proportion of the total number of merged reads generated for that specimen. This was used to determine the specimens' status as positive or negative, based on the number of parasite-derived reads detected in parasite-negative blood specimens, as per Flaherty et al. [18]. To investigate the ability of the sequenced amplicons to differentiate between various parasite taxa, a cluster dendrogram was generated from the haplotypes generated from various specimens. Haplotypes were aligned in Geneious and exported as a fasta file. The alignment was imported into R, and a distance matrix was generated including gaps in the distance measurement, using the "seqinr" R package. This matrix was clustered using hierarchical agglomerative nesting (AGNES) in the R package "cluster," version 2.0.6. AGNES was performed using Ward's clustering method with all other parameters set to default. The resulting tree was visualized using the R package "ggtree."

Limit of detection and comparison to qPCR for *P. falciparum*

Cultured *P. falciparum* 3D7 parasites (primarily ring-stage) were washed once with RPMI and subsequently spiked into whole blood to simulate a *P. falciparum* infection. Thick and thin smears of these simulated infections were then fixed with methanol and stained with 2.5% Giemsa stain for assessment of parasitemia. Concurrently, the total number of white blood cells (WBC) in the cell preparation was determined using a Beckman Coulter Ac-T Diff Hematology Analyzer. The cell preparation, which was determined to be at a parasitemia of 58,000 parasites/ μ L, was serially diluted to generate *P. falciparum* samples in whole blood ranging from 58,000 to 0.0058 parasites/ μ L. Serial dilutions were processed in triplicate, using the method described here, comparing different restriction digestion procedures: only Digest 1,

only Digest 2, or both digestion steps. Each of these reaction conditions was applied to the same specimen aliquots in triplicate and sequenced on three separate MiSeq sequencing runs. Resultant reads were analyzed bioinformatically as described above and compared to a single replicate analyzed using the method described by Flaherty et al. [18]. The significance of differences between all conditions was assessed using a 2-way ANOVA with Tukey's multiple comparisons posttest. The impact of performing only D1, only D2, and the double digestion procedure was also assessed on specimens positive for all other parasite species examined in this study to ensure that the observed impact of the three digestion procedures on *P. falciparum* could also be generalized to other parasites (as described above). DNA extracted from these same serial dilutions of *P. falciparum* in whole blood was also tested using a routine, duplex, species-specific qPCR designed to detect and differentiate both *P. falciparum* and *P. vivax* as previously described [29]. For this assay, 5 μ L of DNA extract was provided as template in a total volume of 25 μ L [29]. Two *P. falciparum* clinical samples that had tested positive during routine testing at the CDC molecular diagnostic laboratory were de-identified and used as positive controls. A no-template water control was used as a negative control. Samples were run in triplicate, and an average CT for each dilution was calculated and compared with the nested PCR TADS method. While we were aware that the specimens contained only *P. falciparum*, the assay was prepared exactly as described including a FAM probe for *P. falciparum* and a CY5 probe for *P. vivax*.

Results

UPDx with double digestion maximizes reduction of host-derived reads

A substantial reduction in human-derived reads was observed in samples processed via the nested UPDx method with two restriction digests prior to TADS in comparison with the method described by Flaherty et al., which involved restriction digestion of the DNA extract only followed by a single PCR amplification (Fig. 2) [18]. Samples that underwent both Digest 1 (D1: digestion before the first PCR) and Digest 2 (D2: digestion between the first and second PCR) contained fewer host-derived reads than samples that underwent D1 or D2 alone (Fig. 2). Samples that underwent only D2 contained fewer host-derived reads than those subjected to D1 alone, and specimens subjected to D2 alone had fewer host-derived reads than samples processed by the method described by Flaherty et al. [18] (Fig. 2). We compared the earlier amplicon sequencing method [18] to our UPDx approach for 16 blood parasites (Fig. 2), including nine apicomplexans (various *Plasmodium* and *Babesia* species), four kinetoplastids (*Leishmania*

Table 2 UPDx (double digest) results obtained for clinical validation specimens confirmed positive for parasites by other diagnostic methods

Parasite in sample (specimen #)	Reads assembled	Total reads	% Reads assembled	% Similarity to sequence	GB accession
<i>Plasmodium falciparum</i> (Specimen 1)	7457	15,861	47.01	100.0% similar to <i>Plasmodium falciparum</i>	XR_002966654.1 ^b
	4820	15,861	30.39	100.0% similar to <i>Plasmodium falciparum</i>	XR_002273081.2
	239	15,861	1.51	99.4% similar to <i>Plasmodium falciparum</i>	XR_002966654.1 ^b
<i>Plasmodium vivax</i> (Specimen 2)	6584	11,546	57.02	100.0% similar to <i>Plasmodium vivax</i>	XR_003001206.1
	2820	11,546	24.42	100.0% similar to <i>Plasmodium vivax</i>	LT635616.1
	208	11,546	1.80	100.0% similar to <i>Plasmodium vivax</i>	U93234.1
<i>Plasmodium ovale</i> (Specimen 3)	11,617	15,295	76.00	100.0% similar to <i>Plasmodium ovale wallikeri</i>	MG255222.1
	122	15,295	0.798	100.0% similar to <i>Plasmodium ovale wallikeri</i>	KY073344.1
<i>Plasmodium malariae</i> (Specimen 4)	8754	12,200	71.75	100.0% similar to <i>Plasmodium malariae</i>	LT594631.1
<i>Plasmodium knowlesi</i> (Specimen 5)	10,119	11,999	84.33	100.0% similar to <i>Plasmodium knowlesi</i>	MF370109.1
<i>Babesia microti</i> (Specimen 6)	9554	10,770	88.71	100.0% similar to <i>Babesia microti</i>	LC314658.1
<i>Babesia divergens</i> (Specimen 7)	11,847	13,546	87.5	100.0% similar to <i>Babesia</i> sp.	MG944238.1
<i>Babesia duncani</i> (Specimen 8)	11,011	12,591	87.5	100.0% similar to <i>Babesia duncani</i>	HQ289870.1
<i>Babesia divergens</i> -like variant MO1 (Specimen 9)	10,842	12,323	88.0	100.0% similar to <i>Babesia</i> sp.	MG944238.1
<i>Trypanosoma cruzi</i> (clinical–Specimen 10)	1882	18,229	10.32	100.0% similar to <i>Demodex</i> sp. ^a	MH891494.1
	111	18,229	0.609	99.4% similar to <i>Demodex</i> sp. ^a	MH891494.1
<i>Trypanosoma cruzi</i> (HIV+ clinical–Specimen 11)	1824	14,897	12.24	100.0% similar to <i>Trypanosoma cruzi</i>	AF288661.1
	158	14,897	1.06	100.0% similar to <i>Trypanosoma cruzi</i>	CP015675.1
<i>Trypanosoma cruzi</i> (culture–Specimen 12)	8852	15,471	57.2	100.0% similar to <i>Trypanosoma cruzi</i>	CP015675.1
	182	15,471	1.18	99.4% similar to <i>Trypanosoma cruzi</i>	CP015675.1
<i>Trypanosoma brucei</i> (Specimen 13)	7928	18,334	43.24	100.0% similar to <i>Trypanosoma brucei brucei</i>	XR_002989632.1
	1361	18,334	7.42	99.4% similar to <i>Trypanosoma brucei brucei</i>	XR_002989632.1
<i>Leishmania infantum</i> (Specimen 14)	11,735	17,418	67.373	100.0% similar to <i>Leishmania</i> sp.	GQ332359.1
<i>Leishmania donovani</i> (Specimen 15)	11,781	14,653	80.400	100.0% similar to <i>Leishmania</i> sp.	GQ332359.1
	110	14,653	0.751	99.4% similar to <i>Leishmania</i> sp.	GQ332359.1
<i>Brugia malayi</i> (Specimen 16)	1042	13,117	7.944	100.0% similar to <i>Dirofilaria repens</i>	MG780293.1
<i>Loa loa</i> (Specimen 17)	2003	17,439	11.49	100.0% similar to <i>Wuchereria bancrofti</i>	AY843436.1
<i>Mansonella perstans</i> (Sample 18)	180	15,196	1.18	100.0% similar to <i>Filarioidea</i> sp.	KT907503.1

Note: A 100% match to a sequence does not necessarily indicate that this is the species detected. The amplicon sequenced is sufficient for a species diagnosis for only some parasites. Please refer to Fig. 5 for details

^a*Demodex* are mites that commonly inhabit the pores in normal human skin and are considered non-pathogenic, we suspect that this represents potential contamination with skin debris from the venipuncture site. *T. cruzi* was not detected in this specimen

^bExamples where multiple hits to the same accession were observed. These are both shown because the percentage similarity to these accessions is different for these two haplotypes

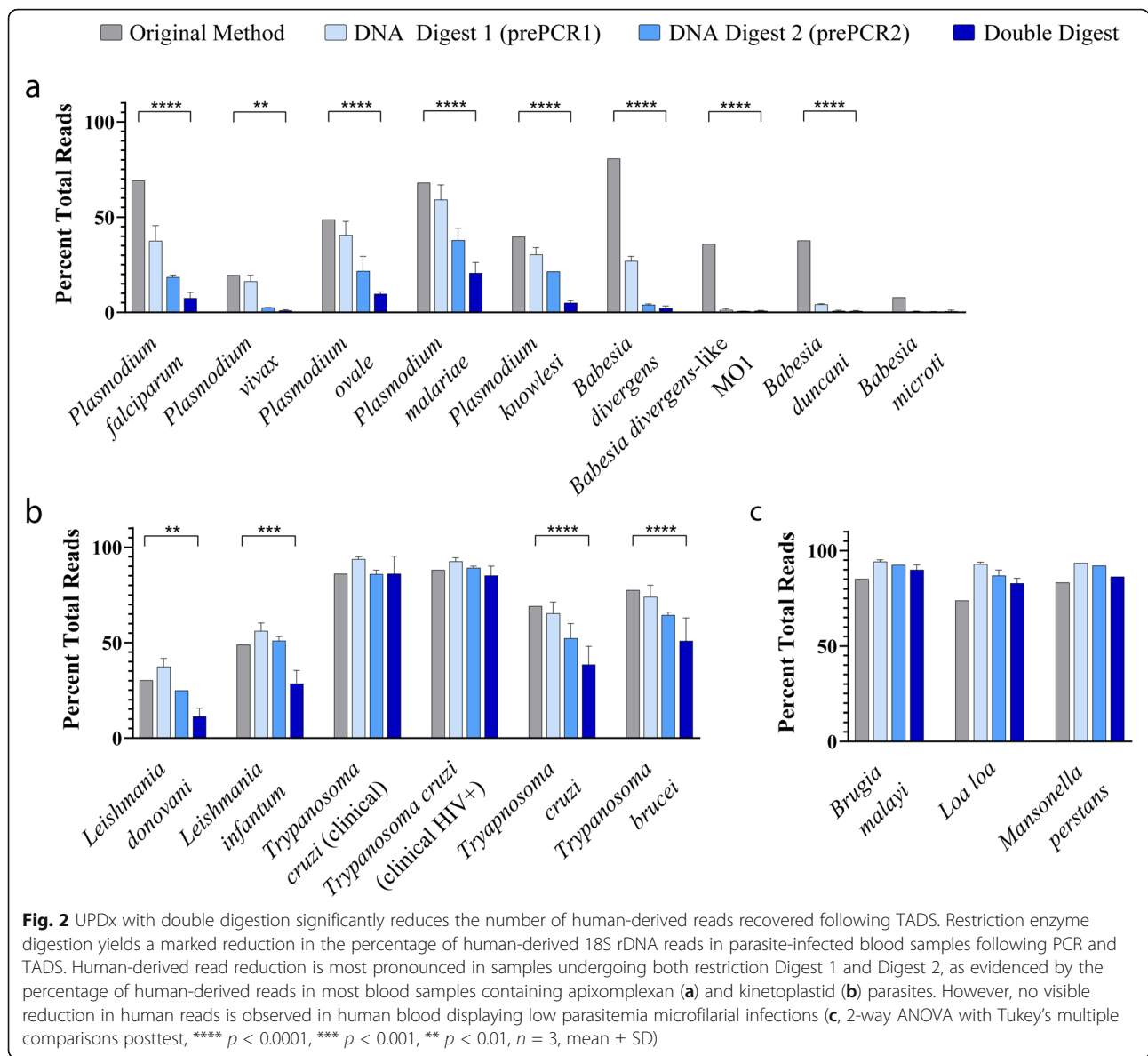
Table 3 Positive detection of multiple parasites in a single specimen using UPDx

Sample name	Spiked analytes	% Similarity to reference sequences	Parasite abbreviation: accession # (% reads assembled ^a)	Total % reads assembled
Mix 1	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (2.3%), XR_002273081.2 (2.1%)	4.4
	<i>P. vivax</i>	100.0% similar to <i>Plasmodium vivax</i>	Pv: LT635616.1 (15.8%), XR_003001206.1 (43.9%)	59.7
Mix 2	<i>P. ovale</i>	100.0% similar to <i>Plasmodium ovale wallikeri</i>	Po: MG255222.1 (34.1%), KY073344.1 (1.5%)	35.6
	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (13.9%), XR_002273081.2 (9.3%)	23.2
Mix 3	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (23.1%), XR_002273081.2 (15.1%)	38.2
	<i>P. malariae</i>	100.0% similar to <i>Plasmodium malariae</i>	Pm: LT594631.1 (15.9%)	15.9
Mix 4	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (5.6%), XR_002273081.2 (4.5%)	10.1
	<i>P. knowlesi</i>	100.0% similar to <i>Plasmodium knowlesi</i>	Pk: MF370109.1 (55.0)	55.0
Mix 5	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (1.5%), XR_002273081.2 (1.3%)	2.8
	<i>P. vivax</i>	100.0% similar to <i>Plasmodium vivax</i>	Pv: LT635616.1 (10.0%), XR_003001206.1 (27.2%), U93234.1 (0.6%)	37.8
	<i>P. ovale wallikeri</i>	100.0% similar to <i>Plasmodium ovale wallikeri</i>	Po: MG255222.1 (5.4%)	5.4
	<i>P. malariae</i>	100.0% similar to <i>Plasmodium malariae</i>	Pm: LT594631.1 (1.6%)	1.6
	<i>P. knowlesi</i>	100.0% similar to <i>Plasmodium knowlesi</i>	Pk: MF370109.1 (16.8)	16.8
Mix 6	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (22.4%), XR_002273081.2 (12.5%)	34.9
	<i>T. cruzi</i>	100.0% similar to <i>Trypanosoma cruzi</i>	Tc: CP015675.1 (20.5%), AF288661.1 (1.2%)	21.7
Mix 7	<i>P. vivax</i>	100.0% similar to <i>Plasmodium vivax</i>	Pv: LT635616.1 (17.4%), XR_003001206.1 (43.8%)	61.2
	<i>T. cruzi</i>	100.0% similar to <i>Trypanosoma cruzi</i>	Tc: CP015675.1 (2.5%)	2.5
Mix 8	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (27.0%), XR_002273081.2 (15.0%)	42.0
	<i>T. brucei</i>	100.0% similar to <i>Trypanosoma brucei brucei</i>	Tb: XR_002989632.1 (10.3%)	10.3
Mix 9	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (27.0%), XR_002273081.2 (15.0%)	52.2
	<i>L. loa</i>	100.0% similar to <i>Loa loa</i> and <i>Wuchereria bancrofti</i> ^b	Ll: AY843436.1 (0.91%)	0.91
Mix 10	<i>P. falciparum</i>	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (34.6%), XR_002273081.2 (18.4%)	49.2
	<i>B. malayi</i>	100.0% similar to filarial nematode sequences ^c	Bm: MG780293.1 (0.55%)	0.55
Natural mixed infection	N/A	100.0% similar to <i>Plasmodium falciparum</i>	Pf: XR_002966654.1 (0.80%), XR_002273081.2 (0.80%)	1.60
		100.0% similar to <i>Plasmodium malariae</i>	Pm: LT594631.1 (2.2%)	2.2

^aPercentage of reads assembled to produce a haplotype that obtained a 100% similarity BLASTN hit to the sequence associated with the accession numbers provided

^bThe sequence generated cannot distinguish between *Loa loa* and *Wuchereria bancrofti*

^cThe sequence generated from this amplicon cannot differentiate between several filarial nematodes. For example, our reference database included a sequence from *Dirofilaria repens* though it is important to note that the nucleotide sequence is identical to *Brugia malayi* at the 18S rDNA region captured by UPDx



donovani, *Leishmania infantum*, *Trypanosoma cruzi* and *Trypanosoma brucei*, and three filarial nematodes (*Brugia malayi*, *Loa loa*, and *Mansonella perstans*).

For eight of nine simulated apicomplexan blood parasite infections, we observed a statistically significant reduction ($p < 0.01$, 2-way ANOVA) in the percentage of host-derived reads detected using UPDx (~ 1% host-derived reads) compared to the original method (~ 20% host-derived reads) (Fig. 2a). For seven of these apicomplexan infections, the reduction was highly significant ($p < 0.0001$), where between ~ 30 and ~ 70% of recovered reads were derived from the host using the earlier method versus between ~ 1 and ~ 20% for UPDx (Fig. 2a). A *Babesia microti* infection did not yield a significant difference between the number of host-derived reads recovered when the

specimen was tested using both methods, yet a reduction was still observed for UPDx; less than 1% of reads were host-derived using UPDx, while ~ 8% of reads were host-derived using the original method (Fig. 2a). For the kinetoplastids, a similar significant reduction in host-derived reads was observed in most cases using UPDx compared to the earlier method, except for some clinical *T. cruzi* infections, where no significant difference was observed (Fig. 2b). Similarly, for each of three infections caused by different species of filarial nematode, no significant difference was observed between our UPDx approach and the original method (Fig. 2c). Ultimately, for the majority of infections tested, the proportional reduction in host reads following nested PCR with double digestion (specimens undergoing both D1 and D2) was significantly

greater than the proportional reduction in host reads described by Flaherty and colleagues using their method [18] (Fig. 2). Similarly, in the majority of cases, subjecting specimens to the nested UPDx method described here combining both D1 and D2 also significantly increased the number of parasite-derived reads detected following TADs when compared to specimens subjected to nested UPDx with D1 or D2 alone and compared to specimens subjected to the original method described by Flaherty et al. [18] (Fig. 3).

UPDx has a limit of detection comparable to qPCR

Assessment of our UPDx assay’s limit of detection (LOD), by applying the method to a 10-fold serially diluted blood sample panel containing *Plasmodium falciparum* of known parasitemia (range 58×10^3 –0

parasites/ μ L), indicated that the number of reads mapping to *P. falciparum* reference sequences post-sequencing was highest for the double-digested samples at all concentrations (Fig. 4a, b). The LOD for double-digested samples was approximately 0.58 parasites/ μ L which is 10-fold lower than the LOD for samples subjected to only D1 or D2 alone and approximately 10-fold more sensitive than the method described by Flaherty et al. [18] (Fig. 4a). For one of the three replicates diluted to 0.58 parasites/ μ L analyzed using the double digestion UPDx protocol, no parasite reads were detected. Despite this, a direct comparison of the performance of a qPCR routinely used within the Parasitic Diseases Branch at CDC [29] applied to the same serially diluted samples also obtained a positive amplification curve for only 2 of 3 replicates at a concentration of 0.58 parasites/ μ L (Fig. 4c). Furthermore,

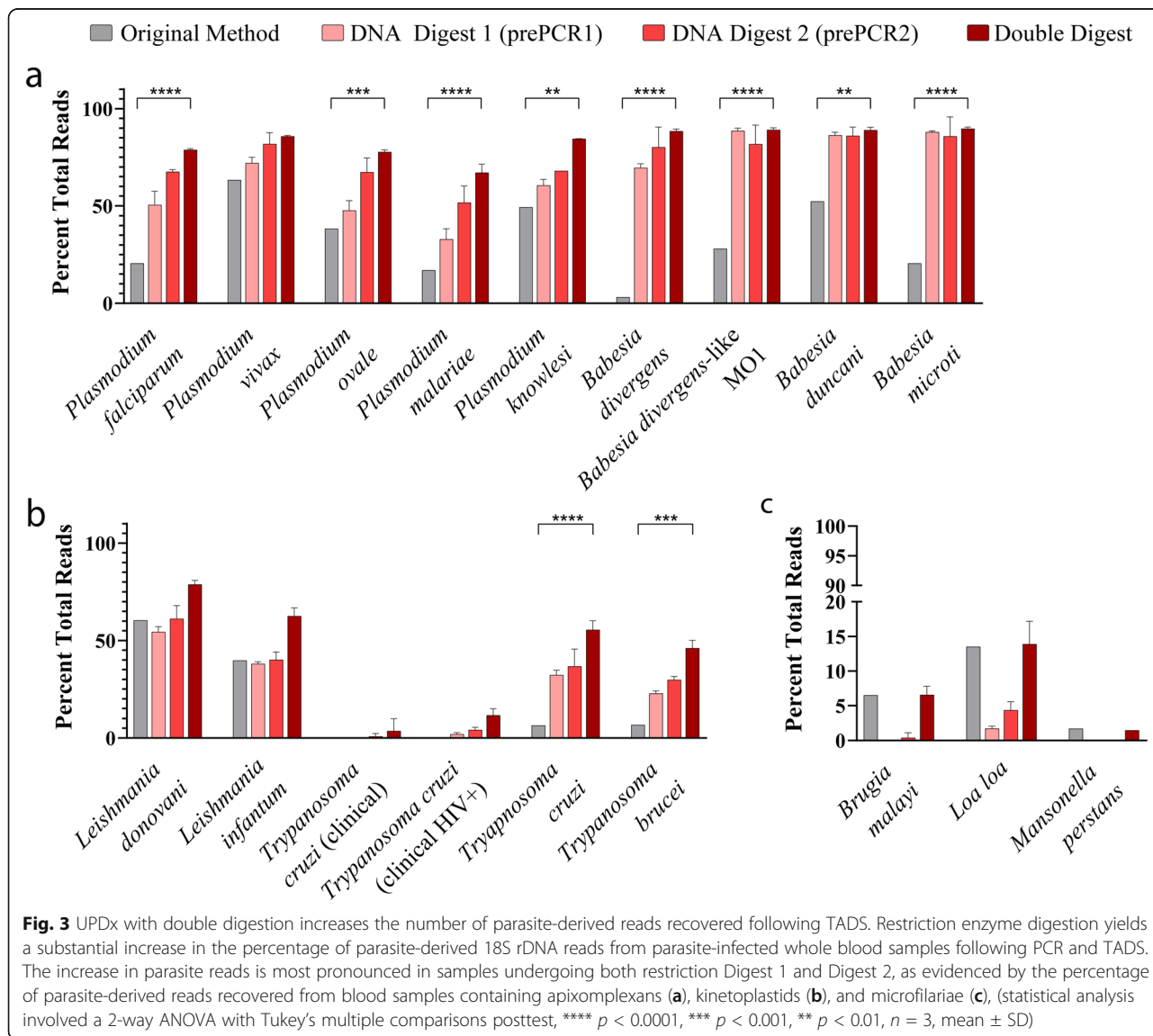
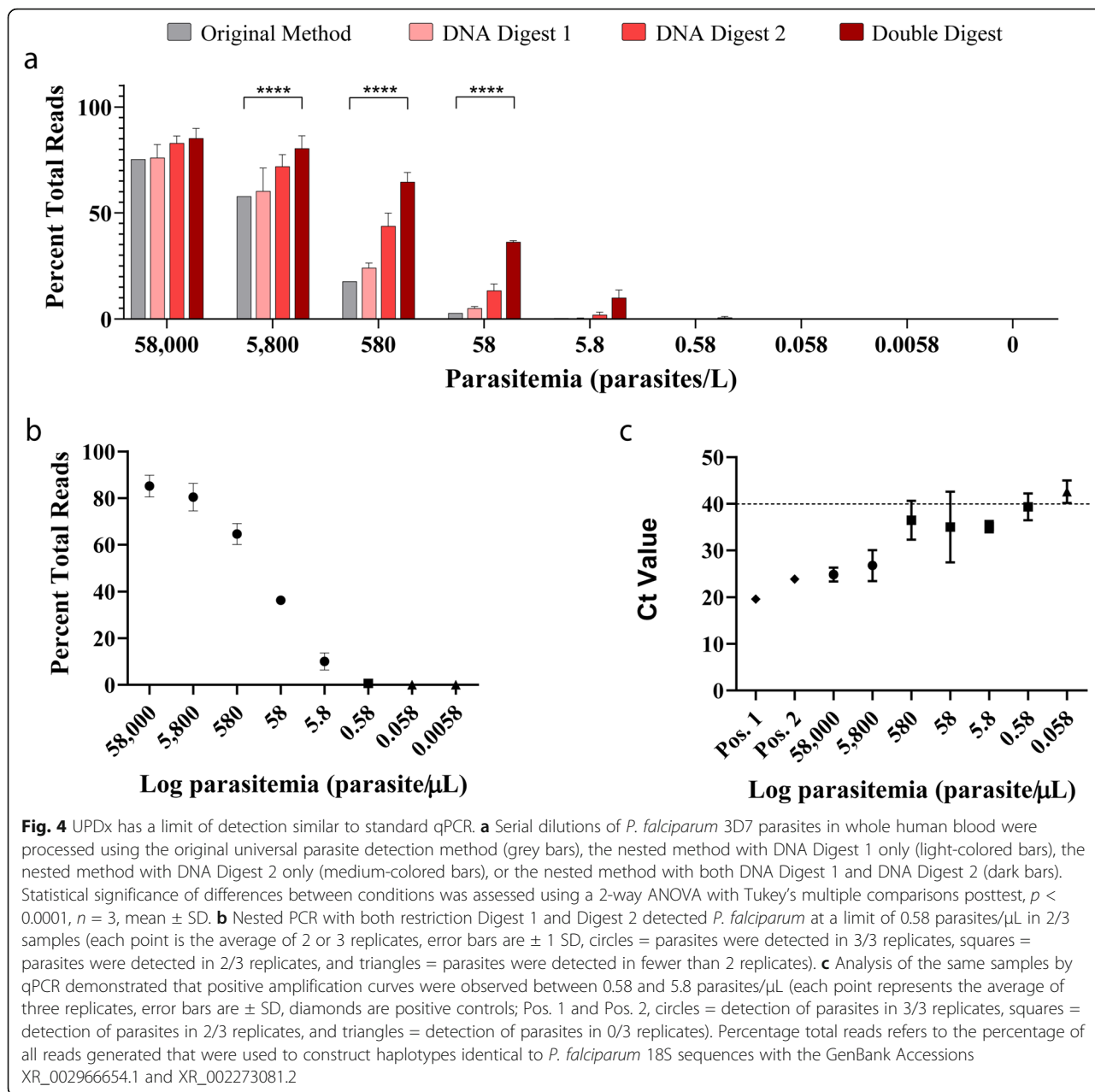


Fig. 3 UPDx with double digestion increases the number of parasite-derived reads recovered following TADs. Restriction enzyme digestion yields a substantial increase in the percentage of parasite-derived 18S rDNA reads from parasite-infected whole blood samples following PCR and TADs. The increase in parasite reads is most pronounced in samples undergoing both restriction Digest 1 and Digest 2, as evidenced by the percentage of parasite-derived reads recovered from blood samples containing apicomplexans (a), kinetoplastids (b), and microfilariae (c), (statistical analysis involved a 2-way ANOVA with Tukey’s multiple comparisons posttest, **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, $n = 3$, mean \pm SD)



the qPCR assay generated a positive amplification curve for only 2 of 3 replicates at all concentrations less than 5800 parasites/ μ L, while all three replicates were consistently positive using UPDx at all concentrations greater than 0.58 parasites/ μ L, indicative of a more consistent performance. At a parasite concentration of 5.8 parasites/ μ L using the UPDx approach, between 400 and 500 parasite-derived reads were detected in each of the three replicates, representing an unambiguous positive result from UPDx; a cutoff of 20 reads was determined for these experiments as previously described [18]. Similarly, for the two positive

UPDx results obtained at a parasite concentration of 0.58 parasites/ μ L, between 100 and 300 parasite reads were detected, which is also indicative of a clear positive result. Additionally, a single replicate at a parasite concentration of 0.058 parasites/ μ L obtained a positive result using UPDx, though only 31 parasite-derived reads were detected. A plot of the log-transformed data from double-digested specimens also supported a limit of detection of 0.58 parasite/ μ L, allowing us to make a conservative LOD estimate of near or below 1 parasite/ μ L for our UPDx method (Fig. 4b).

UPDx detects and differentiates several parasites commonly found in human blood

For the simulated infections comprised of various species of parasites, alone or in combination (Tables 1, 2, 3, and 4), all expected parasites were detected with the exception of some specimens expected to contain *T. cruzi* (Tables 2 and 4). This included all major human-infecting *Plasmodium* species (single-species and mixed infections), as well as simulated combinations of *Plasmodium* species, human-infecting kinetoplastids (*T. cruzi* and *Trypanosoma brucei*), and filarial nematodes (*Loa loa* and *Brugia malayi*) (Tables 2, 3, and 4). A clinical blood sample morphologically diagnosed by the CDC Parasite Reference Diagnostic Laboratory as a natural mixed-species malaria infection (*P. falciparum* and *Plasmodium malariae*) tested positive for *P. falciparum* and *P. malariae*, as expected (Table 3). In a separate sequencing experiment, five blood samples from acute *T. cruzi* patients, five from reactivation patients, and one artificial *T. cruzi* sample (whole blood spiked with 2 μ L *T. cruzi* culture) were also tested. These specimens were included in this analysis as they had previously been confirmed as positive for *T. cruzi* by the CDC Parasite Reference Diagnostic Laboratory using a previously published qPCR assay (Qvarnstrom 2012). Specimens from patients experiencing *T. cruzi* reactivation each tested negative using the UPDx assay while four of five acute *T. cruzi* samples tested positive (Table 4). The sequences generated using our UPDx assay differentiated all *Plasmodium* species most commonly infecting humans (Fig. 5). Filarial nematodes could be separated into three groups, with finer granularity than the family level (*Onchocercidae*), but not to the genus or species level (Fig. 5). The majority of the human-infecting *Leishmania* species cannot be distinguished based on

numerous *Leishmania* 18S sequences in GenBank, though *T. cruzi* and *T. brucei* are clearly distinguished. The three subspecies of *T. brucei* (*rhodesiense*, *gambiense*, and *brucei*) cannot be distinguished from each other nor from several related trypanosomes of veterinary importance included in our BLASTN database (Fig. 5). The three types of *Babesia* most commonly found in humans can be distinguished using our assay: *Babesia microti*, *B. duncani*, and *B. divergens*, with the caveat that sequences belonging to *B. divergens* and the *B. divergens*-like MO1 are identical (Fig. 5).

Discussion

Arguably, the challenge that is most prohibitive to the widespread application of TADS and other NGS technologies to routine parasite diagnosis is biological in nature. The word “parasite” is a generic term with no taxonomic basis, and its use encompasses a broad range of eukaryotic taxa including certain helminths, protozoa, and some arthropods. The terms “helminth,” “protozoan,” and “arthropod” are also highly generic, with some taxa haphazardly forced together into these groups despite sometimes possessing tenuous evolutionary relationships to each other at best. Consequently, designing PCR assays that specifically amplify “parasite” housekeeping loci while avoiding amplification of vertebrate host DNA is extremely difficult. This is particularly challenging when one considers that for most biological specimens, the mass of host DNA in total DNA extracts will vastly outweigh the mass of parasite DNA. This problem is less so for bacterial pathogens that belong to a different domain of life than the eukaryotes they infect and are therefore sufficiently divergent to avoid off-target amplification of vertebrate DNA.

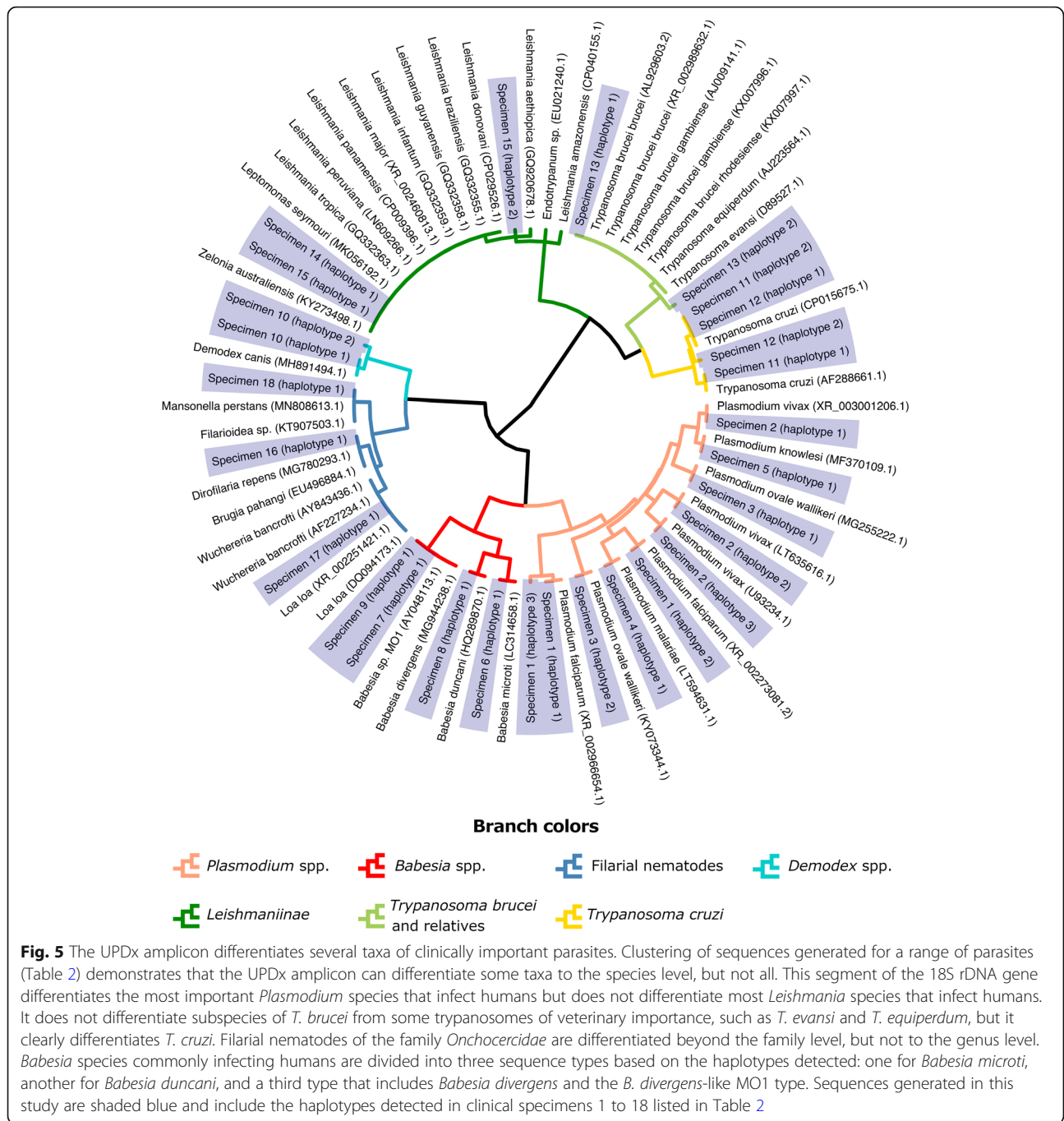
The present work expands on the observations of Flaherty et al. [18] regarding the benefits of host DNA restriction digestion prior to PCR by introducing several improvements. Firstly, to improve upon the assays’ LOD, we introduced an additional PCR amplification step by converting the assay into a nested PCR approach. The improvement in LOD resulting from this modification is a consequence of the increased number of PCR cycles arising from two rounds of amplification (PCR1 and PCR2) and the introduction of the second digestion step between PCR1 and PCR2, providing a second opportunity to deplete the relative mass of amplifiable host-derived 18S rDNA. This led to a 10-fold increase in LOD compared to the previous method [18], with a LOD comparable to a routinely used qPCR for *P. falciparum*. Furthermore, our UPDx assay performed more consistently than the qPCR assay used here, which failed to detect parasites in one of three replicate samples at all concentrations at or below 580 parasites/ μ L of blood and obtained negative results for all replicates at a

Table 4 Detection of acute clinical *T. cruzi* infections using UPDx with double digestion

Sample	Tc reads ^a	Total reads	% Reads	Result
Culture ^b	4496	12,999	34.6%	+
Acute1	64	8549	0.7%	+
Acute2	0	4900	0.0%	–
Acute3	9189	10,731	85.6%	+
Acute4	10,864	13,428	80.9%	+
Acute5	9152	11,161	82.0%	+
Reactivation1	0	7902	0.0%	–
Reactivation2	0	6372	0.0%	–
Reactivation3	0	9592	0.0%	–
Reactivation4	0	9822	0.0%	–
Reactivation5	0	8680	0.0%	–

^aThis is the sum of reads used to construct haplotypes that were identical to *T. cruzi* reference sequences AF288661.1 and/or CP015675.1

^bPositive control



concentration of 0.058 parasites/μL. In contrast, UPDx only failed to detect parasites in one of three replicates at a concentration of 0.58 parasites/μL; all three replicates were positive at concentrations above this. The UPDx assay also detected parasites in one replicate at a concentration of 0.058 parasites/μL and failed to detect parasites at a concentration of 0.0058 parasites/μL. The LOD of our UPDx assay is therefore comparable to other recently published qPCR assays, which possess LODs ranging from 0.03 to 0.3 parasites per microliter

of blood [35–37]. In addition to its consistent performance and comparatively low LOD, this assay has the added benefit that all blood parasites present in a sample may be detected.

Instances where our UPDx assay did not result in significant host read reduction compared to control experiments were primarily restricted to infections caused by the microfilariae and *T. cruzi*, wherein initial parasitemia is often exceptionally low. With each PCR cycle, there is a theoretical doubling of template

molecules assuming 100% PCR efficiency with no competing/spurious side reactions. For infections with very low parasitemias, the negative impacts of excessive host DNA are more pronounced. Restriction digestion reduces the availability of amplifiable host-derived template DNA, as we clearly demonstrate here, yet even the digested host DNA still competes with parasite DNA amplification because the PCR priming sites are still present and intact on digested host DNA. Consequently, digested host DNA will participate in a competing (albeit inefficient) side reaction, whereby primer is extended after binding to digested host DNA. However, this extension occurs without complete DNA molecule doubling, as it only occurs along a single strand in one direction until it terminates at the restriction digest cut sites. While this anticipated side reaction is incredibly inefficient, under circumstances of lower-level parasitemias, the number of parasite-derived amplicon molecules may never reach a relative frequency that sufficiently outcompetes these spurious side reactions. A solution might be to increase the cycle number for PCR1 and/or PCR2, understanding that the high-fidelity Q5 polymerase we utilize has an exceptionally low error rate; 99.7% of amplicon molecules will have the correct sequence for a 250 base pair amplicon, even after 60 cycles (<https://pcrfidelityestimator.neb.com/#/>). While this may increase the assays' LOD, an increase in PCR cycle number may also make the assay more susceptible to contamination. In addition to increasing the PCR cycle number, the LOD could be further improved by reducing the degree of multiplexing during library preparation (i.e., multiplexing fewer specimens per library). Here, we multiplexed 48 specimens in each library, resulting in approximately 10,000 to 20,000 pairs of reads per specimen, facilitating a LOD within the general range of qPCR assays available for *Plasmodium* sp. (discussed above). Reducing the number of specimens multiplexed per library could improve the LOD, which may be helpful for detecting infections of a low parasitemia. We do recognize however that this would increase the cost of the assay on a per specimen basis, which is approximately \$80.00 US per specimen excluding the cost of labor, assuming roughly 40 specimens are multiplexed on the same sequencing run.

Reported error rates for Illumina sequencing are less than 4 erroneous bases for every 1000 bases sequenced [38, 39]. The workflow described here includes a read QC step that removes adapter sequence and quality trims the ends of reads. Any reads less than 50 bases long following trimming are then discarded. Our results support that the read QC process reduced the number of errors in the data that were retained for analysis

relative to reported Illumina error rates. This is reflected in the data presented in Table 2, where we report several haplotypes that are approximately 99% identical to their nearest BLAST hit and likely represents sequencing errors. Importantly, these reads only comprise about 2% of the overall number of parasite-derived reads in these specimens following read QC. Furthermore, this single base difference will not confound the diagnosis because the expected amplicons for the parasites examined here typically differ by numerous SNPs and indels, even for the *Plasmodium* species. Therefore, a single SNP induced by a sequencing error will still allow investigators to accurately identify the sequence, leading them to an appropriate diagnosis.

Our method detected several major parasites found in human blood. Using simple conventional PCR and restriction enzyme technology, followed by in-house or outsourced Illumina MiSeq NGS, any laboratory can diagnose all human blood parasites using one test, requiring minimal sample volume, and with a LOD comparable to qPCR. Our UPDx assay and others like it, therefore, have the potential to address the increasing challenge of maintaining difficult-to-learn parasite morphology competency among laboratory staff in endemic regions [40, 41], a task also made difficult in non-endemic regions where blood parasites may only occasionally be encountered in returned travelers [42, 43]. Due to the relatively high cost of NGS, UPDx in the form described here is most appropriate for adoption by state and national parasitic disease reference laboratories, to which state and regional laboratories often refer diagnostically challenging cases. However, the possibility of adapting the assay to more compact Illumina platforms (i.e., MiniSeq and/or iSeq) could make the assay more accessible to other laboratories. While these platforms produce shorter reads than the MiSeq—a maximum of 150 base pair paired end reads—this is still compatible with the short amplicon utilized for UPDx. In addition, cost reduction by introducing adapter and index sequences to amplicons during the UPDx PCR steps is currently being explored, as this will remove the need for expensive, laborious, and time-consuming Illumina library preparations. This would also make the assay drastically cheaper and less complicated to prepare and reduce diagnostic turnaround times.

In this study, we utilize the same cutoff system described by Flaherty et al. [18], which utilized a “hard minimum” cutoff of 20 parasite-derived reads (below which a specimen is considered negative), and an adaptable “sliding maximum” that must be calculated each time a library is sequenced and requires that each sequencing run include at least three known negative blood specimens. This sliding maximum was introduced to control for index cross-talk resulting from specimen

multiplexing, which leads to a variable number of parasite-derived reads being assigned to the negative control specimens when the data are de-multiplexed. The number of parasite-derived reads assigned to the negatives as a result of this “cross-talk” varies depending on the specimens included in the library preparation. For example, for a library where 35 of 40 specimens are positive for *P. falciparum*, the negative control specimens will almost certainly be assigned a small number of *P. falciparum* reads. Alternatively, if every specimen in a run is negative, index cross-talk is still occurring, although very few (or no) negative control specimens will contain parasite-derived reads, because the only source of these reads would be the positive control specimen included in each run. For further detail, please refer to the work of Flaherty and colleagues [18]. In our experience, the distinction between the commonly occurring low-level index cross-talk and a reagent contamination event is clear when the data are examined. Following a true reagent contamination event, the negative control specimens (and indeed many specimens) contain hundreds to thousands of “off-target” parasite-derived reads, requiring preparation of a new library, whereas index cross-talk will—on average—result in approximately 20 off-target parasite-derived reads in the negative control specimens [18].

Assays similar to the UPDx system have been described. Cannon et al. [44] developed a pan-parasite assay requiring multiplexing of thirteen primer pairs that distinguishes a broad range of parasites, including the microsporidia. The amplicons generated by this assay range from between 200 and 450 bases (depending on the taxon) which is comparable to UPDx (~ 200 bases for all taxa) [44]. An important difference between the UPDx assay and the one described by Cannon et al. is the use of a single primer set for UPDx that captures a smaller amplicon to improve amplification efficiency. To achieve an analytical sensitivity (LOD) similar to qPCR (as was deemed necessary for a diagnostic test), UPDx also incorporates a nested PCR step. Given CDC’s role as a public health agency, the analytical sensitivity was an important consideration for UPDx, as was obtaining a sufficient level of discriminatory power—at least enough to inform clinical decisions. The assay described by Cannon et al. [44] may possess additional discriminatory power compared to the UPDx assay as it generates larger amplicons, perhaps at the expense of its analytical sensitivity—though this requires experimental substantiation—and assay simplicity, given it required multiplexing of several targets. Regardless, the issue of multiplexing multiple targets may become less challenging with the introduction of novel multiplexing technologies, such as the CleanPlex protocol utilized by Tessema et al. [45] to amplify 100 targets from dried

blood spots for characterization of *P. falciparum* genotypes. Schwabl and colleagues were able to multiplex 203 primer pairs within the same PCR reaction prior to Illumina sequencing [46], though this number of targets is likely unnecessary for a diagnostic assay and would almost certainly have a negative impact on the assays’ LOD. In any case, in line with these developments, the addition of different molecular targets to the assay described here is being explored for future iterations, to provide additional discriminatory power for certain taxa (discussed below) where this is limited by the short length of our UPDx amplicon.

One limitation of the UPDx assay is that it does not differentiate every parasite to the species level and cannot differentiate certain taxa to the genus level; the level of differentiation is taxon dependent. For example, the main malaria parasites of humans can each be differentiated to the species level, with the added benefit that these species each possess at least two distinct 18S rDNA types, providing additional granularity (Table 2 and Fig. 5). *Babesia* were divided among 3 groups: one that includes *B. microti*, another including *B. duncani*, and the third containing *B. divergens* and *B. divergens*-like MO1 parasites. While we cannot exclude the possibility that spurious infections caused by rare, zoonotic *Babesia* species cannot be distinguished from any of these three groups, the main *Babesia* species found in humans can be differentiated from each other. The filarial nematodes can be differentiated beyond the family level (*Onchocercidae*) but not to the genus level; *Wuchereria* and *Loa* produce amplicons that are indistinguishable. Furthermore, a second group of identical sequences included a sequence from *Brugia pahangi* that was identical to a sequence from a clinical specimen containing *B. malayi* (Table 2, specimen 16), and a sequence from *Dirofilaria repens*. Amplicons generated for *Mansonella perstans* (Table 2, sample 18) are identical to other *Filarioidea* sequences including a sequence from *Dipetalonema* sp. (GenBank accession DQ531723.1), which is, however, of veterinary significance but was included here for comparison. These parasites yielding indistinguishable amplicons may be differentiated either by their differing geographical distribution or using a specific PCR for any sympatric species.

The differentiation potential of UPDx is probably lowest for the kinetoplastids, the most clinically important of these being various *Leishmania* species, *T. cruzi*, and *T. brucei*, which are the causative agents of leishmaniasis, Chagas disease, and African Sleeping Sickness (Human African Trypanosomiasis), respectively. Using the amplicon sequences captured by UPDx (Fig. 5), most human-infecting *Leishmania* species cannot be distinguished. However, these parasites also cannot be differentiated morphologically, so downstream molecular

testing is required for a species-level diagnosis when *Leishmania* is observed microscopically in culture or tissue biopsies. We note that other monoxenous trypanosomatids of the subfamily *Leishmaniinae* (e.g., *Leptomonas*) cannot be differentiated from *Leishmania* species by UPDx, though the *Leishmaniinae* that infect humans are almost exclusively members of the genus *Leishmania*, aside from rare circumstances where infections caused by monoxenous species have been reported [47]. The *Leishmaniinae* are unambiguously distinguishable from the other kinetoplastids infecting humans—*T. brucei* and *T. cruzi*. Furthermore, *Trypanosoma cruzi* is clearly separated from *T. brucei*, although *T. b. gambiense* and *T. b. rhodesiense* cannot be distinguished. Similarly, the sequence generated for *T. cruzi* distinguishes it from the closely related *Trypanosoma rangeli* based on *T. rangeli* sequences available in GenBank (accessions: XR_003828669.1, KJ742907.1, FJ900242.2, AY491767.1, AJ012416.1, AJ012414.1, AJ009160.1, AF065157.1), each of which possess a single SNP and an indel differentiating them from the homologous *T. cruzi* amplicon sequences. The UPDx haplotype for *T. brucei* is identical to that predicted for other trypanosomes of primarily veterinary importance, such as *T. b. brucei*, *T. equiperdum*, and *T. evansi* (Fig. 5). Although *T. b. brucei* and *T. evansi* may very rarely cause human disease [48, 49], these are considered to be aberrant and often self-limiting events [48]. In most cases, these trypanosomes can be differentiated based on allopatric geographical range and differing clinical features. Despite this limitation, the amplicon captured by UPDx can potentially discriminate between a broad range of parasite taxa including numerous protozoa, nematodes, cestodes, and trematodes based on preliminary analyses of sequences available in GenBank (Supplementary file S1, Appendix B).

The size of the external nested amplicon (approximately 2 kb) represents another limitation of this assay owing to the difficulty in identifying priming sites conserved across the many diverse parasite taxa. It is acknowledged that PCR amplification efficiency is lower for longer amplicons, and when the assay is applied to DNA extracts of a lesser quality, such as extracts from formalin-fixed specimens, or DNA extracts that were incorrectly stored, a slight loss of sensitivity may be observed. However, in instances where DNA is extracted from correctly stored, unpreserved (non-formalin-fixed) specimens, the assays' performance should reflect the results described here.

Despite its limitations, the UPDx assay described here shows great potential for routine diagnostic use. It offers a single test for universal blood parasite detection that is comparable in LOD to qPCR, making it more amenable to routine use than the assay previously described by

Flaherty et al. [18]. Aside from detecting multiple parasite species, the diagnostic result is objective; separation of some taxa is not always achieved, but the assay is not subject to false positive results that may result from spurious off-target amplifications, for example. A major benefit of UPDx is that no prescient knowledge of the potential causative agent is required, which is helpful in complex clinical scenarios where the appropriate path for a differential diagnosis might be elusive. The UPDx primer sets should also amplify parasitic agents from mammals, birds, and reptiles. Furthermore, the restriction sites utilized are seemingly conserved across all vertebrates. We confirmed via extensive sequence alignments that even across different vertebrate classes the expected internal amplicon is between 98 and 100% identical possessing one or both of the necessary BamHI and/or XmaI restriction sites, and the PstI restriction site was present in the majority of reptile, bird, amphibian, fish (cartilaginous and bony), and mammalian 18S sequences examined. Consequently, UPDx will likely be generalizable to a large number of animal species.

Conclusions

This work expands on the observations of Flaherty et al. [18] regarding the benefits of host DNA restriction digestion prior to PCR by introducing several novel modifications intended to improve the utility of TADS for parasite detection and characterization of parasite communities in biological matrices derived from vertebrates. These modifications improved the detection limit for human blood parasites by approximately 10-fold, compared to the earlier assay, to a level that is comparable with real-time PCR. As a consequence of this improved detection limit, the potential adoption of UPDx to routine diagnostic settings is currently being explored. Due to the complexity of UPDx and the relatively high cost of NGS, this assay is currently most appropriate for adoption by state and national parasitic disease reference laboratories, to which state and regional laboratories often refer “difficult” cases that are diagnostically challenging; it is in these circumstances where UPDx offers a clear benefit. This assay is currently being validated for routine use at the Centers for Disease Control and Prevention, and New York State (Wadsworth) Parasitic Disease Reference Laboratories. With continued development including novel library preparation strategies to reduce setup time and running costs, our UPDx has the potential to become increasingly amenable to the routine diagnosis of parasitic pathogens commonly found in human blood.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-020-00939-1>.

Additional file 1.

Abbreviations

CDC: Centers for Disease Control and Prevention;
EDTA: Ethylenediaminetetraacetic acid; NGS: Next-generation sequencing;
TADS: Targeted amplicon deep sequencing; UPDx: Universal parasite diagnostic; LOD: Limit of detection; PCR: Polymerase chain reaction;
qPCR: Quantitative PCR

Acknowledgements

We would like to thank the following people for donating specimens for use in this study: Henry Bishop (CDC, PDB), Andrew Moorhead (UGA, FR3), Marcos de Almeida (CDC, PDB), Vita Cama (CDC, PDB), Amy Kong (CDC Malaria Branch), Stella Chenet (Institute of Public Health, Chile), Maria Isabel Jercic (Institute of Public Health, Chile), Theresa Benedict (CDC, PDB), Stephen Hajduk (UGA). We sincerely thank Mathilde Gondard for depositing the Illumina data in the NCBI SRA database under the UPDx BioProject.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

Authors' contributions

BRF contributed to assay design, optimized the experimental design, performed experiments, analyzed the data, generated figures, and prepared manuscript drafts. JB contributed to assay design, performed bioinformatic analyses, generated figures, analyzed data, and edited manuscript drafts. ML performed experiments, data analysis, library preparation for sequencing, and edited manuscript drafts. ET contributed to the experimental design, project conception, and project supervision and edited manuscript drafts. RSB conceived the project, obtained funding, supervised the study, sourced and/or collected specimens for testing, assisted with experimental design, and prepared manuscript drafts. All authors read and approved the final manuscript.

Funding

Funding for this work was provided by the Centers for Disease Control and Prevention Advanced Molecular Detection (AMD) initiative and the CDC's Parasitic Diseases Branch (PDB) within the Division of Parasitic Diseases and Malaria (DPDM).

Availability of data and materials

All sequence data generated as part of this study is available in the SRA database accessible through NCBI. These have been deposited under BioProject accession number PRJNA437674. Specimens relevant to the present study submitted under this BioProject include the text "nUPDx method validation" in their title.

Ethics approval and consent to participate

Ethics approval for the use of anonymized, de-identified, non-reidentifiable blood samples as non-engaged research was granted by Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria Human Subjects Review, approval number 2016-314. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

Consent for publication

Not applicable.

Competing interests

The authors of this manuscript have no competing interests to disclose.

Author details

¹Parasitic Diseases Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA. ²Oak Ridge Institute for Science and Education, Oak Ridge, TN, USA. ³Synergy America Inc., Duluth, GA, USA. ⁴Malaria Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA. ⁵School of Health and Life Sciences, Federation University, Ballarat, Australia.

Received: 30 April 2020 Accepted: 14 October 2020

Published online: 02 January 2021

References

1. Belak S, Karlsson OE, Leijon M, Granberg F. High-throughput sequencing in veterinary infection biology and diagnostics. *Rev Sci Tech*. 2013;32(3):893–915.
2. Talundzic E, Ravishankar S, Kelley J, Patel D, Plucinski M, Schmedes S, Ljolje D, Clemons B, Madison-Antenucci S, Arguin PM, et al. Next-generation sequencing and bioinformatics protocol for malaria drug resistance marker surveillance. *Antimicrob Agents Chemother*. 2018;62(4):e02474–17.
3. Armstrong GL, MacCannell DR, Taylor J, Carleton HA, Neuhaus EB, Bradbury RS, Posey JE, Gwinn M. Pathogen genomics in public health. *N Engl J Med*. 2019;381(26):2569–80.
4. Granberg F, Balint A, Belak S. Novel technologies applied to the nucleotide sequencing and comparative sequence analysis of the genomes of infectious agents in veterinary medicine. *Rev Sci Tech*. 2016;35(1):25–42.
5. Van Borm S, Belak S, Freimanis G, Fusaro A, Granberg F, Hoper D, King DP, Monne I, Orton R, Rosseel T. Next-generation sequencing in veterinary medicine: how can the massive amount of information arising from high-throughput technologies improve diagnosis, control, and management of infectious diseases? *Methods Mol Biol*. 2015;1247:415–36.
6. Wilson MR, O'Donovan BD, Gelfand JM, Sample HA, Chow FC, Betjemann JP, Shah MP, Richie MB, Gorman MP, Haji-Ali RA, et al. Chronic meningitis investigated via metagenomic next-generation sequencing. *JAMA Neurol*. 2018;75(8):947–55.
7. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, Federman S, Stryke D, Briggs B, Langelier C, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med*. 2019;380(24):2327–40.
8. Saha S, Ramesh A, Kalantar K, Malaker R, Hasanuzzaman M, Khan LM, Mayday MY, Sajib MSI, Li LM, Langelier C, et al. Unbiased metagenomic sequencing for pediatric meningitis in Bangladesh reveals neuroinvasive chikungunya virus outbreak and other unrealized pathogens. *mBio*. 2019;10(6):e02877–19.
9. Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect*. 2018;24(4):335–41.
10. Yang Y, Hu X, Min L, Dong X, Guan Y. Balamuthia mandrillaris-related primary amoebic encephalitis in China diagnosed by next generation sequencing and a review of the literature. *Lab Med*. 2020;51(2):e20–6.
11. Wilson MR, Shanbhag NM, Reid MJ, Singhal NS, Gelfand JM, Sample HA, Benkli B, O'Donovan BD, Ali IK, Keating MK, et al. Diagnosing Balamuthia mandrillaris encephalitis with metagenomic deep sequencing. *Ann Neurol*. 2015;78(5):722–30.
12. Haston JC, Rostad CA, Jerris RC, Milla SS, McCracken C, Pratt C, Wiley M, Prieto K, Palacios G, Shane AL, et al. Prospective cohort study of next-generation sequencing as a diagnostic modality for unexplained encephalitis in children. *J Pediatric Infect Dis Soc*. 2020;9(3):326–33.
13. Xie M, Zhou Z, Guo S, Li Z, Zhao H, Deng J. Next-generation sequencing specifies *Angiostrongylus eosinophilic* meningoencephalitis in infants: two case reports. *Medicine (Baltimore)*. 2019;98(35):e16985.
14. Grumaz S, Stevens P, Grumaz C, Decker SO, Weigand MA, Hofer S, Brenner T, von Haeseler A, Sohn K. Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med*. 2016;8(1):73.
15. Brenner T, Decker SO, Grumaz S, Stevens P, Bruckner T, Schmoek T, Pletz MW, Bracht H, Hofer S, Marx G, et al. Next-generation sequencing diagnostics of bacteremia in sepsis (Next GeneSIS-Trial): study protocol of a prospective, observational, noninterventional, multicenter, clinical trial. *Medicine (Baltimore)*. 2018;97(6):e9868.
16. Somasekar S, Lee D, Rule J, Naccache SN, Stone M, Busch MP, Sanders C, Lee WM, Chiu CY. Viral surveillance in serum samples from patients with acute liver failure by metagenomic next-generation sequencing. *Clin Infect Dis*. 2017;65(9):1477–85.
17. Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, Egger R, Dimitrov B, Juan YF, Sheu J, et al. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS One*. 2019;14(6):e0218318.
18. Flaherty BR, Talundzic E, Barratt J, Kines KJ, Olsen C, Lane M, Sheth M, Bradbury RS. Restriction enzyme digestion of host DNA enhances universal detection of parasitic pathogens in blood via targeted amplicon deep sequencing. *Microbiome*. 2018;6(1):164.

19. Gomez CA, Budvytiene I, Zemek AJ, Banaei N. Performance of targeted fungal sequencing for culture-independent diagnosis of invasive fungal disease. *Clin Infect Dis*. 2017;65(12):2035–41.
20. Early AM, Daniels RF, Farrell TM, Grimsby J, Volkman SK, Wirth DF, MacInnis BL, Neafsey DE. Detection of low-density *Plasmodium falciparum* infections using amplicon deep sequencing. *Malar J*. 2019;18(1):219.
21. Huggins LG, Koehler AV, Ng-Nguyen D, Wilcox S, Schunack B, Inpankaew T, Traub RJ. A novel metabarcoding diagnostic tool to explore protozoan haemoparasite diversity in mammals: a proof-of-concept study using canines from the tropics. *Sci Rep*. 2019;9(1):12644.
22. Barratt JLN, Lane M, Taludzic E, Richins T, Robertson G, Formenti F, Pritt B, Verocai G, Nascimento de Souza J, Mato Soares N, et al. A global genotyping survey of *Strongyloides stercoralis* and *Strongyloides fuelleborni* using deep amplicon sequencing. *PLoS Negl Trop Dis*. 2019;13(9):e0007609.
23. Pafco B, Cizkova D, Kreisinger J, Hasegawa H, Vallo P, Shutt K, Todd A, Petzelkova KJ, Modry D. Metabarcoding analysis of stronglylid nematode diversity in two sympatric primate species. *Sci Rep*. 2018;8(1):5933.
24. Vickova K, Kreisinger J, Pafco B, Cizkova D, Tagg N, Hehl AB, Modry D. Diversity of *Entamoeba* spp. in African great apes and humans: an insight from Illumina MiSeq high-throughput sequencing. *Int J Parasitol*. 2018;48(7):519–30.
25. Cooper C, Keatley S, Northover A, Gofton AW, Brigg F, Lymbery AJ, Pallant L, Clode PL, Thompson RCA. Next generation sequencing reveals widespread trypanosome diversity and polyparasitism in marsupials from Western Australia. *Int J Parasitol Parasites Wildl*. 2018;7(1):58–67.
26. Lalremruata A, Jeyaraj S, Engleitner T, Joanny F, Lang A, Belard S, Mombongoma G, Ramharter M, Kreamsner PG, Mordmuller B, et al. Species and genotype diversity of *Plasmodium* in malaria patients from Gabon analysed by next generation sequencing. *Malar J*. 2017;16(1):398.
27. Greenberg MVC, Bourchis D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 2019;20(10):590–607.
28. Flaherty BR, Ho TG, Schmidt SH, Herberg FW, Peterson DS, Kennedy EJ. Targeted inhibition of *Plasmodium falciparum* calcium-dependent protein kinase 1 with a constrained J domain-derived disruptor peptide. *ACS Infect Dis*. 2019;5(4):506–14.
29. Rougemont M, Van Saanen M, Sahli R, Hinrikson HP, Bille J, Jaton K. Detection of four *Plasmodium* species in blood from humans by 18S rRNA gene subunit-based and species-specific real-time PCR assays. *J Clin Microbiol*. 2004;42(12):5636–43.
30. Bonnet S, Jouglin M, Malandrin L, Becker C, Agoulon A, L'Hostis M, Chauvin A. Transstadial and transovarial persistence of *Babesia divergens* DNA in *Ixodes ricinus* ticks fed on infected blood in a new skin-feeding technique. *Parasitology*. 2007;134(Pt 2):197–207.
31. de Almeida ME, Steurer FJ, Koru O, Herwaldt BL, Pieniazek NJ, da Silva AJ. Identification of *Leishmania* spp. by molecular amplification and DNA sequencing analysis of a fragment of rRNA internal transcribed spacer 2. *J Clin Microbiol*. 2011;49(9):3143–9.
32. Fink DL, Kamgno J, Nutman TB. Rapid molecular assays for specific detection and quantitation of *Loa loa* microfilaremia. *PLoS Negl Trop Dis*. 2011;5(8):e1299.
33. Prince-Guerra JL, Cama VA, Wilson N, Thiele EA, Likwela J, Ndakala N, Muzinga Wa Muzinga J, Ayebazibwe N, Ndjakani YD, Pitchouna NA, et al. Comparison of PCR methods for *Onchocerca volvulus* detection in skin snip biopsies from the Tshopo Province, Democratic Republic of the Congo. *Am J Trop Med Hyg*. 2018;98(5):1427–34.
34. Bassene H, Sambou M, Fenollar F, Clarke S, Djiba S, Mourembou G, Alioune LY, Raoult D, Mediannikov O. High prevalence of *Mansonella perstans* filariasis in Rural Senegal. *Am J Trop Med Hyg*. 2015;93(3):601–6.
35. Meerstein-Kessel L, Andolina C, Carrio E, Mahamar A, Sawa P, Diawara H, van de Vegte-Bolmer M, Stone W, Collins KA, Schneider P, et al. A multiplex assay for the sensitive detection and quantification of male and female *Plasmodium falciparum* gametocytes. *Malar J*. 2018;17(1):441.
36. Hofmann N, Mwingira F, Shekalaghe S, Robinson LJ, Mueller I, Felger I. Ultra-sensitive detection of *Plasmodium falciparum* by amplification of multi-copy subtelomeric targets. *PLoS Med*. 2015;12(3):e1001788.
37. Amaral LC, Robortella DR, Guimaraes LFF, Limongi JE, Fontes CJF, Pereira DB, de Brito CFA, Kano FS, de Sousa TN, Carvalho LH. Ribosomal and non-ribosomal PCR targets for the detection of low-density and mixed malaria infections. *Malar J*. 2019;18(1):154.
38. Minoche AE, Dohm JC, Himmelbauer H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol*. 2011;12(11):R112.
39. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 2012;13:341.
40. Mathison BA, Couturier MR, Pritt BS. Diagnostic identification and differentiation of microfilariae. *J Clin Microbiol*. 2019;57(10):e00706–19.
41. Worges M, Whitehurst N, Saye R, Ndiaye D, Yamo E, Yukich J. Performance outcomes from Africa-based malaria diagnostic competency assessment courses. *Am J Trop Med Hyg*. 2019;100(4):851–60.
42. Mace KE, Arguin PM, Tan KR. *Malaria Surveillance - United States, 2015*. *MMWR Surveill Summ*. 2018;67(7):1–28.
43. Klion AD. Filarial infections in travelers and immigrants. *Curr Infect Dis Rep*. 2008;10(1):50–7.
44. Cannon MV, Bogale H, Rutt L, Humphrys M, Korpe P, Duggal P, Ravel J, Serre D. A high-throughput sequencing assay to comprehensively detect and characterize unicellular eukaryotes and helminths from biological and environmental samples. *Microbiome*. 2018;6(1):195.
45. Tessema SK, Hathaway NJ, Teyssier NB, Murphy M, Chen A, Aydemir O, Duarte EM, Simone W, Colborn J, Saute F, et al. Sensitive, highly multiplexed sequencing of microhaplotypes from the *Plasmodium falciparum* heterozygote. *J Infect Dis*. 2020;jiaa527. <https://doi.org/10.1093/infdis/jiaa527>.
46. Schwabl P, Sánchez JM, Costales JA, Ocaña S, Segovia M, Carrasco HJ, Hernández C, Ramírez JD, Lewis MD, Grijalva MJ, et al. Genome-wide locus sequence typing (GLST) of eukaryotic pathogens. *bioRxiv*. 2020;2020.03.24.003590. <https://doi.org/10.1101/2020.03.24.003590>.
47. Kaufer A, Ellis J, Stark D, Barratt J. The evolution of trypanosomatid taxonomy. *Parasit Vectors*. 2017;10(1):287.
48. Truc P, Buscher P, Cuny G, Gonzatti MI, Jannin J, Joshi P, Juyal P, Lun ZR, Mattioli R, Pays E, et al. Atypical human infections by animal trypanosomes. *PLoS Negl Trop Dis*. 2013;7(9):e2256.
49. Van Vinh CN, Buu Chau L, Desquesnes M, Herder S, Phu Huong Lan N, Campbell JJ, Van Cuong N, Yimming B, Chalermwong P, Jittapalpong S, et al. A clinical and epidemiological investigation of the first reported human infection with the zoonotic parasite *Trypanosoma evansi* in Southeast Asia. *Clin Infect Dis*. 2016;62(8):1002–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

