

Investigating the etiologies of non-malarial febrile illness in Senegal using metagenomic sequencing

Received: 22 August 2023

Accepted: 4 January 2024

Published online: 25 January 2024

 Check for updates

Zoë C. Levine ^{1,2,3}, Aita Sene^{4,5}, Winnie Mkandawire^{1,6}, Awa B. Deme⁵, Tolla Ndiaye^{4,5}, Mouhamad Sy^{4,5}, Amy Gaye^{4,5}, Younouss Diedhiou^{4,5}, Amadou M. Mbaye^{4,5}, Ibrahima M. Ndiaye^{4,5}, Jules Gomis^{4,5}, Médoune Ndiop ⁷, Doudou Sene⁷, Marietou Faye Paye¹, Bronwyn L. MacInnis ^{1,8}, Stephen F. Schaffner ^{1,8,9}, Daniel J. Park ¹, Aida S. Badiane^{4,5}, Andres Colubri^{1,6}, Mouhamadou Ndiaye^{4,5}, Ngayo Sy¹⁰, Pardis C. Sabeti ^{1,8,9,11,13} , Daouda Ndiaye^{4,5,13}  & Katherine J. Siddle ^{1,12,13} 

The worldwide decline in malaria incidence is revealing the extensive burden of non-malarial febrile illness (NMFI), which remains poorly understood and difficult to diagnose. To characterize NMFI in Senegal, we collected venous blood and clinical metadata in a cross-sectional study of febrile patients and healthy controls in a low malaria burden area. Using 16S and untargeted sequencing, we detected viral, bacterial, or eukaryotic pathogens in 23% (38/163) of NMFI cases. Bacteria were the most common, with relapsing fever *Borrelia* and spotted fever *Rickettsia* found in 15.5% and 3.8% of cases, respectively. Four viral pathogens were found in a total of 7 febrile cases (3.5%). Sequencing also detected undiagnosed *Plasmodium*, including one putative *P. ovale* infection. We developed a logistic regression model that can distinguish *Borrelia* from NMFI with similar presentation based on symptoms and vital signs (F1 score: 0.823). These results highlight the challenge and importance of improved diagnostics, especially for *Borrelia*, to support diagnosis and surveillance.

Febrile illness is a significant cause of morbidity and mortality in West Africa. While malaria remains the most common single pathogen causing febrile illness, its incidence has decreased sharply in the last two decades; in Senegal, for example, control measures decreased malaria incidence from 122 per 1000 in 2006 to 59 per 1000 in 2021^{1,2}. With declining incidence and increased use of malaria rapid diagnostic tests, the importance of non-malarial febrile illness (NMFI) has become

more apparent. Unlike malaria, however, many of the pathogens causing NMFI are not the target of robust surveillance programs, and rapid diagnostic tests (RDTs), the backbone of both clinical care and surveillance in peripheral care settings, are not available³.

In the absence of comprehensive surveillance efforts that capture multiple pathogen types, appropriate public health interventions are hindered by our limited understanding of the full landscape of

¹Broad Institute of Harvard and MIT, Cambridge, MA, USA. ²Harvard Graduate Program in Biological and Biomedical Science, Boston, MA, USA. ³Harvard/MIT MD-PhD Program, Boston, MA, USA. ⁴Department of Parasitology, Cheikh Anta Diop University Dakar, Dakar, Senegal. ⁵Centre International de Recherche et de Formation en Génomique Appliquée et de la Surveillance Sanitaire, Dakar, Senegal. ⁶University of Massachusetts Medical School, Worcester, MA, USA. ⁷Programme National de lutte contre le Paludisme, Ministère de la Santé, Dakar Fann, Senegal. ⁸Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. ⁹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. ¹⁰Service de Lutte Anti Parasitaire, Thies, Senegal. ¹¹Howard Hughes Medical Institute, Chevy Chase, MD, USA. ¹²Department of Molecular Microbiology and Immunology, Brown University, Providence, RI, USA. ¹³These authors jointly supervised this work: Pardis C. Sabeti, Daouda Ndiaye, Katherine J. Siddle. ✉ e-mail: pardis@broadinstitute.org; daouda.ndiaye@igass.org; katherine_siddle@brown.edu

common causes of NMFI at the community level. Untargeted sequencing, also known as metagenomic sequencing (mNGS), is a powerful tool for detection of microbial nucleic acids in clinical samples without a priori knowledge of a pathogen^{4,5} and is increasingly used for surveillance in regions at high risk of emerging and reemerging disease^{6–8}. As mNGS sequences all RNA or DNA in a sample, these techniques are typically less sensitive for detecting any single pathogen than targeted approaches (e.g., PCR), due to the abundance of the host relative to pathogen nucleic acids^{4,5}. In order to achieve higher sensitivity and reduce the cost of deep sequencing, mNGS can be applied to cell-free fluids, which have lower host backgrounds, but this can limit the detection of intracellular pathogens. Untargeted approaches are complemented by more targeted strategies, such as 16S sequencing for bacterial pathogens, that specifically amplify pathogen nucleic acids, reducing required sequencing depth and simplifying interpretation. This technique has been applied to detect bacterial bloodstream infections^{9,10} and tick-borne bacterial illness¹¹.

A broad range of pathogens across kingdoms can cause NMFI, but surveillance studies often focus on one or a few pathogens or are limited to hospitalized patients and, therefore, are limited in their ability to detect all causes of disease. Surveillance studies of severely ill hospitalized patients have identified bacterial pathogens, including several vaccine-preventable illnesses such as *Streptococcus pneumoniae* and *Neisseria meningitidis*^{12,13}. Arboviruses, such as Dengue virus and Chikungunya virus have been detected in time-limited outbreaks in Senegal^{14,15}. Community and clinic-based studies of patients with fever have revealed bacterial zoonoses are a common cause of ambulatory febrile illness in Senegal, including tick-borne relapsing fever, Rickettsioses, Q fever, and *Bartonella*^{16–21}. However, no unbiased surveillance of outpatient febrile illness has been done in Senegal to date.

This study aimed to determine and characterize causes of ambulatory NMFI in Thiès, Senegal, a peri-urban community with overall low malaria incidence where malaria transmission dynamics have been deeply characterized²². We collected plasma samples from a population of acutely febrile patients presenting to the Service de Lutte Anti Parasitaire (SLAP) outpatient clinic and healthy controls from the surrounding neighborhoods during the dry and rainy seasons of 2018 and 2019. We aimed to provide insights on clinical presentations of NMFI that can guide providers at the point of care and genomic characterization to inform the design of new detection tools for clinical diagnosis and public health surveillance.

Results

Characterization of NMFI reveals viral, bacterial, and eukaryotic pathogens

We first characterized the pathogens in plasma samples collected from acutely febrile patients suspected of malaria and healthy controls across the dry (febrile: $n = 100$, healthy: $n = 54$) and rainy (febrile: $n = 104$, healthy: $n = 50$) seasons in 2019 using both untargeted and targeted approaches (Supplementary Fig. 1a, Supplementary Fig. 1b). All febrile patients received a malaria RDT; 39.4% (41/104) of patients tested positive for malaria in the wet season and no cases were detected in the dry season. Febrile cases were roughly equally split between children and adults (54% ≥ 18 yrs). Age distribution was similar ($p = 0.98$, Mann–Whitney, two-sided) across case and control groups, but there were more male febrile cases (Male: $n = 111$, Female: $n = 93$) and more female controls (Male: $n = 47$, Female: $n = 57$, $p = 0.15$ Fisher exact two-sided, Supplementary Fig. 1a).

We detected viral, bacterial, and eukaryotic pathogens by sequencing. To detect viral pathogens, we performed RNA-mNGS. We also evaluated the ability of RNA-mNGS to detect non-viral pathogens, including malaria and fungi. To detect bacterial infections, we sequenced the v1–2 region of the 16S rRNA gene, which permitted us to classify the bacterial taxa present in samples with high bacterial load [see methods]. We detected at least one pathogen in the plasma in 23%

(38/163) of RDT-negative acutely febrile patients and co-infections in 7% (3/41) of RDT-positive acutely febrile patients (Fig. 1b). The most common febrile pathogen was *Borrelia*, which was found across the dry ($n = 8$) and rainy ($n = 10$) seasons and in both RDT-negative ($n = 17$) and RDT-positive ($n = 1$) patients. We also detected bacterial infections with *Rickettsia* and *Arcobacter* (Fig. 1d). Bacterial pathogens found as co-infections included two *Borrelia/Rickettsia* co-infections, two *Borrelia/Plasmodium* co-infections, and three *Plasmodium/Rickettsia* co-infections (Fig. 1e). We did not detect any viral/bacterial or viral/*Plasmodium* coinfections.

We identified four known vertebrate viruses: Dengue virus (DENV, $n = 2$), Hepatitis B virus (HBV, $n = 2$), Parvovirus B19 ($n = 2$), and Human immunodeficiency virus 1 (HIV-1, $n = 1$) as well as a human virus not currently believed to cause disease, Human pegivirus 1 (HPgV-1, $n = 4$) (Fig. 1c). Four febrile patients and one healthy control exhibited a high proportion of fungal RNA reads in the plasma, as compared to healthy controls (>99th percentile reads/million raw reads (rpm) for healthy controls), but these reads did not map to any specific fungal pathogen (Supplementary Fig. 2). We did not detect any vertebrate viruses in healthy controls, but we did detect *Borrelia* ($n = 1/35$ samples sequenced) and *Rickettsia* ($n = 1/35$) (Supplementary Table 1).

RNA-mNGS identifies viral pathogens known to circulate in Senegal

We next considered the genetic diversity of detected viral pathogens to determine their relationship to other circulating strains in West Africa. Phylogenetic analysis of complete DENV, HBV, and Parvovirus B-19 genomes indicated that in all three cases, the two genomes were different genotypes and thus not closely related (Supplementary Table 2). Notably, there was a DENV outbreak across Senegal, including Thiès, in 2018. Whole genome phylogenetic analysis revealed the closest relatives to the 2019 DENV3 genome from this study were DENV3 genomes from patients presenting to the SLAP clinic during that outbreak¹⁴ (Supplementary Fig. 3a). Conversely, the DENV1 genome in this study was more closely related to other genomes from West Africa than to the DENV1 genomes from the 2018 outbreak (Supplementary Fig. 3a). The HIV-1 genome could not be assembled or genotyped due to low read count (mean coverage 0.2), but reads mapped across the reference genome including in the *gag*, *pol*, and *env* genes.

While HPgV-1 is known to infect asymptomatic individuals across the world, we only detected HPgV-1 in febrile patients. The prevalence of HPgV-1 has not been well studied in Senegal; we found infection in only 1.3% (4/307) of plasma samples sequenced, lower than the 4–11% reported in blood-donor surveys from Sub-Saharan Africa^{23,24}. All four genomes were clustered with publicly available genomes from human hosts in Sierra Leone, Uganda, and Cameroon and belonged to Genotype I (Supplementary Fig. 3b), the most commonly circulating HPgV-1 genotype in West Africa²⁵.

In order to identify divergent viral species missed at the read level, we performed a translated nucleic acid search of de novo contigs, which identified RNA-dependent RNA Polymerase (RdRp) sequences for two candidate novel viruses, one in the *Naraviridae* family and one in the *Reoviridae* family. We detected a *Reoviridae* RdRp sequence across 18 febrile patients, 12 healthy controls, and 4 non-template controls, suggesting a likely contaminant (Supplementary Fig. 4). However, given that *Reoviruses* have been isolated from ill patients²⁶, we further investigated these sequences. The *Reovirus* RdRp sequences from this study clustered together but were distant from mammalian *Orthoreoviruses*. Given the distance from mammalian *Orthoreovirus* species and presence in non-template controls, these sequences were likely from a contaminated reagent rather than true human infections. We also identified a *Naraviridae* RdRp sequence across 5 febrile patients

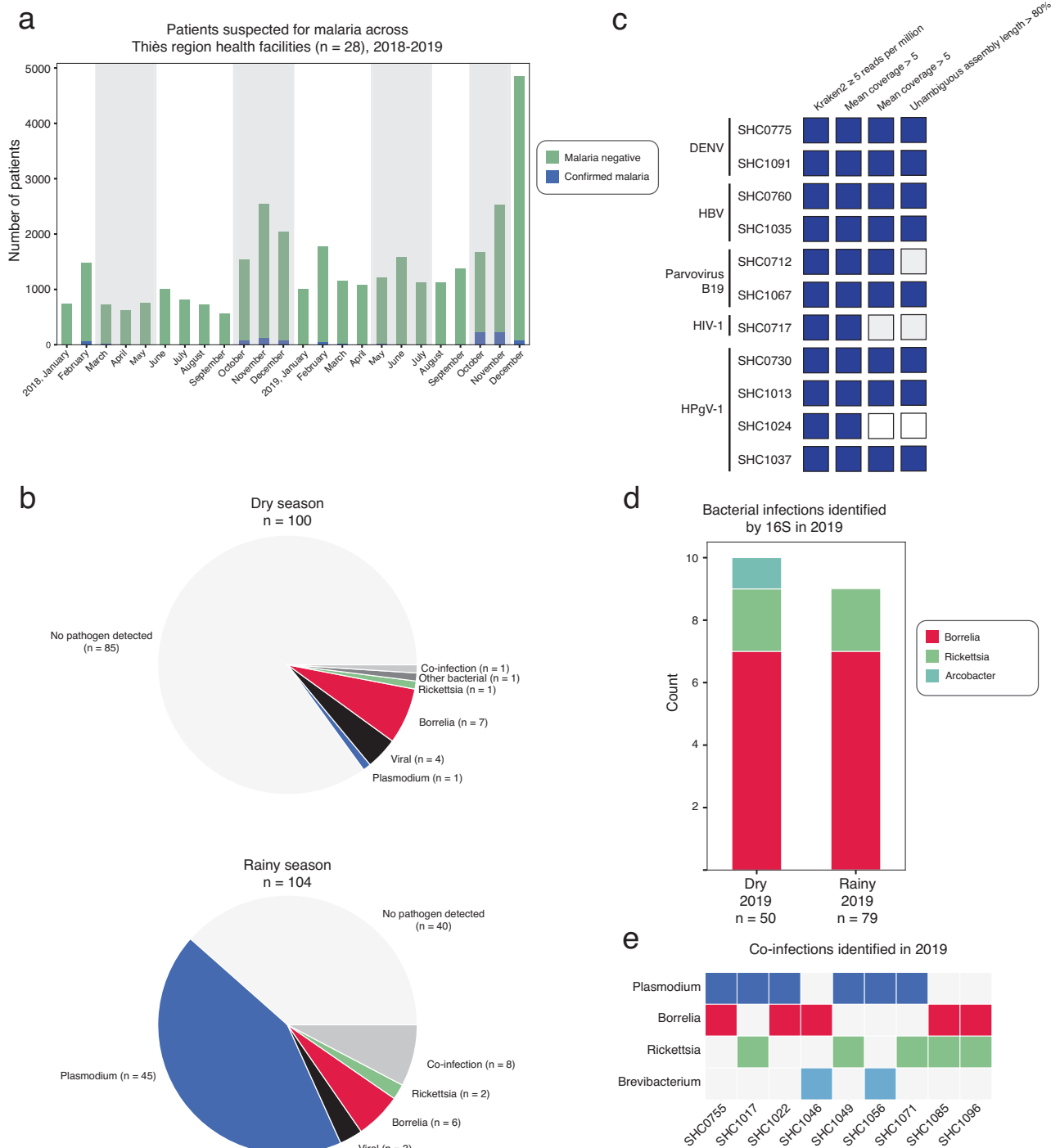


Fig. 1 | Overview of the burden of malaria and NMFI in Thiès, Senegal and NMFI pathogens identified in 2019. a Patients suspected of malaria across 28 health facilities in the Thiès region (data provided by Program National de lutte contre le Paludisme); gray shading indicates periods of sample collection for this study. **b** Types of pathogens detected in febrile patients in dry and rainy seasons in 2019.

Plasmodium detected by RDT or RNA-mNGS, bacterial infections detected by 16S amplicon sequencing, and viral infections detected by RNA-mNGS. **c** Viral infections detected by RNA-mNGS. **d** Bacterial infections detected by 16S sequencing and **e** Co-infections detected in febrile patients from 2019. Source data are provided as a Source Data file.

and 1 healthy control (Supplementary Fig. 5), none of whom had any other identifiable pathogen. Because viruses in the *Narnaviridae* family, which infect plants and fungi, have not been reported as human pathogens and we detected this novel *Narnavirus* species in both febrile cases and a healthy control, the role of this virus in disease, if any, is unclear.

Untargeted RNA sequencing detects *Plasmodium* cases missed by RDT

If RNA-mNGS is to be employed as a tool for the diagnosis of febrile illness in malaria-endemic regions, we also need to understand the ability of RNA-mNGS to detect malaria as compared to clinical diagnostics. We know that RDTs have limited sensitivity, detect only certain

Plasmodium species, and are susceptible to false negatives in genetically diverse parasites, so we evaluated our RNA-mNGS reads for evidence of *Plasmodium* infections.

RNA-mNGS detected abundant *Plasmodium* nucleic acids (>550 rpm) in 6.1% (10/163) of RDT-negative acutely febrile patients and one healthy control (1/104), suggesting these were false negatives. Among RDT positive patients, there was a wide range in the abundance of *Plasmodium* reads (mean 1.7×10^4 rpm, range $1.5 \times 10^2 - 2.6 \times 10^5$ rpm) (Supplementary Fig. 6a). Overall, 63% of RDT positive patients (26/41) and 77% of smear-positive patients (17/22) were also positive by untargeted RNA sequencing (>550 rpm) (Supplementary Fig. 6c). In patients with detectable parasitemia by blood smear examination, *Plasmodium* reads did not correlate strongly with parasite density (Pearson's $R = 0.49$, Supplementary Fig. 6b).

In order to determine why 11 samples were negative by RDT but positive by RNA-mNGS, we first confirmed the results with a pan-*Plasmodium* qPCR. We found that 8/11 samples were positive by pan-*Plasmodium* qPCR, and the cycle threshold (CT) correlated with *Plasmodium* RNA-mNGS rpm (Supplementary Fig. 6d). Next, we determined whether any samples belonged to species other than *Plasmodium falciparum*, the target species of the RDTs, using both RNA-mNGS reads and a *P. falciparum*-specific qPCR. As previous work has demonstrated, Kraken2 is inaccurate for species-level classification of parasites²⁷; we used DIAMOND to identify one sample with a contig (771 bp) that matched perfectly (100% nucleotide identity, 100% coverage) to *Plasmodium ovale* cytochrome oxidase subunit I (Accession: KPO50416.1), suggesting a likely *P. ovale* infection. This sample was positive on the pan-*Plasmodium* qPCR assay but negative on the *P. falciparum* qPCR, confirming it as a case of non-*falciparum* malaria (Supplementary Table 3, Supplementary Fig. 6e).

Additionally, we assessed expression levels of histidine-rich protein 2 (PfHRP-2), the target antigen of the RDTs used in this study. Deletions in *pfhrp2* have been detected in Senegal and demonstrated to cause decreased antigen expression and consequently false negative RDT results^{28,29}. The number of RNA-mNGS rpm aligned to *pfhrp2* was significantly different between patients who were RDT positive and RNA-mNGS positive (RDT+/mNGS+) and patients who were RDT negative but *Plasmodium* RNA-mNGS positive (RDT-/mNGS+) (Supplementary Fig. 6f, Mann Whitney two-sided, $p = 0.038$). To assess for the presence of deletions at common breakpoints, we amplified two targets, one in exon 2 of *pfhrp2* gene and one in a flanking gene, which can also be included in deletions³⁰. Amplification of both targets was intact in all qPCR-confirmed *P. falciparum* infections (7/7, Supplementary Fig. 6g), suggesting none of these parasites had deletions in these targets consistent with the low frequency of *pfhrp2* deletion in Senegal, previously estimated at 2.4%²⁹.

16S sequencing confirms a high burden of relapsing fever *Borrelia* and spotted fever *Rickettsia*

As bacteria were the most common cause of NMFI in our 2019 study population, we extended our 16S sequencing to detect bacterial infections across the complete study population from 2018 to 2019 and further investigated the species causing the disease. Among all high bacterial load samples that underwent 16S sequencing [see methods], we detected *Borrelia* in 15.5% (33/213) and *Rickettsia* in 3.8% (8/213) of febrile patients (Fig. 1d, Supplementary Fig. 7c). 16S sequencing also identified other bacterial genera containing known pathogens, including *Brevibacterium*, *Arcobacter*, and *Veillonella*, across both years (Supplementary Fig. 7c). However, given the limited taxonomic resolution of v1-2 16S sequencing, we were unable to determine whether these sequences represented pathogenic species or harmless commensals.

Phylogenetic clustering of *Borrelia* v1-2 16S sequences showed that all *Borrelia* sequences from this study were similar to each other and fell within the relapsing fever (RF) group, and the majority were

most similar to *B. crocidurae* (Supplementary Fig. 8a). RF *Borrelia* circulate worldwide and are known to cause febrile illness in Senegal¹⁶⁻²⁰. The *Rickettsial* v1-2 16S rRNA gene sequences from our study clustered within the Spotted Fever group, most closely related to *R. felis* (Supplementary Fig. 8c). Spotted fever *Rickettsia*, including *R. felis*²¹ and *R. africae*²⁷ have been previously detected in febrile Senegalese patients.

To validate the 16S v1-2 phylogeny, a subset of 6 *Borrelia* positive samples were typed by amplicon sequencing of the 16S-23S intergenic spacer (IGS)³¹. All sample sequences were *Borrelia crocidurae*, in agreement with the 16S v1-2 phylogeny (Fig. 2c, Supplementary Fig. 8b). Sequences from this study were more similar to published sequences from humans and *Ornithodoros sonrai* in Southern Senegal (unpublished), and *Ornithodoros sonrai* ticks in Mali³² than sequences from *Ornithodoros erraticus* in Tunisia³³ and Morocco³⁴ (Fig. 2c).

Sensitive detection of *Borrelia* by 16S and RNA-mNGS

We assessed the ability of RNA-mNGS to detect bacterial infections based on samples that underwent both sequencing methods. *Borrelia* RNA-mNGS reads were detected in 68% (13/19) of 16S positive samples and 100% (17/17) of pan-*Borrelia* qPCR positive samples (Supplementary Fig. 7i). Conversely, *Rickettsia* RNA-mNGS reads were not detected in any *Rickettsia* positive samples. Acute RF *Borrelia* infection is known to cause high titers of bacteria in the blood during febrile episodes, while *Rickettsia* is an obligate intracellular pathogen and, therefore, has low titers in the cell-free plasma³⁵. To assess the extent to which plasma titers impacted the sensitivity of RNA-mNGS to tick-borne bacterial pathogens, we quantified bacterial abundance by the percent of total v1-2 16S sequences from a given sample classified as *Borrelia* or *Rickettsia*, respectively¹¹. We observed *Borrelia* abundances ranging from 5% to 98%, while *Rickettsia* abundances were lower, ranging from 5% to 16%. RNA-mNGS only detected *Borrelia* in samples with abundance >20% (Supplementary Fig. 7f). Taken together, these data suggest that plasma RNA-mNGS can detect high titer bacterial pathogens but will miss less abundant intracellular species.

We compared the sensitivity of our 16S sequencing to qPCR and Giemsa stained blood smear examination, the current gold standard for clinical diagnosis of RF *Borrelia*³⁶. We examined smears from a subset of RF *Borrelia* patients. 28.6% (4/14) of examined smears were positive; all 4 positive patients had a high RF *Borrelia* load ($\geq 90\%$, Mann-Whitney two-sided $p = 0.002$ compared to 16S positive/smear negative) (Supplementary Fig. 7b, d). Although blood smear examination is the most widely used diagnostic, qPCR diagnostics are known to have higher sensitivity^{36,37}. Using a pan-*Borrelia* qPCR assay previously used for the detection of *Borrelia* in patient blood samples³⁸, 79% (27/34) of *Borrelia* samples were confirmed by qPCR, while six samples that were negative by 16S sequencing were positive by qPCR. Bacterial load measured by qPCR and 16S abundance correlated well (Pearson's $R = 0.72$, Supplementary Fig. 7g). While 16S sequencing was done only on a subset of samples, pan-*Borrelia* qPCR was performed on all febrile cases and healthy controls in 2019 and all febrile cases in 2018. An additional five cases of *Borrelia* were identified in febrile patients by qPCR in samples that did not undergo 16S sequencing.

Borrelia infection presents similarly to other febrile illnesses but can be distinguished by key features

We sought to assess the clinical syndrome associated with qPCR-confirmed *Borrelia* infections, compared to RDT-confirmed malaria and non-*Borrelia* NMFI ("other febrile"), to guide differential diagnosis at the point of care. *Borrelia* infections occurred in a consistent proportion of the study population in the two years (Fisher exact $p = 0.61$) and in both seasons (Fisher exact $p = 0.24$). *Borrelia* infections occurred across all ages (t -test compared to all other febrile $p = 0.45$) (Fig. 2b) and at similar rates in male and female patients (Fisher exact

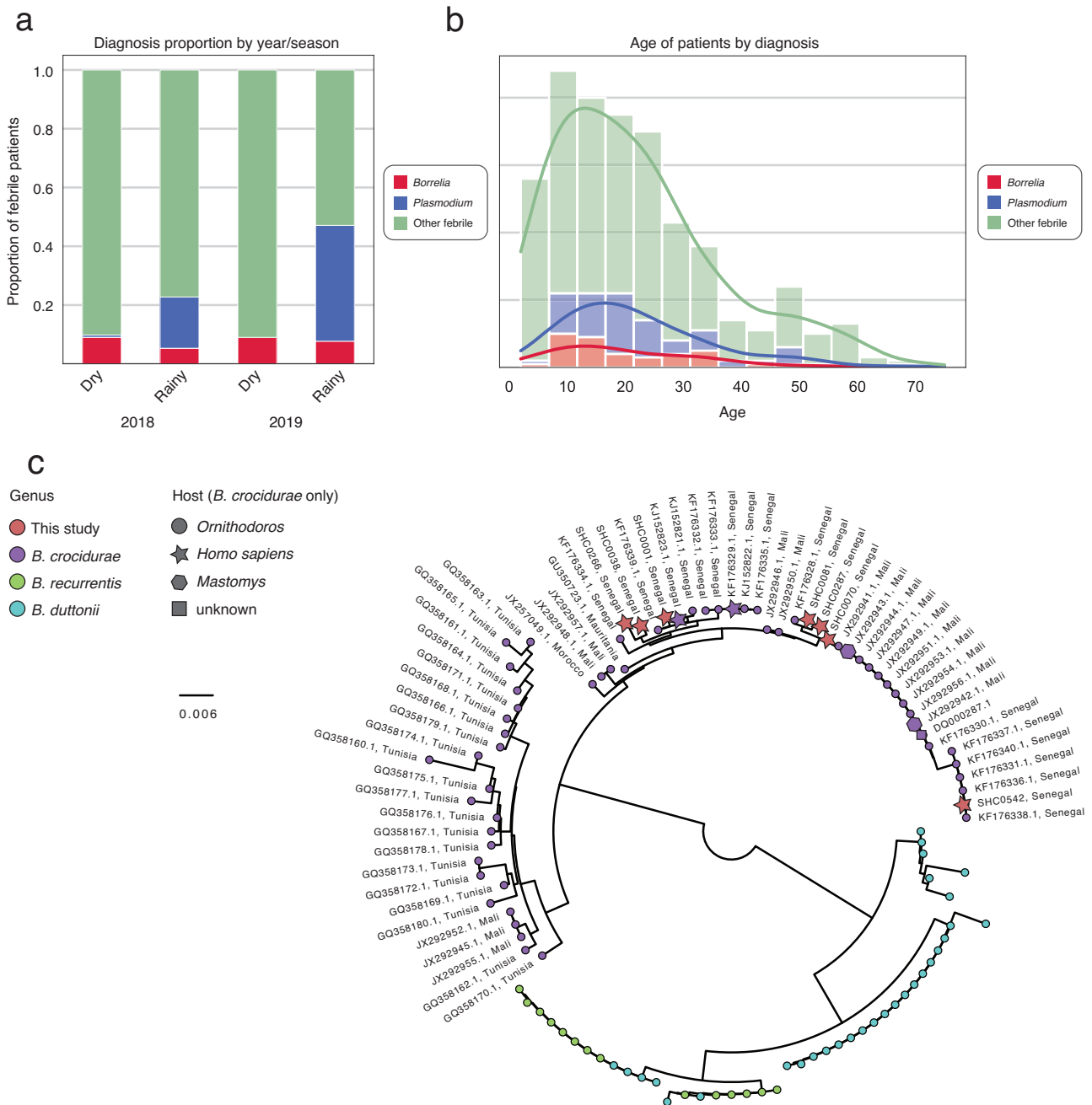


Fig. 2 | Characterization of qPCR-confirmed *Borrelia* and RDT-confirmed *Plasmodium* across both years. **a** Proportion of the study population testing positive for *Borrelia*, *Plasmodium*, or neither (Other febrile) across dry and rainy seasons in 2018 and 2019. **b** Age distribution of *Borrelia*, *Plasmodium*, and other febrile patients. **c** Maximum likelihood phylogenetic tree (midpoint rooted) of IGS

sequences from this study (red) in the context of available reference sequences for *B. crocidurae* (purple), *B. duttonii* (cyan), and *B. recurrentis* (green). For *B. crocidurae* isolates, the location of collection is indicated in the label, and host species are indicated with node shape as available. Source data are provided as a Source Data file.

$p = 0.87$), while male patients were more likely to test positive for malaria (Fisher exact $p = 1.6e - 7$).

A high proportion of *Borrelia*-positive patients reported generalized symptoms that were common across all febrile patients, including headache (100%), body aches (59%), and dizziness (46%), as reported in previous studies of RF *Borrelia* (Fig. 3a)³⁶. *Borrelia* infection was distinguished by vomiting (49%), which was observed more often in *Borrelia*-positive patients (stat = 2.8, $p = 0.0026$, Fisher Exact two-sided). *Borrelia*-positive patients were less likely than other febrile patients to report sore throats (7.7%, stat = 0.21, $p = 0.0039$, Fisher Exact two-sided). Notably, RF *Borrelia* can invade the central nervous

system in severe cases; while we did not observe neurological symptoms such as seizures or loss of consciousness, one *Borrelia*-positive patient reported eye pain, a potential symptom of neuroinvasive infection not observed in any other febrile patients^{36,39}. A high proportion of *Borrelia*-positive patients reported contact with a febrile person (38%), contact with rats (57%), and prior travel (32%), but these exposures were common across malaria and other febrile patients and were not associated with *Borrelia* infection in particular (Fig. 3b).

Borrelia patients showed significant differences in temperature and blood glucose. On average, *Borrelia*-positive patients had a higher fever (mean = 38.6 °C, SD = 1.1 °C) than other febrile patients (mean =

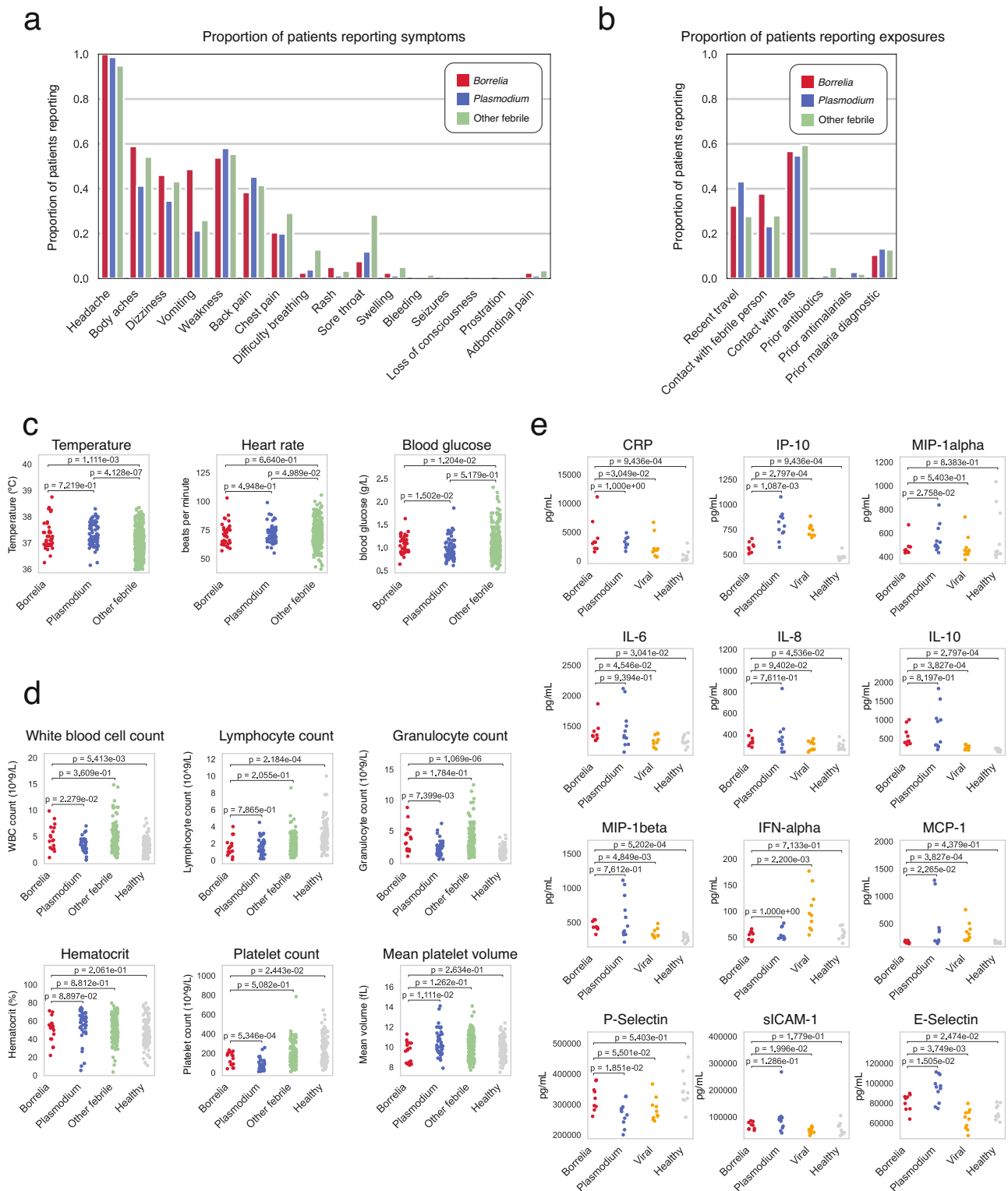


Fig. 3 | Signs, symptoms, and exposures in *Borrelia* positive, *Plasmodium* positive, and other febrile patients. Clinical characteristics across both years of *Borrelia* ($n = 39$), *Plasmodium* ($n = 75$), and other febrile ($n = 412$) patients, including **a** symptoms reported by patients, **b** exposures reported by patients, and **c** vital signs. For 2019 only, **d** blood cell counts in *Borrelia* ($n = 17$), *Plasmodium* ($n = 41$), other febrile ($n = 146$), and healthy ($n = 104$). **e** Serological profile of a subset of

patients with *Borrelia* ($n = 9$), *Plasmodium* ($n = 11$), viral infection ($n = 10$) compared to healthy patients ($n = 10$); cytokines/chemokines with at least one significant difference between groups shown [see Source Data for results for all cytokines/chemokines tested]. All p -values represent the result of the Mann–Whitney–Wilcoxon two-sided test. Source data are provided as a Source Data file.

38.0 °C, SD=1.0 °C, Mann–Whitney two-sided $p = 0.0011$). *Borrelia* patients (mean=1.07 g/L, sd=0.19) were also hypoglycemic compared to other febrile patients (Other febrile: mean=1.22 g/L, sd=3.97, Mann–Whitney two-sided $p = 1.204e-02$) (Fig. 3c).

We assessed whether the addition of a complete blood cell count with differential could help distinguish RF *Borrelia* from other febrile diseases. *Borrelia*-positive patients exhibited abnormal blood counts, including lymphopenia (mean 1.5×10^9 cells/L, std 1.04),

granulocytosis (mean 7.47×10^9 cells/L, std 4.23), and thrombocytopenia (148×10^9 cells/L, std 66) (Fig. 3d). However, this hematologic response was not significantly different from the response observed in other febrile patients. On a subset of patients, we measured 20 cytokines and chemokines to further distinguish *Borrelia* from other infections (Fig. 3e). *Borrelia*-positive patients were distinct from viral infections in elevated CRP (Mann–Whitney two-sided, $p = 0.030$), IL-10 ($p = 0.00038$), and MIP-1Beta ($p = 0.0048$) and decreased IP-10 ($p = 0.00028$), IFN-alpha ($p = 0.0022$), and MCP-1 ($p = 0.00038$) (Fig. 3e).

Given the highly overlapping presentation of the disease, we developed a weighted logistic regression model to distinguish *Borrelia* infection from other NMFI based on clinical symptoms, vital signs, and key demographic information. We evaluated the performance of the model with bootstrapping and found the model was able to predict *Borrelia* infection with high performance (recall: 0.861, 95% CI [0.837–0.910], precision: 0.792, 95% CI [0.696–0.892], and F1-score: 0.823, 95% CI [0.766–0.900]) (Fig. 4a). The model identified features that were significant in univariate analysis, including vomiting, temperature, blood glucose, and sore throat. Common symptoms that were observed across all febrile patients, such as headache, body ache, and dizziness, were not useful predictors on their own but were associated with an increased risk of *Borrelia* when evaluated in the context of other symptoms and exposures. To test whether additional laboratory tests might further increase the performance of the model, we incorporated complete blood count with differential (CBC) values. The addition of CBC values moderately increased model performance (Supplementary Fig. 9), but the difference was not significant, and the test dataset was small ($n = 163$). When blood biomarkers were incorporated, decreased platelet count was an important predictor of *Borrelia*.

Discussion

In this study, we employed untargeted RNA–mNGS and targeted 16S sequencing to understand the causes of acute febrile illness at an ambulatory clinic in Thiès, Senegal. As the first unbiased investigation of causes of fever in Senegal, our findings highlight the importance of looking across many pathogen types simultaneously to understand their contributions—both individually and as co-infections—to disease and enable comparison of the strengths of genomic tools as well as clinical and epidemiological data to support disease characterization.

Our results confirm that arthropod-borne bacterial pathogens, particularly relapsing fever *Borrelia*, are the major identifiable causes of NMFI but remain underdiagnosed. The frequency of these pathogens is broadly consistent with previous findings in targeted studies from other regions of Senegal^{17–20}. *Borrelia spp.* associated with human and zoonotic infection have been identified in ticks across West Africa^{16,36,40,41}, but human surveillance has not been done.

Our broad approaches and sampling illuminated co-infections and missed diagnoses that may be clinically relevant. Both *Borrelia* and *Rickettsia* occurred frequently as co-infections in our study. *Borrelia/Plasmodium* co-infections have been previously detected in Senegal⁴², and murine model evidence suggests co-infection increases the risk of severe malarial illness⁴³. *Rickettsia* was detected more often as a co-infection with *Plasmodium* ($n = 3$) or *Borrelia* ($n = 2$) than as a stand-alone infection ($n = 3$). *Plasmodium/R. felis* co-infections have been previously identified in Senegal⁴⁴, and *Anopheles gambiae*, the primary malaria vector in Senegal, may be able to transmit *R. felis*⁴⁵. Alternatively, Labruna and Walker proposed that *R. felis* may not be a pathogen itself but rather a symbiont of parasites that infect humans, such as protozoa and helminths³⁵. Whether *Rickettsia* is contributing to pathogenesis or simply co-transmitting in the *Rickettsia/Plasmodium* co-infections detected in this study remains an open question. More research is needed to understand the interaction between *Plasmodium* and *R. felis*. Given the common occurrence of bacterial/*Plasmodium* co-infections and the potential for negative outcomes without proper treatment of both pathogens, it is important that both surveillance and diagnostic approaches do not stop at detection of the first pathogen.

Despite using multiple methods and searching across kingdoms, we did not find any pathogen in over 70% of NMFI cases in this study. Although our enrollment criteria focused on suspicion of malaria, many of these patients reported sore throat and difficulty breathing (2019: 35%, 40/113) or abdominal pain (4.4%, 5/113). Given that only cell-free plasma was sequenced, our approach may have missed infections restricted to other body compartments, including the respiratory tract and digestive tract, and may have limited our ability to detect intracellular pathogens. Though we did detect some intracellular pathogens (e.g., *Rickettsia*), we did not observe other intracellular bacterial pathogens, particularly *Salmonella spp.*, which are a common cause of febrile illness in hospitalized patients in Sub-Saharan Africa^{12,13,46}. Studies from elsewhere in Sub-Saharan Africa suggest a lower incidence of bacteremia in outpatients consistent with our

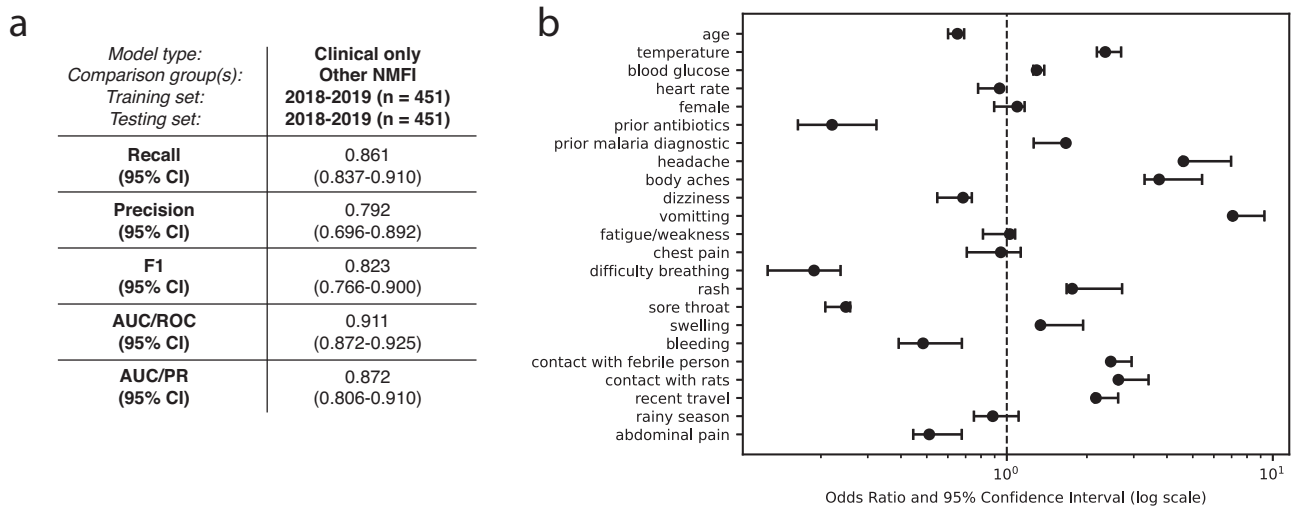


Fig. 4 | Results of best performing weighted logistic regression model to distinguish *Borrelia* infection from other NMFI. a Model performance metrics of best performing weighted logistic regression model to distinguish *Borrelia* infection from other NMFI using clinical data, including demographics, symptoms,

exposures, and vital signs (trained on 2018-2019 data, tested on the same data with bootstrapping and cross-validation, $n = 451$) and **b** mean odds ratios (center point) and 95% confidence interval (error bars) for all features retained in the final model after feature selection. Source data are provided as a Source Data file.

study^{12,47–49}. However, we cannot rule out that other intracellular bacterial infections were present at low titers in the blood and, therefore not detected by plasma 16S and RNA–mNGS. Additionally, fever can persist beyond the acute viremic or bacteremic phase, limiting our ability to detect the causative pathogen. Lastly, noncommunicable diseases, such as cancer or autoimmune disease, can also cause fever. In clinical samples, negative mNGS results are common, but it can be difficult to determine whether these represent false negatives due to limited sensitivity or true negatives⁷. As the burden of noncommunicable disease rises in West Africa⁵⁰, the approach to diagnosing febrile illness must consider non-infectious etiologies.

Our results also demonstrate the difficulty of using RNA–mNGS to detect pathogens with different nucleic acids and cellular compositions, limiting our ability to do truly comprehensive surveillance. Several of the bacterial infections in this study population, particularly those with low abundance in the blood, were detected by 16S but missed by RNA–mNGS. Similarly, we were not able to resolve putative fungal hits to a suitable level of taxonomic resolution; to capture fungal pathogens, extraction methods should be optimized for lysis of fungal cells, and more targeted methods such as ITS sequencing could be used. We observed wide variations in the number of *Plasmodium* reads among RDT-positive patients and a weak correlation with parasite density, suggesting that factors beyond parasite burden affect the number of reads recovered. There are also still many challenges in interpreting RNA–mNGS data. In particular, RNA–mNGS is highly susceptible to contamination, such as the novel *Reoviridae* RdRp sequences we detected across febrile cases, healthy controls, and non-template controls. This type of contamination, which likely differs between sample types and labs, can only be evaluated when negative controls are processed alongside samples⁵¹.

We find that combinations of clinical signs and symptoms can increase suspicion for *Borrelia* and support targeting diagnostic testing and clinical care. Due to the lack of available point-of-care tests, many clinics in LMICs rely on malaria RDTs as the primary diagnostic for febrile illness and may give blanket antibiotic treatment to RDT-negative patients⁵². Concerns have been raised about this practice³, including driving antibiotic resistance and exposing patients without evidence of bacterial infection to unnecessary risks or side effects. Even in patients with a bacterial infection, tetracycline antibiotic treatment—the recommended therapeutic for *Borrelia* and *Rickettsial* infections—poses the risk of Jarisch–Herxheimer reaction (JHR), a severe inflammatory response to spirochete lysis⁵³. Without accurate diagnosis of spirochetal infections, including *Borrelia*, patients and providers cannot effectively anticipate and monitor for this complication. We identify a number of clinical and immunological features that could bolster support for a *Borrelia* diagnosis. Our weighted logistic regression model indicates that when larger study populations and diverse data types are aggregated, *Borrelia* can be differentiated from other NMFI with reasonable accuracy. The addition of laboratory measures, such as CBC, did not substantially improve differential diagnosis. Further, data from a subset of our study population supports the hypothesis that a chemokine panel including TRAIL, IP-10, and CRP, previously employed for the detection of bacterial infections in hospitalized children⁵⁴, could apply more broadly and help distinguish bacterial infections such as *Borrelia* from other causes of febrile illness in the ambulatory setting. However, none of these approaches are currently actionable as point-of-care tests, and their lower precision would still necessitate confirmatory diagnostic testing.

The challenge of quickly and accurately differentiating *Borrelia* infections underscores the need for improved diagnostics. The diagnostic gold standard for *Borrelia* remains the microscopic examination of blood smears, which is highly dependent on a trained practitioner and has low sensitivity in our study and others³⁶. Further, smears cannot distinguish species. Louse-borne relapsing fever agent *B. recurrentis* has a higher fatality rate and risk of JHR than tick-borne

relapsing fever species, but with smear alone, clinicians cannot distinguish which disease they are treating unless the history of vector exposure is known. In Lyme Disease, molecular surveillance has shown that different species have different clinical presentations and severities⁵⁵. Species-specific surveillance at scale could reveal analogous differences between relapsing fever *Borrelia*. Our study showed that molecular methods, including 16S sequencing and mNGS, can detect and distinguish *Borrelia* infections but are not practical in the clinical setting. Increased availability of qPCR assays and improved point-of-care nucleic acid diagnostics could enable the detection and treatment of *Borrelia* across differently resourced clinical settings and improve our understanding of the geographic range and diversity of this pathogen.

Decreased malaria transmission coupled with changes in the range and incidence of other pathogens—influenced in part by global climatic changes—is shifting the burden of pathogens causing febrile disease⁵⁶. New technologies have greatly increased our ability to detect both established and emerging pathogens. Fully understanding this landscape will require surveillance systems utilizing untargeted approaches that have been validated for viral, bacterial, and eukaryotic pathogens and diverse sample types. In tandem, for *Borrelia* and other pathogens known to frequently cause NMFI, the development and availability of diagnostic tests that are cheap, rapid and sensitive will be key to enable appropriate clinical treatment at the point of care and support deeper investigations of the pathophysiology of disease.

Methods

Sample collection

We performed a cross-sectional study of febrile and healthy individuals. Febrile cases were selected from patients presenting to the SLAP outpatient clinic in Thiès, Senegal, during the collection period under local IRB (SEN15/46) and Harvard IRB (IRB19-0023). During times when the study personnel were onsite, all patients who met the following inclusion criteria and gave consent were enrolled: (1) Febrile symptoms within the 3 days up to and including the day of presentation, (2) Age 2–75 years, and (3) Ambulatory with no signs of severe malaria (glucose <2.2 mM, hemoglobin <5 gms/dL). Healthy controls were recruited via a call for participants. All volunteers who met the following inclusion criteria and consented were enrolled: (1) No febrile symptoms within the 3 days up to and including the day of presentation and (2) Age 2–75 years. In each season, enrollment continued until approximately the desired number of participants was reached. In order to detect pathogens present at 1% or greater prevalence, we aimed for 200 febrile cases and 200 healthy controls in each season in 2018. In 2019, to avoid the overrepresentation of healthy individuals, we aimed for 100 febrile cases and 50 healthy controls in each season.

Informed consent was obtained for all enrollees (febrile and healthy); for minors under 18 years of age or individuals unable to provide their own consent, the consent of a parent or legal guardian was obtained. The study team explained to potential participants that their participation was strictly voluntary and that they could withdraw from the study at any time without any penalties or consequences, and translations were made for potential participants (or their parent/legal guardian) who do not understand or cannot read the language in which the consent form was produced. A subset of DENV-positive samples from 2018 were previously sequenced and published¹⁴ and were excluded from this study other than being used as a viral outgroup in the chemokine/cytokine analyses. At the time of enrollment, a structured interview including personal information, demographics, and self-reported symptoms was completed, vital signs were measured, and blood was drawn by a trained practitioner. *P. falciparum* RDT (Bioline Malaria Ag P.f., Abbott), thick blood smear, and thin blood smear were performed, and blood glucose (HemoCue 201/301) and hemoglobin (HemoCue Glucose 201) were measured.

Sample processing

Blood samples were stored on ice upon collection (maximum 8 h). In the lab, blood samples were split into two aliquots. In total, 200 μ L of the first aliquot was used for an automated blood count with differential (Mindray BC 20 s), and 200 μ L was used for a blood spot (Whatman filter paper). The remaining aliquot was centrifuged at 800 \times g for 10 min to separate plasma, buffy coat, and red blood cells. 140 μ L of plasma was transferred to a tube containing 560 μ L of inactivation buffer (AVL, Qiagen). The remaining plasma was stored without inactivation. Red blood cells, buffy coat, plasma, and inactivated plasma were flash-frozen and stored at -80°C until further processing. Total nucleic acids were extracted with Qiagen QIAmp Viral Mini RNA according to the manufacturer's protocol. Total nucleic acids were split into two aliquots and treated with Lucigen RNase I or Ambion TURBO DNase to obtain purified DNA and RNA, respectively. DNA was extracted from dried blood spots with QIAmp DNA Blood Mini Kit (QIAGEN) according to the manufacturer's protocol.

Untargeted RNA sequencing

Library preparation and sequencing. Untargeted RNA sequencing libraries were prepared as described previously⁵⁷. Briefly, cDNA synthesis was performed from DNase-treated RNA using random primers (Invitrogen) and SSIV. Libraries were prepared from cDNA using Nextera XT (Illumina) and UD index primers (IDT/Illumina) with the following modifications to account for low cDNA input: the volume of ATM was decreased from 1 μ L to 0.5 μ L per 10 μ L reaction and PCR cycles were increased to 17 cycles. Libraries were purified using AMPure XP (Agilent), quantified using KAPA Biosystems Universal Library Quantification, and pooled equally for 75 bp paired-end sequencing on NovaSeq SP to obtain at least 2 million reads per sample.

Taxonomic classification of RNA-mNGS reads. Untargeted RNA sequencing reads were processed using viral-ngs v2.1.33.16 (<https://github.com/broadinstitute/viral-ngs>)⁵⁸. Briefly, reads were demultiplexed, adapter sequences were trimmed, sliding-window quality filtering was performed, and human reads were filtered out using the viral-ngs demux_only [https://dockstore.org/workflows/github.com/broadinstitute/viral-pipelines/demux_only:master] and classify_single [https://dockstore.org/workflows/github.com/broadinstitute/viral-pipelines/classify_single:master?tab=info] workflows. Cleaned reads were uploaded to NCBI SRA under BioProject PRJNA662334 (Accession numbers: SRR24622550-SRR24622641, SRR24995052-SRR24995258). Samples with insufficient reads (<2 million) were removed from further analysis ($n = 1$, SHC1064, Accession: SRR24995248).

Taxonomic classification of human-depleted reads was performed using Kraken2⁵⁹ with the PlusPF database (<https://benlangmead.github.io/aws-indexes/k2>, downloaded 12-13-2022). Kraken2 results were thresholded to consider only viral genera with at least 5 reads/million raw reads. Results were filtered to remove viral taxa detected in non-template controls (at least one non-template control) and healthy patients (at least 2 healthy patients). Further, viral genera with non-vertebrate hosts were filtered out.

To verify Kraken2 classifications, we performed a protein-sequence similarity search using DIAMOND v2.0.15⁶⁰. Cleaned, deduplicated reads were used for de novo contig assembly with SPAdes and DIAMOND-blastx was run on all de novo contigs >100 bp long with the complete nr database (downloaded December 2022). The least common ancestor of the top e-value hits was identified using a custom script (lakras/bio-helper-scripts/blast/retrieve_top_blast_hits_LCA_for_each_sequence.pl). A classification was considered verified by DIAMOND if the cumulative length of contigs with a high identity match (mean identity of top e-value hits $\geq 95\%$) to the expected viral genus was at least 1 kb. After DIAMOND-blastx verification, raw reads were aligned to the NCBI Virus RefSeq and assessed for evenness and depth of coverage.

To assess for the presence of divergent viral taxa that are not well represented in the database and may be missed by Kraken2, DIAMOND-blastx results were filtered for viral hits with low identity (20–80% amino acid identity). Results were filtered for viral families with a per-sample cumulative de novo contig length ≥ 500 bp. *Narnaviridae* and *Reoviridae* sequences were aligned with all available RdRp sequences for each family in nr using MAFFT v1.5.0 and trimmed with trimAl⁶¹. Maximum likelihood phylogenetic trees were generated in IQ-TREE v1.6.12⁶² with bootstrapping ($n = 1000$) and visualized in FigTree v1.4.4⁶³.

Viral genome assembly and phylogenetic analysis. For virus-positive samples, reference-guided de novo assembly was performed using viral-ngs v2.1.33.16⁵⁸ assemble de novo workflow [https://dockstore.org/workflows/github.com/broadinstitute/viral-pipelines/assemble_denovo:master?tab=info]. The NCBI Virus RefSeq, as well as any >80% complete genomes available from Senegal, were provided as references. For samples with near-complete or complete (>80%) genomes, genotype was determined with GenomeDetective (<https://www.genomedetective.com/>), DENV, HBV) or by the phylogenetic tree (Parvovirus B-19, HPgV-1). Phylogenetic trees were generated with all >80% complete references from Africa downloaded from NCBI viruses (DENV1, DENV3, HBV), or all global >80% complete genomes available in NCBI virus (Parvovirus B-19, HPgV-1). Sample sequences and reference sequences were aligned using MAFFT v1.5.0 v7, and alignments were trimmed with trimAl v1.4.rev15 to remove any bases with coverage in <80% of sequences. Maximum likelihood phylogenetic trees were generated in IQ-TREE v1.6.12 with a GTR-gamma substitution model and bootstrapping ($n = 1000$). Visualizations were generated with FigTree v1.4.4⁶³.

Plasmodium RNA-mNGS analysis. In order to quantify *Plasmodium* rpm, the number of reads classified as *Plasmodium* at the genus level by Kraken2 was divided by the raw read count in millions. The cut-off for considering a sample positive for *Plasmodium* by mNGS was determined by calculating the 99th percentile for healthy control samples, 550 *Plasmodium* rpm. Cleaned deduplicated reads were aligned to the *P. falciparum* *pfhrp2* [https://plasmodb.org/plasmo/app/record/gene/PF3D7_0831800] with the viral-ngs align_and_count workflow [https://dockstore.org/workflows/github.com/broadinstitute/viral-pipelines/align_and_count_multiple_report:master].

Fungal RNA-mNGS analysis. In order to detect possible fungal infections, we calculated the rpm classified as Fungi by Kraken2 at the kingdom level and determined the 99th percentile for healthy controls, 3399 rpm. We identified 5 samples (4 febrile, 1 control) with high fungal reads. We attempted taxonomic classification of these reads with Kraken2 at the genus level and saw hits across multiple genera for many samples (Supplementary Fig. 2b). To try to improve fungal classification, we performed a nucleic acid search of all de novo contigs from these 5 samples with megablast against nt (Supplementary Fig. 2b) and a translated nucleic acid search with DIAMOND-blastx (as described above). For both searches, we found the least common ancestor of the top hits (as described above) and filtered hits in the kingdom Fungi (taxid: 4751) with >90% identity and >30% coverage.

PCR and qPCR assays

All qPCRs were performed on the QuantStudio 6 Flex ReadTime PCR (Applied Biosystems). Primer sequences and sources are described in Supplementary Table 4. All qPCRs were performed with a standard curve using a synthetic DNA standard (see Supplementary Table 5). Samples were tested in triplicate and considered positive if all three wells had a cycle threshold of less than 40.

Total bacterial load was quantified from extracted plasma DNA with primers targeting the VI-2 region with Power SYBR Green PCR

MasterMix (Thermo Fisher) under the following conditions: 300 nM each V1–2-F/R primer, 95 °C hold for 10 min, 40 cycles of 95 °C for 15 s, 50 °C for 1 min, 75 °C for 30 s. Pan-*Borrelia* qPCR was performed on DNA extracted from plasma with Power SYBR Green PCR MasterMix (Thermo Fisher) under the following conditions: 400 nM each *Borrelia*-F/R primer, 95 °C hold for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 1 min. Pan-*Plasmodium* qPCR was performed on DNA extracted from dried blood spots (DBS) with Power SYBR Green PCR MasterMix (ThermoFisher) under the following conditions: 600 nM each Spp-F/R primer, 95 °C hold for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 1 min. *Plasmodium falciparum* specific qPCR was performed on DNA extracted from DBS with TaqMan Fast Advanced Master Mix (Applied Biosystems) under the following conditions: 450 nM each Fal-F/R primer and 125 nM Fal probe, 95 °C hold for 20 s, 40 cycles of 95 °C for 1 s, 60 °C for 30 s on the fast setting.

Amplification of pflrp2 exon 2 was performed on DNA extracted from DBS with Q5 HighFidelity 2× Master Mix (NEB) under the following conditions: 500 nM each Pf3D7_0831800-F/R primer, 98 °C hold for 30 sec, 40 cycles of 98 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s, followed by 72 °C hold for 2 min. Amplification of the flanking gene was performed with Q5 HighFidelity 2× Master Mix (NEB) under the following conditions: 500 nM each Pf3D7_0831700-F/R, 98 °C hold for 30 s, 40 cycles of 98 °C for 10 sec, 64 °C for 30 s, 72 °C for 30 s, followed by 72 °C hold for 2 min. PCR products were visualized on E-Gel EX 2% Gel (Thermo Fisher) with the E-Gel 1 kb plus DNA ladder (Thermo Fisher), and images were captured with FluorChem E imager (ProteinSimple).

16S sequencing

Library preparation and sequencing. Samples with high bacterial load (V1–2 qPCR CT < 31.5) were selected for 16S sequencing (2019: febrile $n=129$, healthy $n=35$, 2018: febrile $n=84$). Libraries were generated by amplification with tailed universal primers Tail-V1–2F and Tail-V1–2R (Supplementary Table 4) targeting variable regions 1–2. Amplification was performed on 5 µl of RNase-treated DNA using Q5 high fidelity polymerase (NEB) with forward and reverse primers at 100 nM each with the following cycling conditions: 98 °C for 30 s; 35 cycles of 95 °C for 15 s and 63 °C for 2 min, 4 °C hold. PCR products were cleaned using Ampure XP (Agilent) with a 0.7× ratio to remove primer dimers. Adapters and barcodes (BroadDuplex Seq) were added with a second PCR reaction using Q5 polymerase and the following cycling conditions: 98 °C for 30 s; 18 cycles of 95 °C for 15 s, 60 °C for 15 s, 72 °C for 30 s, followed by 72 °C for 5 min, 4 °C hold. A subset of libraries was visualized with hsDNA BioAnalyzer (Agilent), and all libraries were quantified with KAPA Biosystems Universal Library Quantification Kit before pooling equally across samples and 250 bp paired-end sequencing on Illumina MiSeq v2 with 40% PhiX, given the low diversity of the single-amplicon library.

Analysis. Sequencing reads were demultiplexed using viral-ngs v2.1.33.16⁵⁸, and demultiplexed FASTQ files were imported into qiime2 v2022.2.0⁶⁴ for further analysis. Demultiplexed reads were uploaded to NCBI SRA under BioProject PRJNA662334. Briefly, after sliding window quality filtering and adapter trimming with cut-adapt (minimum-length 20), paired reads were joined with vsearch (tuncqual 15, minlen 35, minovlen 10, maxdiffs 3), and ASVs were generated with Deblur (trim-length 280). ASVs were taxonomically classified by blastn search of ASV sequences against the NCBI 16S rRNA db (downloaded June 2023). The least common ancestor of the top e-value hits was determined using a custom script (lakras/bio-helper-scripts/blast/retrieve_top_blast_hits_LCA_for_each_sequence.pl). Results were filtered to include only ASVs that had identity $\geq 94.5\%$ to the closest hit, as this has been proposed as a rational cutoff for 16S rRNA diversity within genera⁶⁵. Abundance was calculated by dividing the number of ASVs in each genus by the total number of ASVs in each sample.

Abundance data was filtered to show only taxa accounting for greater than 5% of ASVs. Taxa detected in non-template controls ($n \geq 1$) and healthy controls ($n \geq 2$) were filtered out.

For *Borrelia* and *Rickettsia* positive samples, sample ASVs were aligned with MAFFT v1.5.0⁶⁶ (Geneious v2022.0.2) with curated contextual sequences for relapsing fever *Borrelia* spp. and *Rickettsia* spp. known to infect humans. When available for a given species, the high-quality 16S sequencing from the NCBI 16S rRNA BioProject [<https://www.ncbi.nlm.nih.gov/bioproject/33175>] was used. For species not included in the NCBI 16S rRNA BioProject, all available sequences from the SilvaNRRef small subunit database were downloaded, sequenced with complete coverage of the v1–2 region were selected, and a representative sequence was selected randomly for each species. Alignments were trimmed manually in Geneious v2022.0.2 to include only the v1–2 region, and a maximum-likelihood phylogenetic tree was generated using IQ-TREE v1.6.12 with bootstrapping ($n=1000$) and annotated with FigTree v1.4.4⁶³.

IGS sequencing

The intergenic spacer (IGS) region was amplified using a nested-PCR. First, the region was amplified from 5 µl of extracted DNA with primers targeting the IGS region (IGS-outer F/R, see Supplementary Table 4) using Q5 polymerase and the following cycling conditions: 98 °C for 30 s; 35 cycles of 94 °C for 30 s, 66 °C for 30 s, 74 °C for 60 s, followed by 74 °C for 2 min, 10 °C hold. After a 0.75× AMPure XP cleanup to remove excess primer, a nested PCR was performed (IGS-inner F/R, see Supplementary Table 4) using Q5 polymerase and the following cycling conditions: 98 °C for 30 s; 30 cycles of 94 °C for 30 s, 67 °C for 30 s, 74 °C for 60 s, followed by 74 °C for 2 min, 10 °C hold. Amplified product was cleaned with AMPure XP (0.75×), visualized with BioAnalyzer TapeStation and bi-directional Sanger sequencing was performed (Azenta). Paired Sanger sequencing traces were analyzed in Geneious v2022.0.2 to generate a consensus sequenced and aligned with all available reference sequences for *B. crocidurae*, *B. duttonii*, and *B. recurrentis* using MAFFT v1.5.0. A maximum-likelihood phylogenetic tree was generated using IQ-TREE v1.6.12 and annotated with FigTree v1.4.4⁶³.

Blood smear examination

Thick and thin blood smears were fixed with methanol and stained with 3 mL 10% Giemsa for 10 min. Stained smears were examined under a brightfield microscope for evidence of *Borrelia*. Slides were first scanned at low (40×) magnification and then at high (100×) magnification with oil immersion to quantify organisms per field (representative image, Supplementary Fig. 7b).

Serology

We assessed levels of common cytokines and chemokines using the Luminex platform on plasma from a subset of patients and controls. Specifically, we selected 10 healthy controls and 30 cases representing 10 each with a confirmed diagnosis of; malaria (based on malaria RDT administered at enrollment), *Borrelia* (based on pan-*Borrelia* qPCR³⁸), or viral infection. Samples with a viral infection were all confirmed to have Dengue virus and are a subset of those cases previously described elsewhere³⁴. Cases were selected randomly within those meeting each criteria. All samples were collected in 2018. We used the Inflammation 20-Plex Human ProcartaPlex™ Panel (Invitrogen) according to the manufacturer's instructions. Two of the assays (IFN-gamma and GM-CSF) did not generate results due to errors during data collection. We additionally performed a 2-plex assay for Human CRP and TRAIL (using CRP and TRAIL ProcartaPlex™ simplex kits from Invitrogen); two markers previously found to differentiate bacterial infections in hospitalized children⁵⁴, as these were not included on the predefined panel. All samples were run in duplicate on a Luminex MAGPIX® instrument in Senegal. We confirmed that standards showed the

expected values and calculated the average across duplicates, which we reported in the text.

Multivariate model

We developed a weighted logistic regression model that considered 46 candidate predictors based on clinical symptoms, vital signs, and key demographic information to predict *Borrelia* infection. Before modeling, the dataset underwent the following preprocessing steps: imputation of missing values with the Multivariate Imputation by Chained Equations (MICE) algorithm from the fancyimpute⁶⁷ Python library, one-hot encoding of categorical features, and standardization of continuous features using the StandardScaler from the scikit-learn library. Given that normal heart rate varies significantly with age, heart rates were binned and treated as a categorical variable. Heart rate categories were high, low, or normal based on the Pediatric Advanced Life Support guidelines as follows: for ages 3–5, normal range of 80–120; ages 6–11, normal range of 75–118; for age 12+, normal range 60–100⁶⁸.

The sample size was relatively small, with 526 patients and only 38 q-PCR-confirmed cases of *Borrelia* infection. Given this inherent class imbalance in the dataset, a combination of techniques was employed to address this challenge. The Synthetic Minority Oversampling Technique (SMOTE) was applied to generate synthetic samples for the minority class (q-PCR confirmed *Borrelia* case)⁶⁹. This technique mitigates class imbalance by interpolating between existing samples and creating synthetic instances of the minority class. During model training, class weights were also applied to adjust the influence of each class. The class weights were inversely proportional to the class frequencies, with higher weights assigned to the minority class.

A logistic regression model was chosen for its interpretability and ability to estimate the probabilities of class membership. After a univariate screening for statistically significant predictors using a chi-squared test (p -value < 0.05), several Feature Selection techniques were applied to construct the final set of predictors in the model. Before training the model, feature importance was assessed using a combination of three different feature selection methods; mutual information classification⁷⁰, Recursive Feature Elimination⁷¹, and Lasso L1 regularization^{72,73} to ensure consistent elimination of redundant variables. The selected features were used in the final weighted logistic regression model to analyze the performance of *Borrelia* prediction. The regularization parameter from Lasso was tuned using cross-validation of different alpha values, and the best alpha value was used to train the final model and evaluate its performance on the test set.

Model performance was evaluated with bootstrapping and five-fold cross-validation to quantify optimism and generalizability. The optimism-corrected metrics include F1-score, precision, recall, area under the ROC curve (AUC-ROC), and area under the precision-recall curve (AUC-PR). Odds ratios and their corresponding confidence intervals for all the model coefficients were calculated to assess the impact of predictor variables on the odds of the outcome. All predictive analyses were conducted using the Python programming language. The scikit-learn library was utilized for preprocessing, model training, and evaluation. Statistical analyses and visualization of odds ratios and confidence intervals were done using the R statistical package.

Inclusion and ethics statement

This study was designed, implemented, and analyzed in close collaboration between US-based researchers and Senegalese researchers at Cheikh Anta Diop University/CIGASS, the SLAP outpatient clinic, and the National Malaria Control Program. All study team members are included as authors, and their contributions are detailed in the CRediT statement. The protocol was approved by the Harvard IRB as well as the IRB of the Ministry of Health in Senegal, and all co-authors who worked directly with samples in Senegal or the U.S. received

appropriate biosafety training. Sequencing and analysis were conducted at the Broad Institute due to the higher capacity for high-depth sequencing. Sample aliquots are stored at both CIGASS and the Broad Institute, and all sequencing data and sample-associated data are accessible to study teams at both locations. Relevant research from other local teams is cited as appropriate.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

RNA-mNGS short read sequence data and 16S v1–2 amplicon sequence data have been deposited in NCBI's SRA databases under BioProject PRJNA662334. Viral genomes have been deposited in NCBI's Genbank (see Supplementary Table 2 for BioSample identifiers). Source data are provided in this paper.

Code availability

The Python source code for the multivariate *Borrelia* prediction model is available at github.com/colobio/borrelia-diagnosis-prediction-models. The custom scripts used to retrieve the least common ancestor of top blast/DIAMOND hits are available at [lakras/bio-helper-415-scripts/blast/retrieve_top_blast_hits_LCA_for_each_sequence.pl](https://github.com/lakras/bio-helper-415-scripts/blob/main/blast/retrieve_top_blast_hits_LCA_for_each_sequence.pl).

References

1. Thwing, J. et al. Declines in malaria burden and all-cause child mortality following increases in control interventions in Senegal, 2005–2010. *Am. J. Trop. Med. Hyg.* **97**, 89–98 (2017).
2. World Bank. Incidence of malaria (per 1,000 population at risk). (2021).
3. WHO informal consultation on fever management in peripheral health care settings: a global review of evidence and practice. <https://www.who.int/publications-detail-redirect/9789241506489>.
4. Houldcroft, C. J., Beale, M. A. & Breuer, J. Clinical and biological insights from viral genome sequencing. *Nat. Rev. Microbiol.* **15**, 183–192 (2017).
5. Gu, W., Miller, S. & Chiu, C. Y. Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* **14**, 319–338 (2019).
6. Mwakibete, L. et al. Metagenomic next-generation sequencing to characterize potential etiologies of non-malarial fever in a cohort living in a high malaria burden area of Uganda. *PLoS Glob. Public Health* **3**, e0001675 (2023).
7. Oguzie, J. U. et al. Metagenomic surveillance uncovers diverse and novel viral taxa in febrile patients from Nigeria. *Nat. Commun.* **14**, 4693 (2023).
8. Ramesh, A. et al. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS ONE* **14**, e0218318 (2019).
9. Decuyper, S. et al. Diagnosis of bacterial bloodstream infections: a 16S metagenomics approach. *PLoS Negl. Trop. Dis.* **10**, e0004470 (2016).
10. Watanabe, N. et al. Detection of pathogenic bacteria in the blood from sepsis patients using 16S rRNA gene amplicon sequencing analysis. *PLoS ONE* **13**, e0202049 (2018).
11. Kingry, L. et al. Targeted metagenomics for clinical detection and discovery of bacterial tick-borne pathogens. *J. Clin. Microbiol.* **58** (2020).
12. Maze, M. J. et al. The epidemiology of febrile illness in sub-Saharan Africa: implications for diagnosis and management. *Clin. Microbiol. Infect.* **24**, 808–814 (2018).
13. Elven, J. et al. Non-malarial febrile illness: a systematic review of published aetiological studies and case reports from Africa, 1980–2015. *BMC Med.* **18**, 279 (2020).

14. Gaye, A. et al. Genomic investigation of a dengue virus outbreak in Thiès, Senegal, in 2018. *Sci. Rep.* **11**, 10321 (2021).
15. Sow, A. et al. Chikungunya outbreak in Kedougou, Southeastern Senegal in 2009–2010. *Open Forum Infect. Dis.* **5**, ofx259 (2018).
16. Vial, L. et al. Incidence of tick-borne relapsing fever in west Africa: longitudinal study. *Lancet* **368**, 37–43 (2006).
17. Sokhna, C. et al. Point-of-care laboratory of pathogen diagnosis in rural Senegal. *PLoS Negl. Trop. Dis.* **7**, e1999 (2013).
18. Mediannikov, O. et al. *Borrelia crocidurae* infection in acutely febrile patients, Senegal. *Emerg. Infect. Dis.* **20**, 1335–1338 (2014).
19. Abat, C. et al. Implementation of syndromic surveillance systems in two rural villages in Senegal. *PLoS Negl. Trop. Dis.* **10**, e0005212 (2016).
20. Ndiaye, E. H. I. et al. Tick-borne relapsing fever Borreliosis, a major public health problem overlooked in Senegal. *PLoS Negl. Trop. Dis.* **15**, e0009184 (2021).
21. Socolovschi, C. et al. *Rickettsia felis*-associated unruptive fever, Senegal. *Emerg. Infect. Dis.* **16**, 1140–1142 (2010).
22. Sy, M. et al. *Plasmodium falciparum* genomic surveillance reveals spatial and temporal trends, association of genetic and physical distance, and household clustering. *Sci. Rep.* **12**, 938 (2022).
23. Casteling, A. et al. GB virus C prevalence in blood donors and high risk groups for parenterally transmitted agents from Gauteng, South. *Afr. J. Med. Virol.* **55**, 103–108 (1998).
24. Compston, L. I. et al. Prevalence of persistent and latent viruses in untreated patients infected with HIV-1 from Ghana, West. *Afr. J. Med. Virol.* **81**, 1860–1868 (2009).
25. Lauck, M. et al. GB virus C coinfections in west African Ebola patients. *J. Virol.* **89**, 2425–2429 (2015).
26. Ouattara, L. A. et al. Novel human reovirus isolated from children with acute necrotizing encephalopathy. *Emerg. Infect. Dis.* **17**, 1436–1444 (2011).
27. Frickmann, H. et al. Metagenomic sequencing for the diagnosis of *Plasmodium* spp. with different levels of parasitemia in EDTA blood of malaria patients—a proof-of-principle assessment. *Int. J. Mol. Sci.* **23** (2022).
28. Deme, A. B. et al. Analysis of *pfhrp2* genetic diversity in Senegal and implications for use of rapid diagnostic tests. *Malar. J.* **13**, 34 (2014).
29. Agaba, B. B. et al. Systematic review of the status of *pfhrp2* and *pfhrp3* gene deletion, approaches and methods used for its estimation and reporting in *Plasmodium falciparum* populations in Africa: review of published studies 2010–2019. *Malar. J.* **18**, 355 (2019).
30. Cheng, Q. et al. *Plasmodium falciparum* parasites lacking histidine-rich protein 2 and 3: a review and recommendations for accurate reporting. *Malar. J.* **13**, 283 (2014).
31. Bunikis, J. et al. Typing of *Borrelia relapsing fever* group strains. *Emerg. Infect. Dis.* **10**, 1661–1664 (2004).
32. Schwan, T. G. et al. Endemic foci of the tick-borne relapsing fever spirochete *Borrelia crocidurae* in Mali, West Africa, and the potential for human infection. *PLoS Negl. Trop. Dis.* **6**, e1924 (2012).
33. Bouattour, A. et al. *Borrelia crocidurae* infection of *Ornithodoros erraticus* (Lucas, 1849) ticks in Tunisia. *Vector Borne Zoonotic Dis.* **10**, 825–830 (2010).
34. Diatta, G. et al. Epidemiology of tick-borne borreliosis in Morocco. *PLoS Negl. Trop. Dis.* **6**, e1810 (2012).
35. Labruna, M. B. & Walker, D. H. *Rickettsia felis* and changing paradigms about pathogenic rickettsiae. *Emerg. Infect. Dis.* **20**, 1768–1769 (2014).
36. Jakab, Á., Kahlig, P., Kuenzli, E. & Neumayr, A. Tick borne relapsing fever—a systematic review and analysis of the literature. *PLoS Negl. Trop. Dis.* **16**, e0010212 (2022).
37. Fotso Fotso, A. & Drancourt, M. Laboratory diagnosis of tick-borne african relapsing fevers: latest developments. *Front. Public Health* **3**, 254 (2015).
38. Kingry, L. C. et al. Surveillance for and discovery of borrelia species in US patients suspected of tickborne illness. *Clin. Infect. Dis.* **66**, 1864–1871 (2018).
39. Lim, L. L. & Rosenbaum, J. T. *Borrelia hermsii* causing relapsing fever and uveitis. *Am. J. Ophthalmol.* **142**, 348–349 (2006).
40. Trape, J.-F. et al. The epidemiology and geographic distribution of relapsing fever borreliosis in West and North Africa, with a review of the *Ornithodoros erraticus* complex (Acari: Ixodida). *PLoS ONE* **8**, e78473 (2013).
41. Cutler Sally J., Idris Jibrin M., Ahmed Akeem O. & Elelu N. *Ornithodoros savignyi*, the Tick Vector of ‘*Candidatus Borrelia kalaharica*’ in Nigeria. *J. Clin. Microbiol.* **56** <https://doi.org/10.1128/jcm.00532-18> (2018).
42. Diallo, M. A. et al. *Plasmodium falciparum* malaria co-infection with tick-borne relapsing fever in Dakar. *Malar. J.* **16**, 24 (2017).
43. Normark, J. et al. Maladjusted host immune responses induce experimental cerebral malaria-like pathology in a murine *Borrelia* and *Plasmodium* co-infection model. *PLoS ONE* **9**, e103295 (2014).
44. Mediannikov, O. et al. Common epidemiology of *Rickettsia felis* infection and malaria, Africa. *Emerg. Infect. Dis.* **19**, 1775–1783 (2013).
45. Dieme, C. et al. Transmission potential of *Rickettsia felis* infection by *Anopheles gambiae* mosquitoes. *Proc. Natl Acad. Sci. USA* **112**, 8088–8093 (2015).
46. Reddy, E. A., Shaw, A. V. & Crump, J. A. Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis. *Lancet Infect. Dis.* **10**, 417–432 (2010).
47. D’Acremont, V. et al. Beyond malaria-causes of fever in outpatient Tanzanian children. *N. Engl. J. Med.* **370**, 809–817 (2014).
48. Hildenwall, H. et al. Causes of non-malarial febrile illness in outpatients in Tanzania. *Trop. Med. Int. Health* **21**, 149–156 (2016).
49. Mahende, C. et al. Bloodstream bacterial infection among outpatient children with acute febrile illness in north-eastern Tanzania. *BMC Res. Notes* **8**, 289 (2015).
50. El Bcheraoui, C. et al. Burden of disease in francophone Africa, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Glob. Health* **8**, e341–e351 (2020).
51. Jurasz, H., Pawłowski, T. & Perlejewski, K. Contamination issue in viral metagenomics: problems, solutions, and clinical perspectives. *Front. Microbiol.* **12**, 745076 (2021).
52. Hopkins, H. et al. Impact of introduction of rapid diagnostic tests for malaria on antibiotic prescribing: analysis of observational and randomised studies in public and private healthcare settings. *Br. Med. J.* **356**, j1054 (2017).
53. Butler, T. The Jarisch-Herxheimer reaction after antibiotic treatment of spirochetal infections: a review of recent cases and our understanding of pathogenesis. *Am. J. Trop. Med. Hyg.* **96**, 46–52 (2017).
54. Papan, C. et al. A host signature based on TRAIL, IP-10, and CRP for reducing antibiotic overuse in children by differentiating bacterial from viral infections: a prospective, multicentre cohort study. *Clin. Microbiol. Infect.* **28**, 723–730 (2022).
55. Radolf, J. D., Strle, K., Lemieux, J. E. & Strle, F. Lyme disease in humans. *Curr. Issues Mol. Biol.* **42**, 333–384 (2021).
56. Rocklöv, J. & Dubrow, R. Climate change: an enduring challenge for vector-borne disease prevention and control. *Nat. Immunol.* **21**, 479–483 (2020).
57. Matranga, C. B. et al. Unbiased deep sequencing of RNA viruses from clinical samples. *J. Vis. Exp.* (2016) 10.3791/54117.
58. viral-ngs. viral-ngs v2.1.33.16. <https://viral-pipelines.readthedocs.io/en/v2.1.33.16/description.html> (2023).
59. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
60. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).

61. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
62. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44**, W232–W235 (2016).
63. Rambaut. *FigTree*. (2018).
64. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
65. Yarza, P. et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645 (2014).
66. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
67. Rubinsteyn, A. & Feldman, S. fancyimpute: Version 0.0. 16. *Version v0.0.16*. May (2016).
68. American Heart Association. *Pediatric Advanced Life Support Provider Manual*. (2016).
69. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
70. Hancer, E., Xue, B. & Zhang, M. Differential evolution for filter feature selection based on information theory and feature ranking. *Knowl. Based Syst.* **140**, 103–119 (2018).
71. Guyon, I. & Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3**, 1157–1182 (2003).
72. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **58**, 267–288 (1996).
73. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

Acknowledgements

We thank the Program National de Lutte contre le Paludisme (National Malaria Control Program) for providing malaria surveillance data and Fama and Astou for support with data entry and organization. We thank Lydia Krasilnikova for advice and custom scripts for BLAST analysis and Flavia Negrete for support with data uploads and sharing. We thank Jacob Lemieux and Gordon Adams for providing valuable guidance on *Borrelia* sequencing. Support for this project was provided in part by the National Institutes of Allergy and Infectious Diseases through the H3Africa and Center for Research in Emerging Infectious Disease programs (U54HG007480, U01HG007480, and U01AI151812 to C.T.H. and P.C.S.), and the Genomic Center for Infectious Disease (U19AI110818 to P.C.S.). This work is made possible by support from Flu Lab and a cohort of generous donors through TED's Audacious Project, including the ELMA Foundation, MacKenzie Scott, the Skoll Foundation, and Open Philanthropy. Z.C.L. was supported by the National Institute of General Medical Sciences (T32GM007753 and T32GM144273). K.J.S. received support from the American Society of Tropical Medicine and Hygiene Centennial Award. P.C.S. is an investigator supported by the Howard Hughes Medical Institute (HHMI). The content is solely the responsibility of the authors and does not necessarily represent the official views of the

National Institute of General Medical Sciences, National Institutes of Allergy and Infectious Diseases, or the National Institutes of Health.

Author contributions

According to the CRediT format, author contributions are as follows: Conceptualization: A.B.D., A.S.B., Z.C.L., N.S., P.C.S., K.J.S., and D.N.; Methodology: A.B.D., A.S.B., Z.C.L., N.S., P.C.S., K.J.S., and D.N.; Software: D.J.P.; Validation: P.C.S., K.J.S., and D.N.; Formal analysis: Z.C.L., K.J.S., W.M., and A.C.; Investigation: Z.C.L., A.S., A.G., T.N., M.S., Y.D., A.M., and I.M.N.; Data Curation: J.G., M.Ndiop, D.S.; Writing—original draft: Z.C.L. Writing—review & editing: K.J.S., S.F.S., B.L.M, and P.C.S. with input from all authors; Visualization: Z.C.L., W.M.; Supervision: P.C.S., K.J.S., and D.N.; Project administration: M.F.P., M.Ndiaye; Funding acquisition: P.C.S., D.N., K.J.S., and D.J.P.

Competing interests

P.C.S. is a co-founder of, shareholder in, and consultant to Sherlock Biosciences, Inc. and Delve Bio, as well as a Board member of and shareholder in Danaher Corporation. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-44800-7>.

Correspondence and requests for materials should be addressed to Pardis C. Sabeti, Daouda Ndiaye or Katherine J. Siddle.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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 Author(s) 2024