

Investigating drug resistance markers,
evolution and co-infections of pathogens
of public health concern using portable
molecular detection and sequencing tools

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Salome Hosch

Basel, 2023

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Erstbetreuerin: Prof. Dr. Claudia Daubenberger

Zweitbetreuerin: Prof. Dr. Jennifer Keiser

Externer Experte: Prof. Dr. Cristian Koepfli

Basel, den 20.06.2023

Prof. Dr. Marcel Mayor

Dekan

Table of contents

Table of contents	i
Acknowledgements.....	iv
Summary	v
List of abbreviations.....	xi
1. Introduction	1
1.1. Causes of acute febrile illnesses in sub-Saharan Africa	1
1.2. Overview of SARS-CoV-2.....	1
1.2.1. Epidemiology and biology of SARS-CoV-2.....	1
1.2.2. Origin of SARS-Cov-2	3
1.2.3. Evolution of SARS-CoV-2	3
1.2.4. Public Health interventions.....	4
1.2.5. Molecular surveillance of SARS-CoV-2.....	4
1.3. Overview of <i>Plasmodium</i> spp.	5
1.3.1. Parasite life cycle.....	5
1.3.2. Epidemiology.....	7
1.3.3. Malaria prevention and treatment	8
1.3.4. Malaria diagnosis	9
1.4. Overview of filarial nematodes.....	10
1.4.1. Parasite characteristics and life cycle	10
1.4.2. Epidemiology.....	11
1.4.3. Diagnosis	13
1.5. Next-generation sequencing technologies	13
1.5.1. Whole genome sequencing in microbiology.....	15
1.5.2. Targeted sequencing of 16S ribosomal RNA gene.....	16
1.5.3. Metagenomics sequencing	16
1.6. One Health surveillance of bacterial pathogens and antimicrobial resistance	17

1.6.1.	Overview of <i>Campylobacter jejuni</i>	17
1.6.2.	Overview of <i>Salmonella enterica</i>	18
1.7.	Aims of this thesis	20
2.	Genomic surveillance of SARS-CoV-2 in Equatorial Guinea.....	21
2.1.	Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable peakPCR Platform	22
2.2.	Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions	33
2.3.	Genomic Surveillance Enables the Identification of Co-infections with Multiple SARS-CoV-2 Lineages in Equatorial Guinea.....	46
2.4.	The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance	54
3.	Improved monitoring of parasites that evade diagnosis, treatment or vaccination.....	77
3.1.	Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment.....	77
3.2.	Characterising co-infections with <i>Plasmodium</i> spp., <i>Mansonella perstans</i> or <i>Loa loa</i> in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests	91
3.3.	Development and evaluation of Plasmopod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance	110
3.4.	PHARE: A bioinformatics pipeline for compositional profiling of multiclonal <i>Plasmodium falciparum</i> infections from long-read Nanopore sequencing data	129
4.	One Health surveillance of foodborne bacteria.....	157
4.1.	Whole Genome Sequencing for One Health Surveillance of Antimicrobial Resistance in Conflict Zones: A Case Study of <i>Salmonella</i> spp. and <i>Campylobacter</i> spp. in the West Bank, Palestine.....	158
5.	Discussion.....	171
5.1.	ONT MinION is a state of the art next-generation sequencing platform	171
5.2.	Molecular surveillance of emerging viruses	172
5.3.	Molecular monitoring of diagnostic escape in multiclonal infections.....	174

5.4.	Molecular monitoring of <i>P. falciparum</i> malaria drug resistance markers.....	176
5.5.	One Health pathogen surveillance using whole genome sequencing.....	177
5.6.	Molecular diagnosis of fever causing viruses, bacteria and parasites.....	179
6.	Conclusion and future directions.....	180
7.	References	182
8.	Appendix	216

Acknowledgements

First and foremost I would like to thank my supervisor Prof Dr. Claudia Daubenberger for giving me the opportunity to pursue my PhD as part of the Clinical Immunology Unit at the Swiss Tropical and Public Health Institute. I am thankful for the support she continues to show me, for her helpful and constructive suggestions, for her trust in me and for the encouragement to pursue research.

I am obliged to Prof. Dr. Jennifer Keiser for supporting me as my second supervisor and the time she put into reviewing this thesis. I am grateful to Prof. Dr. Cristian Koepfli for being my external expert and reviewing my thesis. I would like to thank Prof. Dr. Günther Fink for chairing my PhD defense.

I would like to thank Dr. Tobias Schindler for all his support and mentorship. I am thankful for the current and former members of “Team Molecular” in the Clinical Immunology Unit, Philipp Wagner, Johanna Giger, Nele von Planta, Sarah Rubin, Dr. Etienne Guirou and Dr. Charlene Aya Yoboue for their great work in shared projects. A big thank you to all members of the Clinical Immunology Unit for their support and the interesting science-related and fun discussions.

I would like to express my deepest gratitude to all collaborators who made this PhD project a success. Thanks to Dr. Philippe Bechtold, Dr. Michele Gregorini, Denise Siegrist, Dr. Olivier Engler, Ulrich Vickos, Dr. Maximilian Mpina, Elizabeth Nyakurungu, Dr. Christian Nsanzabana, Dr. Said Abukhattab, Dr. Jan Hattendorf, Prof. Dr. Jakob Zinsstag and all the other people who have contributed to any of the projects.

Heartfelt thanks to my close friends and family for their support and encouragement throughout the PhD. You know who you are.

Summary

Genome sequencing has become an important tool for characterization and surveillance of infectious disease outbreaks. Sequences of novel pathogens can be compared to existing annotated sequences deposited in curated databases. Analysing the genetic diversity of pathogens in a population over time can be used to estimate how fast a disease spreads in combination with the speed and direction of pathogen evolution. If sequencing is done in real time, it can rapidly guide public health measures. In the example of SARS-CoV-2, the availability of whole genome sequences combined with epidemiological data can alert on the spread of novel virus variants that might escape either PCR based detection or vaccine induced immunity. Sequencing can furthermore be used to determine the drug resistance pattern of a pathogen if based on distinct genes or allelic variants of genes. Continuous monitoring of drug resistance markers in circulating *Plasmodium falciparum* strains for example is essential to guide changes of first-line drug treatment recommendations. Whole genome sequences of bacteria contain information to detect sources of contamination, trace transmission chains and detect molecular markers of antimicrobial resistance. To complement sequencing, samples can be screened for the presence of distinct pathogens, specific genes or single nucleotide polymorphisms (SNPs) by PCR-based technologies.

Portable qPCR platform such as the diaxxoPCR device have different advantages over standard qPCR platforms. Acquisition costs are significantly lower, reagents are independent of a cold chain, as they are preloaded and freeze-dried on cartridges and RT-qPCR results are available in 30 minutes. Approaches for simplified extraction from various sample types are being developed to improve widespread and decentralized access to molecular testing.

The MinION is a portable third-generation sequencing device developed by Oxford Nanopore Technologies (ONT) that can generate results with a turnaround time of few hours and with simple library preparation. It is a perfect tool for disease monitoring in resource limiting settings like sub-Saharan Africa, since it does not require capital investment or elaborate laboratory infrastructure. Moreover, the generated data can be analysed in real time, not requiring internet connectivity. ONT sequencing has been used to monitor Ebola and Zika virus outbreaks in real time in the field. ONT sequencing approaches have further been used for unbiased identification of pathogens in patient samples using metagenomics approaches. To decrease costs and increase throughput, targeted approaches have been developed, that amplify microbial DNA using universal primers that are then subjected to ONT sequencing.

The overall aim of this thesis was to develop and apply molecular detection and sequencing approaches as well as bioinformatics analysis tools for molecular surveillance of infectious diseases in Central Africa and the Middle East with high throughput and information depth.

In the first part of this thesis, we aimed to describe the evolution of SARS-CoV-2 in Equatorial Guinea using a combination of reverse transcription quantitative PCR (RT-qPCR) assays to detect mutations of interest and amplicon-based whole genome sequencing (WGS). Part one was structured around four manuscripts.

Manuscript 1: Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable *peak*PCR Platform.

Different SARS-CoV-2 variants of concern (VOC) have a distinct combination of mutations in the Spike gene. This manuscript describes the development of RT-qPCR assays using primers and probes with locked nucleic acids for increased specificity to detect two SNPs (E484K and N501Y) and one deletion (HV69/70 Δ) in the Spike gene. The assays can run on a standard qPCR platform and have been transferred to the portable *peak*PCR platform with preloaded freeze-dried cartridges. The assays were simple to use, rapid and low-cost. Cell culture-derived viral RNA of four SARS-CoV-2 lineages was used for evaluation of the RT-qPCR assays. The diagnostic performances of both platforms were comparable, with 100% analytical specificity and a limit of detection of 10 viral copies per microliter. Evaluation with 59 clinical samples resulted in perfect agreement between RT-qPCR and WGS derived mutation profile.

Manuscript 2: Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions.

This manuscript describes a SARS-CoV-2 lineage carrying many Spike gene SNPs (S477N, E484K and P681H) and deletions (HV69/70 Δ , Y144 Δ , LLA241/243 Δ) previously detected in VOC Alpha and Beta, but lacking the Spike N501Y SNP. The lineage was first described in an outbreak in Lithuania, a cluster of isolates with the Spike E484K SNP was subjected to sequencing and found to carry an unusual combination of SNPs and deletions. This SARS-CoV-2 variant was named B.1.620 and multiple genomes sequenced in Europe were derived from travellers returning from Cameroon. Phylogeographic analysis revealed a likely Central African origin and multiple introductions into other continents. Our SARS-CoV-2 spike gene mutation specific RT-qPCR assays, described in manuscript 1, were used for genomic surveillance in Equatorial Guinea and we found one isolate belonging to lineage B.1.620 in a person returning from Cameroon, representing the earliest collection date of a B.1.620

genome. The Spike gene SNPs found in lineage B.1.620 are likely to induce escape from antibody-mediated immunity. In Lithuania, B.1.620 was more commonly found in vaccine breakthrough infections compared to its population prevalence than other SARS-CoV-2 lineages.

Manuscript 3: Genomic Surveillance Enables the Identification of Co-infections With Multiple SARS-CoV-2 Lineages in Equatorial Guinea.

In this manuscript we describe the molecular epidemiological situation of SARS-CoV-2 in Equatorial Guinea between March 2020 and August 2021. Three waves of infections and related hospitalisations took place that were dominated by distinct SARS-CoV-2 lineages. The majority of sequences from the first wave in April to July 2020 belong to lineage B.1.192, a wild type-like SARS-CoV-2 lineage. The second wave in early 2021 was caused by the introduction of Beta VOC. In July 2021, Delta VOC was first detected in Equatorial Guinea and caused a massive third wave. In November 2021 we implemented mutation specific RT-qPCR assays (described in manuscript 1) to complement WGS and enhance genomic surveillance. We first detected a SARS-CoV-2 isolate with E484K and N501Y SNPs in isolates collected in November 2020, the combination that became dominant in January 2021. We implemented a RT-qPCR assay for the L452R SNP, a marker for Delta VOC in the local diagnostic laboratory in Equatorial Guinea and were able to inform the public health authorities about the probable introduction of Delta VOC within a week of sample collection. We detected qPCR signals indicative of co-infections of two distinct lineages in 2.1% of the samples analysed by mutation specific RT-qPCR assays. WGS of one co-infection case collected in August 2021 showed the presence of all lineage-defining mutations of Beta and Delta with mixed reads in an asymptomatic 59-year old Equatoguinean man who had been vaccinated twice with Sinopharm COVID-19 vaccine.

Manuscript 4: The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance.

This manuscript describes the milestone of reaching 100,000 SARS-CoV-2 genomes collected from Africa and the continents approach to tracking the pandemic. The number of African countries with local sequencing facilities increased massively in the first two years of the pandemic and regional sequencing networks were created. Almost all African countries had SARS-CoV-2 genomes deposited publicly, but 16 countries did not have local sequencing capacities. Turnaround time of locally sequenced SARS-CoV-2 genomes was significantly shorter than for outsourced sequencing to laboratories located outside of Africa. Interestingly, nearly half of the African SARS-CoV-2 genomes were sequenced using ONT. The first wave of infections was caused by B.1 and B.1.1 lineages on the whole continent. The second wave of infections was dominated by Alpha VOC in West, North and parts of Central Africa and Beta VOC in Southern and East Africa, co-circulation of multiple VOC only

occurred in few countries. Variants of interest (VOI) circulated in certain regions at high levels. As an example, Eta significantly contributed to the second and third wave in West Africa. Delta and Omicron VOC spread across the whole continent. South Africa and Nigeria each sequenced about 1% of their cases, enabling the timely detection of Beta and Omicron VOC and ETA VOI, respectively.

In the second part of this thesis we aimed to improve the monitoring of parasites that evade diagnosis, treatment or vaccination. We used DNA from malaria rapid diagnostic tests (RDTs) to investigate their diagnostic performance and to characterise *Plasmodium* spp. and filarial nematodes circulating on Bioko Island. We developed a simplified protocol to detect malaria parasite nucleic acids from dried blood spots using a portable qPCR-platform. We developed a bioinformatics pipeline to detect single nucleotide polymorphisms associated with antimalarial drug resistance in multiclonal *P. falciparum* infections. The second part is based on four manuscripts.

Manuscript 5: Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment.

This manuscript presents the analysis of nucleic acids extracted from used RDTs for the presence of *Plasmodium* spp. to evaluate the diagnostic performance of RDTs used in the 2018 malaria indicator survey on Bioko Island, Equatorial Guinea. We tested 2865 RDTs, of which 1800 were recorded malaria negative and 1050 were recorded malaria positive. Using RT-qPCR as a gold standard, the overall sensitivity and specificity of RDTs were 90% and 85% respectively. Only 4.7% of negative RDTs were false-negative, whereas 28.4% of positive RDTs were false-positive. False-negative RDT had lower parasite densities and *P. malariae* and *P. ovale* spp. were more prevalent than in true-positive RDTs. False-positive RDTs were associated with the recent use of antimalarial drugs and higher socio-economic status. Moderate to severe anaemia and residing in the rural Bioko Sur province decreased the odds of having a false-positive RDTs. We observed 11.1% of tested *P. falciparum* likely having a *P. falciparum*-specific histidine rich protein 3 (*pfhrp3*) gene deletion. Masked *pfhrp2* and *pfhrp3* gene deletions, where at least one strain in a mixed *P. falciparum* infection carried a deletion, were identified in 16.6% of tested *P. falciparum*.

Manuscript 6: Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests.

In this manuscript, we repurposed the nucleic acids extracted from used RDTs described in manuscript 5 for molecular detection of the blood-dwelling, highly neglected filarial nematodes *Mansonella*

perstans and *Loa loa*. Both species were simultaneously screened for in a multiplex qPCR assay. Overall, 6.6% of RDTs were positive for *M. perstans* and 1.5% for *L. loa*. The *M. perstans* positivity rate was higher in males, residing in rural districts and with lower socio-economic status and also increased with age. *L. loa* had higher positivity rates in the rural Bioko Sur province and in adults, but no difference between males and females was observed. Co-infections between *L. loa*, *M. perstans* and *Plasmodium* spp. were as frequent as expected based on the relative positivity rates in each age group and were most often found in Malabo.

Manuscript 7: Development and evaluation of PlasmoPod: A cartridge-based nucleic acid amplification test for rapid and decentralized malaria diagnosis and surveillance.

This manuscript describes the development of a cartridge-based RT-qPCR test for rapid detection and quantification of malaria parasites called PlasmoPod. A simplified nucleic acid extraction from dried blood spots (DBS) was performed by boiling a single 3 mm diameter DBS punch in Chelex for three minutes. After centrifugation the supernatant was loaded into a ready-to-use cartridge containing dried primers, probes and RT-qPCR reagents. The RT-qPCR on the rapid and mobile PCR cyclers was completed within 30 minutes. Tests using purified DNA from cultured *P. falciparum*, malaria-free human blood and other pathogens resulted in perfect specificity and high sensitivity of the PlasmoPod with a calculated detection limit of 0.02 *P. falciparum* parasites per microliter. PlasmoPod identified 81.4% of samples positive by a reference RT-qPCR on a standard laboratory device using nucleic acids extracted from RDTs of asymptomatic individuals carrying malaria parasites during the malaria indicator survey in Equatorial Guinea, detection probability depended on the number of input target molecules. Using clinical samples of symptomatic children, PlasmoPod had a sensitivity of 93.6% and a specificity of 100% compared to a conventional column-based nucleic acid extraction method and the highly sensitive 18S rRNA RT-qPCR on a standard qPCR device. PlasmoPod outperformed thick blood smear microscopy and RDT for detection of *Plasmodium* parasites in this study.

Manuscript 8: PHARE: A bioinformatics pipeline for compositional profiling of multiclonal *Plasmodium falciparum* infections from long-read Nanopore sequencing data.

Presented in the manuscript is the development of a bioinformatics pipeline to detect all haplotypes of distinct genes in *P. falciparum* multiclonal infections sequenced with ONT. We evaluated the pipeline with the well-known drug resistance marker genes dihydrofolate reductase (*pfdhfr*), dihydropteroate synthase (*pfdhps*) and Kelch 13 (*pfk13*). We amplified the full length drug resistance marker genes from mixtures of different ratios of *P. falciparum* laboratory strains with well characterized SNPs. A minor haplotype had to be present in at least 9% of reads to distinguish from false haplotypes. The pipeline detected the correct haplotypes in mixtures of two and four *P.*

falciparum laboratory strains. We tested the pipeline on data from twelve blood samples of febrile children collected at a single site in the Central African Republic. In this clinical dataset, six different haplotypes were found that are defined by five SNPs in *pfdhps*, while two haplotypes were found in *pfdhfr* based on one SNP. All SNPs identified in the *pfdhfr* and *pfdhps* genes are associated with antimalarial drug resistant phenotypes. One SNP not associated with antimalarial drug resistance was detected in the *pfk13* gene.

In the third part of this thesis we aimed to evaluate the use of bacterial WGS for the surveillance of foodborne bacteria in the white meat production chain using a One Health approach.

Manuscript 9: Whole Genome Sequencing for One Health Surveillance of Antimicrobial Resistance in Conflict Zones: A Case Study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine.

In this manuscript we characterise *Campylobacter jejuni* and *Salmonella enterica* collected during a One Health cross-sectional study along the white meat production chain in Palestine. The positivity rates for *C. jejuni* were higher than for *S. enterica* in chicken manure, chicken meat and human stool. Chicken manure for *C. jejuni* isolation was collected from the same farms in two consecutive years. Additionally we included *C. jejuni* isolates from hospitalized gastroenteritis patients collected in 2022. WGS revealed that isolates collected during the same year were highly similar independent of sample origin and belonged to the same sequence type. The average nucleotide identities (ANI) of whole genome sequences compared between the two years were only 97.84%. Isolates collected in 2021 had clearly different drug resistance gene patterns than those collected in 2022. The majority of *S. enterica* isolates collected in 2021 were highly similar with ANI above 99.99% and were assigned to *S. enterica* serotype Muenchen. They were almost identical to an emerging multidrug resistant clinical isolate carrying a megaplasmid recently described from Israel and contained the same drug resistance genes. Our work highlights the importance of WGS for monitoring of pathogens and the utility of easily collectable environmental samples for One Health surveillance.

List of abbreviations

ACT	Artemisinin-based combination therapy
AMR	Antimicrobial resistance
ANI	Average nucleotide identity
ASIC	Application-specific integration circuit
CAR	Central African Republic
CI	Confidence interval
CNV	Copy number variation
COVID-19	Coronavirus disease 2019
Cq	Quantification cycle
DBS	Dried blood spot
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of the Congo
ENAR	Extraction of Nucleic Acids from RDTs
GISAID	Global Initiative on Sharing All Influenza Data
HIV	Human immunodeficiency virus
HRP2	Histidine-rich protein 2
iNTS	Invasive non-typhoidal salmonella
IPTp	Intermittent preventive treatment of malaria in pregnancy
IRS	Indoor residual spraying
iSRS	Integrated surveillance response system
ITN	Insecticide-treated net
LDH	Lactate dehydrogenase
LF	Lymphatic filariasis
LNA	Locked nucleic acid

LOD	Limit of detection
MDA	Mass drug administration
MERS-CoV	Middle East respiratory syndrome coronavirus
MIS	Malaria indicator survey
MLST	Multilocus sequence typing
mRNA	Messenger ribonucleic acid
NA	Nucleic acids
NAAT	Nucleic acid amplification test
NGS	Next-generation sequencing
ONT	Oxford Nanopore Technologies
PacBio	Pacific Biosciences
PCR	Polymerase chain reaction
Pf	Plasmodium falciparum
P spp	Plasmodium spp.
qPCR	Quantitative polymerase chain reaction
RBD	Receptor binding domain
RDT	Rapid diagnostic test
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
RT-qPCR	Reverse transcription qPCR
SARS-CoV	Severe acute respiratory syndrome coronavirus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SNP	Single-nucleotide polymorphism
SP	Sulfadoxine-pyrimethamine
STR	Short tandem repeat

SV	Structural variant
TBS	Thick blood smear
UTR	Untranslated region
VOC	Variant of concern
VOI	Variant of interest
WGS	Whole genome sequencing
WHO	World Health Organisation

1. Introduction

1.1. Causes of acute febrile illnesses in sub-Saharan Africa

Common syndromes of acute infectious diseases include signs and symptoms of acute respiratory infection and/or involvement of the gastrointestinal tract often accompanied by acute febrile illness (Feikin et al., 2011). Acute febrile illness is caused by many different pathogens, making the distinction of bacterial, parasitic and viral infections challenging based on clinical diagnosis alone. Appropriate drug treatment regimens targeting the disease causing pathogen should be administered rapidly to avoid complications and to prevent drug resistance development. A study conducted in Tanzania diagnosed 1005 febrile children aged between two months and ten years for the potential cause of fever using a range of laboratory tests (D'Acremont et al., 2014). Taking into account multiple diagnoses, viral infection was found in 70.5% of fever cases, bacterial infections in 22.0% and parasitic disease in 10.9%. The cause remained undetermined in 3.2% of fever cases (D'Acremont et al., 2014). Clearly, treatment decision in primary health care based purely on clinical symptoms is insufficient to guide adequate treatment decisions (van de Maat et al., 2021).

A study in Burkina Faso found that empirical treatment with antimalarial drugs in children with acute febrile illness was common even when rapid diagnostic tests (RDTs) were available and that the majority of viral infections were treated with antibiotics (Kaboré et al., 2021). They also described malaria co-infections in 21.3% of invasive bacterial infections (Kaboré et al., 2021). Overall, in low and middle income countries the prescription of antibiotics to children is very high and might contribute to the emergence of antibiotic drug resistance in bacterial infections (Fink et al., 2020). Better decision making processes particularly in the primary health care sector in combination with next-generation, specific, fast and decentralized testing approaches for a diverse set of pathogens would alleviate this threat possibly saving lives (Kaboré et al., 2021).

1.2. Overview of SARS-CoV-2

1.2.1. Epidemiology and biology of SARS-CoV-2

In December 2019, a novel coronavirus causing pneumonia was first reported from the Chinese town Wuhan which then rapidly spread worldwide in the following weeks and months (Zhu et al., 2020). The first human-to-human transmission was estimated to have occurred between mid-October and mid-November 2019 in Hubei Province, China (Pekar et al., 2021). On 11th March 2020, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causing coronavirus disease 2019 (COVID-19) was declared a pandemic. The World Health Organisation (WHO) reported 7.3 billion SARS-CoV-2 cases and 6.7 million COVID-19 deaths between January 2020 and December 2022. Male sex, pre-existing

comorbidities such as hypertension and Diabetes mellitus, obesity and older age are risk factors for progression to severe COVID-19 disease (Zhang et al., 2023). Common symptoms of SARS-CoV-2 infections include fever, dry cough, loss of smell, fatigue, myalgia and dyspnoea (Docherty et al., 2020). Common routes of human to human transmission are through droplets and aerosols (Salzberger et al., 2021). The mean incubation period for wildtype SARS-CoV-2 strains is 6.6 days, variants of concern (VOC) Delta and Omicron have shortened incubation periods of 4.4 days and 3.4 days, respectively (Wu et al., 2022). Asymptomatic SARS-CoV-2 infections account for about one third of all laboratory confirmed cases, with higher numbers of asymptomatic cases in children and higher numbers of symptomatic infections in people suffering from comorbidities (Sah et al., 2021).

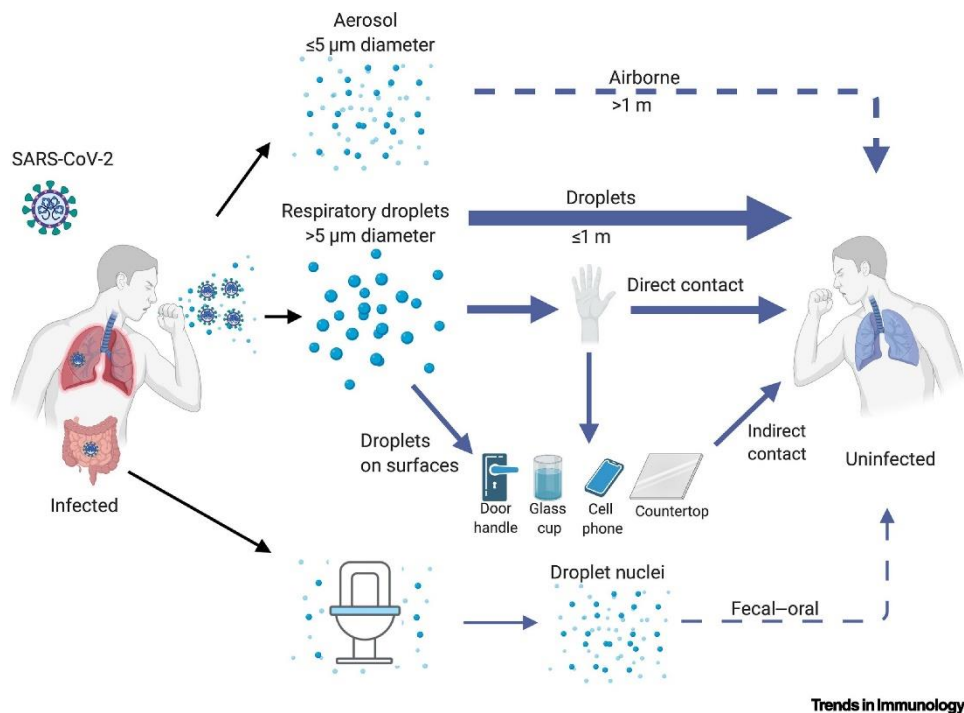


Figure 1. Transmission routes of SARS-CoV-2 between humans. Infectious viruses are found in the respiratory tract and the gastrointestinal tract. Different modes of transmission for close contacts are respiratory droplets, direct contact and indirect contact through contaminated surfaces. Airborne transmission through inhaled aerosols can occur over longer distances. The faecal-oral route is an additional transmission route. Figure taken from Harrison *et al.* (Harrison et al., 2020).

SARS-CoV-2 belongs to the genus Betacoronavirus and the subgenus Sarbecovirus (Wu et al., 2020). Coronaviruses are enveloped, positive-stranded RNA viruses with genome sizes of about 30 kilobases (Su et al., 2016). Genomic organisation is similar among Coronaviruses, consisting of a 5' untranslated region (UTR), replicase complex containing two open reading frames with several non-structural proteins (ORF1ab), spike, envelope, membrane, nucleocapsid, and 3' UTR (Fehr and Perlman, 2015).

Recombination and genomic plasticity is common in coronaviruses, due to their large unsegmented RNA genome (Latinne et al., 2020). Before 2020, six coronavirus species that caused respiratory disease in humans were known. Four of them are well adapted to humans, occur worldwide and cause common cold, namely HCoV-229E, HCoV-OC43, HCoV-NL63 and HCoV-HKU1 (Su et al., 2016). Only two of them were highly pathogenic and caused large outbreaks, namely the Severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV) (Drosten et al., 2003, Zaki et al., 2012).

1.2.2. Origin of SARS-Cov-2

SARS-CoV-2 is most likely of zoonotic origin and related coronaviruses have been found in *Rhinolophus* bat species across Southeast Asia (Zhou et al., 2021). The similarity between SARS-CoV-2 and the closely related bat coronavirus RaTG13 isolated from *Rhinolophus affinis* in Yunnan province in 2013 is 96.2% at the nucleotide level, however the conservation in the receptor binding domain (RBD) is below 90% (Zhou et al., 2020). Bat coronaviruses isolated from *Rhinolophus* spp. in Laos in 2020 have similar RBD to SARS-CoV-2 and a high binding affinity to the human cellular angiotensin-converting enzyme 2 receptor, but lack the furin cleavage site (Temmam et al., 2022). Other SARS-related coronaviruses with high similarity in the receptor binding domain have been isolated from pangolins, which were proposed as intermediate hosts (Lam et al., 2020). Multiple recombination breakpoints were identified in *Sarbecovirus* genomes and SARS-CoV-2 fragments were assigned to multiple bat and pangolin coronavirus strains (Temmam et al., 2022). A wide range of animals is susceptible to infection with SARS-CoV-2, but no clear intermediate host between bats and humans has been identified (Maurin et al., 2021).

1.2.3. Evolution of SARS-CoV-2

Little diversifying selection took place between December 2019 and October 2020, indicating that most adaptive changes for human to human transmission probably occurred before SARS-CoV-2 emerged in the human population (MacLean et al., 2021). The D614G Spike mutation which emerged in March 2020 in Europe led to increased viral loads in infected patients and rapidly became the dominant SARS-CoV-2 variant worldwide (Korber et al., 2020). Three divergent SARS-CoV-2 lineages that shared the N501Y Spike mutation in addition to the D614G Spike mutation emerged in late 2020 in different parts of the world and caused a second wave of infections (Martin et al., 2021). Lineage B.1.1.7, which was later named Alpha VOC was first identified in the United Kingdom (Rambaut et al., 2020b), lineage B.1.351 (Beta VOC) was first identified in South Africa (Tegally et al., 2021) and P.1 (Gamma VOC) was first identified in Brazil (Faria et al., 2021). All three lineages have a set of lineage defining amino acid substitutions and deletions, many of which arose independently in multiple lineages, that resulted in increased virus transmissibility, evasion of immune responses or increase in

virulence (Martin et al., 2021). Persistent SARS-CoV-2 infections in immunocompromised patients can lead to highly mutated virus variants (Corey et al., 2021). A third wave was caused by lineage B.1.617.2 (Delta VOC), which emerged in India in late 2020 and had reduced sensitivity to neutralizing antibodies (Mlcochova et al., 2021). In November 2021, SARS-CoV-2 lineage B.1.1.529 (Omicron VOC) was detected in South Africa, causing a fourth wave of infections and within three weeks Omicron was identified in 87 countries (Viana et al., 2022). Omicron VOC has been divided in over one hundred sublineages, with distinct transmission and immune evasion capabilities (Xia et al., 2022). The original Omicron lineage BA.1 is highly divergent from previous lineages, has 30 amino acid substitutions in the Spike gene alone, 13 of which are at previously conserved sites (Martin et al., 2022). These substitutions represent a shift in antigenicity and causes a significant reduction in antibody neutralizing activity in vaccinated and convalescent individuals (Cameroni et al., 2022). Early in 2022, SARS-CoV-2 Omicron lineage BA.2 began outcompeting BA.1. BA.2 is more transmissible, more pathogenic and escapes immunity induced by BA.1 infection (Yamasoba et al., 2022). Comparing the fitness advantage of VOC to the original A lineage of SARS-CoV-2, Omicron lineage BA.2 has an almost nine-fold and Omicron lineage BA.1 a 7.5-fold fitness advantage, much higher than Delta which has a three-fold fitness advantage (Obermeyer et al., 2022). Recombinant viruses with genomic elements of different VOC including Delta and Omicron or with BA.1 and BA.2 have been observed (Chakraborty et al., 2022)

1.2.4. Public Health interventions

Lockdown measures, quarantine and isolation increased the doubling time of COVID-19 cases from two to four days in China (Lau et al., 2020). Public health measures such as regular handwashing, mask wearing and physical distancing decreased the incidence of COVID-19 (Talic et al., 2021). Within one year of the pandemic, vaccines had been developed using traditional methods such as inactivated virus and protein subunit vaccines, but also viral vector and nucleic acid-based (mRNA) vaccine approaches were followed (Graña et al., 2022). By the end of 2022, 5.5 billion people worldwide have received at least one dose of any COVID-19 vaccine and 13.1 billion COVID-19 vaccine doses have been administered in total (Mathieu et al., 2021). A meta-analysis found overall above 70% vaccine efficacy against severe disease for two adenovirus vector vaccines and above 80% for two mRNA-based vaccines for up to six months after vaccination (Feikin et al., 2022).

1.2.5. Molecular surveillance of SARS-CoV-2

The gold standard for diagnosis is molecular detection of SARS-CoV-2 nucleic acids by reverse transcription quantitative polymerase chain reaction (RT-qPCR) in respiratory samples (Böger et al., 2021). Rapid antigen tests are useful to shorten time to result for symptomatic patients less than a week after symptom onset (Dinnes et al., 2021). Antibody tests have a high accuracy three weeks to 100 days after infection, making them most useful for seroprevalence surveys (Fox et al., 2022).

Genomic surveillance has been critical for monitoring the evolution of SARS-CoV-2 and for detection of new variants (Knyazev et al., 2022). The majority of SARS-CoV-2 whole genomes have been generated with amplicon sequencing using Illumina or Nanopore technology (Chiara et al., 2021). Over 15 Million SARS-CoV-2 genomic sequences and their metadata have been shared on the online platform Global Initiative on Sharing All Influenza Data (GISAID) (Shu and McCauley, 2017). Up-to-date phylogenetic trees are publicly available on the interactive visualization platform Nextstrain (Hadfield et al., 2018). Rambaut et al. proposed a dynamic nomenclature for SARS-CoV-2 lineages based on machine learning (Rambaut et al., 2020a). Over 90% of all sequences from 2021 were submitted by high-income countries, while only 0.1% were submitted by low-income countries (Ohlsen et al., 2022). The time from sample collection to submission was five times shorter for high-income countries compared to low-income countries, with a median lag time of 20 days and 98 days respectively (Ohlsen et al., 2022). Global genomic surveillance capacity increased over time, the number of sequences generated during the first eight weeks of the spread of a new VOC doubled from Alpha to Delta and again to Omicron (Ohlsen et al., 2022).

1.3. Overview of *Plasmodium* spp.

1.3.1. Parasite life cycle

Malaria is an infectious disease caused by eukaryotic single cell organisms belonging to apicomplexan parasites *Plasmodium* spp.. Six species cause malaria in humans, namely *P. falciparum*, *P. malariae*, *P. ovale curtisi*, *P. ovale wallikeri* and *P. vivax* (Ashley et al., 2018). In South East Asia, *P. knowlesi* causes malaria in long-tailed macaques but was found to increasingly become a zoonotic disease with clinical presentations in humans (Chin et al., 1965, Singh et al., 2004). Molecular analysis revealed that *P. ovale* spp. malaria is caused by two closely related sympatric species, *P. ovale curtisi* and *P. ovale wallikeri* (Sutherland et al., 2010). *P. falciparum* is the dominant and most lethal malaria species in Africa (World Health Organization, 2022c). Malaria is transmitted by infected female *Anopheles* mosquitoes during feeding on the human host. *Plasmodium* spp. have a complex life cycle, shown in Figure 2. A malaria-infected female *Anopheles* deposits an estimated 1-100 motile sporozoites into the dermis during a blood meal (Beier et al., 1991). The extracellular sporozoites migrate through the dermis to reach the blood stream that transports them to the liver where they invade hepatocytes (Flores-Garcia et al., 2018). Depending on *Plasmodium* species, strain and host immunity the parasite goes through the asymptomatic liver stage in one to two weeks (Ashley et al., 2018). After asexual replication ten thousands of liver merozoites are released in the bloodstream and invade erythrocytes (Prudêncio et al., 2006). Dormant liver stages called hypnozoites in *P. vivax* and *P. ovale* spp. cause relapses weeks to months after infection (White, 2011). The repeated asexual blood stage cycles take about 48 hours for *P. falciparum*, *P. vivax*, and *P. ovale* spp., while it is 72 hours for *P. malariae* and

only 24 hours for *P. knowlesi* (White et al., 2014). At the end of each asexual blood stage cycle, schizonts burst and release up to 32 single merozoites into the bloodstream, leading to an approximately eight-fold parasite multiplication rate in susceptible individuals (Simpson et al., 2002). Malaria infection followed by the destruction of infected and uninfected erythrocytes, in combination with the host immune response cause malarial disease (White et al., 2014). A small proportion of blood stage merozoites undergoes sexual differentiation and develops into male and female gametocytes that transmit malaria to mosquitoes (Josling and Llinás, 2015). In the mosquito gut male gametocytes undergo mitosis to form eight microgametes that fertilize the activated female gametocyte (Guttery et al., 2022). The diploid zygote transforms into an ookinete, invades the midgut wall, and develops into an oocyst that produces hundreds of sporozoites that migrate to the salivary glands, to complete the life cycle (Guttery et al., 2022).

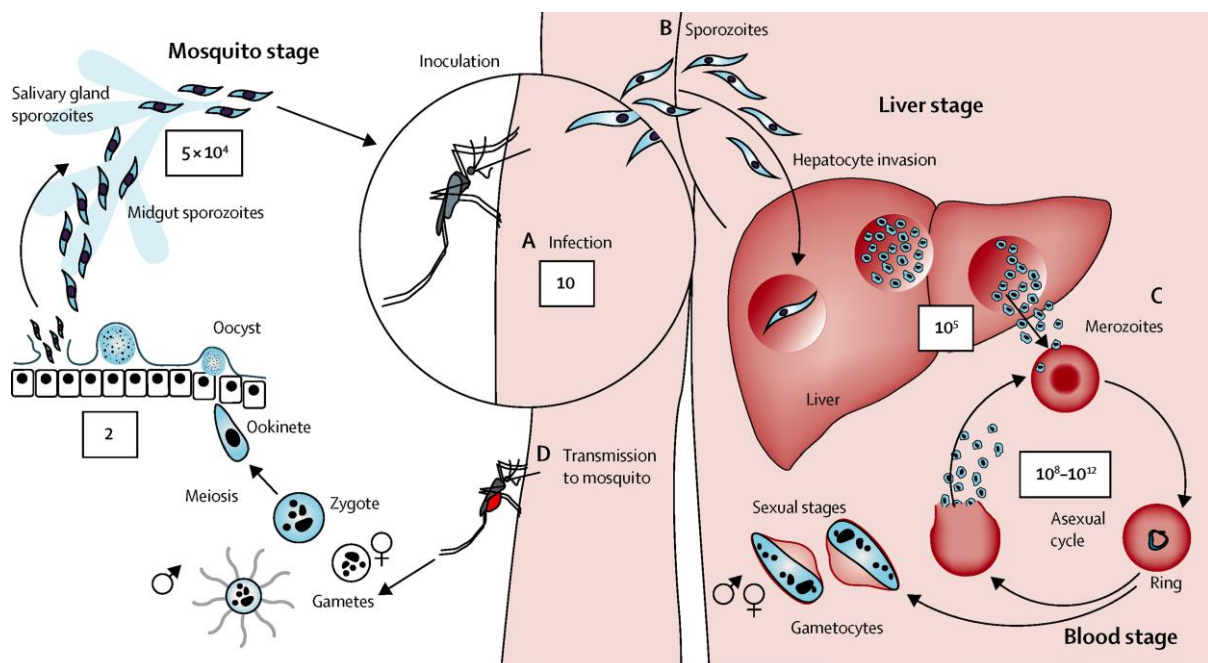


Figure 2. The life cycle of *Plasmodium* spp. parasites. The life cycle starts when (A) a malaria-infected female *Anopheles* injects motile sporozoites into a human during a blood meal. (B) Sporozoites migrate to the liver where they infect hepatocytes, undergo asexual replication and mature into liver schizonts. (C) From the infected hepatocytes, merozoites are released into the bloodstream where they infect erythrocytes. The asexual blood-stage cycle results in exponential parasite replication causing malaria symptoms. A number of merozoites develop into female and male gametocytes, starting the sexual erythrocytic stages. (D) A mosquito takes up a blood meal containing gametocytes. In the mosquito's gut the gametes fuse and undergo multiple steps to develop into sporozoites. Figure taken from White *et al.* (White et al., 2014).

1.3.2. Epidemiology

With an estimated 247 million cases in 2021 and almost half of the world's population at risk, malaria is one of the top three infectious diseases worldwide (World Health Organization, 2022c). The main burden of malaria affects vulnerable populations such as children below the age of five years and primigravidae women in resource poor settings (World Health Organization, 2022c). Children below the age of five years accounted for 76.8% of total malaria deaths in 2021 (World Health Organization, 2022c). The global *P. falciparum* prevalence in children aged 2 to 10 years is shown in Figure 3. The WHO African Region carried 95% of global malaria cases and four countries including Nigeria (27%), the Democratic Republic of the Congo (DRC) (12%), Uganda (5%) and Mozambique (4%) harboured half of the African malaria cases (World Health Organization, 2022c). Between 2000 and 2015 the malaria mortality rate was reduced by roughly fifty percent and continued to decrease to 568,000 deaths in 2019. Due to prevention and treatment interruptions caused by the COVID-19 pandemic and a change in mortality estimations, the number of malaria deaths increased again to 625,000 and 619,000 in the years 2020 and 2021, respectively (World Health Organization, 2022c).

Uncomplicated malaria causes non-specific symptoms such as fever, chills, headache, cough and diarrhoea and the mortality rate is 0.1% if effective treatment is administered promptly (Ashley et al., 2018). Severe malaria more commonly manifests with severe anaemia, convulsions and hypoglycaemia in children, while acute renal failure and jaundice are more common in adults (White, 2022). Cerebral malaria with coma and metabolic acidosis increase significantly the odds of death among patients with severe malaria in all age groups (Dondorp et al., 2008). The mortality rate of severe malaria in African children and Asian adults treated intravenously with the artemisinin derivate artesunate was 8.5% and 15%, respectively (Dondorp et al., 2005a, Dondorp et al., 2010).

Children in malaria-endemic settings with stable transmission acquire sufficient immunity to protect them from life-threatening disease after one to two episodes of severe malaria at a young age (Gupta et al., 1999). Older children still have mild malaria symptoms while adults are mostly asymptomatic parasite carriers (Marsh and Kinyanjui, 2006). The WHO estimated 42 million pregnant women in sub-Saharan Africa being exposed to malaria in 2021, of which 13.3 million had malaria infections (World Health Organization, 2022c). Globally, 8.2% of stillbirth is attributable to malaria in pregnancy, in sub-Saharan Africa the estimate is 19.7% (Lawn et al., 2016). In semi-immune women in stable transmission settings, malaria infections during pregnancy are mostly asymptomatic but can result in low birthweight and severe anaemia, with the highest risk in primigravid women (Fried and Duffy, 2017). Intermittent preventive treatment of malaria in pregnancy (IPTp) has been adapted in 35 African countries. In 2021, 55% of pregnant women received at least one dose of chemoprophylaxis

at an antenatal care visit, averting 457,000 neonates with low birthweight (World Health Organization, 2022c).

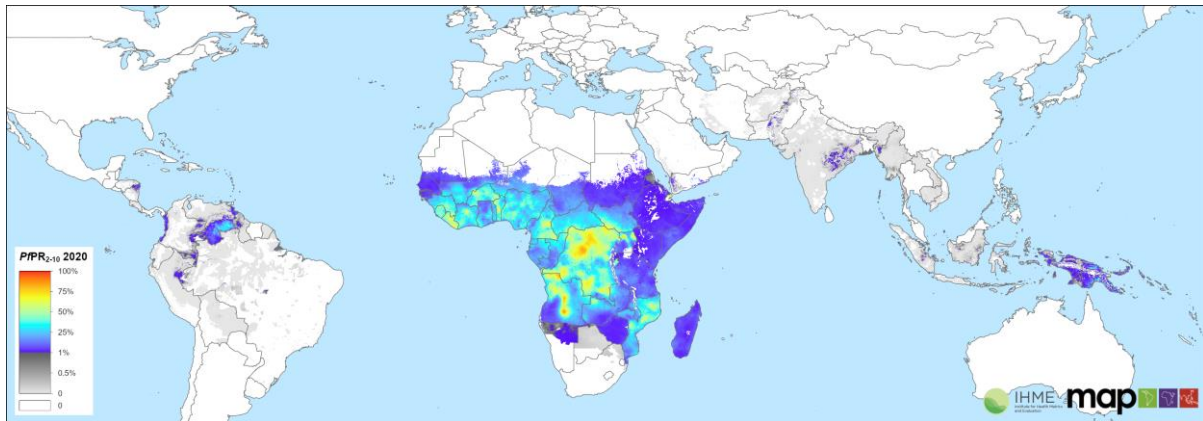


Figure 3. Global *P. falciparum* infection prevalence in 2020. The Malaria Atlas Project (MAP, <https://map.ox.ac.uk/>) was used to map the *P. falciparum* parasite rate, in 2-10 year old children with data collected in 2020.

1.3.3. Malaria prevention and treatment

Successful malaria control consists of a combination of rapid access to diagnostics, effective treatment and vector control. Insecticide-treated nets (ITN) and indoor residual spraying (IRS) with long-lasting insecticides are used to prevent malaria transmission. More than 2.5 billion ITNs were distributed globally between 2004 and 2021, 87% of them in sub-Saharan Africa (World Health Organization, 2022c). Of 663 million clinical malaria cases averted between 2000 and 2015 in Africa, 68% can be attributed to ITNs, 22% to artemisinin-based combination therapy (ACT) and only 10% to IRS (Bhatt et al., 2015). Incidence of uncomplicated and severe malaria episodes was reduced by 45% and 44% respectively in participants using ITNs compared to not sleeping under a bednet (Pryce et al., 2018). Adding non-pyrethroid-like IRS to ITS decreased the malaria parasite prevalence by an additional 39% compared to ITS alone (Pryce et al., 2022). Unfortunately, pyrethroid resistance is widespread in *Anopheles* spp. across Africa (Ranson et al., 2011). Resistance to all four insecticide chemical classes currently in use (pyrethroids, organophosphates, carbamates and organochlorines) was confirmed for *Anopheles* spp. from all WHO regions (World Health Organization, 2022c).

Resistance to chloroquine evolved in the late 1950s in two independent foci in Colombia and at the Cambodia-Thailand border (Payne, 1987). The first chloroquine resistant *P. falciparum* in Africa were described in Tanzania and Kenya in the late 1970s and within a decade were widespread across the continent (Peters, 1987). Geographic variants of chloroquine resistant *P. falciparum* have combinations of four to ten mutations in the chloroquine resistance transporter gene (Ecker et al.,

2012). Chloroquine resistance increased malaria morbidity and mortality in children massively, due to the lack of an effective and cheap alternative drug (Trape et al., 1998). A combination of the two antifolate drugs pyrimethamine and sulfadoxine (SP) became the first line treatment when chloroquine was determined as ineffective (Bloland et al., 1993). However, resistance to SP developed rapidly and was associated with a stepwise accumulation of mutations in the two genes dihydrofolate reductase and dihydropteroate synthase (Sibley et al., 2001). SP is still used for IPTp and has been shown to prevent adverse birth outcomes but does not clear blood stage parasitaemia in pregnant women infected with SP resistant *P. falciparum* (Kayiba et al., 2021). The antimalarial properties of artemisinin, extracted from *Artemisia annua*, was first described in the 1970s by Chinese scientists (1979). Artemisinin derivatives reduce the parasite load by a factor of 10,000 per replication cycle, having a higher killing rate than any other known antimalarial (White and Olliaro, 1996). Due to the short half-life of artemisinin derivatives, combination therapy with a slow acting partner drug is recommended to prevent treatment failure and resistance development (White, 2004). Artemisinin resistance, characterized by slow parasite clearance at the ring stage after appropriate treatment dosing, was first reported from Cambodia in 2009 (Dondorp et al., 2009). Artemisinin resistance has become widespread in the Greater Mekong Subregion and more recently it was reported from Africa (Ashley et al., 2014, Balikagala et al., 2021). Different strategies have been proposed to prolong the efficacy of ACT, such as extending the treatment duration, alternating between different ACT regimens or adding a third antimalarial drug to the ACT (Hanboonkunupakarn et al., 2022).

1.3.4. Malaria diagnosis

The gold standard for malaria diagnosis remains microscopy of Giemsa stained whole blood films. Thick blood smears provide detection and also allow quantification of parasitaemia, while thin blood smears additionally allow morphological species identification (Moody, 2002). Experienced microscopists can detect parasite densities of 50 *Plasmodium* spp. per microliter of blood, however routine diagnostic laboratories achieve up to ten times lower sensitivities (Moody, 2002). Microscopy can underestimate parasitaemia, as it is unable to detect sequestered *P. falciparum* (Dondorp et al., 2005b). RDTs have almost completely replaced microscopy in resource-limited settings, with 3.5 billion RDTs sold globally between 2010 and 2021 (World Health Organization, 2022c). RDTs most commonly detect histidine-rich protein 2 (HRP2) specific for *P. falciparum* and lactate dehydrogenase (LDH) and aldolase for pan-*Plasmodium* detection (Mouatcho and Goldring, 2013). HRP2-based RDT are more sensitive and have a slightly higher accuracy than LDH-based RDTs but show lower specificity (Li et al., 2017). HRP2-based RDTs have a limit of detection of 100 *P. falciparum* per microliter of blood (Varo et al., 2021). False-positive RDTs after recent malaria treatment are common due to the persistence of HRP2 in the blood for up to six weeks (Dalrymple et al., 2018). *P. falciparum* lacking the

pfhrp2 gene can cause false-negative HRP2-based RDTs (Gendrot et al., 2019). Molecular assays, such as quantitative PCR (qPCR) of multi-copy DNA targets or RT-qPCR targeting total nucleic acids can reach sensitivities of 0.1 *Plasmodium* spp. per microliter of blood (Hofmann et al., 2015, Kamau et al., 2011). Sub-microscopic infections accounted for 49% of malaria infections detected by qPCR, with significantly higher proportions in low-transmission settings versus high-transmission settings (Okell et al., 2009). The current limitations for widespread use of qPCR are the need for laboratory equipment, electricity, cold chain, trained personnel and the high costs per test (Varo et al., 2021).

1.4. Overview of filarial nematodes

1.4.1. Parasite characteristics and life cycle

Eight species of filarial nematodes infect humans, the most well-known are *Wuchereria bancrofti*, *Brugia malayi* and *B. timori* causing lymphatic filariasis (LF) and *Onchocerca volvulus* causing river blindness (Taylor et al., 2010). Filarial nematodes share similar life cycles, shown in

Figure 4. Depending on the species, the long threadlike adult worms live in tissues, body cavities or lymphatic vessels (Bogitsh et al., 2019a). Females are ovoviviparous, the microfilariae migrate to the blood vessels after deposition (Bogitsh et al., 2019a). Microfilariae can survive for years in the blood stream, and are taken up by an insect vector eventually (Bogitsh et al., 2019a). In the insect vector, two moults take place during larval development and after three weeks infective filariform larvae can be transmitted to the human host (Bogitsh et al., 2019a). During migration to the definitive site of infection, larvae metamorphose into adult worms (Bogitsh et al., 2019a). *W. bancrofti*, *B. malayi*, *B. timori* and *O. volvulus* depend on the presence of the bacterial *Wolbachia* endosymbionts, that are absent in *Loa loa* and *Mansonella perstans* (Taylor et al., 2005). Adult *W. bancrofti*, *B. malayi* and *B. timori* reside in the lymphatic system causing LF (Cross, 1996). *W. bancrofti* occurs in many subtropical and tropical countries and is responsible for 90% of cases, *B. malayi* is restricted to Southeast Asia and *B. timori* only occurs in South-eastern Indonesia (Taylor et al., 2010). Female worms of *W. bancrofti* can live for up to eight years and produce millions of microfilariae that migrate to peripheral blood vessels to be ingested by mosquitoes during a blood meal (Taylor et al., 2010). *Aedes* spp., *Anopheles* spp. and *Culex* spp. mosquitoes can transmit LF in different parts of the world (Knopp et al., 2012). Adult *O. volvulus* reside in subcutaneous tissues, their microfilariae migrate through the skin and eyes and cause skin disease and blindness (Cross, 1996). Blackflies of the genus *Simulium* that are intermediate hosts of *O. volvulus* lay their eggs in fast-flowing water (Bogitsh et al., 2019b). Adult *L. loa* worms reside in the subcutaneous tissues and the microfilariae can be found in the blood, other body fluids and in the lungs (Padgett and Jacobsen, 2008). Tabanid flies of the genus *Chrysops* transmit *L. loa* (Boussinesq, 2006). Clinical symptoms of loiasis are worm migration through the conjunctiva of

the eyes and Calabar swellings frequently located on the forearms, which are subcutaneous oedemas often accompanied by itching (Boussinesq, 2006). Human mansonellosis is caused by the three species *M. perstans*, *Mansonella streptocerca*, and *Mansonella ozzardi* (Simonsen et al., 2014). Adult *M. perstans* reside in the serous body cavities, the microfilariae circulate in the blood and are transmitted by biting midges of the genus *Culicoides* (Simonsen et al., 2011).

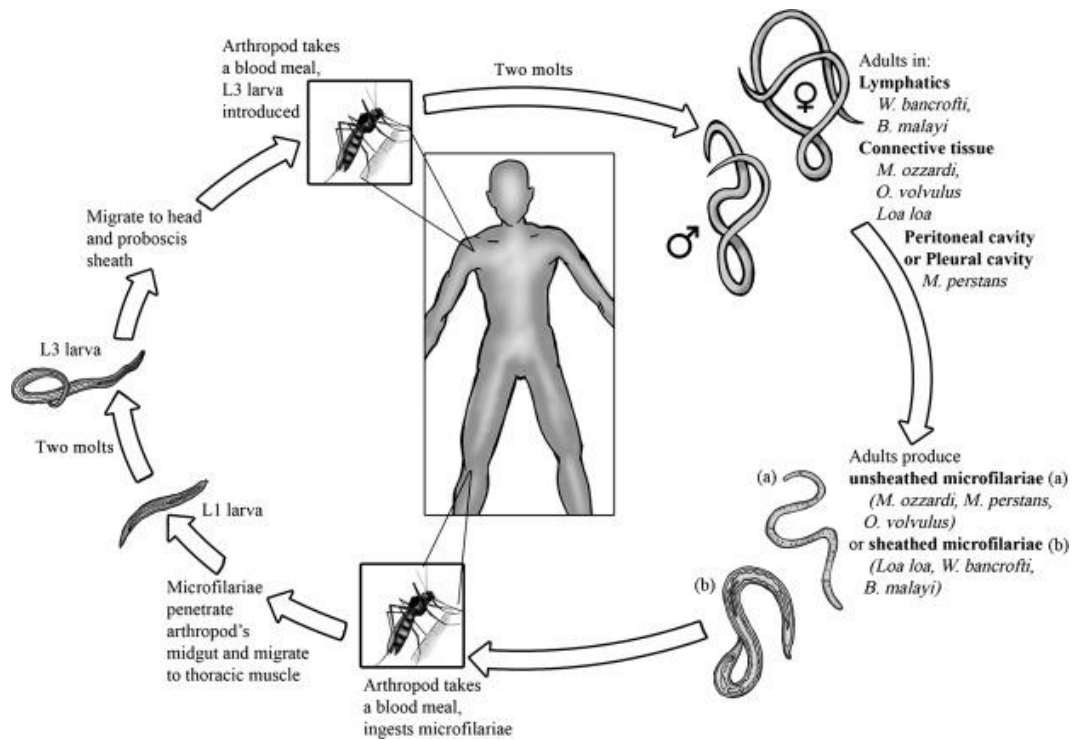


Figure 4. Life cycle of filarial nematodes. Microfilariae hatch in the uterus of female worms. After larviposition, the microfilaria migrate to the blood stream where they can survive for several years. Upon ingestion by a suitable arthropod vector, they develop into infective filariaform larvae and migrate to the proboscis sheath. During the next blood meal the larvae gain access to the human circulatory system. Larvae develop into adult worms and migrate to the definite site of infection- Figure taken from Bogitsh *et al.* (Bogitsh et al., 2019b)

1.4.2. Epidemiology

Filarial nematodes, like other neglected tropical diseases, mainly affect rural, poorer populations in tropical and subtropical areas of the world (Engels and Zhou, 2020). In 2021, 72 countries were endemic for LF, with 885 million people at risk (World Health Organization, 2022b). Due to annual mass drug administration (MDA), the number of individuals with LF decreased from 199 million in 2000 to 51 million in 2018 (Cromwell et al., 2020). In 2018, the majority of infected individuals (71.5%) lived in Southeast Asia, but prevalence estimates were also high in Central Africa and Coastal West

Africa (Cromwell et al., 2020). LF causes acute adeno-lymphangitis characterized by fever and inflammation of lymph nodes, as well as painful swellings of affected body areas (Palumbo, 2008). About one third of infected people develop chronic filarial disease, with the resulting lymphatic vessel damage leading to lymphedema, elephantiasis of limbs and hydrocele in males (Taylor et al., 2010). In 2021, onchocerciasis was endemic in 30 countries in sub-Saharan Africa, South America and the Arabic Peninsula with 245 million people at risk, over 99% of the burden was in 26 African countries (World Health Organization, 2022a). In 2008, 25.7 million people were infected with *O. volvulus*, 4.2 million people suffered from severe itching and one million people had impaired vision or blindness (World Health, 2010).

In ten Central and West African countries, 14.4 million people live in high risk areas and 15.2 million people in intermediate risk areas for loiasis, as estimated from the reported frequency of eye worm history (Zouré et al., 2011). More than 10 million people are estimated to be infected with *L. loa* (Metzger and Mordmüller, 2014). At community level, 40% prevalence of eye worm history corresponds to 20% prevalence of microfilariae, 5% prevalence of high intensity *L. loa* infection and 2% prevalence of very high intensity *L. loa* infections (Takougang et al., 2002). High *L. loa* microfilaraemia has been associated with increased risk of death and the population-attributable fraction of mortality for loiasis was 14.5% (Chesnais et al., 2017). Infections with adult *L. loa* but no detectable microfilariae in the blood stream are common, making diagnosis difficult (Dupont et al., 1988). Ivermectin treatment for onchocerciasis control programs has been associated with encephalopathy in people with high *L. loa* microfilarial loads (Gardon et al., 1997, Boussinesq et al., 2003). To prevent serious adverse events, a test-and-not-treat strategy has been evaluated in Cameroon, where 95.5% of participants could be treated with ivermectin and only 2.4% were excluded due to *L. loa* microfilarial loads above the risk threshold (20,000 microfilariae per millilitre of blood) (Kamgno et al., 2017). *L. loa* microfilaraemia prevalence and intensity are higher in males than in females and increase with age (Whittaker et al., 2018). Mansonellosis is widely distributed in Africa, Central and South America and the Caribbean and is the most neglected human filariasis despite being likely the most prevalent (Ta-Tang et al., 2018). In 2007, an estimated 114 million people were infected with *M. perstans* in 33 sub-Saharan African countries with 581 million inhabitants (Simonsen et al., 2011). Most infections with *M. perstans* are asymptomatic and no specific clinical picture has been associated with mansonellosis (Simonsen et al., 2011). Transient swellings, pruritis fever, joint pain and eosinophilia have been observed in *M. perstans* infections (Simonsen et al., 2014). A potential new *Mansonella* species, termed *Mansonella* sp. "DEUX" was detected in febrile Gabonese children (Mourembou et al., 2015). In two endemic communities in Uganda, *M. perstans* microfilaraemia

prevalence increased rapidly during childhood and remained above 50% for adults, with higher prevalence in men than women and increasing microfilaria intensity with age (Asio et al., 2009).

1.4.3. Diagnosis

L. loa shows a diurnal periodicity, while *W. bancrofti*, *B. malayi* and *B. timori* show a nocturnal periodicity and *M. perstans* does not show periodicity of microfilariae in peripheral blood (Knopp et al., 2012). Therefore blood collection for diagnosis should be carried out at peak levels of microfilaraemia for the species investigated (Alhassan et al., 2015). Filtration of peripheral blood using 3- μ m or 5- μ m pore membrane for concentration of microfilariae, subsequent staining and light microscopy is the standard diagnostic tool for filarial detection (Moody and Chiodini, 2000). Species differentiation is done morphologically by size and the absence or presence of a sheath and the nuclei position in the microfilariae (Moody and Chiodini, 2000). Two filarial antigen tests have been developed that detect circulating *W. bancrofti* antigen from serum or fingerprick blood. They are easy to perform and do not require sampling at night (Weil et al., 1997, Weil et al., 2013). The *W. bancrofti* immunochromatographic card test was shown to be false-positive in some individuals with high *L. loa* microfilaraemia (Bakajika et al., 2014). Lateral flow strips detecting antibodies to both *W. bancrofti* and *O. volvulus* antigens can be used to assess past and present infections and are more useful for post-MDA surveillance in children (Steel et al., 2015). PCR and qPCR based assays exist for sensitive and specific detection and differentiation of filarial species, but are more expensive, have a longer turnaround time and are restricted to well-equipped laboratories (Alhassan et al., 2015).

1.5. Next-generation sequencing technologies

Clinical microbiology laboratories use next-generation sequencing (NGS) for three main applications, these are whole genome sequencing (WGS), targeted sequencing of distinct genes and metagenomics sequencing (Mitchell and Simner, 2019). A number of second- and third-generation sequencing platforms exist that use different methods and each have their own advantages and disadvantages (Hilt and Ferrieri, 2022).

Second-generation sequencing platforms rely on amplification of fragmented DNA and subsequent sequencing by either ligation using ligases or sequencing by synthesis using polymerases (Metzker, 2010). The single-nucleotide addition approach works by adding each of the four nucleotides iteratively and measuring whether each nucleotide was incorporated, homopolymer regions are indicated by an increase of signal strength (Goodwin et al., 2016). The first NGS instrument was called 454 pyrosequencing where the incorporation of a base triggered an enzymatic cascade that resulted in a bioluminescence signal (Margulies et al., 2005). The single-nucleotide addition approach used by Ion Torrent is based on measuring pH changes, as the incorporation of one base leads to the release

of a single H⁺ ion (Rothberg et al., 2011). The cyclic reversible termination sequencing approach used by Illumina works by incorporating one fluorescent nucleotide in every cycle, imaging and cleavage of the fluorescent dye (Metzker, 2010). Second-generation sequencing platforms have limited read lengths of up to a few hundred base pairs and sequencing run lengths of hours to days (Goodwin et al., 2016). Cyclic reversible termination sequencing has error rates of 0.1% with mainly substitutions, while single-nucleotide addition sequencing has error rates of 1%, mainly due to insertions and deletions in homopolymer regions (Goodwin et al., 2016).

Third generation sequencing platforms rely on single-molecule sequencing technologies, they have longer read lengths but suffer from lower raw read accuracy compared to second generation sequencing platforms (Schadt et al., 2010). Pacific Biosciences (PacBio) uses a technology based on an array of zero-mode wavelength nanostructures (Levene et al., 2003). The single-molecule real-time sequencing approach measures light emitted from labelled nucleotides incorporated by a single DNA polymerase in a zero-mode wavelength nanostructure (Eid et al., 2009). The template is a single-stranded circular DNA that can be sequenced multiple times by the same polymerase to improve the accuracy (Rhoads and Au, 2015). High-fidelity PacBio sequencing was able to generate reads with 99.8% accuracy and a mean length of 13.5 kilobases to assemble a human genome with 28x coverage using 89 Gigabases of sequencing data (Wenger et al., 2019).

Nanopore sequencing technology is based on protein nanopores that are embedded in a membrane. (Clarke et al., 2009). Voltage is applied and changes in ionic current are measured as single-stranded nucleic acids pass through the nanopore (Clarke et al., 2009). Figure 5 shows the approach developed by Oxford Nanopore Technologies (ONT). A motor protein controls translocation speed and unwinds double-stranded DNA or DNA-RNA duplexes (Cherf et al., 2012). Continuous engineering and refinement of the nanopore and motor protein by ONT improved read length, yield and sequencing quality (Wang et al., 2021b). Basecalling algorithms rely on machine learning and have improved with each iteration (Rang et al., 2018). Recently, it was shown that species-specific basecalling models improve sequencing accuracy, as lineages differ in their methylation motifs and patterns (Ferguson et al., 2022). The only limitation to read length in ONT sequencing is the length of DNA molecules, read lengths of one million bases have been reported (Jain et al., 2018). Direct RNA sequencing is also possible, but error rates are high, as was shown with Influenza virus RNA sequencing, where read level accuracy was 85% and consensus sequence accuracy was 98.5% (Keller et al., 2018). DNA methylation patterns can be identified from raw ONT data and have been shown useful for epigenetic studies (Lee et al., 2020). The latest flow cell R10.4.1 had a modal read accuracy of 98.1% and did not require polishing with short reads for assembly of near-complete bacterial genome (Sereika et al., 2022).

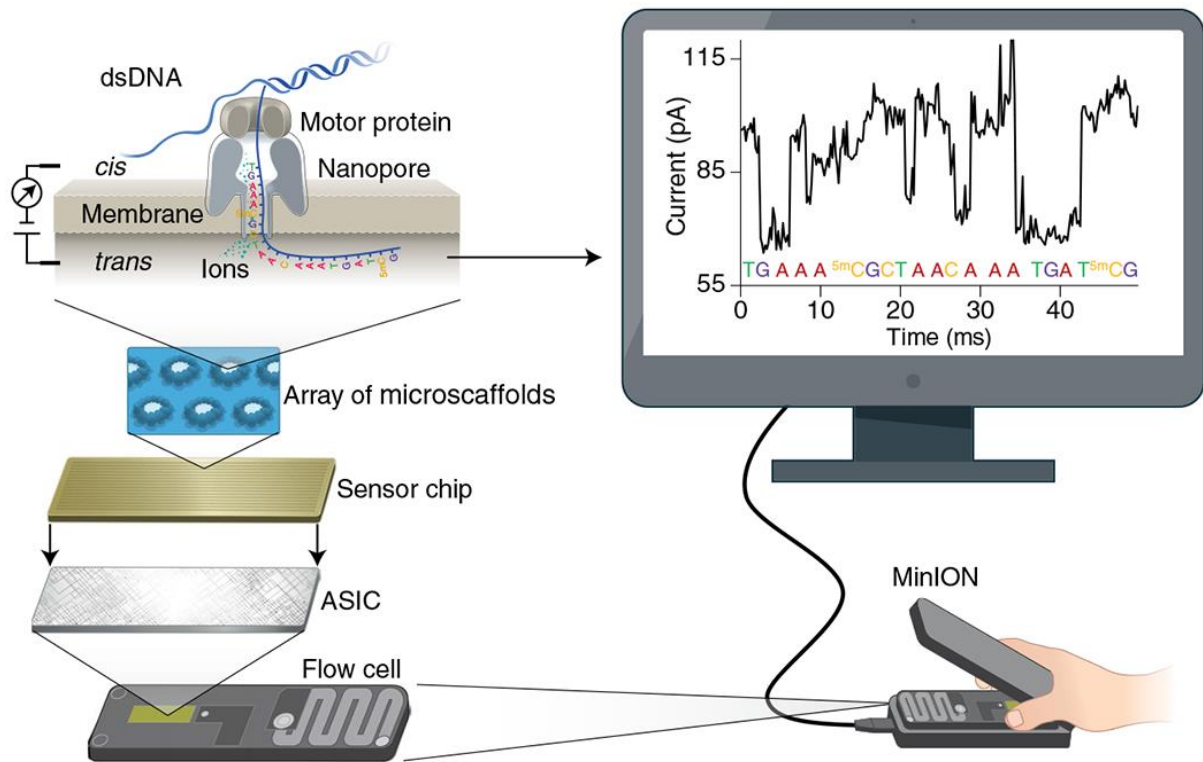


Figure 5. Principle of the ONT MinION sequencer. An ONT MinION flow cell contains 512 channels with four nanopores in each channel, resulting in a total of 2,048 nanopores. Nanopores are embedded in an electrically resistant polymer membrane in arrays of microcaffolds that are connected to a sensor chip. Each channel is controlled by the application-specific integration circuit (ASIC). A constant voltage is applied across the membrane, with the *trans* side being positively charged and ionic current passes through the nanopore. A motor protein unwinds the double-stranded DNA molecule. The negatively charged single-stranded DNA ratchets through the nanopore driven by the voltage. A characteristic current change is measured as the nucleotides pass through the nanopore. Figure taken from Wang *et al.* (Wang *et al.*, 2021b).

1.5.1. Whole genome sequencing in microbiology

Clinical specimen often contain multiple bacterial species and normal, physiological body flora, potentially confounding WGS results and antimicrobial resistance (AMR) prediction (Mitchell and Simner, 2019). Culture and isolation of the organism of interest before DNA extraction is required for WGS of most bacteria, however severe limitations exist for organisms that are currently difficult or impossible to culture *in vitro* (Mitchell and Simner, 2019). Traditional molecular biology typing techniques such as multilocus sequence typing (MLST) are well established, cheap and fast, however they lack discriminatory power to differentiate closely related pathogens (Fournier *et al.*, 2007). Epidemiological typing based on WGS can detect the transmission of health-care associated

pathogens, identify sources of infection for food-borne pathogens and track transmission chains (Schürch and Siezen, 2010). In contrast, the value and application of WGS in antimicrobial susceptibility testing is limited due to the high sensitivity, robustness and low price of phenotypic susceptibility testing and the limited link between phenotype and genotype (Köser et al., 2012). The presence of resistance genes and mutations can predict resistance to some antibiotics, but if the genetic basis of resistance is unknown, more complex statistical models are required (Su et al., 2019). Machine learning algorithms can predict complex AMR patterns from WGS data (Anahtar et al., 2021). For slow growing bacteria like the *Mycobacterium tuberculosis* complex, phenotypic tests take weeks, but resistance genes and mutations are well known, showing high concordance between WGS and phenotypic resistance testing (Quan et al., 2018). WGS is able to detect drug resistant subpopulations in viral infections, as was shown in Human Immunodeficiency Virus (HIV) where low-frequency drug resistance mutations undetectable by Sanger sequencing could be identified with NGS (Tzou et al., 2018).

1.5.2. Targeted sequencing of 16S ribosomal RNA gene

The enrichment or selection of organisms of interest from complex samples is most commonly done by PCR amplification of distinct marker genes for bacteria and fungi and by probe enrichment for viruses (Hilt and Ferrieri, 2022). For bacterial identification, the 16S ribosomal RNA (rRNA) gene is amplified, while for fungal identification ribosomal ITS or 28S rRNA genes are amplified and used for amplicon deep sequencing (Janda and Abbott, 2007, Wagner et al., 2018). Identification of difficult to culture bacteria is facilitated with 16S rRNA sequencing, which is able to detect species in an agnostic manner, as was shown for tick-borne bacteria such as *Borrelia burgdorferi* or *Rickettsia rickettsii* (Kingry et al., 2020). While Illumina sequencing generates more accurate reads, it is unable to sequence the full length 16S rRNA gene, whereas ONT sequencing can cover the full length 16S rRNA gene but has higher error rates. Both technologies are able to accurately classify the organism in question to the genus level, but have limited discriminatory power at the species level (Winand et al., 2020). A recent study has shown the improved accuracy to the species level assignment and better replicability of ONT for 16S rRNA sequencing compared to Illumina and therefore recommends ONT for microbiome studies (Szoboszlay et al., 2023).

1.5.3. Metagenomics sequencing

Metagenomics NGS allows unbiased detection of pathogen nucleic acids directly from clinical specimens (Gu et al., 2019). It can simultaneously detect nucleic acids from viruses, bacteria, fungi and parasites as well as host biomarkers (Mitchell and Simner, 2019). The vast majority of nucleic acids in clinical specimen derive from the host, limiting the number of nonhuman reads in direct metagenomics sequencing (Gu et al., 2019). Different methods exist for the depletion of host nucleic

acids (Hasan et al., 2016, He et al., 2010). The main advantage of metagenomics sequencing is its ability to identify a wide range of pathogens and even discover novel pathogens without the need for cultivation or prior knowledge (Chiu, 2013). Metagenomics ONT sequencing was able to detect pathogens from body fluids of critically ill patients within 6 hours, while Illumina results were available after 23 hours (Gu et al., 2021). A combination of metagenomics sequencing and conventional testing including serology improved diagnosis of neurologic infections, led to adjustment of treatment and enhanced patient care (Wilson et al., 2019).

1.6. One Health surveillance of bacterial pathogens and antimicrobial resistance

Microbes circulating in animal reservoirs are more likely to spill over to other wildlife, livestock or humans, causing zoonosis, if large-scale environmental, agricultural, or demographic shifts occur (Morse et al., 2012). The evolutionary transformation from a pathogen exclusively infecting animals to a pathogen exclusively infecting humans has been defined by five stages with increasing transmission to and between humans in each stage (Wolfe et al., 2007). Many of the major infectious diseases are not human-exclusive pathogens and the animal origin of some pathogens is still unknown (Wolfe et al., 2007). AMR is driven by different factors, the highest impact arising from antimicrobial misuse or overuse in humans and animals (Holmes et al., 2016a). AMR spreads between humans, animals and the environment through different direct and indirect pathways (Woolhouse and Ward, 2013). As the majority of antimicrobial classes are used in animal and human sectors, it is important to minimize antimicrobial use for prophylaxis and avoid use for growth promotion in animal production systems to prevent AMR development (Collignon and McEwen, 2019). Additionally, sanitation and hygiene need to be improved to prevent the spread of AMR once developed (Collignon and McEwen, 2019). One Health approaches collaborate between academic disciplines and society to improve health at the human–animal–environment interface, such as surveillance of known and novel zoonotic diseases, AMR, food safety and climate change (Zinsstag et al., 2023).

1.6.1. Overview of *Campylobacter jejuni*

Campylobacter jejuni is a motile, spiral-shaped Gram-negative bacterium that belongs to the family *Campylobacteraceae* (Debruyne et al., 2008). The genome consists of a single circular chromosome with a size of 1.6 Megabases and a genomic guanine-cytosine (GC) content of 30.6% (Parkhill et al., 2000). *C. jejuni* colonize the gut mucosa in the jejunum and ileum (Blaser and Engberg, 2008). Almost 20 *Campylobacter* species have been shown to infect humans, the majority of infections are caused by *C. jejuni* (Man, 2011). Two subspecies of *C. jejuni* exist, *C. jejuni* subsp. *jejuni* and *C. jejuni* subsp. *doiley* (Man, 2011). Diarrhoea, abdominal pain, myalgia, vomiting and blood in faeces are symptoms of *Campylobacter* enteritis (Blaser and Engberg, 2008). Long-term complications include reactive

arthritis, Guillain-Barre´ syndrome and postinfectious irritable bowel syndrome (Blaser and Engberg, 2008). A low infection dose is sufficient for *Campylobacter* enteritis and the incubation period ranges from one to seven days with a mean of three days (Blaser and Engberg, 2008). Diarrhoeal disease is usually self-limiting, but immunocompromised and at risk individuals receive treatment with fluoroquinolones or macrolides to prevent progression to severe infections (Dai et al., 2020). *Campylobacter* have developed resistance to these first-line antibiotics, requiring novel treatment options and preventive interventions (Dai et al., 2020). *Campylobacter* outbreaks in the United States were most commonly caused by foodborne transmission and to a lesser degree by waterborne transmission or animal contact, while human to human transmission was uncommon (Taylor et al., 2013). In the United States, the majority of foodborne outbreaks were caused by dairy products, while undercooked poultry was responsible for most outbreaks in Europe (Greig and Ravel, 2009). *Campylobacter* detection is done through bacterial culture and biochemical testing or molecular methods such as genus- or species-specific PCR or 16S rRNA sequencing (Kaakoush et al., 2015). In 2010, *Campylobacter* spp. were the most common cause of foodborne illness, accounting for 95.6 million cases of diarrheal disease and 21,000 deaths globally (Havelaar et al., 2015).

1.6.2. Overview of *Salmonella enterica*

Salmonella enterica is a rod-shaped Gram-negative facultative intracellular bacterium that belongs to the family *Enterobacteriaceae* (Knodler and Effenbein, 2019). Depending on the serotype, *S. enterica* invades the intestinal epithelium in the ileum and colon after ingestion and causes gastroenteritis or disseminates to systemic sites and causes sepsis (Knodler and Effenbein, 2019). The genome consists of a single circular chromosome with a size of 4-5 Megabases (Knodler and Effenbein, 2019). Over 1,500 serotypes of *S. enterica* subsp. *enterica* exist, that are classified by antigenically diverse surface antigens (Grimont and Weill, 2007).

Salmonella serotypes Typhi, Paratyphi A, B and C cause enteric fever, a systemic disease with unspecific symptoms such as fever, flu-like and abdominal symptoms (Harris and Brooks, 2020). *Salmonella* Typhi is restricted to the human host, shedding of bacteria in stool or urine can be temporary or chronic, transmission occurs through the faecal-oral route, with contaminated water or food as the main source of infection (Crump, 2019). An estimated 14.3 million cases and 135,900 deaths occurred globally in 2017, a 41% decline in deaths from 1990 (Stanaway et al., 2019b). Incidence rates were highest among children, case fatality rates were higher among children and older people in lower-income countries and South Asia accounted for almost 70% of deaths (Stanaway et al., 2019b). The isolation of *S. enterica* from bone-marrow culture is considered the gold standard and is more sensitive than blood culture (Gilman et al., 1975). Antibody and antigen tests have limited specificity and sensitivity and are cross-reactive with other infections (Parry et al., 2011). Enteric fever

vaccines, access to clean water and hygiene measures during food preparation are effective at preventing *S. enterica* infections (Qamar et al., 2022). Antibiotic treatment depends on the local AMR profile, disease severity and available resources (Qamar et al., 2022).

Non-typhoidal *S. enterica* typically cause acute self-limiting gastroenteritis. Diarrhoea is usually non-bloody and may be accompanied by fever, nausea, vomiting and abdominal cramping (Pegues and Miller, 2015). Non-typhoidal *S. enterica* have human and animal hosts, transmission is mostly foodborne, with poultry being the most common source (Pegues and Miller, 2015). Isolation of *S. enterica* from stool confirms diagnosis (Haeusler and Curtis, 2013). Non-typhoidal *S. enterica* accounted for 78.7 million cases of diarrheal disease and 59,000 foodborne deaths in 2010 (Havelaar et al., 2015). Additionally, non-typhoidal *S. enterica* can cause bacteraemia, leading to a febrile illness termed invasive non-typhoidal salmonella (iNTS) disease (Crump et al., 2015). An estimated 535,000 cases of iNTS disease occurred worldwide in 2017, sub-Saharan Africa accounted for 78.8% of all cases (Stanaway et al., 2019a). In 2017, iNTS was responsible for 77,500 deaths, with higher case fatality rates among children under the age of 5 years, elderly people and HIV positive people and in areas of low sociodemographic development (Stanaway et al., 2019a). Risk factors for iNTS disease include HIV, malaria, acute malnutrition and sickle-cell disease (Crump et al., 2015).

1.7. Aims of this thesis

The overall goal of this thesis was to develop and apply novel molecular diagnostic and bioinformatics analysis tools with potential for decentralized and mobile testing for the surveillance of fever causing viruses, bacteria and parasites in Central Africa and the Middle East.

In part one, the aim was to describe the evolution of SARS-CoV-2 in Equatorial Guinea using molecular tools with three objectives:

- i) Develop a sensitive, robust and rapid diagnostic tool for the identification of SARS-CoV-2 mutations associated with variants of concern
- ii) Describe the epidemiology of SARS-CoV-2 using whole genome sequencing
- iii) Understand SARS-CoV-2 epidemiology over time in context of the African continent

In part two, the aim was to improve the molecular monitoring of malaria and filarial parasites that evade diagnosis or treatment in Equatorial Guinea with the following objectives:

- i) Assess the sensitivity and specificity of malaria rapid diagnostic tests collected in the framework of a malaria indicator survey
- ii) Determine the epidemiology of filarial nematode and malaria co-infections
- iii) Evaluate a novel qPCR platform for rapid and sensitive detection of malaria suitable for decentralized testing
- iv) Detect minority clones carrying drug resistance conferring mutations in mixed *Plasmodium falciparum* infections

In part three the aim was to deploy ONT MinION for whole genome sequencing of environmental and foodborne bacterial pathogens by using a One Health approach for surveillance of foodborne pathogens in Palestine.

2. Genomic surveillance of SARS-CoV-2 in Equatorial Guinea

This chapter contains the following publications:

Bechtold *et al.* **Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable *peak*PCR Platform.** *Analytical Chemistry* 2021

Dudas *et al.* **Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions.** *Nature Communications* 2021

Hosch *et al.* **Genomic Surveillance Enables the Identification of Co-infections With Multiple SARS-CoV-2 Lineages in Equatorial Guinea.** *Frontiers in Public Health* 2022

Tegally *et al.* **The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance.** *Science* 2022

2.1. Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable peakPCR Platform

Published in Analytical Chemistry, 2021

Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable *peakPCR* Platform

Philippe Bechtold, Philipp Wagner, Salome Hosch, Denise Siegrist, Amalia Ruiz-Serrano, Michele Gregorini, Maxmillian Mpina, Florentino Abaga Ondó, Justino Obama, Mitoha Ondo'o Ayekaba, Olivier Engler, Wendelin J. Stark, Claudia A. Daubenberger, and Tobias Schindler*



Cite This: *Anal. Chem.* 2021, 93, 16350–16359



Read Online

ACCESS |



Metrics & More

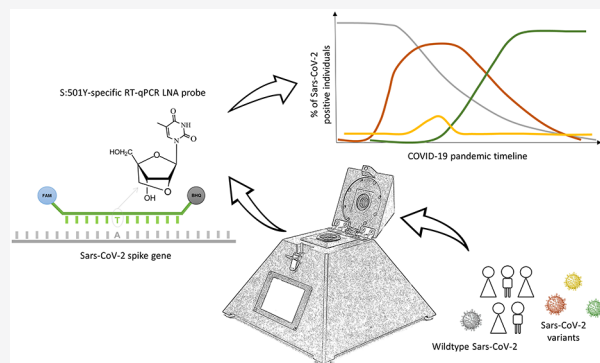


Article Recommendations



Supporting Information

ABSTRACT: The need for tools that facilitate rapid detection and continuous monitoring of SARS-CoV-2 variants of concern (VOCs) is greater than ever, as these variants are more transmissible and therefore increase the pressure of COVID-19 on healthcare systems. To address this demand, we aimed at developing and evaluating a robust and fast diagnostic approach for the identification of SARS-CoV-2 VOC-associated spike gene mutations. Our diagnostic assays detect the E484K and N501Y single-nucleotide polymorphisms (SNPs) as well as a spike gene deletion (HV69/70) and can be run on standard laboratory equipment or on the portable rapid diagnostic technology platform *peakPCR*. The assays achieved excellent diagnostic performance when tested with RNA extracted from culture-derived SARS-CoV-2 VOC lineages and clinical samples collected in Equatorial Guinea, Central-West Africa. Simplicity of usage and the relatively low cost are advantages that make our approach well suitable for decentralized and rapid testing, especially in resource-limited settings.



INTRODUCTION

More than a year after the World Health Organization (WHO) declared the severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) outbreak a Public Health Emergency of International Concern, coronavirus disease 2019 (COVID-19) has caused more than 3.7 million deaths.¹ Globally, public health systems are severely impacted and are further challenged by the emergence of SARS-CoV-2 variants carrying mutations that are of concern (VOC).² Molecular diagnostic tools, particularly reverse transcription-quantitative polymerase chain reaction (RT-qPCR) for viral RNA detection and next-generation sequencing (NGS) for molecular monitoring SARS-CoV-2 genetic diversity at the whole genome level, have proven critical for public health decision-making.³ Investigating SARS-CoV-2 genomes by NGS to track transmission chains, understand transmission dynamics, and rapidly identify mutations that potentially have an impact on transmissibility, morbidity, and mortality, as well as potential escape of diagnostic tools or vaccine-induced immunity have become an integral part of public health measures during this pandemic.⁴

Since the first whole genome sequence (WGS) analysis of SARS-CoV-2 has been published in January 2020,⁵ the virus has been continuously sequenced, characterized, and data made publicly available through global initiatives such as

Global Initiative on Sharing All Influenza Data (GISAID). More than 1.8 million SARS-CoV-2 sequences have been publicly shared via GISAID, and numerous mutations in the gene encoding the spike protein have been identified.⁶ For example, the D614G variant has been shown to increase the viral load of infected patients and has replaced the original variant since June 2020 around the globe.⁷ More recently, SARS-CoV-2 lineages characterized by a combination of multiple mutations in the spike gene have emerged independently in different regions of the world. The SARS-CoV-2 lineages B.1.1.7 (also known as the Alpha variant, VOC 202012/01 or 501Y.V1), B.1.351 (also known as the Beta variant or 501Y.V2), and P.1 (also known as the Gamma variant, B.1.1.28.1 or 501Y.V3) were the first VOCs identified.⁸ The Alpha variant (lineage B.1.1.7) was first described in mid-December 2020 in the United Kingdom, and the mutation appears to have substantially increased transmissibility and has

Received: June 4, 2021

Accepted: September 29, 2021

Published: December 1, 2021

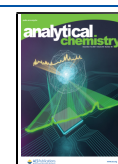


Table 1. Primer and Probe Combinations Developed for SARS-CoV-2 VOC Identification and Discrimination

assay	primer name	oligo sequence (5'-3')	modifications
HV69/70 assay: 21765–21770 ^a deletion in alpha VOC	HV69/70_F	TCA ACT CAG GAC TTG TTC TTA CCT	
	HV69/70_R	TGG TAG GAC AGG GTT ATC AAA C	
	Δ69/70 (mut)	TCC ATG CTA TCT CTG GGA CCA	FAM - BHQ1
	HV69/70 (wt)	ACA TGT CTC TGG GAC CAA TGG	YYE ^b - BHQ1
N501Y assay: A23063T ^a SNP in alpha, beta and gamma VOCs	spike_gene_LNA_F	C+TA TCA GGC +CGG TAG CAC +AC	
	spike_gene_LNA_R	+AGT ACT ACT ACT CTG TAT +GGT TGG +T	
	501Y_LNA (mut)	C+CC A+CT +t+AT G+GT +G	FAM - BHQ1
	N501_LNA (wt)	C+CC A+CT +A+AT G+GT +G	YYE ^b - BHQ1
E484K assay: G23011A ^a SNP in beta and gamma VOCs	Spike_gene_LNA_F	C+TA TCA GGC +CGG TAG CAC +AC	
	Spike_gene_LNA_R	+AGT ACT ACT ACT CTG TAT +GGT TGG +T	
	484K_LNA (mut)	TGG +T+GT TaA A+GG T	FAM - BHQ1
	E484_LNA (wt)	TGG +T+GT TGA A+GG T	YYE ^b - BHQ1

^aGenome position according to MN908947.3 (SARS-CoV-2 isolate Wuhan Hu-1). LNA nucleotides are indicated with + in front of the nucleotide. The SNP associated with VOCs is indicated as lower case and bold nucleotides. ^bYYE = Yakima Yellow (VIC and HEX dye alternative).

quickly developed into the dominant variant circulating in the UK and beyond.⁹ The Beta variant (lineage B.1.351) was identified in December 2020, which emerged most likely in South Africa and is also associated with higher transmissibility.¹⁰ The Gamma variant (lineage P.1) was identified in January 2021 in Manaus, the largest city in the Amazon region of Brazil.¹¹ In January 2021, this region experienced a resurgence of COVID-19 despite the reported high seroprevalence of antibodies against SARS-CoV-2 in this population.^{12,13} During the preparation and revision of this manuscript, the WHO had designated the emerging lineage B.1.617.2 as the Delta VOC.

To rapidly detect and continuously monitor the appearance, introduction, and spread of (novel) VOCs, the level of molecular surveillance needs to be increased globally. The gold standard of genomic surveillance, NGS, allows for unbiased identification of mutations, but is limited by its relatively slow sample-to-result turnaround time and level of laboratory infrastructure and scientific expertise required. Furthermore, the relatively high costs of NGS increase the financial burden on establishing a widespread VOC-tracking strategy, a limiting factor particularly for resource-limited settings. Therefore, mutation-specific PCR-based approaches that are more cost-efficient and allow for high-throughput screening of a significant proportion of SARS-CoV-2-positive individuals were developed.^{14,15} To identify the transmission dynamics of VOCs in settings with limited sequencing capabilities, we have designed rapid and cost-efficient RT-qPCR assays detecting relevant mutations in the spike protein of SARS-CoV-2. The N501Y mutation is found in the Alpha, Beta, and Gamma VOCs, while the E484K mutation is restricted to Beta and Gamma variants. The Alpha variant is characterized by an additional spike gene deletion (ΔHV69/70). To further simplify, decentralize, and speed up the process of VOC identification, we transformed our assay to a portable and inexpensive qPCR device, named *peakPCR*.¹⁶ The device can complete up to 20 RT-qPCR reactions in less than 40 min. Important characteristics of *peakPCR* are the relatively low cost of production and the simplicity of usage. By using cartridges that are preloaded with lyophilized RT-qPCR reagents, the user interaction is reduced to loading the sample onto the cartridge. Furthermore, the preloaded *peakPCR* cartridges can easily be shipped and stored at room temperature, making cold chains superfluous. Here, we report the development of a new approach for rapid, robust, and

decentralized identification of SARS-CoV-2 VOCs that can be both run on standard laboratory RT-qPCR equipment and the portable and rapid diagnostic technology platform *peakPCR*.

■ MATERIALS AND METHODS

SARS-CoV-2 Cell Culture Supernatants and Clinical Samples for Assay Evaluation.

The cultivation of SARS-CoV-2 was carried out in a Biosafety level 3 laboratory and conducted under appropriate safety conditions. Three different VOC lineages of SARS-CoV-2, namely, B.1.1.7, B.1.351, and P.1 provided from the University Hospital of Geneva, Laboratory of Virology, were grown on VeroE6/TMPRSS2 cells obtained from the Centre for AIDS Reagents (National Institute for Biological Standards and Control).^{17,18} The day before infection, VeroE6/TMPRSS2 cells were seeded at 2×10^6 cells per T75 flask in Dulbecco's modified Eagle medium (DMEM) (Seraglob, Switzerland) supplemented with 10% fetal bovine serum (FBS) (Merck, Germany) and 2% SEeticin (Seraglob, Switzerland). On the day of infection, the cells reached about 70–90% confluency. The growth medium was removed and replaced with 5 mL of infection medium (DMEM + 2% FBS + 2% SEeticin). Cells were inoculated with 70 μL of SARS-CoV-2 swab material and incubated for 1 h at 37 °C, 5% CO₂, and >85% humidity. After adsorption, 10 mL of the infection medium was added to each flask. Cells were observed for cytopathic effects (CPE) for 3–6 days using an EVOSTM FL digital inverted microscope. When CPE reached 40–100%, the supernatant was collected, cleared from cell debris by centrifugation (10 min at 500 g), and samples were aliquoted and frozen. TCID₅₀ was determined on VeroE6/TMPRSS2 cells. Virus was inactivated with Qiazol, and RNA was extracted with RNeasy Plus Universal Mini Kit (Qiagen, Germany).

As a positive control and for initial assay evaluation, a 1869 bp long synthetic SARS-CoV-2 spike gene fragment (genome position 21,557–23,434 bp), based on the sequence of B.1.1.7, was synthesized (Eurofin Genomics, Ebersberg, Germany; the sequence is provided in Figure S1). Using a serial dilution of the synthetic spike gene, a calibration curve ranging from 0.05 to 50,000,000 copies/μL was prepared (data provided in Figure S2). The initial viral copy number per μL (cp/μL) of the cell culture-derived RNA from SARS-CoV-2 was estimated using the calibration curve's y -intercept and its slope. Serial dilutions of the RNA extracted from the culture supernatants of SARS-CoV-2 isolates Wuhan Hu-1, B.1.1.7, B.1.351, and

P.1, ranging from 0.1 to 1,000,000 cp/ μ L, was prepared and used to evaluate the assays' performance on both RT-qPCR platforms. Additionally, a SARS-CoV-2 RT-qPCR diagnostic assay, targeting the envelope (E) gene, published by the Institute of Virology at Charité (Berlin, Germany), was used as a positive control for viral RNA on both platforms.¹⁹

Designing HV69/70-Deletion-, E484K-, and N501Y-Specific RT-qPCR Assays. We developed assays targeting the HV69/70-deletion, the E484K-, and N501Y-single-nucleotide polymorphisms (SNPs). For standard RT-qPCR platforms like the Bio-Rad CFX96 device, multiplex assays were developed. The multiplex assays are able to detect both sequence variations, the wildtype and mutated, in a single RT-qPCR reaction. For the rapid identification of mutations of interest for mobile and rapid RT-qPCR platforms, such as the *peak*PCR device, only the mutated sequence variation is detected, and no multiplex amplification is performed. SNP discrimination was enhanced by using primers and probes containing locked nucleic acids (LNAs). Sequence analysis and primer design were performed using the Geneious Prime 2021.0.3 software. All oligos, including the LNAs, were synthesized at Microsynth AG (Balgach, Switzerland), and details are provided in Table 1.

SARS-CoV-2 HV69/70-, E484K-, and N501Y-Specific RT-qPCR Assays. The HV69/70-, E484K-, and N501Y-specific assays were performed using the Bio-Rad CFX96 real-time PCR System (Bio-Rad Laboratories, California, USA). A RT-qPCR run was completed within 1 h and 10 min using the following thermal profile: reverse transcription step at 50 °C for 5 min; polymerase activation at 95 °C for 20 s; and 45 cycles of 3 s at 95 °C and 30 s at 61 °C. Each reaction consisted of 2 μ L of RNA and 8 μ L of reaction master mix containing 1 \times TaqMan Fast Virus 1-Step Master Mix (Thermo Fisher Scientific, Leiden, The Netherlands) and the corresponding 1 \times primer/probe mixture consisting of 0.4 μ M primers and 0.2 μ M probes. All RT-qPCR assays were run in duplicates with appropriate controls. The mutated sequences were detected by Fluorescein (FAM)-labeled probes and the wildtype sequences by Yakima yellow (YYE)-labeled probes in multiplex reactions. Data analysis of the RT-qPCR data was conducted using the CFX Maestro Software (Bio-Rad Laboratories, California, USA). RT-qPCR amplification efficiencies were calculated based on the slope of the standard curve, as described elsewhere.²⁰

The HV69/70-, E484K-, and N501Y-specific RT-qPCR assays were transferred to the *peak*PCR platform (Diaggio AG, Zurich, Switzerland) on which FAM-labeled probes detected the mutated sequence variations only. In order to simplify the testing procedure for the user, the *peak*PCR aluminum sample holders (herein referred to as cartridges) were preloaded with all necessary reagents in the freeze-dried form. Lyophilized cartridges were loaded with 4.4 μ L of sample, sealed off with 1.2 mL of paraffin oil (Sigma-Aldrich, Germany), and run on the *peak*PCR device using the following program: reverse transcription step at 50 °C for 5 min; initial denaturation at 95 °C for 60 s; and 45 cycles of 6 s at 95 °C and 30 s at 62 °C. The total runtime of a *peak*PCR experiment was 37 min. *Peak*PCR data were analyzed using the *peak*PCR dataAnalysis 1.0 software (Diaggio AG, Zurich, Switzerland). No drop in performance was observed when lyophilized reagents were used compared to nonlyophilized standard RT-qPCR reagents (Figure S3).

Evaluation of Diagnostic Performance with Clinical Samples. Clinical evaluation was conducted using RNA extracted from SARS-CoV-2-positive samples collected in Equatorial Guinea. Sample collection and analysis was done as part of a research collaboration with the Equato-Guinean Ministry of Health and Social Welfare and was enabled by several presidential emergency decrees. All patient data were fully anonymized, and publication was approved by the National Technical Committee for the Response and Monitoring of the Novel Coronavirus (Comité Técnico Nacional de Respuesta y Vigilancia del Nuevo Coronavirus), which is charged with preventing, containing, controlling, tracking, and evaluating the development and evolution of COVID-19 in Equatorial Guinea.

MinION SARS-CoV-2 Whole Genome Sequencing. A total of 59 SARS-CoV-2-positive samples from Equatorial Guinea were selected for reconfirmation using WGS by MinION (Oxford Nanopore Technologies, Oxford, UK) according to the open-source ARTIC protocol (<https://artic.network/ncov-2019>). Sample preparation for MinION sequencing was based on the ARTIC Network nCoV-2019 sequencing protocol v2²¹ and v3.²² The RNA samples were diluted in nuclease-free water according to their cycle threshold value in the diagnostic RT-qPCR. (C_q <15: 1:100 dilution, C_q 15–18: 1:10 dilution) for cDNA synthesis, for which either SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, USA) or LunaScript RT SuperMix (New England BioLabs, USA) was used with random hexamer primers. A total of 218 primer pairs covering the whole virus genome were used for PCR amplification.²³ The ligation sequencing kit (Oxford Nanopore Technologies, UK) was used for library preparation. Sequencing was conducted on a FLO-MIN106 (R9.4.1) flow cell. Base calling was performed in real time on a MinION Mk1c using MinKNOW version 20.10.6. The ARTIC Network bioinformatics protocol was followed for data analysis.²⁴ Consensus sequences were generated with the Wuhan Hu-1 isolate (GenBank accession number MN908947.3) as a reference sequence. Variants were called using Nanopolish and Medaka. Lineage assignment was done using the pangolin tool.²⁵ All sequences are deposited in GISAID.

RESULTS

Design of RT-qPCR Assays for the Rapid Identification of SARS-CoV-2 VOCs. Three mutation-specific RT-qPCR assays based on TaqMan chemistry were designed. The first assay targets the 6 bp deletion in the spike gene, leading to the loss of two amino acids at positions 69 and 70 within the spike protein (HV69/70 assay). This deletion is found in the Alpha VOC, but not in Beta or Gamma VOCs. Universal primers amplify a 102-bp (wildtype) or 96-bp (mutant) amplicon. Based on the presence or absence of the deletion, either a FAM-labeled probe or YYE-labeled probe binds, and the resulting fluorescence is detected. The second assay targets a nonsynonymous SNP in the spike gene (A23011G), leading to an amino-acid exchange at positions 484 (E484K). The E484K mutation is only present in Beta and Gamma VOCs. In a multiplex reaction, a YYE-labeled probe detects the wildtype and a FAM-labeled probe the mutated sequence. The third assay targets a nonsynonymous SNP in the spike gene (A23063T), leading to an amino-acid exchange at positions 501 (N501Y). The N501Y mutation is present in Alpha, Beta, and Gamma VOCs. Similar to the E484K assay, a YYE-labeled

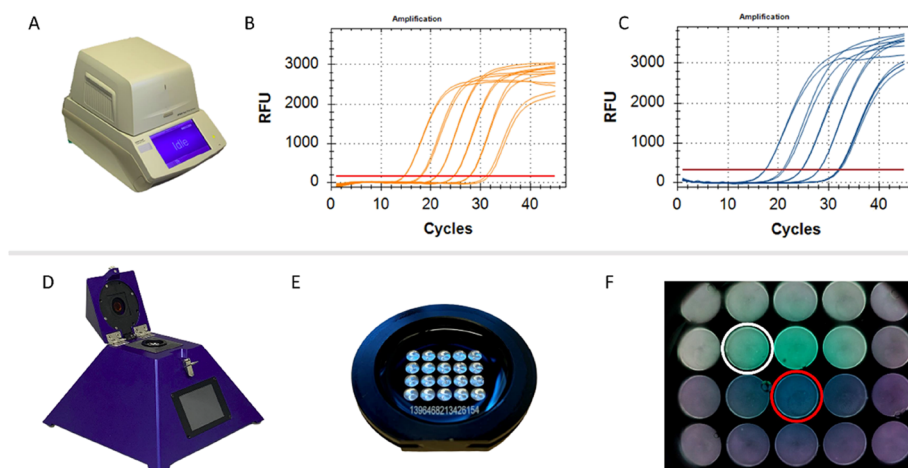


Figure 1. RT-qPCR platforms for SARS-CoV-2 VOC identification. (A) Standard RT-qPCR device Bio-Rad CFX96. (B) Detection of N501Y-wt in serial dilution of Wuhan Hu-1 lineage, ranging from 1 to 10,000 cp/ μ L using the YYE channel of the Bio-Rad CFX96 instrument. (C) Detection of N501Y-mut in serial dilution of P.1 lineage, ranging from 1 to 1,000,000 cp/ μ L using the FAM channel of the Bio-Rad CFX96 instrument. (D) Portable and rapid diagnostic platform *peakPCR*. (E) Ready-to-use cartridges with preloaded lyophilized RT-qPCR reagents. (F) Photograph after cycle 45, detecting fluorescence in each well with a CCD sensor. The depicted well marked with a white circle contains a positive signal after RT-qPCR amplification, while the negative sample, marked with a red circle, did not display amplification of a PCR product.

probe detects the wildtype and a FAM-labeled probe the mutated sequence. A summary of the oligonucleotide sequences used is provided in Table 1. Wildtype sequences are defined as the nucleotide sequences of the original Wuhan Hu-1 isolate published.²⁶

The novel assays were run on two different RT-qPCR platforms in parallel. On the Bio-Rad CFX96 platform (Figure 1A), three sequence-discriminatory assays were run as duplex assays, detecting the wildtype sequence in the YYE channel (Figure 1B) and the mutated sequence in the FAM channel (Figure 1C). As a second technology platform, the *peakPCR* device was selected (Figure 1D), which is a portable and rapid diagnostic technology platform running the RT-qPCR reaction on ready-to-use cartridges (Figure 1E). Fluorescence is detected using the Raspberry Pi Camera Module V2 as an inexpensive charge-coupled detector (CCD) sensor (Figure 1F). For the *peakPCR* device, the multiplex assays were reduced to mutation-specific assays, capable of detecting the mutated sequences only.

Analytical Performance of HV69/70, E484K, and N501Y Assays Using Well-Characterized RNA from SARS-CoV-2 VOCs. Four SARS-CoV-2 lineages, namely, Wuhan Hu-1 (wildtype, non VOC), B.1.1.7 (Alpha VOC), B.1.351 (Beta VOC), and P.1 (Gamma VOC), were used to assess the RT-qPCR efficiency, specificity, and sensitivity of novel mutation-specific assays. We used serial dilutions, ranging from 0.1 to 1,000,000 cp/ μ L, of cell culture-derived viral RNA for assay characterization. The presence and quantity of RNA molecules in these serial dilutions were confirmed by monitoring the pan-Sarbecovirus E-gene amplification (Figure 2A). Mutation-specific assays for HV69/70 (Figure 2B), E484K (Figure 2C), and N501Y (Figure 2D) were run on both platforms, while wildtype-specific assays were solely run on the Bio-Rad CFX96 platform.

The data provided in Figure 2 were used to obtain the RT-qPCR amplification efficiencies, sensitivities, and specificities shown in Figure 3. High amplification efficiencies (>90%) were obtained for the E-gene, the HV69/70 and E484K assays on the Bio-Rad CFX96 platform, and for the E-gene and E484K assays on the *peakPCR* device. All other assays achieved

amplification efficiencies >80%, which is considered moderate (Figure 3A). The analytical sensitivity of the assays was defined as the lowest viral RNA concentration at which mutations are identified in >80% of replicates. We used the detection rate among all four SARS-CoV-2 lineages to identify the limit of detection (LOD) (Figure 3B). For the E-gene, HV69/70, and E484K assay, a detection rate of 100% was achieved at a viral RNA concentration as low as 10 cp/ μ L. At the same concentration for the N501Y assay, 5 out of 6 replicates (83%) were amplified. At the LOD of 10 cp/ μ L (dashed line in Figure 3B), no difference between the two RT-qPCR platforms in terms of sensitivity was observed. Viral RNA concentrations below 10 cp/ μ L cannot be detected, with the exception of the HV69/70 assay run on the Bio-Rad CFX96 device, where 1 cp/ μ L is still reliably detected.

The specificity of all three assays and their ability to distinguish between mutated and wildtype sequences were assessed by testing the assays with RNA from SARS-CoV-2 cell culture supernatants. On both platforms, no signal was observed at any viral RNA concentration if there was not a perfect sequence match of the oligos to the nucleotide sequence to be detected, resulting in a 100% analytical specificity. At a viral RNA concentration of 10,000 cp/ μ L, the HV69/70, E484K, and N501Y genotypes were all correctly identified among wildtype, Alpha, Beta, and Gamma SARS-CoV-2 strains (Figure 3C). The mutation-specific probe of the HV69/70 assay gave a signal only when run with RNA of the Alpha VOC carrying the mutation. The E484K-mutation assay did not result in amplification when run on RNA from wildtype non-VOC Wuhan Hu-1 lineage and Alpha VOC. The N501Y-mutation assay correctly detected all VOCs but not the wildtype non-VOC Wuhan Hu-1 lineage. In summary, the three assays correctly identify lineage-associated mutations with moderate to high RT-qPCR efficiencies in samples with more than 10 cp/ μ L of SARS-CoV-2 RNA. We also demonstrated that these assays can be successfully conducted on the rapid diagnostic platform *peakPCR*, and the performance in terms of sensitivity and specificity does not significantly differ between these two RT-qPCR platforms.

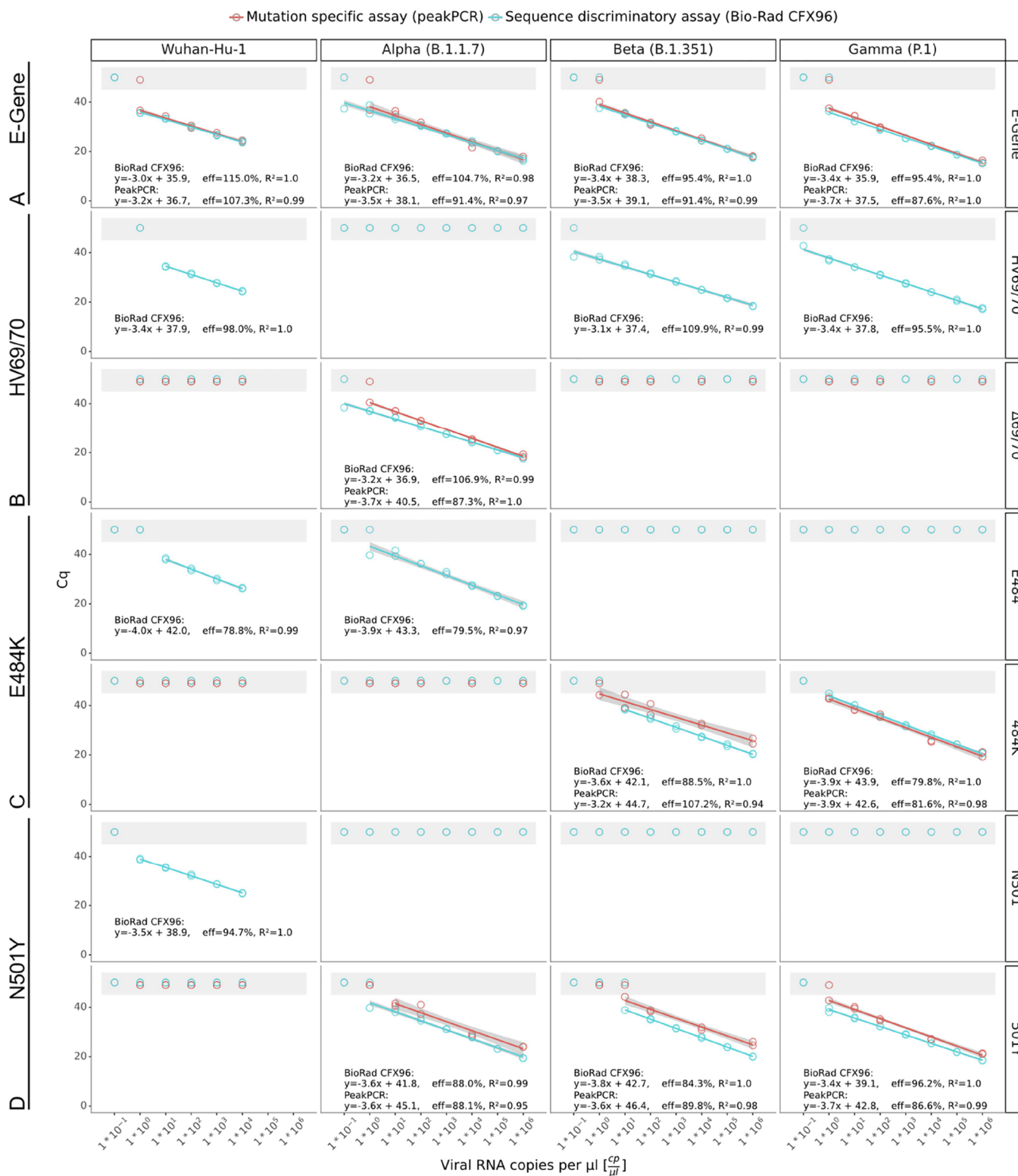


Figure 2. RT-qPCR performance of novel SARS-CoV-2 mutation-specific assays. (A) SARS-CoV-2 E-gene reference assay, (B) HV69/70 assay, (C) E484K assay, and (D) N501Y assay. Each circle represents a technical replicate. Mutation-specific assays were run on both platforms, while wildtype-specific assays were only run on the Bio-Rad CFX96 platform. For the Wuhan Hu-1 lineage, the two highest RNA concentrations of 1,000,000 and 100,000 cp/μL were not available. Tests for performance on the *peakPCR* device used the 1,000,000, 10,000, 100, 10, and 1 cp/μL concentrations. The Cq values for samples without amplification are set arbitrarily to 46 for the *peakPCR* and to 47 for Bio-Rad CFX96 devices. Data points within the gray area are considered negative (Cq values >45).

Clinical Performance of HV69/70, E484K, and N501Y Detecting RT-qPCR Assays. We used 59 clinical samples positive for SARS-CoV-2 collected in Equatorial Guinea

between November 2020 and March 2021 for further evaluation of all three mutation-specific RT-qPCR assays. The outcome of the RT-qPCR assays was compared to

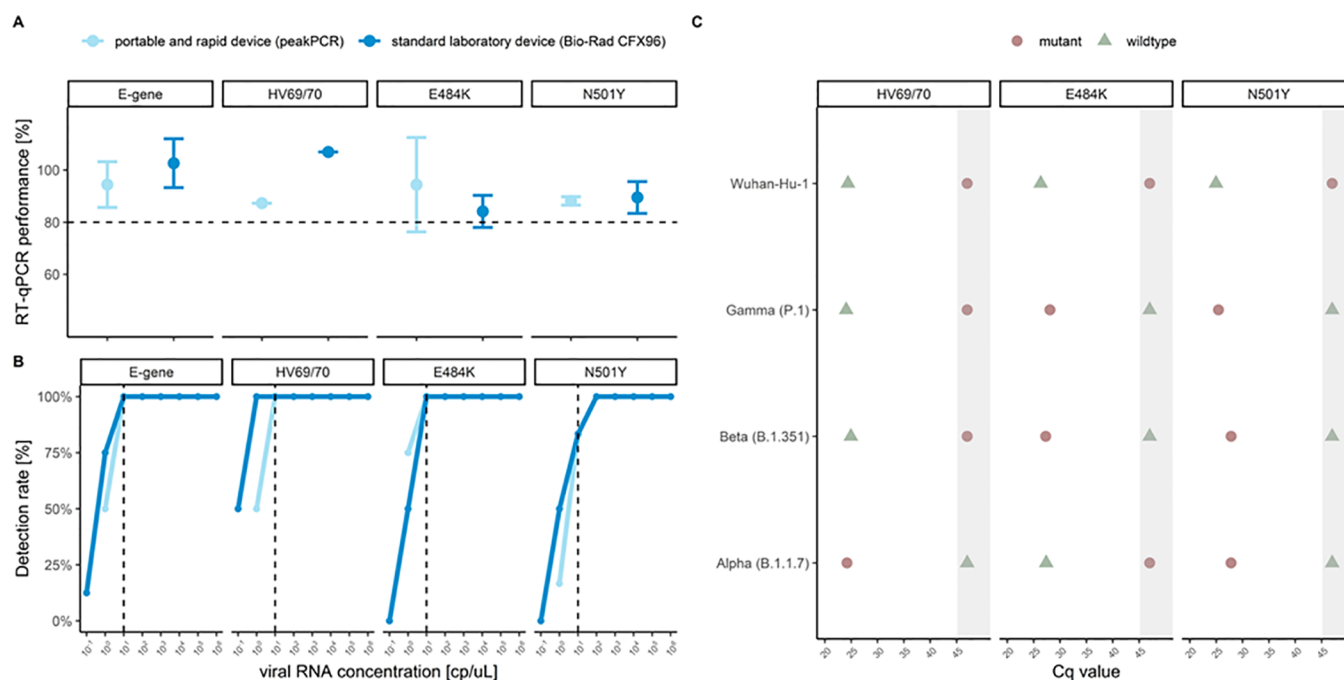


Figure 3. Analytical performance of HV69/70, E484K, and N501Y detecting RT-qPCR assays. (A) RT-qPCR amplification efficiency of the E-gene, HV69/70, E484K, and N501Y assays as determined by serial dilutions of RNA derived from four cell culture supernatant SARS-CoV-2 lineages. Amplification efficiencies >80% are considered moderate (dashed lines). (B) Analytical sensitivity represented by detection rates calculated from all replicates for each viral RNA concentration. The LOD was defined as the lowest concentration at which >80% of replicates were amplified. The dashed line represents the LOD of 10 copies per μL . (C) Analytical specificity for the multiplex sequence-discrimination assays run on the Bio-Rad CFX96 device. The data shown are based on RT-qPCR amplification for viral RNA concentrations of 10,000 cp/ μL . Data points within the gray area are considered negative (Cq values >45).

Table 2. Performance Evaluation of HV69/70, E484K, and N501Y Detecting RT-qPCR Assays Using Clinical Samples

SARS-CoV-2 lineage	Mutation profile ^a	n	HV69/70	$\Delta 69/70$	E484	484K	N501	501Y
wild type ^b	HV69/70, E484, N501	14	14/14	0/14	14/14	0/14	14/14	0/14
alpha (B.1.1.7)	$\Delta 69/70$, E484, 501Y	1	0/1	1/1	1/1	0/1	0/1	1/1
beta (B.1.351)	HV69/70, 484K, 501Y	43	43/43	0/43	0/43	43/43	0/43	43/43
B.1.620	$\Delta 69/70$, 484K, N501	1	0/1	1/1	0/1	1/1	0/1	0/1

^aBased on SARS-CoV-2 WGS. ^bIncludes the following SARS-CoV-2 lineages: B.1, B.1.1, B.1.177, B.1.192, B.1.36.10, B.1.535, B.1.596, B.1.623.

Nanopore MinION-based SARS-CoV-2 WGS data obtained from the same clinical samples (Table 2). The HV69/70 assay identified the spike gene deletion $\Delta 69/70$ correctly in 2 out of 2 samples, while for all other samples, in accordance with WGS, the wildtype HV69/70 genotype was found. The E484K assay accurately identified the 484K SNP in 43 samples with the Beta VOC and in one sample assigned to the B.1.620 lineage. The N501Y assay correctly identified the 501Y mutation in one confirmed Alpha VOC sample and 43 confirmed Beta VOC samples. In summary, the evaluation with clinical samples resulted in a 100% agreement between the novel mutation-specific RT-qPCR assays and SARS-CoV-2 WGS.

Investigating the Introduction and Spread of SARS-CoV-2 Beta VOC in Equatorial Guinea Using Mutation-Specific RT-qPCR Assays. In total, we analyzed 184 SARS-CoV-2-positive samples from Equatorial Guinea using all three mutation-specific RT-qPCR assays collected from November 2020 to March 2021 (Figure 4A). While between November and December 2020, all samples were wildtype for the three spike gene mutations associated with SARS-CoV-2 VOCs, starting from January 2021, more than 85% (102/119) of samples carried the 484K + 501Y mutant combination. The

WGS analysis of a subset of these samples revealed an expansion of the SARS-CoV-2 Beta VOC (lineage B.1.351) in Equatorial Guinea (Figure 4B). Other combinations of mutations of interest were also found, the sample with $\Delta 69/70$ + 501Y was identified as the Alpha variant (B.1.1.7 lineage), and the sample with $\Delta 69/70$ + 484K was identified as the B.1.620 lineage. In summary, these RT-qPCR-based assays enable rapid and cost-effective genotyping of larger numbers of clinical samples, resulting in a more accurate reflection of SARS-CoV-2 epidemiology and their local transmission dynamics.

DISCUSSION

The COVID-19 pandemic is a continuous, unprecedented, global public health crisis with severe economic and social consequences.²⁷ More than 1 year into the pandemic, the emergence of VOCs starts to pose again a serious threat to contain the virus. Rapid and reliable identification of SARS-CoV-2 variants is a critical component of public health interventions to mitigate the further spread of VOCs that might undermine the performance of diagnostic tests and vaccine-induced immunity against this virus.²⁸

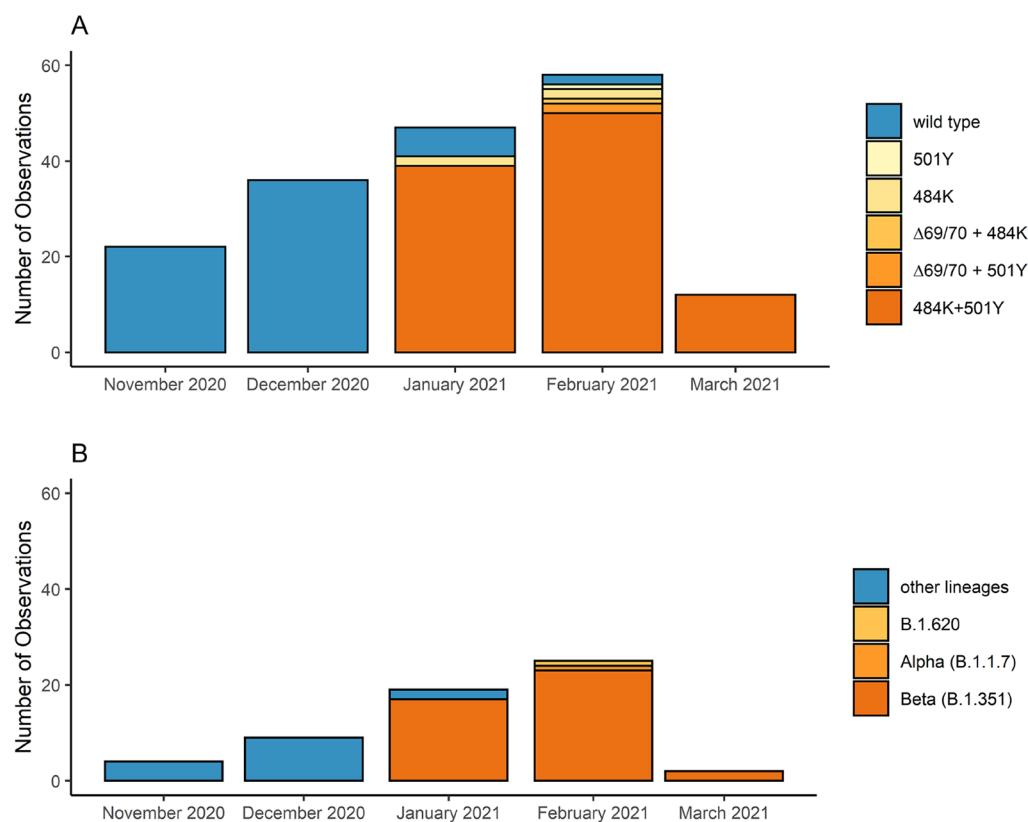


Figure 4. Rapid detection of SARS-CoV-2 VOC-associated mutations using HV69/70, E484K, and N501Y RT-qPCR assays. (A) Identification of spike gene mutations using HV69/70, E484K, and N501Y RT-qPCR assays in 184 clinical samples collected in Equatorial Guinea from November 2020 to March 2021. (B) Identification of SARS-CoV-2 lineages using Nanopore MinION SARS-CoV-2 WGS in 59 clinical samples collected in Equatorial Guinea from November 2020 to March 2021.

We designed, tested, and validated three mutation-specific RT-qPCR assays, detecting the E484K and N501Y SNPs as well as the 6-bp deletion affecting HV69/70, all located in the SARS-CoV-2 spike gene. All assays can be performed under standard RT-qPCR conditions, simplifying integration into existing laboratory environments, and the assays proved to be highly sensitive, specific, and reproducible. We demonstrated the usefulness of such screening assays to rapidly identify potential VOCs using clinical samples. Using our assays, we were able to observe the introduction and spread of the B.1.351 lineage in Equatorial Guinea. It is noteworthy that it took less than 4 weeks for the B.1.351 lineage to become the dominant lineage in this cohort. Using more than one VOC marker enabled us to identify SARS-CoV-2 variants with an unusual combination of mutations, such as the E484K plus the HV69/70 deletion. This sample was later assigned to the SARS-CoV-2 lineage B.1.620, which is most likely of Central African origin and has recently been described in several countries in Europe.²⁹

The approach presented here is well suited for cost-effective, robust, and high-throughput screening of large cohorts. Mutation-specific RT-qPCRs are not intended to replace NGS, but rather complement and extend molecular surveillance programs and focal outbreak monitoring. During the preparation of this manuscript, the WHO had designated the lineage B.1.617.2 as the fourth VOC. The Delta variant (lineage B.1.617.2) was first documented in India, where it has contributed to the surge in cases and has now been detected across the globe.³⁰ Based on the perfect mismatch discrimination of LNA-based assays and the simplicity of this type of

assay in both design and implementation will allow for rapid adaptation of our approach to B.1.617.2 and also to newly emerging VOCs identified in future.

A similar LNA-based approach for the detection of N501Y and HV69/70 has been successfully tested in Canada on 2430 samples.³¹ The results of their in-house assay were concordant with the commercial assay VirSNiP SARS-CoV-2 (TIB Molbiol, Berlin, Germany), which is based on melting-curve analysis. This underlines the possibility of using LNA-based RT-qPCR assays for single-nucleotide discrimination as opposed to using melting-curve analysis. Although the SARS-CoV-2 SNP genotyping by melting-curve analysis is widely used,³² the LNA-based approach has several advantages: it is faster, easier to integrate into existing laboratory workflows, and could be combined with a diagnostic assay, allowing immediate genotyping. Furthermore, LNA-based sequence-discriminatory assays are better suited to identify the presence of more than one lineage of SARS-CoV-2 in a single sample, a phenomenon which was recently observed in Brazil.³³

To reduce the sample-to-result turnaround time in routine, decentralized testing settings, our strategy included the transfer of these assays to a portable, robust, and rapid diagnostic RT-qPCR platform. Starting from extracted RNA, the *peakPCR* platform completed sample analysis in 37 min, which is half of the time required to run the same assay on a standard RT-qPCR platform, while retaining comparable efficiency, specificity, and sensitivity. For the first time, we show that complex RT-qPCR-based genotyping assays can be transferred to the rapid and portable *peakPCR* platform. Apart from the speed and simplicity of usage, the relatively low costs of

equipment and reagents make this platform interesting. The *peakPCR* device costs are an estimated \$2500, a fraction of the price at which commercially available qPCR devices are being sold, and because of the lower reaction volume used in *peakPCR* experiments (4 μ L), the cost per preloaded and freeze-dried 20-well cartridge is kept similar to the reagent costs for a standard RT-qPCR reaction. Future technological developments will focus on simplified sample-preprocessing strategies to replace the RNA extraction step completely. This would allow the placement of this molecular diagnostic platform in peripheral health care settings. Furthermore, the deployment of cartridges preloaded with lyophilized reagents independent from the cold chain for reagent supply is a significant technological advantage that makes this platform well suited for decentralized, rapid molecular testing of infectious diseases, particularly in resource-limited settings.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.1c02368>.

A 1869 bp long synthetic SARS-CoV-2 spike gene fragment based on the sequence of the SARS-CoV-2 Alpha VOC; serial dilution experiments for HV69/70 and NS01Y assays using the synthetic gene fragment of the SARS-CoV-2 Alpha VOC; and comparison of lyophilized reagents with nonlyophilized standard RT-qPCR reagents on the *peakPCR* device. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Tobias Schindler – Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland; University of Basel, 4051 Basel, Switzerland; orcid.org/0000-0002-5961-095X; Email: tobias.schindler@swisstph.ch

Authors

Philippe Bechtold – Institute for Chemical and Bioengineering, ETH Zurich, 8093 Zuerich, Switzerland; Diaxxo AG, 8093 Zuerich, Switzerland
Philipp Wagner – Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland; University of Basel, 4051 Basel, Switzerland
Salome Hosch – Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland; University of Basel, 4051 Basel, Switzerland
Denise Siegrist – Spiez Laboratory, 3700 Spiez, Switzerland
Amalia Ruiz-Serrano – Institute for Chemical and Bioengineering, ETH Zurich, 8093 Zuerich, Switzerland; Diaxxo AG, 8093 Zuerich, Switzerland
Michele Gregorini – Institute for Chemical and Bioengineering, ETH Zurich, 8093 Zuerich, Switzerland; Diaxxo AG, 8093 Zuerich, Switzerland
Maxmillian Mpina – Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland; University of Basel, 4051 Basel, Switzerland; Laboratorio de Investigaciones de Baney, Baney, Equatorial Guinea
Florentino Abaga Ondó – Ministry of Health and Social Welfare, Malabo, Equatorial Guinea
Justino Obama – Ministry of Health and Social Welfare, Malabo, Equatorial Guinea

Mitoha Ondo'o Ayekaba – Ministry of Health and Social Welfare, Malabo, Equatorial Guinea
Olivier Engler – Spiez Laboratory, 3700 Spiez, Switzerland
Wendelin J. Stark – Institute for Chemical and Bioengineering, ETH Zurich, 8093 Zuerich, Switzerland; Diaxxo AG, 8093 Zuerich, Switzerland; orcid.org/0000-0002-8957-7687
Claudia A. Daubenberger – Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland; University of Basel, 4051 Basel, Switzerland

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.analchem.1c02368>

Author Contributions

Conceptualization by S.H., P.B., P.W., and T.S. *peakPCR* platform technology and cartridge design: P.B., A.R.S., M.G., W.J.S. Formal analysis was done by P.W. and T.S. D.S. and O.E. provided the cell culture-derived RNA from SARS-CoV-2 VOCs. Clinical evaluation was conducted by M.P., F.A.O., J.O., and M.O.A. The first draft of the manuscript was written by P.B., C.A.D., and T.S. All authors reviewed and approved the final manuscript. P.B. and P.W. contributed equally to this study.

Funding

Funding for P.B. and M.G. was provided by the Botnar Research Centre for Child Health as part of the Fast Track Call for Acute Global Health Challenges as well as the BRIDGE program by the Swiss National Science Foundation and Innosuisse.

Notes

The authors declare the following competing financial interest(s): P.B., M.G., and W.J.S. are co-inventors on a corresponding patent application around the detection platform and share-holders of the ETH spin-off company Diaxxo AG. T.S. has consulted for Diaxxo AG. All other authors declare no competing interests.

■ REFERENCES

- (1) Dong, E.; Du, H.; Gardner, L. *Lancet Infect. Dis.* **2020**, *5*, 533–534.
- (2) Walensky, R. P.; Walke, H. T.; Fauci, A. S. *JAMA* **2021**, *325*, 1037.
- (3) Oude Munnink, B. B.; Nieuwenhuijse, D. F.; Stein, M.; O'Toole, A.; Haverkate, M.; Mollers, M.; Kamga, S. K.; Schapendonk, C.; Pronk, M.; Lexmond, P.; van der Linden, A.; Bestebroer, T.; Chestakova, I.; Overmars, R. J.; van Nieuwkoop, S.; Molenkamp, R.; van der Eijk, A. A.; GeurtsvanKessel, C.; Vennema, H.; Meijer, A.; Rambaut, A.; van Dissel, J.; Sikkema, R. S.; Timen, A.; Koopmans, M.; The Dutch-Covid-19 response team; Oudehous, G. J. A. P. M.; Schinkel, J.; Kluytmans, J.; Kluytmans-van den Bergh, M.; van den Bijllaardt, W.; Berntvelsen, R. G.; van Rijen, M. M. L.; Schneeberger, P.; Pas, S.; Diederer, B. M.; Bergmans, A. M. C.; van der Eijk, P. A. V.; Verweij, J. J.; Buiting, A. G. N.; Streefkerk, R.; Aldenkamp, A. P.; de Man, P.; Koelemal, J. G. M.; Ong, D.; Paltansing, S.; Veassen, N.; Slevin, J.; Bakker, L.; Brockhoff, H.; Rietveld, A.; Slijkerman Megelink, F.; Cohen Stuart, J.; de Vries, A.; van der Reijden, W.; Ros, A.; Lodder, E.; Verspui-van der Eijk, E.; Huijskens, I.; Kraan, E. M.; van der Linden, M. P. M.; Debast, S. B.; Naiemi, N. A.; Kroes, A. C. M.; Damen, M.; Dinant, S.; Lekkerkerk, S.; Pontesilli, O.; Smit, P.; van Tienen, C.; Godschalk, P. C. R.; van Pelt, J.; Ott, A.; van der Weijden, C.; Wertheim, H.; Rahamat-Langendoen, J.; Reimerink, J.; Bodewes, R.; Duizer, E.; van der Veer, B.; Reusken, C.; Lutgens, S.; Schneeberger, P.; Hermans, M.; Wever, P.; Leenders, A.; ter Waarbeek, H.; Hoebe, C. *Nat. Med.* **2020**, *26*, 1405–1410.

(4) World Health Organization. *Genomic Sequencing of SARS-CoV-2: A Guide to Implementation for Maximum Impact on Public Health*, 8 January 2021; World Health Organization, 2021.

(5) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. *N. Engl. J. Med.* **2020**, *382*, 727–733.

(6) Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; Qin, H.; Wang, M.; Lu, Q.; Li, X.; Sun, Q.; Liu, J.; Zhang, L.; Li, X.; Huang, W.; Wang, Y. *Cell* **2020**, *182*, 1284–1294.e9.

(7) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; McDanal, C.; Perez, L. G.; Tang, H.; Moon-Walker, A.; Whelan, S. P.; LaBranche, C. C.; Sapphire, E. O.; Montefiori, D. C.; Angyal, A.; Brown, R. L.; Carrilero, L.; Green, L. R.; Groves, D. C.; Johnson, K. J.; Keeley, A. J.; Lindsey, B. B.; Parsons, P. J.; Raza, M.; Rowland-Jones, S.; Smith, N.; Tucker, R. M.; Wang, D.; Wyles, M. D. *Cell* **2020**, *182*, 812–827.e19.

(8) Eurosurveillance Editorial Team. *Eurosurveillance* **2021**, *26*, No. 2101211.

(9) Horby, P.; Huntley, C.; Davies, N.; Edmunds, J.; Ferguson, N.; Medley, G.; Semple, C. SAGE Meeting paper 2021/01/21 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961037/NERVTAG_note_on_B.1.1.7_severity_for_SAGE_77__1_.pdf (accessed February 16, 2021).

(10) Tegally, H.; Wilkinson, E.; Giovanetti, M.; Iranzadeh, A.; Fonseca, V.; Giandhari, J.; Doolabh, D.; Pillay, S.; San, E. J.; Msomi, N.; Mlisana, K.; von Gottberg, A.; Walaza, S.; Allam, M.; Ismail, A.; Mohale, T.; Glass, A. J.; Engelbrecht, S.; van Zyl, G.; Preiser, W.; Petruccione, F.; Sigal, A.; Hardie, D.; Marais, G.; Hsiao, M.; Korsman, S.; Davies, M. A.; Tyers, L.; Mudau, I.; York, D.; Maslo, C.; Goedhals, D.; Abrahams, S.; Laguda-Akingba, O.; Alisoltani-Dehkordi, A.; Godzik, A.; Wibmer, C. K.; Sewell, B. T.; Lourenço, J.; Alcántara, L. C. J.; Kosakovsky Pond, S. L.; Weaver, S.; Martin, D.; Lessells, R. J.; Bhiman, J. N.; Williamson, C.; de Oliveira, T. Emergence and Rapid Spread of a New Severe Acute Respiratory Syndrome-Related Coronavirus 2 (SARS-CoV-2) Lineage with Multiple Spike Mutations in South Africa. *medRxiv*, December 22, 2020, doi: DOI: 10.1101/2020.12.21.20248640.

(11) Faria, N. R.; Claro, I. M.; Candido, D.; Moyses Franco, L. A.; Andrade, P. S.; Coletti, T. M.; Silva, C. A. M.; Sales, F. C.; Manuli, E. R.; Aguiar, R. S.; Gaburo, N.; Camilo, C. D. C.; Fraijji, N. A.; Esashi, M. A.; Sabino, E. C.; on behalf of CADDE Genomic Network. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology - Virological <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586> (accessed February 16, 2021).

(12) Sabino, E. C.; Buss, L. F.; Carvalho, M. P. S.; Prete, C. A.; Crispim, M. A. E.; Fraijji, N. A.; Pereira, R. H. M.; Parag, K. V.; da Silva Peixoto, P.; Kraemer, M. U. G.; Oikawa, M. K.; Salomon, T.; Cucunuba, Z. M.; Castro, M. C.; de Souza Santos, A. A.; Nascimento, V. H.; Pereira, H. S.; Ferguson, N. M.; Pybus, O. G.; Kucharski, A.; Busch, M. P.; Dye, C.; Faria, N. R. *Lancet* **2021**, *397*, 452–455.

(13) Buss, L. F.; Prete, C. A.; Abraham, C. M. M.; Mendrone, A.; Salomon, T.; de Almeida-Neto, C.; França, R. F. O.; Belotti, M. C.; Carvalho, M. P. S. S.; Costa, A. G.; Crispim, M. A. E.; Ferreira, S. C.; Fraijji, N. A.; Gurzenda, S.; Whittaker, C.; Kamaura, L. T.; Takecian, P. L.; da Silva Peixoto, P.; Oikawa, M. K.; Nishiya, A. S.; Rocha, V.; Salles, N. A.; de Souza Santos, A. A.; da Silva, M. A.; Custer, B.; Parag, K. V.; Barral-Netto, M.; Kraemer, M. U. G.; Pereira, R. H. M.; Pybus, O. G.; Busch, M. P.; Castro, M. C.; Dye, C.; Nascimento, V. H.; Faria, N. R.; Sabino, E. C. *Science* **2021**, *371*, 288–292.

(14) Wang, H.; Jean, S.; Eltringham, R.; Madison, J.; Snyder, P.; Tu, H.; Jones, D. M.; Leber, A. L. *J. Clin. Microbiol.* **2021**, *59*, No. e0092621.

(15) Vogels, C. B. F.; Breban, M. I.; Ott, I. M.; Alpert, T.; Petrone, M. E.; Watkins, A. E.; Kalinich, C. C.; Earnest, R.; Rothman, J. E.; Goes de Jesus, J.; Morales Claro, I.; Magalhães Ferreira, G.; Crispim, M. A. E.; Brazil-UK CADDE Genomic Network; Singh, L.; Tegally, H.; Anyaneji, U. J.; Network for Genomic Surveillance in South Africa; Hodcroft, E. B.; Mason, C. E.; Khullar, G.; Metti, J.; Dudley, J. T.; MacKay, M. J.; Nash, M.; Wang, J.; Liu, C.; Hui, P.; Murphy, S.; Neal, C.; Laszlo, E.; Landry, M. L.; Muyombwe, A.; Downing, R.; Razeq, J.; de Oliveira, T.; Faria, N. R.; Sabino, E. C.; Neher, R. A.; Fauver, J. R.; Grubaugh, N. D. *PLoS Biol.* **2021**, *19*, No. e3001236.

(16) Gregorini, M.; Mikutis, G.; Grass, R. N.; Stark, W. J. *Ind. Eng. Chem. Res.* **2019**, *58*, 9665–9674.

(17) Matsuyama, S.; Nao, N.; Shirato, K.; Kawase, M.; Saito, S.; Takayama, I.; Nagata, N.; Sekizuka, T.; Katoh, H.; Kato, F.; Sakata, M.; Tahara, M.; Kutsuna, S.; Ohmagari, N.; Kuroda, M.; Suzuki, T.; Kageyama, T.; Takeda, M. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 7001–7003.

(18) Nao, N.; Sato, K.; Yamagishi, J.; Tahara, M.; Nakatsu, Y.; Seki, F.; Katoh, H.; Ohnuma, A.; Shirogane, Y.; Hayashi, M.; Suzuki, T.; Kikuta, H.; Nishimura, H.; Takeda, M. *PLoS One* **2019**, *14*, No. e0215822.

(19) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G.; Haagmans, B. L.; van der Veer, B.; van den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J.-L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P.; Drosten, C. *Eurosurveillance* **2020**, *25*, No. 2000045.

(20) Svec, D.; Tichopad, A.; Novosadova, V.; Pfaffl, M. W.; Kubista, M. *Biomol. Detect. Quantif.* **2015**, *3*, 9.

(21) nCoV-2019 sequencing protocol v2 (GunIt) https://www.protocols.io/view/ncov-2019-sequencing-protocol-v2-bdp7i5rn?version_warning=no (accessed April 6, 2021).

(22) nCoV-2019 sequencing protocol v3 (LoCost) <https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye> (accessed April 6, 2021).

(23) Tyson, J. R.; James, P.; Stoddart, D.; Sparks, N.; Wickenhagen, A.; Hall, G.; Choi, J. H.; Lapointe, H.; Kamelian, K.; Smith, A. D.; Prystajek, N.; Goodfellow, I.; Wilson, S. J.; Harrigan, R.; Snutch, T. P.; Loman, N. J.; Quick, J. Improvements to the ARTIC Multiplex PCR Method for SARS-CoV-2 Genome Sequencing Using Nanopore. *bioRxiv*, September 4, 2020, p. 1. doi: DOI: 10.1101/2020.09.04.283077.

(24) Artic Network nCoV2019-bioinformatics-sop <https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html> (accessed April 6, 2021).

(25) Rambaut, A.; Holmes, E. C.; O'Toole, Á.; Hill, V.; McCrone, J. T.; Ruis, C.; du Plessis, L.; Pybus, O. G. *Nat. Microbiol.* **2020**, *5*, 1403–1407.

(26) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. *Nature* **2020**, *579*, 265–269.

(27) (No Title) https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/briefingnote/wcms_767028.pdf (accessed June 3, 2021).

(28) Gupta, R. K. *Nat. Rev. Immunol.* **2021**, *21*, 340–341.

(29) Dudas, G.; Hong, S. L.; Potter, B.; Niatou-Singa, F. S.; Tombolomako, T. B.; Fuh-Neba, T.; Vickos, U.; Ulrich, M.; Khan, K.; Watts, A.; Olendrait, I.; Snijder, J.; Wijnant, K. N.; Bonvin, A. M. J. J. Travel-Driven Emergence and Spread of SARS-CoV-2 Lineage B.1.620 with Multiple VOC-like Mutations and Deletions in Europe 3. *medRxiv*, 2021, doi: DOI: 10.1101/2021.05.04.21256637.

(30) Threat Assessment Brief: Emergence of SARS-CoV-2 B.1.617 variants in India and situation in the EU/EEA <https://www.ecdc.europa.eu/en/publications-data/threat-assessment-emergence-sars-cov-2-b1617-variants> (accessed June 4, 2021).

(31) Matic, N.; Lowe, C. F.; Ritchie, G.; Stefanovic, A.; Lawson, T.; Jang, W.; Young, M.; Dong, W.; Brumme, Z. L.; Brumme, C. J.; Leung, V.; Romney, M. G. *Emerg. Infect. Dis.* **2021**, *27*, 1673.

(32) Haim-Boukobza, S.; Roquebert, B.; Trombert-Paolantoni, S.; Lecorche, E.; Verdurme, L.; Foulongne, V.; Selinger, C.; Michalakis, Y.; Sofonea, M. T.; Alizon, S. *Emerg. Infect. Dis.* **2021**, *27*, 1496.

(33) Francisco Jr, R. S.; Benites, L. F.; Lamarca, A. P.; de Almeida, L. G. P.; Hansen, A. W.; Gularte, J. S.; Demoliner, M.; Gerber, A. L.; de C Guimarães, A. P.; Antunes, A. K. E.; Heldt, F. H.; Mallmann, L.; Hermann, B.; Ziulkoski, A. L.; Goes, V.; Schallenberger, K.; Fillipi, M.; Pereira, F.; Weber, M. N.; de Almeida, P. R.; Fleck, J. D.; Vasconcelos, A. T. R.; Spilki, F. R. *Virus Res.* **2021**, *296*, No. 198345.

2.2. Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions

Published in Nature Communications, 2021

Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions

Distinct SARS-CoV-2 lineages, discovered through various genomic surveillance initiatives, have emerged during the pandemic following unprecedented reductions in worldwide human mobility. We here describe a SARS-CoV-2 lineage - designated B.1.620 - discovered in Lithuania and carrying many mutations and deletions in the spike protein shared with widespread variants of concern (VOCs), including E484K, S477N and deletions HV69 Δ , Y144 Δ , and LLA241/243 Δ . As well as documenting the suite of mutations this lineage carries, we also describe its potential to be resistant to neutralising antibodies, accompanying travel histories for a subset of European cases, evidence of local B.1.620 transmission in Europe with a focus on Lithuania, and significance of its prevalence in Central Africa owing to recent genome sequencing efforts there. We make a case for its likely Central African origin using advanced phylogeographic inference methodologies incorporating recorded travel histories of infected travellers.

Over a year into the pandemic and with an unprecedented reduction in human mobility worldwide, distinct SARS-CoV-2 lineages have arisen in multiple geographic areas around the world^{1–3}. New lineages are constantly appearing (and disappearing) all over the world and may be designated variant under investigation (VUI) if considered to have concerning epidemiological, immunological or pathogenic properties. So far, four lineages (i.e. B.1.1.7, B.1.351, P.1 and B.1.617.2 according to the Pango SARS-CoV-2 lineage nomenclature^{4,5}) have been universally categorised as variants of concern (VOCs), due to evidence of increased transmissibility, disease severity and/or possible reduced vaccine efficacy. An even broader category termed variant of interest (VOI) encompasses lineages that are suspected to have an altered phenotype implied by their mutation profile.

In some cases, a lineage may rise to high frequency in one location and seed others in its vicinity, such as lineage B.1.177 that became prevalent in Spain and was later spread across the rest of Europe². In others, reductions in human mobility, insufficient surveillance and passage of time allowed lineages to emerge and rise to high frequency in certain areas, as has happened with lineage A.23.1 in Uganda⁶, a pattern reminiscent of holdover H1N1 lineages discovered in West Africa years after the 2009 pandemic⁷. In the absence of routine genomic surveillance at their origin location, diverged lineages may still be observed as travel cases or transmission chains sparked by such in countries that do have sequencing programmes in place. A unique SARS-CoV-2 variant found in Iran early in the pandemic was characterised in this way⁸, and recently travellers returning from Tanzania were found to be infected with a lineage bearing multiple amino acid changes of concern⁹. As more countries launch their own SARS-CoV-2 sequencing programmes, introduced strains are easier to detect since they tend to be atypical of a host country's endemic SARS-CoV-2 diversity, particularly so when introduced lineages have accumulated genetic diversity not observed previously, a phenomenon that is characterised by long branches in phylogenetic trees. In Rwanda, this was exemplified by the detection of lineage B.1.380⁶, which was characteristic of Rwandan and Ugandan epidemics at the time. The same sequencing programme was then perfectly positioned to observe a sweep where B.1.380 was replaced by lineage A.23.1⁶, which was first detected in Uganda¹⁰, and to detect the country's first cases of B.1.1.7 and B.1.351. Similarly, sequencing programmes in Europe were witness to the rapid displacement of pan-European and endemic lineages with VOCs, primarily B.1.1.7 (e.g. Lyngse et al.¹¹).

Given the appearance of VOCs towards the end of 2020 and the continued detection of previously unobserved SARS-CoV-2 diversity, it stands to reason that more variants of interest (VOIs), and perhaps even VOCs, can and likely do circulate in areas of the world where access to genome sequencing is not available nor provided as a service by international organisations. Lineage A.23.1¹⁰ from Uganda and a provisionally designated variant of interest A.VOI.V2⁹ from Tanzania might represent the first detections of a much more diverse pool of variants circulating in Africa. We here describe a similar case in the form of a lineage designated B.1.620 that first caught our attention as a result of what was initially a small outbreak caused by a distinct and diverged lineage previously not detected in Lithuania, bearing multiple VOC-like mutations and deletions, many of which substantially alter the spike protein.

The first samples of B.1.620 in Lithuania were redirected to sequencing because they were flagged by occasional targeted PCR testing for SARS-CoV-2 spike protein mutation E484K repeated on PCR-positive samples. Starting April 2nd 2021, targeted E484K PCR confirmed a growing cluster of cases with this mutation in Anykščiai municipality in Utena county with a total

of 43 E484K⁺ cases out of 81 tested by April 28th (Supplementary Fig. S1). Up to this point, the Lithuanian genomic surveillance programme had sequenced over 10% of PCR-positive SARS-CoV-2 cases in Lithuania and identified few lineages with E484K circulating in Lithuania. During initial B.1.620 circulation in Lithuania the only other E484K-bearing lineages in Lithuania had been B.1.351 (one isolated case in Kaunas county, and 12 cases from a transmission chain centred in Vilnius county) and B.1.1.318 (one isolated case in Alytus county), none of which had been found in Utena county despite a high epidemic sequencing coverage in Lithuania (Supplementary Fig. S2).

An in-depth search for relatives of this lineage on GISAID¹² uncovered a few genomes from Europe initially, though more continue to be found since B.1.620 received its Pango lineage designation which was subsequently integrated into GISAID. This lineage now includes genomes from a number of European countries such as France, Switzerland, Belgium, Germany, England, Scotland, Italy, Spain, Czechia, Norway, Sweden, Ireland, and Portugal, North America: the United States (US) and Canada, and most recently The Philippines and South Korea in Asia. Interestingly, a considerable proportion of initial European cases turned out to be travellers returning from Cameroon. Since late April 2021, sequencing teams operating in central Africa, primarily working on samples from the Central African Republic, Equatorial Guinea, the Democratic Republic of the Congo, Gabon and lately the Republic of Congo have been submitting B.1.620 genomes to GISAID.

We here describe the mutations and deletions the B.1.620 lineage carries, many of which were previously observed in individual VOCs, but not in combination, and present evidence that this lineage likely originated in central Africa and is likely to circulate in the wider region where its prevalence is expected to be high. By combining collected travel records from infected patients entering different European countries, and by exploiting this information in a recently developed Bayesian phylogeographic inference methodology^{13,14}, we reconstruct the dispersal of lineage B.1.620 from its inferred origin in the Central African Republic to several of its neighbouring countries, Europe and the US. Finally, we provide a description of local transmission in Lithuania, France, Spain, Italy, and Germany through phylogenetic and phylogeographic analysis, and in Belgium through the collection of travel records.

Results

B.1.620 carries numerous VOC mutations and deletions. Lineage B.1.620 attracted our attention due to large numbers of unique mutations in B.1.620 genomes from Lithuania in next-clade analyses (its genomes are 18 mutations away from nearest relatives and 26 from reference strain Wuhan-Hu-1), and those genomes initially being assigned to clade 20A, corresponding to B.1 in Pangolin nomenclature^{4,5}. Meanwhile, Pangolin (using the 2021-04-01 version of pangoLEARN) variously misclassified B.1.620 genomes as B.1.177 or B.1.177.57 and occasionally as correct but unhelpful B.1, prior to the official designation of B.1.620 by the Pango SARS-CoV-2 lineage nomenclature team. To this day even after official designation Pangolin still often struggles with B.1.620 sequences and classifies them as various VOCs (often as B.1.1.7) when not used in the new USHER mode and vice versa sometimes classifies non-B.1.620 genomes as B.1.620. Closer inspection of B.1.620 genomes revealed that this lineage carries a number of mutations and deletions that have been previously observed individually in VOCs and VOIs (Fig. 1 and Supplementary Fig. S3), but had not been seen in combination. Despite sharing multiple mutations and deletions with known VOCs (most prominently HV69/70Δ, LLA241/243Δ,

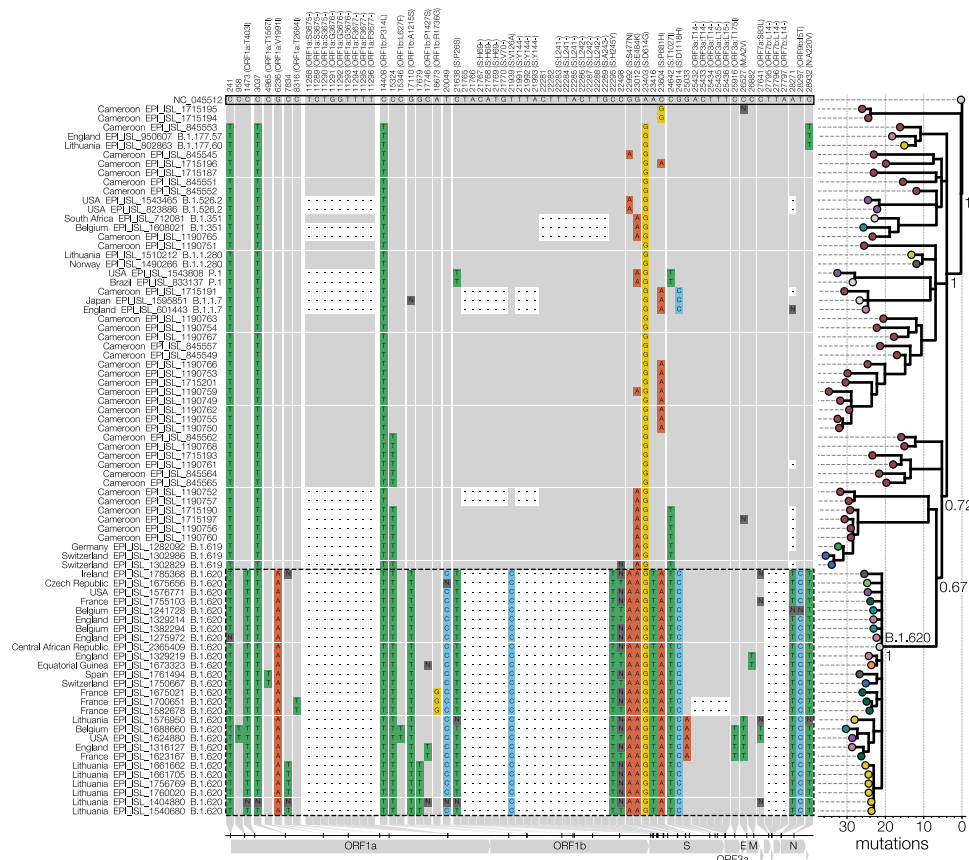


Fig. 1 Lineage-defining SNPs of lineage B.1.620. Only SNPs that differentiate B.1.620 (genomes outlined with a dashed line) from the reference (GenBank accession NC_045512) and that are shared by at least two B.1.620 genomes are shown in the condensed SNP alignment. Sites identical to the reference are shown in grey, changes from the reference are indicated and coloured by nucleotide (green for thymidine, red for adenosine, blue for cytosine, yellow for guanine, dark grey for ambiguities, black for gaps). The first 100 and the last 50 nucleotides are not included in the figure but were used to infer the phylogeny. If a mutation results in an amino acid change, the column label indicates the gene, reference amino acid, amino acid site, and amino acid change in brackets. The phylogeny (branch lengths in the number of mutations) on the right shows the relationships between depicted genomes and was rooted on the reference sequence with coloured circles at the tips indicating the country from which the genome came. Posterior probabilities of nodes leading up to lineage B.1.620 are shown near each node with the long branch leading to lineage B.1.620 labelled as ‘B.1.620’.

S477N, E484K and P681H), lineage B.1.620 does not appear to be of recombinant origin (Supplementary Fig. S4).

Through travel-related cases of B.1.620 discussed later we suspected Cameroon as the immediate source of this lineage and therefore sought to identify close relatives of this lineage there. While genomic surveillance in Cameroon has been limited, the genomes that have been shared on GISAID are quite diverse and informative. A handful appears to bear several mutations in common with lineage B.1.620 and could be its distant relatives (Fig. 1). Synonymous mutations at site 15324 and S:T10271 appear to be some of the earliest mutations that occurred in the evolution of lineage B.1.620, both of which are found in at least one other lineage associated with Cameroon (B.1.619), followed by S:E484K which also appears in genomes closest to lineage B.1.620. Even though the closest genomes to B.1.620 were sequenced from samples collected in January and February, lineage B.1.620 has 23 changes (mutations and deletions) leading up to it compared to the reference. During this study, SARS-CoV-2 genomes collected in January-March 2021 from the Central African Republic were deposited on GISAID, but none of them resembles forebearer or sibling lineages to B.1.620.

B.1.620 is likely to escape antibody-mediated immunity. Like most currently circulating variants, B.1.620 carries the D614G mutation, which enhances infectivity of SARS-CoV-2, likely through

enhanced interactions with the ACE2 receptor by promoting the up-conformation of the receptor-binding domain (RBD)¹⁵. Furthermore, B.1.620 contains P26S, HV69/70Δ, V126A, Y144Δ, LLA241/243Δ and H245Y in the N-terminal domain (NTD) of the spike protein. The individual V126A and H245Y substitutions are still largely uncharacterised to the best of our knowledge, but might be counterparts to the R246I substitution in B.1.351, and the latter may interfere with a putative glycan binding pocket in the NTD¹⁶. All other mutations of B.1.620 in the NTD result in partial loss of neutralisation of convalescent serum and NTD-directed monoclonal antibodies¹⁷. This indicates that these mutations present in B.1.620 may have arisen as an escape to antibody-mediated immunity¹⁸. The spike protein of B.1.620 also carries both S477N and E484K mutations in the RBD, but in contrast to other VOCs not the N501Y or K417 mutations. Like the mutations in the NTD, S477N and E484K individually enable broad escape from antibody-mediated immunity¹⁸. Moreover, deep mutational scanning experiments have shown that these substitutions also increase the affinity of the RBD for the ACE2 receptor¹⁹. Both S477N and E484K occur on the same flexible loop at the periphery of the RBD-ACE2 interface²⁰.

We have modelled the RBD-ACE2 interface with the S477N and E484K substitutions using refinement in HADDOCK 2.4²¹. These models show that both individual substitutions and their combination produce a favourable interaction with comparable scores and individual energy terms to the ancestral RBD

(Supplementary Fig. S5). Whereas S477N may modulate the loop conformation²², E484K may introduce new salt bridges with E35/E75 of ACE2. These results indicate that B.1.620 may escape antibody-mediated immunity while maintaining a favourable interaction with ACE2. The remaining mutations in the spike protein—P681H, T1027I and D1118H—are uncharacterised to the best of our knowledge. Of these, P681H is also located on the outer surface of the spike protein, directly preceding the multibasic S1/S2 furin cleavage site²³. In contrast, T1027I and D1118H are both buried in the trimerisation interface of the S2 subunit²⁴.

While only limited empirical data are available, they seem to agree with the expectation that B.1.620 is likely to be antigenically drifted relative to primary genotypes. A report presented to the Lithuanian government on May 22, 2021²⁵ indicated that amongst 101 sequenced B.1.620 cases at the time, 13 were infections in fully vaccinated individuals, five of whom were younger than 57 years old. Though not systematised properly, sequencing indications for a substantial number of SARS-CoV-2 genomes from Lithuania were available, of which 213 were ‘positive PCR at least 2 weeks after the second dose of vaccine’, of which 195 were B.1.1.7 and 12 were B.1.620. Since detection of the first B.1.620 case on March 15, 2021, in Lithuania ~10,000 SARS-CoV-2 genomes were sequenced to date, 9251 of which were B.1.1.7 and 248 of which were B.1.620. Thus B.1.620 is found 2.4 times more often in vaccine breakthrough cases compared to its population prevalence, whereas for B.1.1.7 this enrichment is only 1.05-fold. Similarly, the frequency of B.1.620 across the five most affected European countries (Lithuania, Germany, Switzerland, France and Belgium) appears relatively stable though at a low level, unlike B.1.1.7 which has been in noticeable decline since April–May (Supplementary Fig. S6), presumably on account of increasing vaccination rates and improving weather in Europe.

Local transmission of B.1.620 in Europe. Local transmission of B.1.620 in Lithuania has been established as a result of monitoring the outbreak in Anykščiai municipality (Utena county, Lithuania) via sequencing and repeat PCR testing of SARS-CoV-2 positive samples for the presence of E484K and N501Y mutations, as well as looking for S gene target failure (SGTF) caused by the HV69 Δ deletion. Genotypes identical to those found initially in Vilnius and Utena counties were later identified by sequencing in Panevėžys and Šiauliai counties, indicating continued transmission of lineage B.1.620 in-country. Interestingly, a single case in Tauragė county, Lithuania, identified by sequencing was a traveller returning from France found to be infected with a different genotype than the main outbreak lineage in Lithuania without evidence of onward transmission via local contact tracing efforts or genomic surveillance.

In addition to an ongoing disseminated outbreak of B.1.620 in Lithuania, genomes of this lineage have been found elsewhere in Europe. Though derived from separate introductions from the one that sparked outbreaks in Lithuania, other B.1.620 genomes from Europe appear to indicate ongoing transmission in Europe, with the clearest evidence of this in Germany and France, where emerging clades are comprised of identical or nearly identical genotypes (Fig. 2). Presenting evidence for local transmission in Europe, B.1.620 genomes from countries like Spain and Belgium (also see next section) were notably picked up by baseline surveillance and thus are likely to represent local circulation, though presumably at much lower levels at the time of writing. Figure 2 shows the aforementioned local transmission clusters in Lithuania, Spain (Vilassar De Mar, province of Barcelona), France (see below), and Germany (state of Bavaria), amongst numerous others.

In France, nine B.1.620 genomes (EPI ISL 1789089 - EPI ISL 1789097) were recently obtained from a large contact tracing

investigation of a single transmission chain. These infections in the municipality of Pontoise (Val d’Oise department, to the northwest of Paris) occurred in adults (ages 24–38) who were all asymptomatic at the time of sampling. Additional infections in Pontoise outside of this cluster occurred in four adults (ages 29–57) and form a monophyletic cluster with the other nine infected individuals (Supplementary Fig. S4). The putative index case for these infections has yet to be determined through contact tracing at the time of writing but these cases clearly point to the B.1.620 lineage circulating in the Val d’Oise department. These infections seem to stem from local ongoing transmission in the Île-de-France region, clustering with two patients ages 1 (sample from a children’s hospital in Paris: Hôpital Necker-Enfants malades) and 69. These infections in Île-de-France in turn cluster with two infections from Le Havre (region of Normandy; 180km from Pontoise), pointing to either a travel event from Normandy to Île-de-France or possible local transmission in the north of France (Supplementary Fig. S4).

B.1.620 likely circulates at high frequency in central Africa. In the absence of routine surveillance at a location, sequencing infected travellers originating from there constitutes the next most efficient way to monitor distinct viral populations. This has been used successfully to uncover cryptic outbreaks of Zika virus in Cuba²⁶ and SARS-CoV-2 in Iran at the beginning of the pandemic¹³. The latter study describes a novel approach to accommodate differences in sampling location and location of infection, and is hence specifically targeted to exploit recorded travel histories of infected individuals in Bayesian phylogeographic inference, rather than arbitrarily assigning the origin of the sample to either location. When we first compiled our B.1.620 genomes dataset we had seven genomes from travellers and six were sampled in the Central African Republic (CAR) near the border with Cameroon, indicating the most plausible geographic region where B.1.620 is circulating widely to be central Africa (Supplementary Fig. S7). Neighbours of countries reporting local B.1.620 circulation (Cameroon, CAR, DRC, Gabon, Equatorial Guinea, and later the Republic of Congo) have either not submitted genomes to GISAID during the study period (Chad, Sudan, South Sudan, Burundi) or have epidemics dominated by SARS-CoV-2 lineages that are not B.1.620 (Supplementary Fig. S8).

The collected individual travel histories themselves point to several independent introductions of B.1.620 into Europe, with documented cases of infected travellers returning from Cameroon to Belgium, France and Switzerland, and from Mali to Czechia (Fig. 3). We note that the metadata for a returning traveller from Cameroon to Belgium (EPI_ISL_1498300) presents evidence of ongoing local transmission within Belgium of B.1.620. Whereas this patient had spent time in Cameroon from the 16th of January until the 7th of February, a positive sample was only collected on the 15th of March, 2021. Even when assuming a lengthy infectious period of up to twenty days²⁷, this patient’s infection can not stem from his prior travel to Cameroon, which indicates an infection with B.1.620 within Belgium and hence stemming from contact within the patient’s community. Additionally, two Belgian patients (EPI_ISL_1688635 and EPI_ISL_1688660) were likely infected by the former’s niece who had travelled with her family to Cameroon and tested positive upon their return to Belgium. These findings are reinforced by more recent samples from Belgium, for which no travel history could be recorded and the patients declared not having left the country.

Using a Bayesian phylogeographic inference methodology that accommodates individual travel histories we were able to reconstruct location-annotated phylogenies at both the continent and country

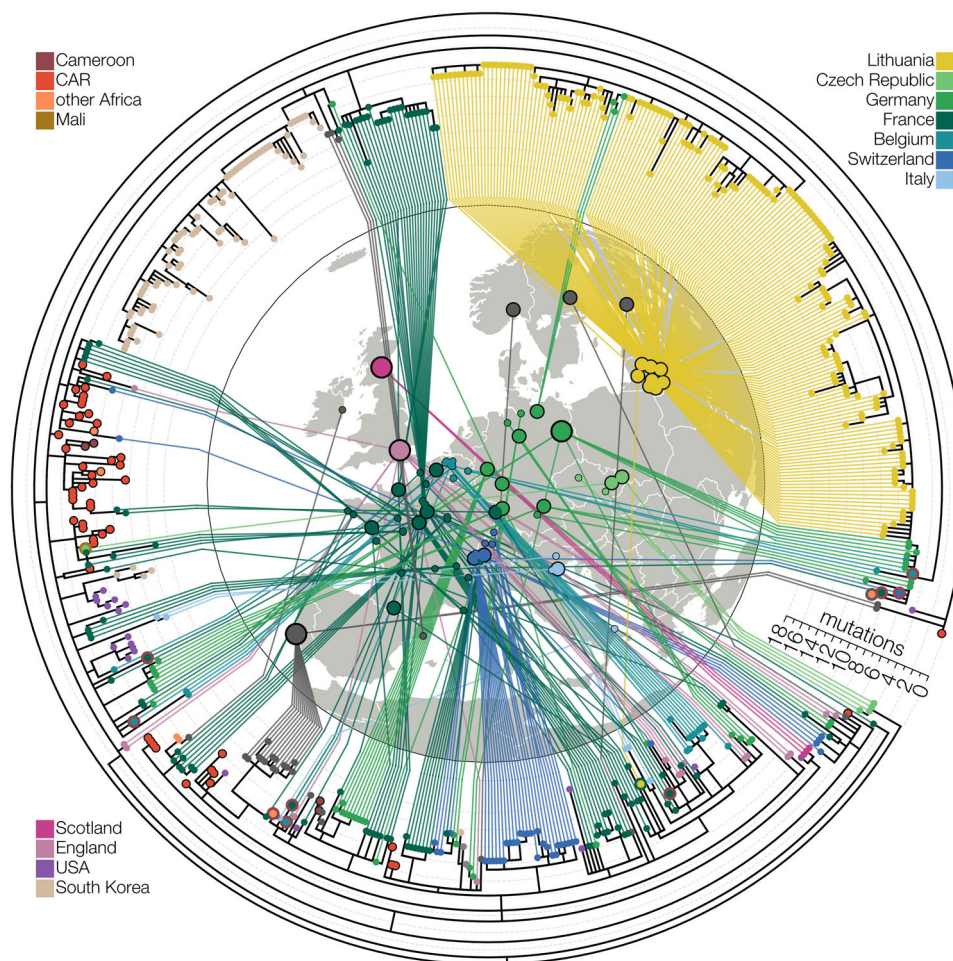


Fig. 2 Maximum-likelihood tree of lineage B.1.620 in Europe. Relationships between B.1.620 genomes, coloured by country of origin (same as Fig. 1) with a thicker coloured outline indicating the country of origin for travel cases. At least ten genomes shown (samples collected in Belgium, Switzerland, France and Equatorial Guinea) are from individuals who returned from Cameroon, one is from a traveller returning from Mali and one Lithuanian case returned from France. Genomes from the Central African Republic (CAR) and Czechia (returning traveller from Mali) are descended from the original B.1.620 genotype, while the genome from Equatorial Guinea is already closely related to genomes found in Europe and happens to be a travel case from Cameroon. Each genome is connected to the available geographic location in Europe with the smallest circles indicating municipality-level precision, intermediate size corresponding to county-level information (centred on county capital) and largest circle sizes indicating country-level information (centred on country capital). Countries are assigned the same colours as in Figs. 1 and 3.

levels. Figure 4A shows the MCC tree of the continent-level phylogeographic analysis, which yields 99.5% posterior support for an African origin of lineage B.1.620. From this inferred African origin, the variant then spread to different European countries via multiple introductions, which is confirmed by our collection of travel history records for individuals returning to these countries. Subsequent country-level phylogeographic analysis—shown in Fig. 4B—points to central Africa as the likely origin of this lineage, with the Central African Republic receiving posterior support of 80.5% and Cameroon 16.8%, taking up 97.3% of the probability mass together. Assuming a Central African Republic origin, the variant is estimated to have spread to Europe via a series of introductions, confirming what was also observed in our recorded travel history records. Interestingly, a single Lithuanian case—a returning traveller from France—does not cluster with the cluster of remaining sequences from Lithuania, illustrative of at least two independent introductions of lineage B.1.620 into Lithuania. Figure 4B also shows multiple separate B.1.620 introduction events from central Africa into the United Kingdom and the United States.

Air passenger flux out of Cameroon and Central African Republic (Fig. 5) shows that many travellers had African

countries as their destination, including many that have not reported any B.1.620 genomes to date. This suggests that B.1.620 could be circulating more widely in Africa and its detection in Europe has mostly occurred in countries with recent active genomic surveillance programmes. Detections of B.1.620 in African states neighbouring Cameroon and Central African Republic (Equatorial Guinea, Gabon, DRC and lately the Republic of Congo), even at low sequencing levels, suggest that B.1.620 may be prevalent in central Africa. We find this apparent rise to high frequency and rapid spread across large areas of Africa noteworthy in light of other findings reported here, namely that currently available B.1.620 genomes appeared suddenly in February 2021 (Fig. 3), are genetically homogeneous (Fig. 2), and to date have no clear close relatives (Fig. 1).

Discussion

In this study, we have presented evidence that a SARS-CoV-2 lineage designated B.1.620, first detected in Europe in late February, is associated with the central African region, where it appears to circulate at high prevalence, and has been introduced

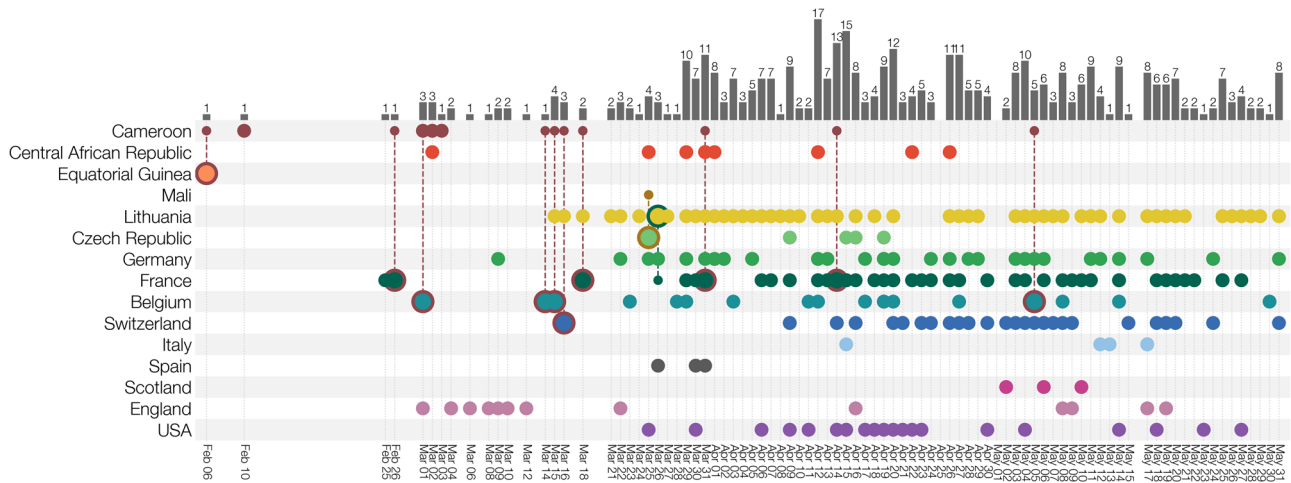


Fig. 3 Known locations and travel history of B.1.620 cases. Collection dates of B.1.620 genomes are shown for each country (rows). Genomes from travellers are outlined with colour indicating travel of origin (e.g. dark red for Cameroon) and connected to a smaller dot indicating which country's diversity is being sampled at the travel destination. Bars at the top indicate the number of genomes of B.1.620 available for a given date across all countries. Countries are assigned the same colours as in Fig. 1.

into Europe, North America, and Asia on multiple occasions. A fair number of known B.1.620 genomes that were sequenced in Europe stem from travel-related cases returning from Cameroon (Fig. 3), and recently sequenced genomes from CAR and Cameroon similarly belong to lineage B.1.620, suggesting that the central African region is likely to be the immediate source of this lineage. Importantly, our findings are quite insensitive to the actual sequence data used. Older datasets we used dating from the end of April 2021 (Supplementary Fig. S9) included only six genomes from CAR and travel cases in Europe coming from Cameroon and yet still confidently identified Cameroon as the immediate origin of lineage B.1.620. Adding more data from CAR (Supplementary Fig. S10) made available later made the Central African Republic the more likely country where B.1.620 circulated prior to spreading elsewhere, but ultimately no country other than CAR and Cameroon are considered as remotely plausible by the model.

Substantially higher passenger flux out of Cameroon compared to CAR (practically an order of magnitude) is a likely explanation for why B.1.620-infected travellers were overwhelmingly coming to Europe from Cameroon. So far the only observation that is difficult to explain is the Czech case returning from Mali, since Mali is over 1000 km away from Cameroon. We consider the introduction of B.1.620 from central Africa to Mali via land routes improbable, since outbreaks caused by B.1.620 have not been observed in Niger and Nigeria, the countries separating the region from Mali. The lack of any B.1.620 genomes from Nigeria in particular, one of the leaders in SARS-CoV-2 genome sequencing on the continent to date, despite higher civil air passenger volumes (Fig. 5) suggests other means of long-distance travel between central Africa and Mali^{28,29}.

In addition to the multiple introductions of the B.1.620 lineage we observe (Fig. 3) and estimate (Fig. 4) in Europe and North America, we also found evidence of local transmission of this lineage in Europe, with clearest evidence in Lithuania (Supplementary Fig. S1) followed by Germany and France (Fig. 3), and finally, Belgium and Catalonia, where B.1.620 genomes were picked up by baseline surveillance and infected individuals did not report having travelled abroad. B.1.620 is worrying for several reasons—its genomes are genetically homogeneous—as it appeared suddenly in February 2021 bearing a large number of VOC-like mutations and deletions in common with multiple VOCs (Supplementary Fig. S3), yet in the absence of any clear

close relatives or sampled antecedents (Fig. 1). The discovery of a novel lineage bearing many mutations of concern and with indications that they are introduced from locations where sequencing is not routine, is concerning and such occurrences may become an alarming norm.

The continued lack of genomic surveillance in multiple areas of the world, let alone equitable access to vaccines to drive transmission down, will continue to undermine efforts to control SARS-CoV-2 everywhere. Without the ability to identify unusual variants, to observe their evolution and learn from it, and to evaluate how vaccine-induced immunity protects against them, any response enacted by individual countries is reactive and, much like the process of evolution that generates variants of concern, short-sighted. The emergence of B.1.1.7 was unprecedented and has had a devastating impact on the state of the pandemic, so it is concerning that similar information gaps in global genomic surveillance still persist to this day. As an example we have shown that B.1.620 lacks intermediate relatives, resulting in a long branch that connects this lineage to the ancestral genotype of B.1. This could be the result of gradual but unsampled evolution, perhaps even far away from central Africa, but it could have also happened due to unusual selection pressures in immunosuppressed individuals³⁰ which is hypothesised for lineage B.1.1.7. The long branch leading to B.1.620 also means that we can not reconstruct the order of mutations that have occurred during the genesis of this lineage and therefore whether some amino acid changes have allowed others to happen by altering the fitness landscape via epistatic interactions³¹. Given the number of VOC-like mutations B.1.620 has, this is a significant loss.

Our work highlights that global inequalities, as far as infectious disease monitoring is concerned, have tangible impacts around the world and that until the SARS-CoV-2 pandemic is brought to heel everywhere, nowhere is safe for long. Additionally, we highlight the importance of collecting and sharing associated metadata with genome sequences, in particular regarding individual travel histories, as well as collection dates and locations, all of which are important to perform detailed phylogenetic and phylogeographic analysis. We only observed one single instance where a GISAID entry was accompanied by travel information and had to request such information for all the samples in our core dataset by contacting each individual lab. Whereas many labs were quick to provide the requested information, we were certainly not able to retrieve all related individual travel histories.

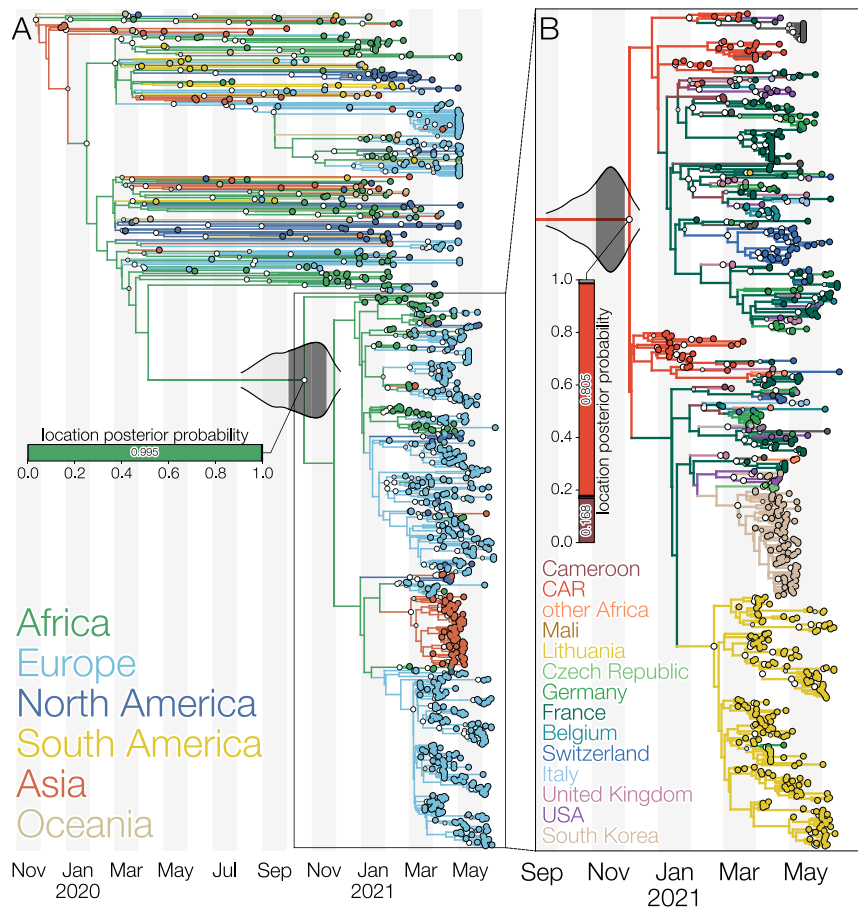


Fig. 4 Maximum clade credibility trees of lineage B.1.620 coloured by reconstructed location using the latest available data as of June 2021. **A** Global phylogeny of SARS-CoV-2 genomes with branches coloured by inferred continent from a Bayesian phylogeographic analysis that makes use of individual travel histories. Lineage B.1.620 is outlined and a horizontal bar shows the posterior probability of its common ancestor existing in a given continent. Africa is reconstructed as the most likely location (posterior probability 0.995) where B.1.620 originated. The 95% and 50% highest posterior density (HPD) intervals for the most recent common ancestor date of lineage B.1.620 are indicated with violin plots centred on the common ancestor. **B** Phylogeny of lineage B.1.620 with branches coloured by inferred country from a Bayesian phylogeographic analysis that makes use of travel histories. A vertical bar shows posterior probabilities of where the common ancestor of B.1.620 existed. In this analysis, Central African Republic (CAR) and Cameroon are reconstructed as the most likely locations (with posterior probabilities of 0.805 and 0.168, respectively) of the common ancestor of lineage B.1.620. Larger white dots at nodes indicate nodes with a posterior probability of at least 95%, while smaller grey circles indicate nodes with a posterior probability of at least 50%. The 95% and 50% highest posterior density (HPD) intervals for the most recent common ancestor date of lineage B.1.620 are indicated with violin plots centred on the common ancestor.

The scientific community therefore still faces the important task of reporting and sharing such critical metadata in a consistent manner, an aspect that has been brought to attention again during the ongoing pandemic^{32,33}.

Methods

Study design. This study was initiated upon detection of SARS-CoV-2 strains in Lithuania bearing spike protein amino acid substitutions E484K, S477N and numerous B.1.1.7-like (HV69/70Δ and Y144Δ) and B.1.351-like (LLA241/243Δ) deletions, amongst others. In Lithuania, repeat PCR testing of SARS-CoV-2 positive samples is occasionally carried out to detect N501Y, E484K and S gene target failure (SGTF) caused by the HV69Δ deletion. Upon detection of E484K-positive cases, samples were redirected to sequencing. Initially identified cases of B.1.620 were mistakenly classified by pangolin as B.1.177 or B.1.177.57, while nextclade³⁴ assigned it to clade 20A rather than the expected 20E (EU1), while highlighting that B.1.620 sequences bore many unique mutations compared to the closest sequence. Searching GISAID for mutations E484K, S477N and HV69/70Δ, which are found in numerous VOCs individually but not in combination, identified additional genomes that contained other mutations and deletions found in B.1.620.

We downloaded all available sequences of this lineage from GISAID in July 2021, and identified members that clearly belonged to this lineage. Said to official lineage designation as B.1.620, most of its genomes could be identified by the presence of spike protein E484K and S477N mutations and the HV69/70Δ deletion. Some of B.1.620 genomes were excluded from phylogenetic analyses because they

were misassembled (e.g. hCoV-19/Belgium/UZA-UA-24912930/2021 is missing deletions characteristic of this lineage but has the mutations) or had too many ambiguous sites (e.g. hCoV-19/France/ARA-HCL021061598501/2021) but we recovered travel information about them regardless as this may prove useful to perform travel history-aware phylogeographic reconstruction¹³.

SARS-CoV-2 whole-genome sequencing. Every sample that tests positive for SARS-CoV-2 by PCR in Lithuania with Ct values < 30 may be redirected by the National Public Health Surveillance Laboratory to be sequenced by the European Centre for Disease Prevention and Control (ECDC), Vilnius University Hospital Santaros Klinikos (VUHSK), Hospital of Lithuanian University of Health Sciences Kauno Klinikos (HLUHSKK), Vilnius University Life Sciences Centre (VULSC) or Lithuanian University of Health Sciences (LUHS). Samples of this particular lineage were sequenced by ECDC using in-house protocols, infrastructure and assembly methods, VUHSK using Illumina COVIDSeq reagents, Illumina MiSeq platform, and assembled with covid-19-signal³⁵, HLUHSKK using Twist SARS-CoV-2 Research Panel reagents, Illumina NextSeq550 platform, and assembled with V-pipe³⁶, LUHS using ARTIC protocol, Oxford Nanopore Technologies MinION platform, and assembled using ARTIC bioinformatics protocol for SARS-CoV-2, and VULSC using ARTIC V3 protocol combined with Invitrogen Collibri reagents, Illumina MiniSeq platform, Illumina DRAGEN COVID Lineage combined with an in-house BLAST v2.10.18-based assembly protocol. Samples from CAR were sequenced using the very same ARTIC V3 protocol as the Lithuanian University of Health Sciences (LUHS).

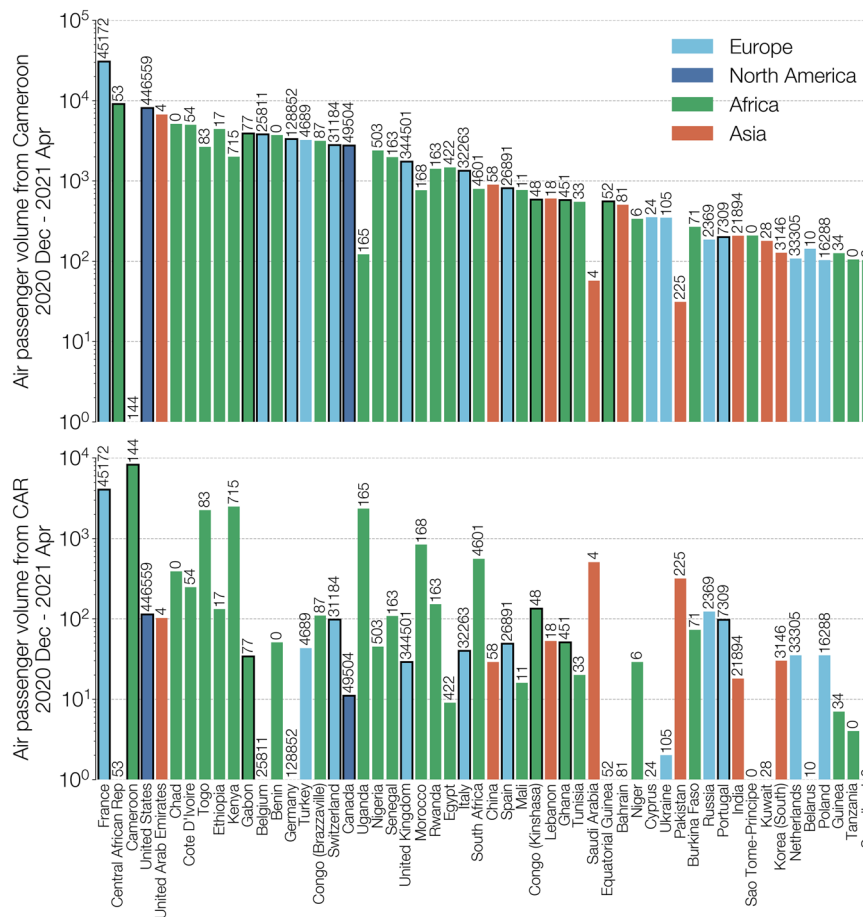


Fig. 5 Total air passenger flows out of Cameroon (top) and Central African Republic (bottom) between December 2020 and April 2021. Destination countries are sorted by total passenger volume arriving from Cameroon and Central African Republic (CAR) combined, coloured by continent (Europe in light blue, North America in dark blue, Africa in green, and Asia in red) and limited to countries where at least 100 passengers have arrived from either Cameroon or CAR between December 2020 and April 2021. Note the nearly order of magnitude greater passenger flux out of Cameroon compared to the Central African Republic (CAR). Numbers above each country's bar indicate the total number of genomes on GISAID from that country since January 1st 2021, according to GISAID's 2021-07-02 metadata release. Bars outlined in black represent countries that have submitted at least one B.1.620 genome as of June 2021.

All SARS-CoV-2 genomes used here were downloaded from GISAID. A GISAID acknowledgement table containing all genome accession numbers is included with this study as Supplementary Data 1.

Associated travel history. When available on GISAID as part of the uploaded metadata, we made use of this associated metadata information and contacted the submitting labs to determine precise travel dates. For all other cases, we retrieved individual travel histories by contacting the submitting labs—who then, in turn, contacted either the originating lab or the patient's general practitioner—for any travel records they may have available. This resulted in travel itineraries for 10 patients, with 7 of these also containing detailed dates for the recorded travel. When a returning traveller visited multiple countries on the return trip, we included all visited countries as possible locations of infection by using an ambiguity code in the phylogeographic analysis^{13,14}. The travel history information collected can be found in Supplementary Table S1. While we were able to retrieve travel history for a fair number of cases, this information is considered private information in certain countries and we were hence unable to retrieve such data for a subset of our sequences.

SARS-CoV-2 genomes from the United Kingdom (UK) make up a sizeable proportion of any phylogenetic and phylogeographic analysis, given significant sequencing efforts by the COVID-19 Genomics UK Consortium. Given the lack of individual travel histories for B.1.620 genomes from England in our dataset, we investigated the passenger volumes from all airports in Cameroon and the Central African Republic to all airports internationally, incorporating volumes from both direct and connecting flights between December 2020 and April 2021, from the International Air Transportation Association (IATA³⁷). These passenger data cover the time frame of our estimated B.1.620 lineage since its origin (see 'Results' section), with the passenger volumes for February having become available at the time of writing as these data need to be retrieved and processed. These air

passenger flux data reveal a very real possibility of missing travel histories from Cameroon for B.1.620 cases in England, given that over 98% (i.e. 852 out of 867) of the passengers from Cameroon to the UK during this time frame had an English airport (London, Manchester or Birmingham) as their final destination. At the time of writing, information on the origin of B.1.620 infections detected in England is not available.

Modelling RBD-ACE2 interaction. We have modelled the RBD-ACE2 interface with the S477N and E484K substitutions using the final refinement step of HADDOCK 2.4²¹. We used the crystal structure of ACE2 (19-615) bound to SARS-CoV-2 RBD (PDB ID: 6m0j²⁰) as a starting point and introduced the substitutions using UCSF ChimeraX³⁸. We used default parameters for refinement with extended molecular dynamics (MD) simulation (steps for heating phase: 200, steps for 300K phase: 2500, steps for cooling phase: 1000).

Phylogenetic and phylogeographic analysis. We combined 614 sequences belonging to lineage B.1.620 with sequences from lineages that have circulated in Lithuania at appreciable levels: B.1.1.7, B.1.1.280, B.1.177.60 and other VOCs that share mutations with lineage B.1.620: B.1.351, P.1 and B.1.526.2. We included high-quality sequences from Cameroon that were closest to lineage B.1.620 as well as the reference SARS-CoV-2 genome NC_045512. Some sequences had clusters of SNPs different from the reference at the ends of the genome, particularly the 5' end. In such cases, the ends of the genomes were trimmed to exclude these regions of likely sequencing or assembly error. This resulted in a core set of 665 genomes, which is visualised in Supplementary Fig. S4, that serves as the starting point for our phylogenetic and phylogeographic analyses. This core set was subsequently combined with 250 randomly selected sequences from the Nextstrain global analysis on April 29, 2021 (<https://nextstrain.org/ncov/global34>) to provide context for the B.1.620 analysis, plus an additional two reference sequences: Wuhan/Hu-1/

2019 and Wuhan/WH01/2019. We filtered these sequences based on metadata completeness and added an additional four Chinese sequences as well as eight non-Chinese sequences from Asia spanning both A and B lineages, in order to balance the representation of different continents in our analyses. These sequences were aligned in MAFFT (FFT-NS-2 setting)³⁹ with insertions relative to reference removed, and 5' and 3' untranslated regions of the genome that were susceptible to sequencing and assembly error trimmed. We employed TempEst⁴⁰ to inspect the dataset for any data quality issues that could result in an excess or shortage of private mutations in any sequences, or would point to assembly or any other type of sequencing issues.

To look for sequences that could resolve the long period of unobserved evolution separating lineage B.1.620 from its closest relatives, we constructed a BLAST nucleotide database⁴¹ of all contemporary SARS-CoV-2 lineages available via GISAID (accessed 2021-07-01, $n = 2,038,838$). We queried this database using a synthetic B.1.620-like sequence containing SNPs and deletions shared by B.1.620 sequences England/CAMC-13B04C1/2021 and France/PDL-IPP07069/2021 that were not present in the reference sequence Wuhan/Hu-1/2019. The synthetic query sequence was primarily comprised of ambiguous nucleotides (N) except for 100 nt surrounding each mutation or deletion characteristic of B.1.620. We checked the top 500 matches to see if the mutations they carry, their pango lineages or phylogenetic placement via IQ-TREE⁴²—using a general time-reversible substitution model with among-site rate variation (GTR+ Γ_4)^{43,44}—could identify sequences closer to lineage B.1.620 than B.1.619. No such sequences were identified.

We performed Bayesian model selection through (log) marginal likelihood estimation to determine the combination of substitution, molecular clock and coalescent models that best fits the data. To this end, we employed generalised stepping-stone sampling (GSS⁴⁵) by running an initial Markov chain of 5 million iterations, followed by 50 path steps that each comprise 100,000 iterations, sampling every 500th iteration. We found that a combination of a non-parametric skygrid coalescent model⁴⁶, an uncorrelated relaxed clock model with underlying lognormal distribution⁴⁷ and a GTR+ Γ_4 substitution model provided the optimal model fit to the data. We employed Hamiltonian Monte Carlo sampling to efficiently infer the skygrid's parameters⁴⁸.

We subsequently performed a discrete Bayesian phylogeographic analysis in BEAST 1.10.5⁴⁹ using a recently developed model that is able to incorporate available individual travel history information associated with the collected samples^{13,14}. Exploiting such information can yield more realistic reconstructions of virus spread, particularly when travellers from unsampled or under-sampled locations are included to mitigate sampling bias. When the travel date for a sample could not be retrieved, we treated the time when the traveller started the journey as a random variable, and specified normal prior distributions over these random variables informed by an estimate of time of infection and truncated to be positive (back-in-time) relative to sampling date. As in previous work^{13,14}, we used a mean of 10 days before sampling based on a mean incubation time of 5 days⁵⁰, a constant ascertainment period of 5 days between symptom onset and testing⁵¹, and a standard deviation of 3 days to incorporate the uncertainty on the incubation time.

In our phylogeographic analysis, we made use of Bayesian stochastic search variable selection (BSSVS) to simultaneously determine which migration rates are zero depending on the evidence in the data and infer ancestral locations, in addition to providing a Bayes factor test to identify migration rates of significance⁵². We first performed a continent-level phylogeographic analysis by aggregating sampling locations as well as the individual travel histories that occurred between continents. To ensure consistent spatial reconstruction regardless of sampling, we fixed the root location of this tree to be in Asia—so as to match the known epidemiology of the COVID-19 pandemic. Conditional on the results of this analysis, we performed a country-level analysis on the B.1.620 lineage and its parental lineage, in order to substantially reduce the computational burden and statistical complexity associated with having 87 sampling locations in a travel history-aware phylogeographic analysis. We made use of the following prior specifications for this analysis: a gamma (shape = 0.001; scale = 1000) prior on the skygrid precision parameter, Dirichlet (1.0, K) priors on all sets of frequencies (with K the number of categories), Gamma prior distributions (shape = rate = 1.0) on the unnormalized transition rates between locations⁵², a Poisson prior (country level: $\lambda = 28$; continent level: $\lambda = 5$) on the sum of non-zero transition rates between locations, a CTMC reference prior on the mean evolutionary rate and as well as on the overall (constant) diffusion rate⁵³. In the country-level analysis, we assumed a normally distributed root height prior on the time of origin of B.1.620's parental lineage, with a mean on the 27th of February 2020 and standard deviation of 2 weeks, as derived from the corresponding internal node's 95% highest posterior density interval in the preceding continent-level analysis. For continent-level analysis 18 independent Markov chains were set up, running for ~50 million states and sampling every 40,000th state. All 18 runs were then combined after removing 10% of the states as burnin, giving a total MCMC length of 810 million states. For country-level analysis 16 independent Markov chains were set up, running for ~3.5 million states and also sampling every 40,000th state. All 16 runs were then combined after removing 10% of the states as burnin, giving a total MCMC length of 50.4 million states. Both continent-level and country-level combined run were inspected using Tracer v1.7⁵⁴ to confirm that effective sample sizes (ESSs) for all relevant parameters were at least 200. We used TreeAnnotator to

construct maximum clade credibility (MCC) trees for both posterior sets of trees and used baltic (<https://github.com/evogytis/baltic>) to visualise it.

In addition to sophisticated phylogeographic analyses, we also depict the raw relationships between SARS-CoV-2 in the core dataset of 665 genomes using substitution phylogenies. Figure 2 and Supplementary Fig. S4 depict maximum-likelihood phylogenies inferred from the core dataset using PhyML⁵⁵ under the HKY+ Γ_4 model of nucleotide substitution^{44,56} which was then rooted on the reference sequence. To occupy less space in Fig. 1 the number of B.1.620 genomes was reduced down to a representative set of 27, and a phylogeny was inferred using MrBayes v3.2⁵⁷ under the HKY+ Γ_4 model of nucleotide substitution^{44,56} and rooted on the reference sequence. MCMC was run for 2 million states, sampling every 1000th state and convergence confirmed by checking that effective sample sizes (ESSs) were above 200 for every parameter.

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

Data availability

SARS-CoV-2 sequence data generated in this study have been deposited in the GISAID database. These sequence data are available under restricted access due to GISAID's Database Access Agreement, access can be obtained by registering an account with GISAID and downloaded via the list of accession used that we provide in the supplementary GISAID acknowledgement table. The processed SARS-CoV-2 genome data in the form of phylogenetic trees are available at <https://github.com/evogytis/B.1.620-in-Europe> or under Zenodo <https://doi.org/10.5281/zenodo.5494346>. The SARS-CoV-2 genome data used in this study are available in the GISAID database under accession codes provided in the supplementary acknowledgement table, <https://github.com/evogytis/B.1.620-in-Europe>, and under Zenodo <https://doi.org/10.5281/zenodo.5494346>. A list of GISAID accessions for genomes used here, as well as phylogenetic trees used in figures, are available at <https://github.com/evogytis/B.1.620-in-Europe> under Zenodo <https://doi.org/10.5281/zenodo.5494346>. To access sequence data from GISAID one has to register an account with <https://www.gisaid.org/>, which involves identifying oneself and agreeing to GISAID's Database Access Agreement.

Code availability

Scripts used to generate figures are available at <https://github.com/evogytis/B.1.620-in-Europe> or under Zenodo <https://doi.org/10.5281/zenodo.5494346>. We provide the XML files to perform the Bayesian phylogeographic reconstructions in BEAST 1.10.5⁴⁹ as Supplementary Data 2.

Received: 26 May 2021; Accepted: 15 September 2021;

Published online: 01 October 2021

References

1. Tegally, H. et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. Preprint at *medRxiv* <https://doi.org/10.1101/2020.12.21.20248640> (2020).
2. Hodcroft, E. B. et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712 (2021).
3. Faria, N.R. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021).
4. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
5. O'Toole, A. et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol.* **7**, veab064 (2021).
6. Butera, Y. et al. Genomic Sequencing of SARS-CoV-2 in Rwanda: evolution and regional dynamics. Preprint at *medRxiv* <https://doi.org/10.1101/2021.04.02.21254839> (2021).
7. Nelson, M. I. et al. Multiyear persistence of 2 pandemic A/H1N1 influenza virus lineages in West Africa. *J. Infect. Dis.* **210**, 121–125 (2014).
8. Eden, J.-S. et al. An emergent clade of SARS-CoV-2 linked to returned travellers from Iran. *Virus Evol.* **6**, veaa027 (2020).
9. de Oliveira, T. et al. A novel variant of interest of SARS-CoV-2 with multiple spike mutations detected through travel surveillance in Africa. Preprint at *medRxiv* <https://doi.org/10.1101/2021.03.30.21254323> (2021).
10. Bugembe, D. L. et al. Emergence and spread of a SARS-CoV-2 lineage A variant (A.23.1) with altered spike protein in Uganda. *Nat. Microbiol.* **6**, 1094–1101 (2021).
11. Lyngse, F. P. et al. Increased Transmissibility of SARS-CoV-2 Lineage B.1.1.7 by Age and Viral Load: Evidence from Danish Households. Preprint at *medRxiv*. <https://doi.org/10.1101/2021.04.16.21255459> (2021).

12. Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, 30494 (2017).
13. Lemey, P. et al. Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nat. Commun.* **11**, 5110 (2020).
14. Hong, S. L., Lemey, P., Suchard, M. A. & Baele, G. Bayesian phylogeographic analysis incorporating predictors and individual travel histories in BEAST. *Curr. Protoc.* **1**, e98 (2021).
15. Yurkovetskiy, L. et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **183**, 739–751.e8 (2020).
16. Buchanan, C.J. et al. Cryptic SARS-CoV2-spike-with-sugar interactions revealed by ‘universal’ saturation transfer analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.04.14.439284> (2021).
17. Wang, P. et al. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747–751.e4 (2021).
18. Liu, Z. et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* **29**, 477–488.e4 (2021).
19. Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295–1310.e20 (2020).
20. Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).
21. van Zundert, G. C. P. et al. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
22. Singh, A., Steinkellner, G., Köchl, K., Gruber, K. & Gruber, C. C. Serine 477 plays a crucial role in the interaction of the SARS-CoV-2 spike protein with the human receptor ACE2. *Sci. Rep.* **11**, 4320 (2021).
23. Hoffmann, M., Kleine-Weber, H. & Pöhlmann, S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **78**, 779–784.e5 (2020).
24. Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6 (2020).
25. Šimaitis, A. Situation report to the Lithuanian government regarding SARS-CoV-2. <https://lrv.lt/uploads/main/documents/files/20210511%20COVID-19%20situacijos%20ap%C5%BEvalga.pdf> (2021).
26. Grubaugh, N. D. et al. Travel surveillance and genomics uncover a hidden zika outbreak during the waning epidemic. *Cell* **178**, 1057–1071 (2019).
27. Byrne, A. W. et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open* **10**, e039856 (2020).
28. EU. EUTM RCA. https://eeas.europa.eu/csdp-missions-operations/eutm-rca_en (2016).
29. EU. EUTM Mali. <https://eutmmali.eu/> (2013).
30. Choi, B. et al. Persistence and evolution of SARS-CoV-2 in an immunocompromised host. *N. Engl. J. Med.* **383**, 2291–2293 (2020).
31. Kemp, S. A. et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **592**, 277–282 (2021).
32. Oude Munnink, B. B. et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat. Med.* **26**, 1405–1410 (2020).
33. Gardner, L., Ratcliff, J., Dong, E. & Katz, A. A need for open public data standards and sharing in light of COVID-19. *Lancet Infect. Dis.* **21**, E80 (2021).
34. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
35. Nasir, J. A. et al. A comparison of whole genome sequencing of SARS-CoV-2 using amplicon-based sequencing, random hexamers, and bait capture. *Viruses* **12**, 895 (2020).
36. Posada-Céspedes, S. et al. V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data. *Bioinformatics* **37**, 1673–1680 (2021).
37. International Air Transport Association. <https://www.iata.org/pages/default.aspx> (2021).
38. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
39. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
40. Rambaut, A. et al. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).
41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. J Mol Biol Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
42. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
43. Tavaré, S. In *Some Mathematical Questions in Biology: DNA Sequence Analysis*. (ed. Waterman, M. S.) 57–86 (American Mathematical Society, Providence (RI), 1986).
44. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).
45. Baele, G., Suchard, M. A. & Lemey, P. Genealogical working distributions for Bayesian model testing with phylogenetic uncertainty. *Syst. Biol.* **65**, 250–264 (2016).
46. Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).
47. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
48. Baele, G., Gill, M. S., Lemey, P. & Suchard, M. A. Hamiltonian Monte Carlo sampling to estimate past population dynamics using the skygrid coalescent model in a Bayesian phylogenetics framework [version 1; peer review: 1 approved, 2 approved with reservations]. *Wellcome Open Res.* **5**, 53 (2020).
49. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
50. Lauer, S. A. et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
51. Lauer, S. A. et al. *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand*. Technical report (Imperial College, London, 2020).
52. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**, e1000520 (2009).
53. Ferreira, M. A. R. & Suchard, M. A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **26**, 355–368 (2008).
54. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
55. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
56. Hasegawa, M., Kishino, H. & Yano, T.-a. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).
57. Ronquist, F. et al. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).

Acknowledgements

We gratefully acknowledge the authors from originating laboratories responsible for obtaining the specimens, as well as submitting laboratories where the genome data were generated and shared via GISAID, on which this research is based. An acknowledgement table with GISAID accession IDs of SARS-CoV-2 genomes used here is included. We thank all involved in the collection and processing of SARS-CoV-2 testing and genomic data, as well as associated metadata on individual travel histories. In particular, we would like to thank Marc Noguera Julian, Elisa Martro Catala, Samuel Cordey, Piet Maes, Keith Durkin, Bruno Verhasselt, Lize Cuyppers, Lien Cattoir, Veerle Matheussen, Vincent Enouf, Sylvie van der Werf, Etienne Simon-Lorière, Tobias Schindler, Vladimira Koudelakova, Gabriel Gonzalez, Ariane Düx, Yanthe Nobel, Livia Patrono, Justas Dapkūnas and Andrew J. Tatem. We would like to thank Richard Neher and Kristian G. Andersen for thoughtful discussions. S.L.H. acknowledges support from the Research Foundation - Flanders (‘Fonds voor Wetenschappelijk Onderzoek - Vlaanderen,’ G0D5117N). B.P. and G.B. acknowledge support from the Internal Fondsen KU Leuven/Internal Funds KU Leuven (Grant No. C14/18/094). G.B. acknowledges support from the Research Foundation - Flanders (‘Fonds voor Wetenschappelijk Onderzoek - Vlaanderen,’ G0E1420N, G098321N). F.H.L. and T.F.N. were supported by WWF and German Research Council’s grant LE1813/14-1 (Great Ape Health in Tropical Africa), Research in CAR took place under permit #098/MRSIT/DIRCAB/CB.20, granted to T.T. by the Ministry of Scientific Research and Technological Innovation. J.S. acknowledges funding from the Dutch Research Council NWO Gravitation 2013 BOO, Institute for Chemical Immunology (ICI; 024.002.009). A.M.J.J.B. acknowledges the support of European Union Horizon 2020 projects BioExcel (823830) and EOSC-Hub (777536) projects. A.A. and C.B. acknowledge the support of the French National Research Institute for Sustainable Development (IRD).

Author contributions

G.D.—conceptualisation, methodology, formal analysis, investigation, resources, data curation, writing—original draft, visualisation, supervision, project administration, funding acquisition; S.L.H.—formal analysis, data curation, writing—original draft; B.I.P.—formal analysis, data curation, writing—original draft; S.C.-S.—investigation, resources, data curation, writing—review and editing; F.S.N.-S.—investigation, resources; T.B.T.—investigation, resources; T.F.-N.—investigation, resources; U.V.—investigation, resources; M.U.—investigation, resources; F.H.L.—investigation, resources, data curation, writing—review and editing; K.K.—investigation, resources, data curation; C.H.—investigation, resources, data curation; A.W.—investigation, resources, data curation; I.O.—formal analysis, investigation, resources, data curation, writing—review and

editing, project administration, funding acquisition; J.S.—formal analysis, investigation, resources, data curation, writing—original draft, visualisation; K.N.W.—formal analysis, investigation, resources, data curation, writing—original draft, visualisation; A.M.J.J.B.—formal analysis, investigation, resources, data curation, writing—original draft, visualisation; P.M.—resources; S.B.—resources; A.A.—resources; M.F.M.—resources; D.M.D.—resources; C.G.—resources; C.B.—resources; A.S.—investigation, resources; M.G.—resources, data curation, writing—review and editing; M.K.—resources; R.N.—resources; L.R.—resources; G.W.K.—resources; J.K.K.—resources; R.J.—resources; I.N.—resources; Ž.Z.—resources; D.G.—resources; K.T.—resources; M.N.—resources; E.V.—resources, writing—review and editing; D.Ž.—resources; A.T.—resources; M.S.—resources; M.S.—resources; G.A.—resources; A.A.A.—resources; E.K.L.—resources; J.-C.M.C.—resources; F.M.M.—resources; E.L.L.—resources; P.M.K.—resources; J.-J.M.T.—resources; M.R.D.B.—resources; R.G.E.—resources; M.C.O.A.—resources; A.B.M.—resources; A.B.D.—resources; D.J.—resources, funding acquisition; S.H.—resources; J.O.—resources; M.O.A.—resources; D.N.—resources, funding acquisition; A.P.—resources; C.D.R.—resources; A.V.—resources; R.U.—resources; A.G.—resources; D.Č.—investigation, resources; V.L.—resources; L.Ž.—project administration; L.G.—investigation, resources, funding acquisition; G.B.—conceptualisation, methodology, formal analysis, investigation, resources, data curation, writing—original draft, supervision, project administration.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-26055-8>.

Correspondence and requests for materials should be addressed to Gytis Dudas or Guy Baele.

Peer review information *Nature Communications* thanks Damien Tully and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Gytis Dudas^{1,2,40}✉, Samuel L. Hong³, Barney I. Potter³, Sébastien Calvignac-Spencer^{4,5}, Frédéric S. Niatou-Singa⁶, Thais B. Tombolomako⁶, Terence Fuh-Neba⁶, Ulrich Vickos^{7,8}, Markus Ulrich⁴, Fabian H. Leendertz⁴, Kamran Khan^{9,10,11}, Carmen Huber⁹, Alexander Watts⁹, Ingrida Olendraitė^{2,12}, Joost Snijder¹³, Kim N. Wijnant¹³, Alexandre M.J.J. Bonvin¹⁴, Pascale Martres¹⁵, Sylvie Behillil^{16,17}, Ahidjo Ayouba¹⁸, Martin Foudi Maidadi¹⁹, Dowbiss Meta Djomsi¹⁹, Celestin Godwe¹⁹, Christelle Butel¹⁸, Aistis Šimaitis²⁰, Miglė Gabrielaitė²¹, Monika Katėnaitė², Rimvydas Norvilas^{2,22}, Ligita Raugaitė², Giscard Wilfried Koyaweda²³, Jephthé Kaleb Kandou²³, Rimvydas Jonikas²⁴, Inga Nasvytienė²⁴, Živilė Žemeckienė²⁴, Dovydas Gečys²⁵, Kamilė Tamušauskaitė²⁵, Milda Norkienė²⁶, Emilija Vasilūnaitė²⁶, Danguolė Žiogienė²⁶, Albertas Timinskas²⁶, Marius Šukys^{24,27}, Mantas Šarauskas²⁴, Gediminas Alzbutas²⁸, Adrienne Amuri Aziza^{29,30}, Eddy Kinganda Lusamaki^{29,30}, Jean-Claude Makangara Cigolo^{29,30}, Francisca Muyembe Mawete^{29,30}, Emmanuel Lokilo Lofiko²⁹, Placide Mbala Kingebeni^{29,30}, Jean-Jacques Muyembe Tamfum^{29,30}, Marie Roseline Darnycka Belizaire³¹, René Ghislain Essomba^{32,33}, Marie Claire Okomo Assoumou^{32,33}, Akenji Blaise Mboringong³², Alle Baba Dieng³⁴, Dovilė Juozapaitė², Salome Hosch³⁵, Justino Obama³⁶, Mitoha Ondo'o Ayekaba³⁶, Daniel Naumovas², Arnoldas Pautienius³⁷, Clotaire Donatien Rafai²³, Astra Vitkauskienė³⁸, Rasa Ugenskienė^{24,27}, Alma Gedvilaitė²⁶, Darius Čereškevičius^{24,25}, Vaiva Lesauskaitė²⁵, Lukas Žemaitis^{25,39}, Laimonas Griškevičius² & Guy Baele^{3,40}✉

¹Gothenburg Global Biodiversity Centre, Gothenburg, Sweden. ²Hematology, Oncology and Transfusion Medicine Center, Vilnius University Hospital Santaros Klinikos, Vilnius, Lithuania. ³Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium. ⁴Epidemiology of Highly Pathogenic Organisms, Robert Koch Institute, 13353 Berlin, Germany. ⁵Viral Evolution, Robert Koch Institute, 13353 Berlin, Germany. ⁶WWF Central African Republic Programme Office, Dzanga Sangha Protected Areas, BP 1053 Bangui, Central African Republic. ⁷Infectious and Tropical Diseases Unit, Department of medicine, Amitié Hospital, Bangui, Central African Republic. ⁸Academic Department of Pediatrics, Clinical immunology and vaccinology, Children's Hospital Bambino Gesù, IRCCS, Rome, Italy. ⁹BlueDot, Toronto, ON M5J 1A7, Canada. ¹⁰Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON M5B 1A6, Canada. ¹¹Division of Infectious Diseases, Department of Medicine, University of Toronto, Toronto, ON M5S 3H2, Canada. ¹²Division of Virology, Department of Pathology, University of Cambridge, Addenbrooke's Hospital Lab, CB2 2QQ Cambridge, UK. ¹³Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for

Biomolecular Research and Utrecht Institute of Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. ¹⁴Bijvoet Centre for Biomolecular Research, Faculty of Science - Chemistry, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. ¹⁵Microbiology, Centre Hospitalier René Dubos, Cergy Pontoise, France. ¹⁶Molecular Genetics of RNA viruses, CNRS UMR 3569, Université de Paris, Institut Pasteur, Paris, France. ¹⁷National Reference Center for Respiratory Viruses, Institut Pasteur, Paris, France. ¹⁸TransVIHMI, Université de Montpellier, IRD, INSERM, 911 Avenue Agropolis, 34394 Montpellier cedex, France. ¹⁹Centre de Recherches sur les Maladies Émergentes, Ré-émérgentes et la Médecine Nucléaire, Institut de Recherches Médicales et D'études des Plantes Médicinales, Yaoundé, Cameroon. ²⁰The Office of the Government of the Republic of Lithuania, Vilnius, Lithuania. ²¹Center for Genomic Medicine, Rigshospitalet, Copenhagen, Denmark. ²²Department of Experimental, Preventive and Clinical Medicine, State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania. ²³Le Laboratoire National de Biologie Clinique et de Santé Publique (LNBCSP), Bangui, Central African Republic. ²⁴Department of Genetics and Molecular Medicine, Hospital of Lithuanian University of Health Sciences Kauno Klinikos, Kaunas, Lithuania. ²⁵Institute of Cardiology, Lithuanian University of Health Sciences, Kaunas, Lithuania. ²⁶Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania. ²⁷Department of Genetics and Molecular Medicine, Lithuanian University of Health Sciences, Kaunas, Lithuania. ²⁸Institute for Digestive Research, Lithuanian University of Health Sciences, Kaunas, Lithuania. ²⁹National Institute for Biomedical Research (INRB), Avenue De la Democratie (Ex Huileries), BP 1197 Kinshasa-Gombe, Democratic Republic of the Congo. ³⁰University of Kinshasa (UNIKIN), BP 127 Kinshasa XI, Democratic Republic of the Congo. ³¹World Health Organization, Central African Republic Office, Bangui, Central African Republic. ³²National Public Health Laboratory, Ministry of Public Health, Yaoundé, Cameroon. ³³Faculty of Medicine and Biomedical Sciences, University of Yaoundé I, Yaoundé, Cameroon. ³⁴World Health Organization, Cameroon Office, Yaoundé, Cameroon. ³⁵Swiss Tropical and Public Health Institute, Basel, Switzerland. ³⁶Ministry of Health and Social Welfare, Malabo, Equatorial Guinea. ³⁷Institute of Microbiology and Virology, Lithuanian University of Health Sciences, Kaunas, Lithuania. ³⁸Department of Laboratory Medicine, Lithuanian University of Health Sciences, Kaunas, Lithuania. ³⁹National Public Health Surveillance Laboratory, Vilnius, Lithuania. ⁴⁰These authors contributed equally: Gytis Dudas, Guy Baele. ✉email: gytis.dudas@gmail.com; guy.baele@kuleuven.be

2.3. Genomic Surveillance Enables the Identification of Co-infections with Multiple SARS-CoV-2 Lineages in Equatorial Guinea

Published in *Frontiers in Public Health*, 2022



Genomic Surveillance Enables the Identification of Co-infections With Multiple SARS-CoV-2 Lineages in Equatorial Guinea

Salome Hosch^{1,2*†}, Maxmillian Mpina^{1,2,3†}, Elizabeth Nyakurungu³, Nelson Silochi Borico³, Teodora Mikumu Alogo Obama³, Maria Carmen Ovona³, Philipp Wagner^{1,2}, Sarah E. Rubin^{1,2}, Ulrich Vickos^{4,5}, Diosdado Vicente Nsue Milang⁶, Mitoha Ondo'o Ayekaba⁶, Wonder P. Phiri⁷, Claudia A. Daubenberger^{1,2*} and Tobias Schindler^{1,2}

OPEN ACCESS

Edited by:

Sanjay Kumar,
Armed Forces Medical College,
Pune, India

Reviewed by:

Hirawati Deval,
ICMR-Regional Medical Research
Centre, India
Jayanthi S. Shastri,
Brihanmumbai Municipal
Corporation, India

*Correspondence:

Salome Hosch
salome.hosch@swisstph.ch
Claudia A. Daubenberger
claudia.daubenberger@swisstph.ch

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Infectious Diseases - Surveillance,
Prevention and Treatment,
a section of the journal
Frontiers in Public Health

Received: 19 November 2021

Accepted: 30 November 2021

Published: 04 January 2022

Citation:

Hosch S, Mpina M, Nyakurungu E,
Borico NS, Obama TMA, Ovona MC,
Wagner P, Rubin SE, Vickos U,
Milang DVN, Ayekaba MO, Phiri WP,
Daubenberger CA and Schindler T
(2022) Genomic Surveillance Enables
the Identification of Co-infections With
Multiple SARS-CoV-2 Lineages in
Equatorial Guinea.
Front. Public Health 9:818401.
doi: 10.3389/fpubh.2021.818401

¹ Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland, ² University of Basel, Basel, Switzerland, ³ Laboratorio de Investigaciones de Baney, Baney, Equatorial Guinea, ⁴ Infectious and Tropical Diseases Unit, Department of Medicine, Amitié Hospital, Bangui, Central African Republic, ⁵ Academic Department of Pediatrics, Clinical Immunology and Vaccinology, Children's Hospital Bambino Gesù, Scientific Institute for Research, Hospitalization and Healthcare (IRCCS), Rome, Italy, ⁶ Ministry of Health and Social Welfare, Malabo, Equatorial Guinea, ⁷ Medical Care Development International, Malabo, Equatorial Guinea

COVID-19 disease caused by SARS-CoV-2 represents an ongoing global public health emergency. Rapid identification of emergence, evolution, and spread of SARS-CoV-2 variants of concern (VOC) would enable timely and tailored responses by public health decision-making bodies. Yet, global disparities in current SARS-CoV-2 genomic surveillance activities reveal serious geographical gaps. Here, we discuss the experiences and lessons learned from the SARS-CoV-2 monitoring and surveillance program at the Public Health Laboratory on Bioko Island, Equatorial Guinea that was implemented as part of the national COVID-19 response and monitoring activities. We report how three distinct SARS-CoV-2 variants have dominated the epidemiological situation in Equatorial Guinea since March 2020. In addition, a case of co-infection of two SARS-CoV-2 VOC, Beta and Delta, in a clinically asymptomatic and fully COVID-19 vaccinated man living in Equatorial Guinea is presented. To our knowledge, this is the first report of a person co-infected with Beta and Delta VOC globally. Rapid identification of co-infections is relevant since these might provide an opportunity for genetic recombination resulting in emergence of novel SARS-CoV-2 lineages with enhanced transmission or immune evasion potential.

Keywords: SARS-CoV-2, co-infection, variant of concern, genomic surveillance, Central-Africa

INTRODUCTION

Whole genome sequencing of SARS-CoV-2 viruses has been widely used since the beginning of the COVID-19 pandemic to facilitate understanding of virus biology and epidemiology. The World Health Organization (WHO) recommends that countries ship at least 5% of their COVID-19 samples to reference sequencing laboratories or keep producing sequencing data if they have the capacity (1). Angola, Burundi, Cameroon, the Central African Republic, Chad, the Democratic Republic of the Congo, the Republic of the Congo, Equatorial Guinea, Gabon, Rwanda, and

São Tomé and Príncipe are the 11 Central-African nations forming the Economic Community of Central African States (ECCAS). Combined, the ECCAS nations have deposited 3,924 SARS-CoV-2 whole genome sequences to GISAID and therefore have sequenced on average 0.9% of all reported cases. The proportion of sequenced cases from Central Africa over the time course of the COVID-19 pandemic reveals that with the exception of March and April 2020, the recommended sequencing rate of 5% could not be achieved.

A cost-efficient alternative to whole genome sequencing are multiplex reverse transcription quantitative polymerase chain reactions (RT-qPCR) assays which detect relevant mutations associated with SARS-CoV-2 variants of concern (VOC) (2–4). Mutation-specific assays can complement genomic surveillance programs, especially in settings where widespread sequencing capabilities are not available. If carefully designed and evaluated these kind of assays show a perfect concordance with whole genome sequencing as reported elsewhere for monitoring distinct VOC (5, 6). Mutation-specific RT-qPCR assays allow identification of patients with co-infections by more than one SARS-CoV-2 lineage simultaneously. The frequency of co-infected humans and their role in promoting SARS-CoV-2 evolution is poorly understood (7). Co-infections might provide an opportunity for genetic recombination between circulating strains resulting in the emergence of novel SARS-CoV-2 lineages (8). Interlineage recombination has been described for SARS-CoV-2 (9) as well as for other closely related viruses of the *Coronaviridae* family (10, 11).

In this brief research report, we describe the experiences and lessons learned from the SARS-CoV-2 genomic surveillance program at the Public Health Laboratory on Bioko Island, which was implemented as part of the national COVID-19 response and monitoring activities in Equatorial Guinea. We describe our approach to identify efficiently and timely SARS-CoV-2 VOC and co-infections in Sub-Saharan Africa, which is highly neglected when it comes to global genomic surveillance activities.

METHODS

COVID-19 Datasets for Central-Africa

The number of confirmed COVID-19 cases and SARS-CoV-2 whole genome sequences from Central-Africa were obtained through the data repositories of “Our World in Data” (OWD) (12) and the “Global Initiative on Sharing All Influenza Data” (GISAID) (13), respectively on November 11th 2021.

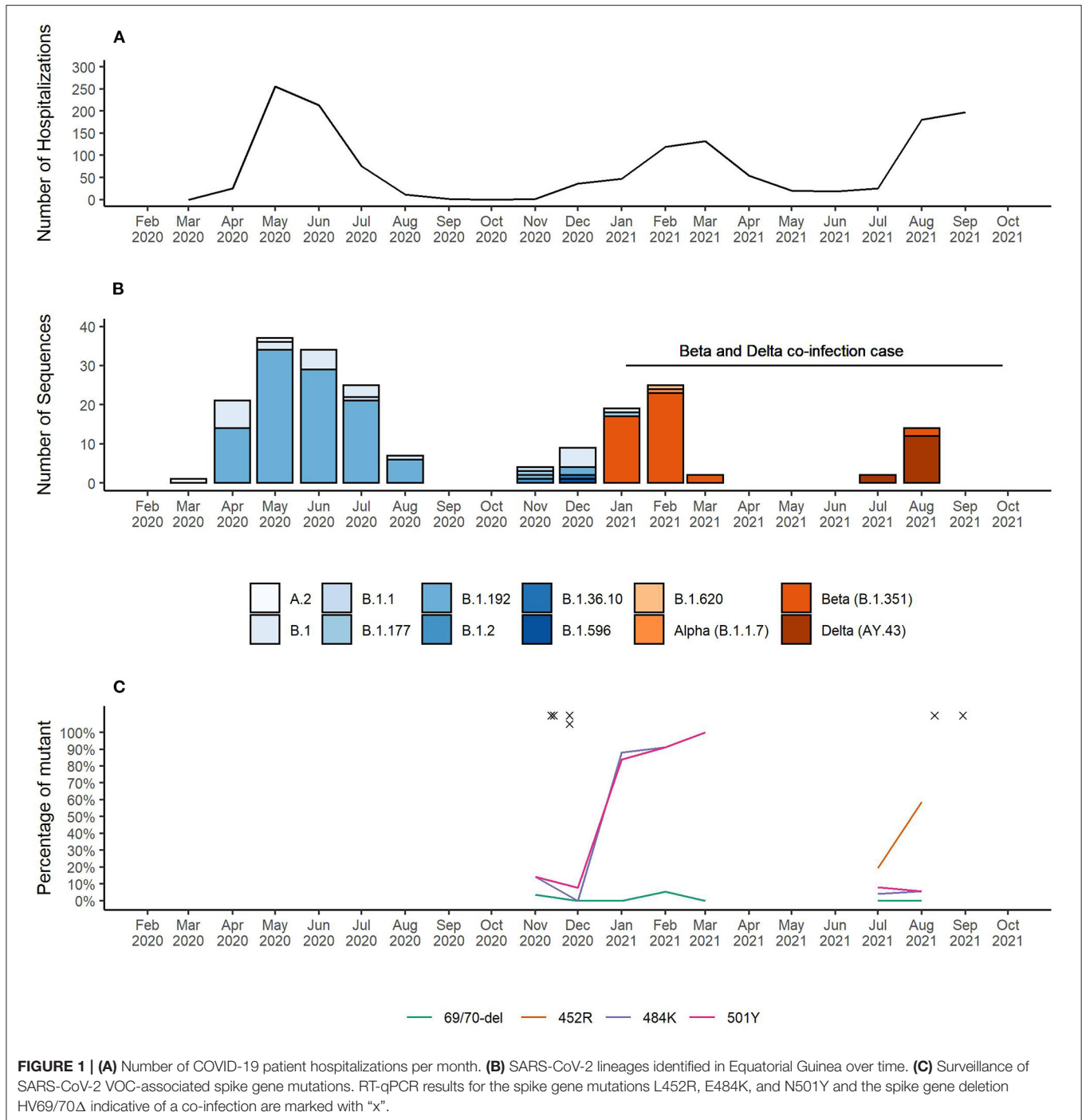
SARS-CoV-2 Genomic Surveillance in Equatorial Guinea

Nasopharyngeal and/or oropharyngeal swab samples were collected between February 2020 and October 2021 under the umbrella of the current Equatorial Guinea SARS-CoV-2 surveillance activities. Samples that are processed at the Public Health Laboratory on Bioko Island include asymptomatic and symptomatic cases, as well as samples derived from contact tracing. Extracted RNA aliquots from positive samples are collected and stored in a local biobank at -80°C . A randomly selected subset of SARS-CoV-2 samples with $C_q < 30$ are

analyzed by spike gene mutation-specific RT-qPCR assays ($n = 281$) and/or by whole genome sequencing ($n = 206$). The spike gene mutation-specific RT-qPCR assays target three spike gene single nucleotide polymorphisms (SNPs) (L452R, E484K, and N501Y) and one spike gene deletion ($\Delta 69/70$) associated with VOCs (5, 14). In each of the four multiplex RT-qPCR assays, probes detecting the wildtype as well as probes detecting the mutant nucleotide sequences are used to enable detection of wildtype and mutated sequence simultaneously. Co-infections, with more than one single SARS-CoV-2 lineage, were defined as samples with more than one genotype in at least two of the spike gene markers analyzed. SARS-CoV-2 whole genome sequencing was conducted using the R9.4.1 flow cell on a MinION Mk1C device (Oxford Nanopore Technologies) based on the ARTIC protocol (15).

RESULTS

As of November 25th 2021, Equatorial Guinea has reported a total of 13,579 confirmed SARS-CoV-2 infections, which resulted in 173 deaths. The majority (77.2%) of SARS-CoV-2 infections were reported from the insular region (Bioko Island) and 23.8% from the continental region (Río Muni) (<https://guineasalud.org/estadisticas/>). Since the first case was identified on March 16th 2020, the country has experienced three distinct epidemic waves characterized by an increase of COVID-19-related hospitalizations (**Figure 1A**). Continuous sequencing of 206 SARS-CoV-2 positive samples revealed that each of these epidemic waves were dominated by a distinct lineage (**Figure 1B**). We showed that the first wave lasting from April to July 2020 was dominated by wildtype-like SARS-CoV-2 lineage B.1.192. The introduction of the Beta VOC (B.1.351) caused a second wave lasting from January to April 2021. The first Delta VOC (AY.43) positive cases in Equatorial Guinea were identified in the first half of July 2021 leading to a severe third wave that is still ongoing at the time of this report. Interestingly, the Alpha VOC was found only in a single sample with no indication of its widespread circulation in Equatorial Guinea. Starting from November 2020, mutation specific RT-qPCR assays were implemented to complement whole genome sequencing for enhanced genomic surveillance (**Figure 1C**). Three spike gene SNPs (L452R, E484K, and N501Y) and one spike gene deletion (HV69/70 Δ) associated with VOCs were monitored. In each of the four multiplex RT-qPCR assays, a probe detecting the wildtype as well as the mutant nucleotide sequences are used. The 484K+501Y combination of spike gene mutation, which is associated with the Beta VOC, were first observed in November 2020 and became dominant by January 2021. This combination was later replaced by viruses with the 452R spike gene mutation, a marker for the Delta VOC. Overall, 2.1% (6/281) of samples analyzed with the spike gene mutation-specific RT-qPCR assays showed a pattern indicating co-infections of two distinct variants (**Table 1**). The four cases identified in November 2020, included three that had a combination indicative of co-infection between a wildtype-like lineage and the Beta VOC (wildtype and mutant specific amplification signals at spike



gene positions E484K and N501Y). The fourth case showed a pattern that corresponds to a co-infection between a wildtype-like and a B.1.620-like lineage (wildtype and mutant specific amplification at spike gene positions 69/70 and E484K). Two additional co-infections were observed in August 2021, both of them with spike gene mutation patterns indicative of Beta and Delta VOC co-infections (wildtype and mutant specific amplification at spike gene positions L452R, E484K, and N501Y).

All six co-infections were identified during the transition phase between two consecutive epidemic waves in which new variants emerged while other variants were still circulating.

The receptor-binding domain of the spike protein constitutes the immunodominant target of 90% of the neutralizing activity present in SARS-CoV-2 immune sera (16). Importantly, all co-infections included one SARS-CoV-2 variant carrying the receptor-binding domain mutation E484K that has

TABLE 1 | SARS-CoV-2 co-infections identified in Equatorial Guinea.

Patient	Date of swab collection	Spike gene mutation specific RT-qPCR assay results				Potential lineages involved in co-infection*
		HV69/70Δ	L452R	E484K	N501Y	
EG-SARS-COV-2-P1	12/11/2020	HV69/70	NA	E484 + 484K	N501 + 501Y	Wildtype (e.g., B.1.192) + Beta VOC (B.1.351)
EG-SARS-COV-2-P2	14/11/2020	HV69/70	NA	E484 + 484K	N501 + 501Y	Wildtype (e.g., B.1.192) + Beta VOC (B.1.351)
EG-SARS-COV-2-P3	25/11/2020	HV69 + 69/70Δ	NA	E484 + 484K	N501	Wildtype (e.g., B.1.192) + B.1.620
EG-SARS-COV-2-P4	25/11/2020	HV69/70	NA	E484 + 484K	N501 + 501Y	Wildtype (e.g., B.1.192) + Beta VOC (B.1.351)
EG-SARS-COV-2-P5	10/08/2021	HV69/70	L452 + 452R	E484 + 484K	N501 + 501Y	Beta VOC (B.1.351) + Delta VOC (AY.43)
EG-SARS-COV-2-P6	30/08/2021	HV69/70	L452 + 452R	E484 + 484K	N501 + 501Y	Beta VOC (B.1.351) + Delta VOC (AY.43)

*Potential lineages involved in co-infections were proposed based on spike gene mutations and known circulating variants at the time.

been associated with immune evasion in polyclonal human antibodies (17, 18).

For the EG-SARS-CoV-2-P6 co-infection case, which had the highest viral load among all co-infected patients (spike gene RT-qPCR assay Cq-value of 20.2), we were able to generate whole genome sequencing data for further investigation. The timeline of this co-infection case is shown in **Figure 2A**. While living under quarantine and after having been in contact with a positive SARS-CoV-2 case, a 59-year-old Equatoguinean man was tested positive by RT-qPCR for SARS-CoV-2 on August 30th 2021. According to his vaccination certificate, he had received the first dose of Sinopharm COVID-19 vaccine (Beijing Bio-Institute of Biological Products Co., Ltd.) on April 20th 2021 and the second dose on May 5th 2021. This person had no history of clinically significant underlying medical conditions, did not report any clinical symptom during the entire period of virus infection and has reported no travel history. As part of our routine contact tracing and SARS-CoV-2 surveillance activities his sample was analyzed for the presence of three spike gene mutations associated with VOCs. The mutation-specific RT-qPCR results for the spike gene mutations are shown for L452R (**Figure 2B**), E484K (**Figure 2C**), and N501Y (**Figure 2D**). For all three SNPs, the wildtype as well as the mutated sequence were amplified simultaneously. An earlier and stronger amplification was observed for the mutations associated with the Beta VOC compared to the Delta VOC. In the HV69/70Δ assay, targeting a spike gene deletion common in the Alpha VOC, only the wildtype target was amplified, as neither Beta nor Delta VOC carry the deletion at positions 69 and 70 of the spike gene (data not shown).

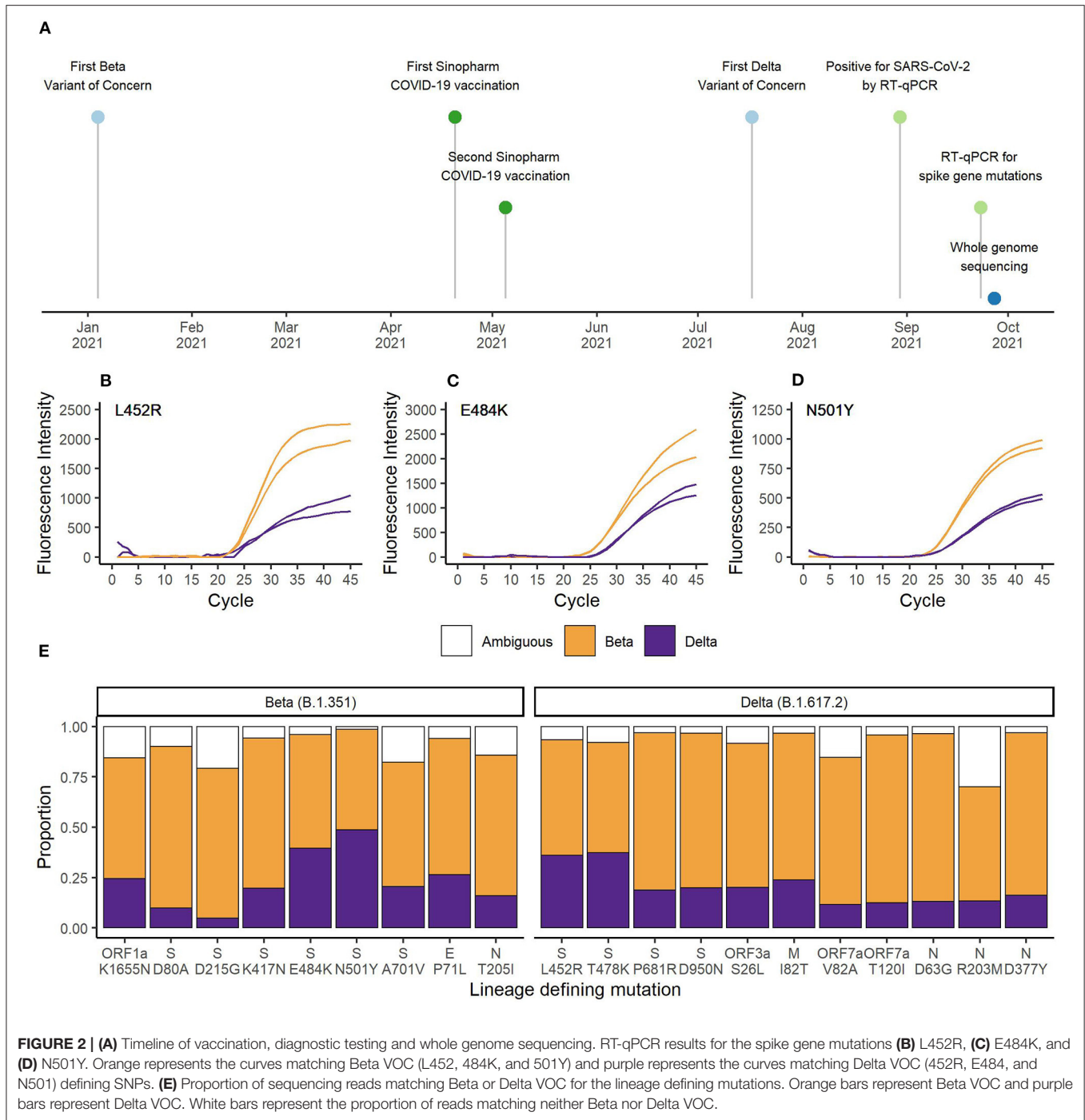
Next, we conducted whole genome sequencing of this particular sample using our MinION sequencing platform. A total of 27,932 reads passed quality control and we achieved an average sequencing depth of 203x. Sequencing quality and depth was sufficient to identify all nine lineage-defining mutations of Beta and 11 of 12 lineage-defining mutations of Delta VOC (19). There was insufficient coverage of the spike gene SNP T19R. The average percentage of reads matching Beta VOC was 69.2% (range 50.0–83.3%) and the average percentage of reads matching Delta VOC was 21.6% (range 4.9–48.6%). Ambiguities represented 9.2% (range 1.4–30.0%) of reads. As shown in **Figure 2E**, the number of

reads containing the mutations associated with Delta VOC was lower compared to the reads with sequences associated with Beta VOC. The higher Beta to Delta VOC ratio of RNA sample input was also observed in the RT-qPCR data (**Figures 2B–D**).

DISCUSSION

Large scale genome sequencing has become a critical part of SARS-CoV-2 surveillance globally. Yet, comprehensive analysis of genome sequences deposited to GISAID highlight stark global disparities with only 6% of genome sequences derived from low and middle income countries (20). In settings with limited access to sequencing capabilities, targeted sampling with low coverage of positive cases might be a possible alternative to representative sampling which would enable population-based surveillance. The Equato-Guinean Public Health Laboratory located in the Baney district of Bioko Island uses a hybrid approach combining rapidly adaptable spike gene mutation-specific RT-qPCR assays and whole genome sequencing to monitor the emergence, evolution, and spread of SARS-CoV-2 lineages in the country. Selection of samples for whole genome sequencing can be guided by the results of the mutation-specific RT-qPCR assays and has successfully led in the past to the discovery of the novel lineage B.1.620 (21). Additionally, results from mutation-specific RT-qPCR assays can be generated in real-time in standard local diagnostic laboratories and therefore genomic surveillance becomes actionable. Before implementing mutation-specific RT-qPCR assays, it took more than 8 weeks between sample collection and providing the information to the Ministry of Health that the Beta VOC is circulating in Equatorial Guinea based on the whole genome sequencing approach. By this time, the Beta VOC was already the dominant lineage. By using the L452R-specific RT-qPCR assay, the time from sample collection to providing the information of potential introduction of the Delta VOC was significantly reduced to <1 week. The first Delta VOC whole genome sequence was generated and uploaded to GISAID 8 weeks later.

We propose that in locations where SARS-CoV-2 sequencing capacity are limited and difficult to sustain, sequencing resources might be better utilized when focusing on phases in-between epidemiological waves which are characterized by



lower infection rates. This might provide essential information on the introduction or emergence of new variants that might soon dominate the ensuing COVID-19 wave. Increase of sequencing efforts during times of high infection rates in the population might not yield novel sequencing information based on the strong dominance of distinct variants at that time.

We describe here the case of an asymptotically infected adult male carrying two distinct lineages of SARS-CoV-2 that are both VOCs. The Beta VOC was first described in South Africa

in October 2020 and has become the dominant lineage in many African countries by March 2021 (22). The Delta VOC was first described in India in October 2020 and has become the dominant lineage worldwide by August 2021 (23). Both VOC lineages have been circulating in Equatorial Guinea at the time when this co-infection was identified. Co-infection events between dominant SARS-CoV-2 lineages have been previously reported from the USA in 0.18% (53/29,993) of sequenced samples (24) and Brazil in 0.61% (9/1,462) of investigated samples (7). To our knowledge,

this is the first time that a co-infection between the Beta and Delta VOCs is reported globally. Recombination among *Coronaviridae*, including SARS-CoV-2, has been described as an important evolutionary mechanism underlying genetic shift (9, 25). Newly recombined viruses might result in increased transmissibility or immune evasion and therefore continuous genomic monitoring of SARS-CoV-2 lineages are warranted (8, 22). Strengthening and continuous support of public health laboratories in Sub-Saharan countries to avoid underreporting of cases and enhance detection of emerging variants is a prerequisite for successful and global SARS-CoV-2 containment (20).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA770861; <https://www.gisaid.org/>, EPI_ISL_1672551–EPI_ISL_1672557, EPI_ISL_1673311–EPI_ISL_1673331, EPI_ISL_1700674–EPI_ISL_1700687, EPI_ISL_1753020–EPI_ISL_1753026, EPI_ISL_1989331, EPI_ISL_1989335, EPI_ISL_1989339, EPI_ISL_2002669–EPI_ISL_2002688, EPI_ISL_4601590–EPI_ISL_4601605, EPI_ISL_648303–EPI_ISL_648379, EPI_ISL_649155–EPI_ISL_649172, EPI_ISL_953402–EPI_ISL_953425.

ETHICS STATEMENT

Sample collection and retrospective sequence analysis was conducted according to the guidelines of the Declaration of Helsinki and approved by the National Technical Committee for the Response and Monitoring of the Novel Coronavirus (Comité Técnico Nacional de Respuesta y Vigilancia del Nuevo

Coronavirus), which is charged with preventing, containing, controlling, tracking, and evaluating the development and evolution of COVID-19 in Equatorial Guinea. Informed consent for publishing the Beta and Delta VOC co-infection case was obtained from the patient. Publication of the SARS-CoV-2 epidemiological, genomic surveillance data and the co-infection case was additionally approved by the Equato-Guinean Ministry of Health and Social Welfare.

AUTHOR CONTRIBUTIONS

TS and CD: conceptualization, supervision, and project administration. SH, MM, PW, SR, and TS: methodology. SH, PW, and TS: software and validation. SH and TS: formal analysis, data curation, and visualization. SH, MM, EN, NB, TO, MO, PW, and UV: investigation. DM, MA, and WP: resources. SH, CD, and TS: writing—original draft preparation. All authors have read and agreed to the submitted version of the manuscript.

FUNDING

The funding for this work was provided through the public-private partnership, the Equatorial Guinea Malaria Vaccine Initiative, supported by the Government of Equatorial Guinea, Ministries of Mines and Hydrocarbons, and Health and Social Welfare, Marathon Equatorial Guinea Production Limited, Noble Energy, Atlantic Methanol, Production Company, and the Equatorial Guinea Liquefied Natural Gas Company. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

REFERENCES

1. *Scaling Up Genomic Sequencing in Africa*. WHO. Regional Office for Africa. Available online at: <https://www.afro.who.int/news/scaling-genomic-sequencing-africa> (accessed November 13, 2021).
2. Boršová K, Paul ED, Kováčová V, Radvánszka M, Hajdu R, Cabanová V, et al. Surveillance of SARS-CoV-2 lineage B.1.1.7 in Slovakia using a novel, multiplexed RT-qPCR assay. *Sci Rep.* (2021) 11:20494. doi: 10.1038/s41598-021-99661-7
3. Vega-Magaña N, Sánchez-Sánchez R, Hernández-Bello J, Venancio-Landeros AA, Peña-Rodríguez M, Vega-Zepeda RA, et al. RT-qPCR assays for rapid detection of the N501Y, 69-70del, K417N, and E484K SARS-CoV-2 mutations: a screening strategy to identify variants with clinical impact. *Front Cell Infect Microbiol.* (2021) 11:434. doi: 10.3389/fcimb.2021.672562/BIBTEX
4. Vogels CBF, Breban MI, Ott IM, Alpert T, Petrone ME, Watkins AE, et al. Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *PLoS Biol.* (2021) 19:e3001236. doi: 10.1371/JOURNAL.PBIO.3001236
5. Bechtold P, Wagner P, Hosch S, Siegrist D, Ruiz-serrano A, Gregorini M, et al. Rapid identification of SARS-CoV-2 variants of concern using a portable peak PCR platform. *Anal Chem.* (2021). doi: 10.1021/acs.analchem.1c02368
6. Wang H, Jean S, Eltringham R, Madison J, Snyder P, Tu H, et al. Mutation-specific SARS-CoV-2 PCR screen: rapid and accurate detection of variants of concern and the identification of a newly emerging variant with spike L452R mutation. *J Clin Microbiol.* (2021) 59:e0092621. doi: 10.1128/JCM.00926-21
7. Dezordi FZ, Resende PC, Naveca FG, do Nascimento VA, de Souza VC, Paixão ACD, et al. Unusual SARS-CoV-2 intra-host diversity reveals lineages superinfection. *medRxiv.* (2021) 2021.09.18.21263755. doi: 10.1101/2021.09.18.21263755
8. Otto SP, Day T, Arino J, Colijn C, Dushoff J, Li M, et al. The origins and potential future of SARS-CoV-2 variants of concern in the evolving COVID-19 pandemic. *Curr Biol.* (2021) 31:R918–29. doi: 10.1016/j.cub.2021.06.049/ATTACHMENT/6F1FD272-67F8-4D33-9F65-7122149111EC/MMC1.PDF
9. Jackson B, Boni MF, Bull MJ, Collier A, Colquhoun RM, Darby AC, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell.* (2021) 184:5179.e8–88.e8. doi: 10.1016/j.cell.2021.08.014
10. Terada Y, Matsui N, Noguchi K, Kuwata R, Shimoda H, Soma T, et al. Emergence of pathogenic coronaviruses in cats by homologous recombination between feline and canine coronaviruses. *PLoS ONE.* (2014) 9:e106534. doi: 10.1371/JOURNAL.PONE.0106534
11. Sabir JSM, Lam TTY, Ahmed MMM, Li L, Shen Y, Abo-Aba SEM, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science.* (2016) 351:81–4. doi: 10.1126/SCIENCE.AAC8608
12. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Coronavirus pandemic (COVID-19). *Our World Data* (2020). Retrieved from: <https://ourworldindata.org/coronavirus>

13. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*. (2017) 22:30494. doi: 10.2807/1560-7917.ES.2017.22.13.30494/CITE/PLAINTEXT
14. Wang H, Miller JA, Verghese M, Sibai M, Solis D, Mfuh KO, et al. Multiplex sars-cov-2 genotyping reverse transcriptase pcr for population-level variant screening and epidemiologic surveillance. *J Clin Microbiol*. (2021) 59:e0085921. doi: 10.1128/JCM.00859-21
15. Tyson JR, James P, Stoddart D, Sparks N, Wickenhagen A, Hall G, et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *bioRxiv*. (2020) 3:1. doi: 10.1101/2020.09.04.283077
16. Piccoli L, Park YJ, Tortorici MA, Czudnochowski N, Walls AC, Beltramello M, et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell*. (2020) 183:1024.e21–42.e21. doi: 10.1016/J.CELL.2020.09.037/ATTACHMENT/648161E5-608D-41F4-A4EA-8561C8F1FD6C/MMC5.PDF
17. Liu Z, VanBlargan LA, Bloyet L-M, Rothlauf PW, Chen RE, Stumpf S, et al. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *bioRxiv*. (2021) 2020.11.06.372037. doi: 10.1101/2020.11.06.372037
18. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe*. (2021) 29:463.e6–76.e6. doi: 10.1016/J.CHOM.2021.02.003/ATTACHMENT/DEF4C5DD-A093-4174-817F-5DB16DCF13A9/MMC3.ZIP
19. Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. (2020) 5:1403–7. doi: 10.1038/s41564-020-0770-5
20. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, Kraemer MUG, et al. Global disparities in SARS-CoV-2 genomic surveillance. *medRxiv Prepr Serv Heal Sci*. (2021) 2021.08.21.21262393. doi: 10.1101/2021.08.21.21262393
21. Dudas G, Hong SL, Potter BI, Calvignac-Spencer S, Niatou-Singa FS, Tombolomako TB, et al. Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions. *Nat Commun*. (2021) 12:1–12. doi: 10.1038/s41467-021-26055-8
22. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science*. (2021) 374:423–31. doi: 10.1126/science.abc4336
23. Chen Z, Azman AS, Chen X, Zou J, Tian Y, Sun R, et al. Landscape of SARS-CoV-2 genomic surveillance, public availability extent of genomic data, and epidemic shaped by variants: a global descriptive study. *medRxiv*. (2021) 2021.09.06.21263152. doi: 10.1101/2021.09.06.21263152
24. Zhou H-Y, Cheng Y-X, Xu L, Li J-Y, Tao C-Y, Ji C-Y, et al. Genomic evidence for divergent co-infections of SARS-CoV-2 lineages. *bioRxiv*. (2021) 2021.09.03.458951. doi: 10.1101/2021.09.03.458951
25. Garvin MR, Prates ET, Romero J, Cliff A, Gazolla JGFM, Pickholz M, et al. The emergence of highly fit SARS-CoV-2 variants accelerated by epistasis and caused by recombination. *bioRxiv*. (2021) 2021.08.03.454981. doi: 10.1101/2021.08.03.454981

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hosch, Mpina, Nyakurungu, Borico, Obama, Ovona, Wagner, Rubin, Vickos, Milang, Ayekaba, Phiri, Daubenberger and Schindler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

2.4. The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance

Published in Science, 2022

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance

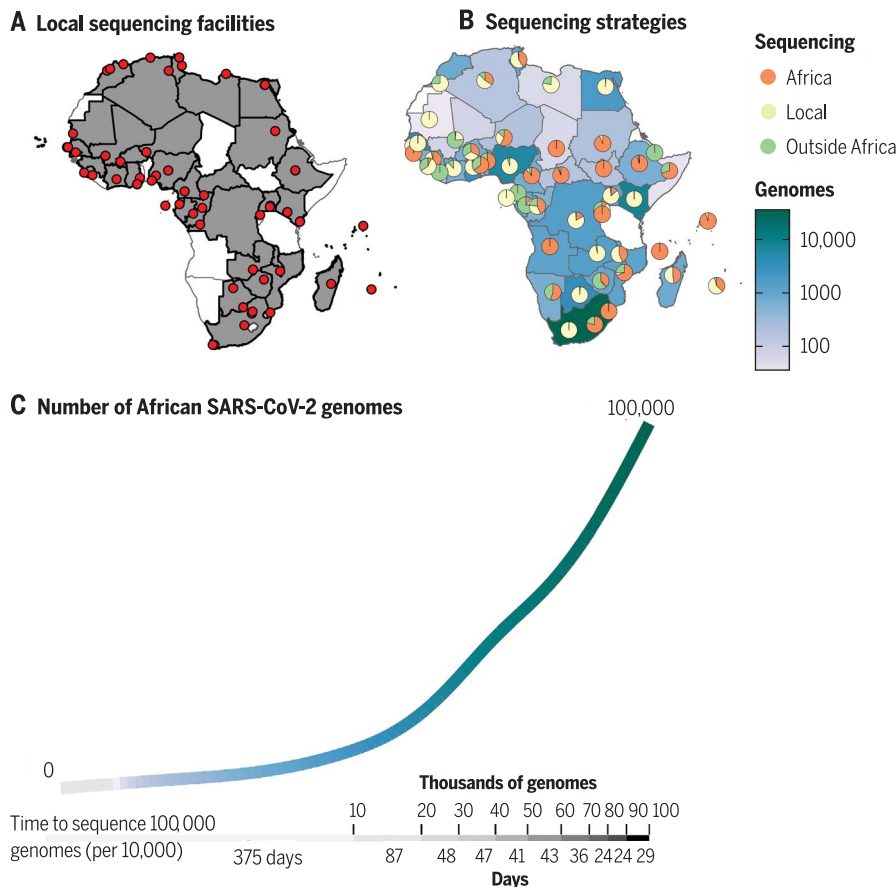
Houriiyah Tegally *et al.**†

INTRODUCTION: Investment in Africa over the past year with regard to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) sequencing has led to a massive increase in the number of sequences, which, to date, exceeds 100,000 sequences generated to track the pandemic on the continent. These sequences have profoundly affected how public health officials in Africa have navigated the COVID-19 pandemic.

RATIONALE: We demonstrate how the first 100,000 SARS-CoV-2 sequences from Africa

have helped monitor the epidemic on the continent, how genomic surveillance expanded over the course of the pandemic, and how we adapted our sequencing methods to deal with an evolving virus. Finally, we also examine how viral lineages have spread across the continent in a phylogeographic framework to gain insights into the underlying temporal and spatial transmission dynamics for several variants of concern (VOCs).

RESULTS: Our results indicate that the number of countries in Africa that can sequence the



Expanse of SARS-CoV-2 sequencing capacity in Africa. (A) African countries (shaded in gray) and institutions (red circles) with on-site sequencing facilities that are capable of producing SARS-CoV-2 whole genomes locally. (B) The number of SARS-CoV-2 genomes produced per country and the proportion of those genomes that were produced locally, regionally within Africa, or abroad. (C) Decreased turnaround time of sequencing output in Africa to an almost real-time release of genomic data.

virus within their own borders is growing and that this is coupled with a shorter turnaround time from the time of sampling to sequence submission. Ongoing evolution necessitated the continual updating of primer sets, and, as a result, eight primer sets were designed in tandem with viral evolution and used to ensure effective sequencing of the virus. The pandemic unfolded through multiple waves of infection that were each driven by distinct genetic lineages, with B.1-like ancestral strains associated with the first pandemic wave of infections in 2020. Successive waves on the continent were fueled by different VOCs, with Alpha and Beta cocirculating in distinct spatial patterns during the second wave and Delta and Omicron affecting the whole continent during the third and fourth waves, respectively. Phylogeographic reconstruction points toward distinct differences in viral importation and exportation patterns associated with the Alpha, Beta, Delta, and Omicron variants and subvariants, when considering both Africa versus the rest of the world and viral dissemination within the continent. Our epidemiological and phylogenetic inferences therefore underscore the heterogeneous nature of the pandemic on the continent and highlight key insights and challenges, for instance, recognizing the limitations of low testing proportions. We also highlight the early warning capacity that genomic surveillance in Africa has had for the rest of the world with the detection of new lineages and variants, the most recent being the characterization of various Omicron subvariants.

CONCLUSION: Sustained investment for diagnostics and genomic surveillance in Africa is needed as the virus continues to evolve. This is important not only to help combat SARS-CoV-2 on the continent but also because it can be used as a platform to help address the many emerging and reemerging infectious disease threats in Africa. In particular, capacity building for local sequencing within countries or within the continent should be prioritized because this is generally associated with shorter turnaround times, providing the most benefit to local public health authorities tasked with pandemic response and mitigation and allowing for the fastest reaction to localized outbreaks. These investments are crucial for pandemic preparedness and response and will serve the health of the continent well into the 21st century. ■

*Corresponding author. Email: Tulio de Oliveira (tulio@sun.ac.za); Eduan Wilkinson (ewilkinson@sun.ac.za)
 †All authors and affiliations appear in the full article online.
 Cite this article as H. Tegally *et al.*, *Science* **378**, eabq5358 (2022). DOI: 10.1126/science.abq5358

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abq5358>

RESEARCH ARTICLE

CORONAVIRUS

The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance

All authors and their affiliations appear at the end of this paper.

Investment in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) sequencing in Africa over the past year has led to a major increase in the number of sequences that have been generated and used to track the pandemic on the continent, a number that now exceeds 100,000 genomes. Our results show an increase in the number of African countries that are able to sequence domestically and highlight that local sequencing enables faster turnaround times and more-regular routine surveillance. Despite limitations of low testing proportions, findings from this genomic surveillance study underscore the heterogeneous nature of the pandemic and illuminate the distinct dispersal dynamics of variants of concern—particularly Alpha, Beta, Delta, and Omicron—on the continent. Sustained investment for diagnostics and genomic surveillance in Africa is needed as the virus continues to evolve while the continent faces many emerging and reemerging infectious disease threats. These investments are crucial for pandemic preparedness and response and will serve the health of the continent well into the 21st century.

What originally started as a small cluster of pneumonia cases in Wuhan, China, more than 2 years ago (1) quickly turned into a global pandemic. COVID-19 is the clinical manifestation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, and by March 2022, there had been more than 437 million reported cases and more than 5.9 million reported deaths (2). Although Africa accounts for the lowest number of reported cases and deaths thus far, with ~11.3 million reported cases and 245,000 reported deaths as of February 2022, the continent has played an important role in shaping the scientific response to the pandemic with the implementation of genomic surveillance and the identification of two of the five variants of concern (VOCs) (3, 4).

Since it emerged in 2019, SARS-CoV-2 has continued to evolve and adapt (5). This has led to the emergence of several viral lineages that carry mutations that either confer some viral adaptive advantages that increase transmission and infection (6, 7) or counter the effect of neutralizing antibodies from vaccination (8) or previous infections (9–11). The World Health Organization (WHO) classifies certain viral lineages as VOCs or variants of interest (VOIs) based on the potential impact they may have on the pandemic, with VOCs regarded as the highest risk. To date, five VOCs have been classified by the WHO; of these, two were first detected on the African continent (Beta and Omicron) (3, 4, 12) and two (Alpha and Delta) (12, 13) have spread extensively on the continent in successive waves. The remaining VOC, Gamma (14), originated in Brazil and had a limited influence in Africa, with only four recorded sequenced cases.

For genomic surveillance to be useful for public health responses, sampling for sequenc-

ing needs to be both spatially and temporally representative. In the case of SARS-CoV-2 in Africa, this means extending the geographic coverage of sequencing capacity to capture the dynamic genomic epidemiology in as many locations as possible. In a meta-analysis of the first 10,000 SARS-CoV-2 sequences generated in 2020 from Africa (15), several blind spots were identified with regard to genomic surveillance on the continent. Since then, much investment has been devoted to building capacity for genomic surveillance in Africa, coordinated mostly by the Africa Centers for Disease Control (Africa CDC) and the regional office of the WHO in Africa (or WHO AFRO) but also provided by several national and international partners, resulting in an additional 90,000 sequences shared over the past year (April 2021 to March 2022). This makes the sequencing effort for SARS-CoV-2 a phenomenal milestone. In comparison, only 12,000 whole-genome influenza sequences (16) and only ~3700 whole-genome HIV sequences (17) from Africa have been shared publicly, even though HIV has plagued the continent for decades.

Here, we describe how the first 100,000 SARS-CoV-2 sequences from Africa have helped describe the pandemic on the continent, how this genomic surveillance in Africa has expanded, and how we adapted our sequencing methods to deal with an evolving virus. We also highlight the impact that genomic sequencing in Africa has had on the global public health response, particularly through the identification and early analysis of new variants. Finally, we also describe here how the Delta and Omicron variants have spread across the continent and how their transmission dynamics were distinct from the Alpha and Beta variants that preceded them.

Results

Epidemic waves driven by variant dynamics and geography

Scaling up sequencing in Africa has provided a wealth of information on how the pandemic unfolded on the continent. The epidemic has largely been spatially heterogeneous across Africa, but most countries have experienced multiple waves of infection (18–29), with substantial local and regional diversity in the first wave and to a lesser extent in the second wave, followed by successive sweeps of the continent with Delta and Omicron (Fig. 1A). In all regions of the continent, different lineages and VOIs evolved and cocirculated with VOCs and, in some cases, contributed considerably to epidemic waves.

In North Africa (Fig. 1B and fig. S1A), B.1 lineages and Alpha dominated in the first and second waves of the pandemic and were replaced by Delta and Omicron in the third and fourth waves, respectively. Interestingly, the C.36 and C.36.3 sublineages dominated the epidemic in Egypt (~40% of reported infections) before July 2021 when they were replaced by Delta (30). Similarly, in Tunisia, the first and second waves were associated with the B.1.160 lineage and were replaced by Delta during the country's third wave of infections. In southern Africa (Fig. 1C and fig. S1C), we see a similar pandemic profile, with B.1 dominating the first wave; however, instead of Alpha, Beta was responsible for the second wave, followed by Delta and Omicron. Another lineage that was flagged for close monitoring in the region was C.1.2 because of its mutational profile and predicted capacity for immune escape (31). However, the C.1.2 lineage did not cause many infections in the region because it was circulating at a time when Delta was dominant. In West Africa (Fig. 1D and fig. S1B), the B.1.525 lineage caused a large proportion of infections in the second and third waves, where it shared the pandemic landscape with the Alpha variant. As with other regions on the continent, these variants were later replaced by the Delta and then the Omicron VOCs in successive waves. In Central Africa (Fig. 1E and fig. S1D), the B.1.620 lineage caused most of the infections between January and June 2021 (32) before systematically being replaced by Delta and then Omicron. Lastly, in East Africa (Fig. 1F and fig. S1E), the A.23.1 lineage dominated the second wave of infections in Uganda (33) and much of East Africa. In all of these regions, minor lineages such as B.1.525, C.36, and A.23.1 were eventually replaced by VOCs that emerged in later waves.

Finally, we directly compared the official recorded cases in Africa with the ongoing SARS-CoV-2 genomic surveillance data (GISAID date of access: 31 March 2022) for a crude estimation of the variants' contributions to cases. We observe that Delta was responsible for an epidemic wave between May and October 2021

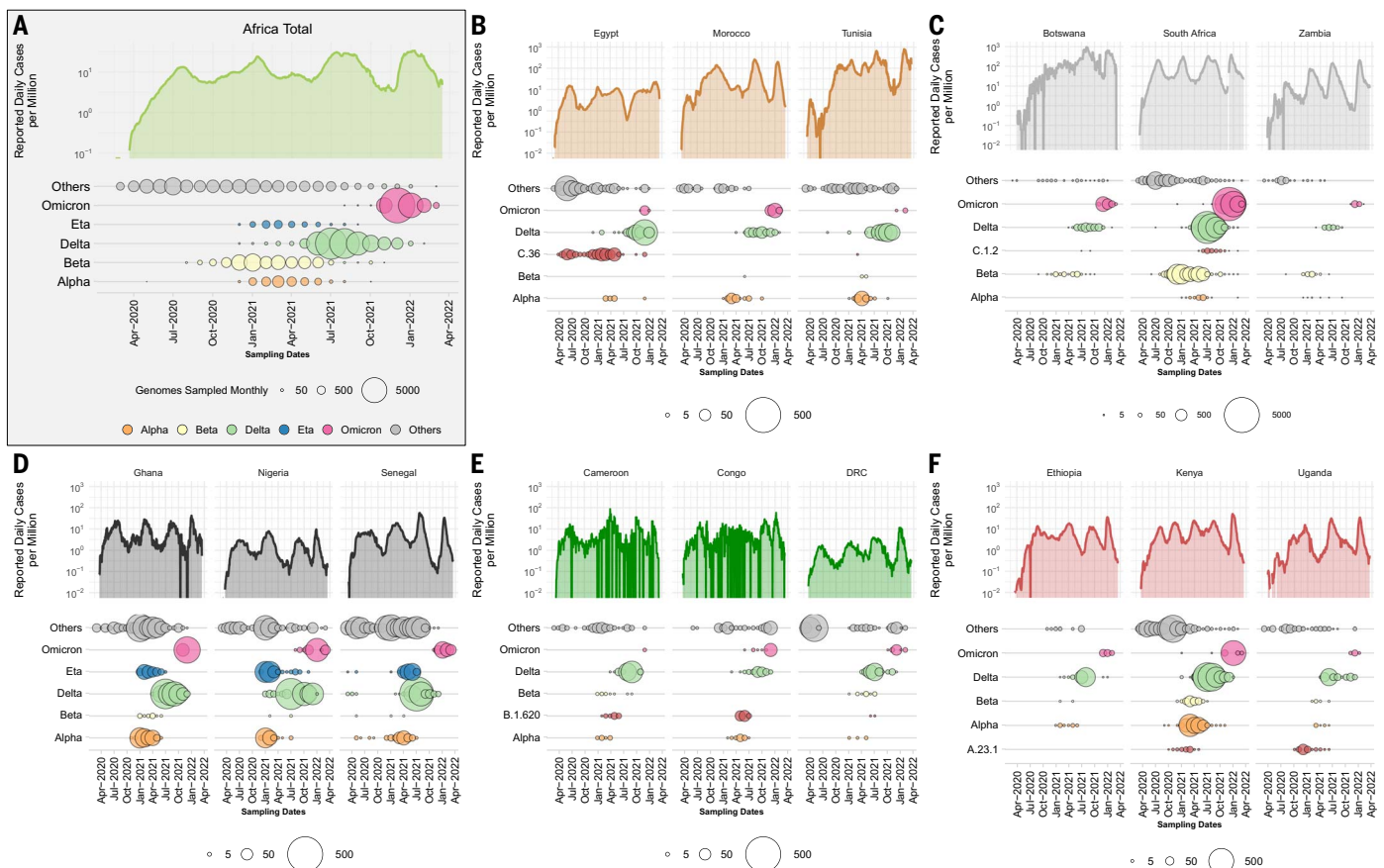


Fig. 1. Epidemiological progression of the COVID-19 pandemic on the African continent. (A) Total reported new case counts per million inhabitants in Africa (data source: Our World in Data; log-transformed) along with the distribution of VOCs, the Eta VOI, and other lineages through time (the size of each circle is proportional to the number of genomes sampled per month for each category). (B to F) Breakdown of reported new cases

(Fig. 1A) and had the greatest impact on the continent, with almost 34.2% of overall infections in Africa possibly attributed to it. Beta was responsible for an epidemic wave at the end of 2020 and beginning of 2021 (Fig. 1A), with 13.3% of infections overall attributed to it. Notably, Alpha, despite being predominant in other parts of the world at the beginning of 2021, had only minimal importance in Africa, accounting for just 4.3% of infections. At the time of writing, the Omicron VOC had contributed to 21.6% of the overall number of sequenced infections. At this time, the Omicron wave was still unfolding globally and in Africa with the expansion of several sublineages (34), such that its full impact is yet to be determined. However, because of increased population immunity (35) from SARS-CoV-2 infection and vaccination (fig. S2), the impact of Omicron on mortality has been less in comparison to the other VOCs, as can be observed by the relatively low death rate in South Africa during the Omicron wave (36). The findings from mapping epidemiological numbers onto genomic sur-

veillance data are reliable as far as the proportional scaling of genomic sampling across Africa with the size and timing of epidemic waves [fig. S3; model estimate (b) = 0.011, standard error (SE) = 0.001, $p < 2 \times 10^{-16}$].

This comes with the obvious caveats that testing and reporting practices have varied widely across the continent along with genomic surveillance volumes throughout the pandemic. Countries in Africa with reported data have tested in proportions from as little as 0.1 daily tests per million population to more than 1000 tests per million (fig. S4). Some countries have consistently tested at high proportions, for example, South Africa, Botswana, Morocco, and Tunisia. Incidentally, these countries have also generally reported more cases per million population, providing an indication that recorded low incidences in other parts of the continent have been underestimates due to low testing rates. However, even for these countries, epidemic numbers are certainly underrepresented and underdetected, given that in several time frames, test positivity rates were still on the

per million (data source: Our World in Data; log-transformed) and monthly sampling of VOCs, regional variant, or lineage of interest and other lineages for three selected countries for North, southern, West, Central, and East Africa, respectively. For each region, a different variant or lineage of interest is shown, relevant to that region (C.36, C.1.2, Eta, B.1.620, and A.23.1, respectively).

higher end, approaching or exceeding 20% (fig. S4), and as concluded by seroprevalence surveys and estimates of true infection burdens in Africa (37, 38). Findings of attributing case numbers of variants must therefore be interpreted in the context of this limitation but can nevertheless provide a qualitative overview of the spatial and temporal dynamics of VOCs in relation to epidemic progression in Africa.

The African regional (table S1) and country-specific (table S2) NextStrain builds also clearly support the changing nature of the pandemic over time. From these builds, we observe a strong association of B.1-like viruses circulating on the continent during the first wave. These “ancestral” lineages were subsequently replaced by the Alpha and Beta variants, which dominated the pandemic landscape during the second wave and were later replaced by the Delta and Omicron variants during the third and fourth waves.

Optimizing surveillance coverage in Africa

By mapping and comparing the locations of specimen sampling laboratories to the sequencing

laboratories, a number of aspects regarding the expansion of genomic surveillance on the continent became clear. First, even though several countries in Africa started sequencing SARS-CoV-2 in the first months of the pandemic, local sequencing capacity was initially limited. However, local sequencing capabil-

ities slowly expanded over time, particularly after the emergence of VOCs (Fig. 2A). The fact that almost half of all SARS-CoV-2 sequencing in Africa was performed using the Oxford Nanopore Technology (ONT), which is relatively low-cost compared with other sequencing technologies and better adapted to modest

laboratory infrastructures, illustrates one component of how this rapid scale-up of local sequencing was achieved (fig. S5). Yet, to rely only on local sequencing would have thwarted the continent's chance at a reliable genomic surveillance program. At the time of writing, 52 of 55 countries in Africa had SARS-CoV-2

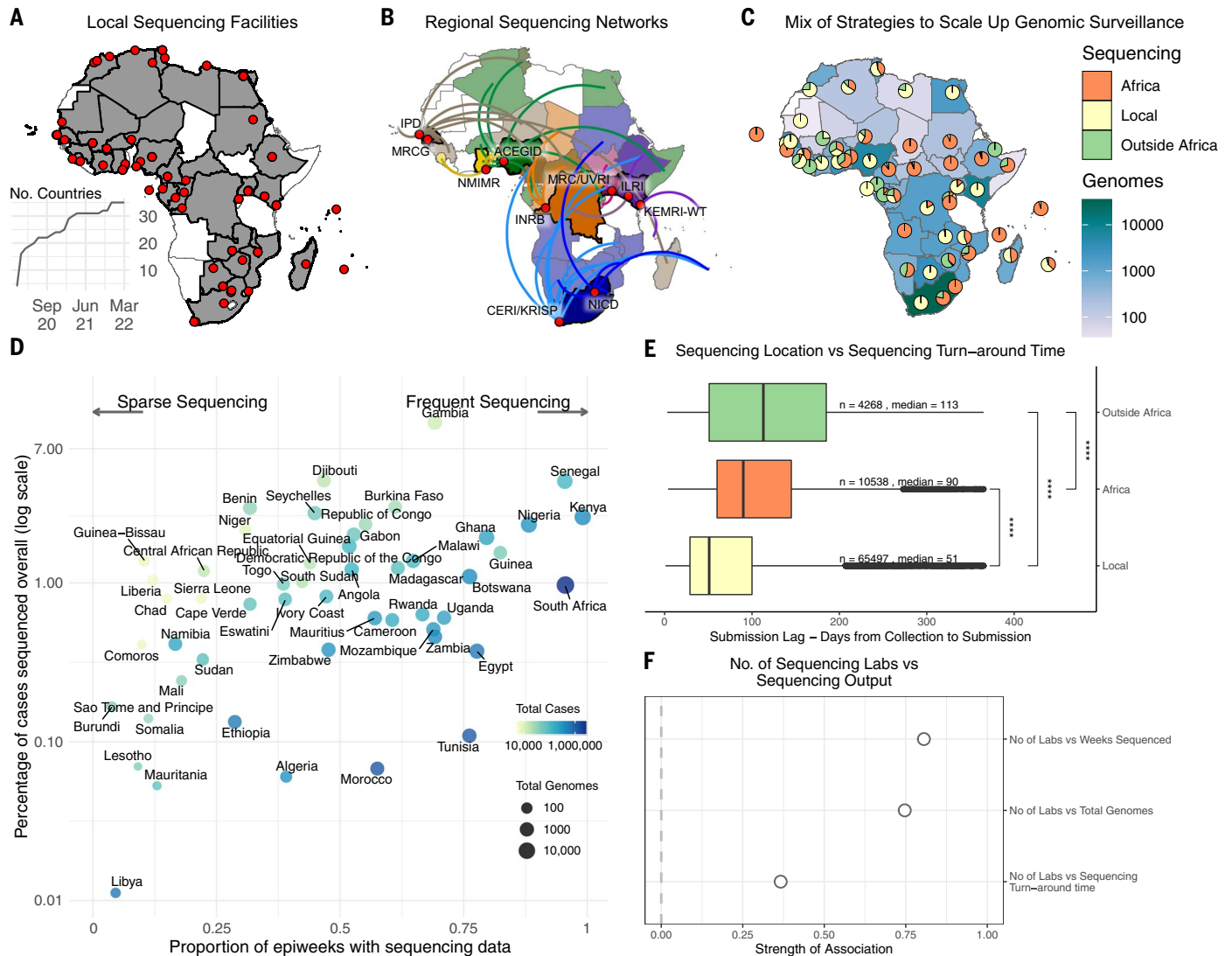


Fig. 2. Sequencing strategies and outputs in Africa. (A) Geographical representation of all countries (shaded in gray) and institutions (red dots) in Africa with their own on-site sequencing facilities. The inset graph shows the number of countries in Africa that are able to carry out sequencing locally over time. (B) Key regional sequencing hubs and networks in Africa showing countries (shaded in bright colors) and institutions (red dots) that have sequenced for other countries (shaded in corresponding light colors and linking curves) on the continent. ACEGID, African Centre of Excellence for Genomics of Infectious Diseases; CERI, Centre for Epidemic Response and Innovation; KEMRI-WT, Kenya Medical Research Institute–Wellcome Trust; KRISP, KwaZulu-Natal Research Innovation and Sequencing Platform; ILRI, International Livestock Research Institute; INRB, Institut National de Recherche Biomédicale; IPD, Institut Pasteur de Dakar; MRC/UVRI, Medical Research Council/Uganda Virus Research Institute; MRCCG, Medical Research Council Unit–The Gambia; NICD, National Institute for Communicable Diseases; NMIMR, Noguchi Memorial

Institute for Medical Research. (C) Geographical representation of the total number of SARS-CoV-2 whole genomes produced over the course of the pandemic in each country, as well as the proportion of those sequences that were produced locally, regionally, or abroad. (D) Correlation of the proportion of COVID-19 positive cases that have been sequenced and the corresponding number of epidemiological weeks since the start of the pandemic that are represented with genomes for each African country. The color of each circle represents the number of cases and its size the number of genomes. (E) Comparison of sequencing turnaround times (lag times from sample collection to sequence submission) for the three strategies of sequencing in Africa, showing a significant difference in the means (**** $p < 0.0001$). The box and whisker plot denotes the lower quartile, the median and upper quartiles (box), the minimum and maximum values (whiskers), and the outliers (black dots). (F) Pearson correlations of the total number of sequencing laboratories per country against key sequencing outputs.

genomes deposited in GISAID; however, there were still 16 countries with no reported local sequencing capacity (Fig. 2A) and undoubtedly many with limited capacity to meet demand during pandemic waves.

To tackle this, three centers of excellence and various regional sequencing hubs were established to maximize the resources available in a few countries to assist in genomic surveillance across the continent. This sequencing is done either as the sole source of viral genomes for those countries (e.g., Angola, South Sudan, and Namibia) or concurrently with local efforts to increase capacity during resurgences (Fig. 2B). Sequencing is further supplemented by a number of countries that use facilities outside of Africa. Ultimately, a mix of strategies from local sequencing, collaborative resource sharing among African countries, and sequencing with academic collaborators outside the continent helped close surveillance blind spots (Fig. 2C). Countries in sub-Saharan Africa, particularly in southern and East Africa, most benefited from the regional sequencing networks, whereas countries in West and North Africa often partnered with collaborators outside of Africa.

The success of pathogen genomic surveillance programs relies on how representative it is of the epidemic under investigation. For SARS-CoV-2, this is often measured in terms of the percentage of reported cases sequenced and the regularity of sampling. African countries were positioned across a range of different combinations of overall proportion and frequency of genomic sampling (Fig. 2D). Although the ultimate goal would be to optimize both of these parameters, a lower proportion of sampling can also be useful if the frequency of sampling is maintained at as high a level as possible. For instance, South Africa and Nigeria, which have both sequenced ~1% of cases overall, can be considered to have successful genomic surveillance programs based on the fact that sampling is representative over time and has enabled the timely detection of variants (Beta, Eta, Omicron).

Additionally, for genomic surveillance to be most useful for rapid public health response during a pandemic, sequencing would ideally be done in real time or in a framework as close as possible to that. We show a general trend of decreasing sequencing turnaround time in Africa (fig. S6), particularly from a mean of 182 days between October and December 2020 to a mean of 50 days over the same period a year later, although this does come with several caveats. First, we measure sequencing turnaround time in the most accessible manner, which is by comparing the date of sampling of a specimen to the date its sequence was deposited in GISAID. Generally, the genomic data potentially informs the public health response more rapidly than reflected here, particularly

when it comes to local outbreak investigations or variant detection. This analysis is also confounded by various factors such as country-to-country variation in these trends (fig. S7), delays in data sharing, and potential retrospective sequencing, particularly by countries that joined sequencing efforts at later stages of the pandemic. The most critical caveat is the fact that sequencing from the most recently collected samples (e.g., over the past 6 months) may still be ongoing. The shortening duration between sampling and genomic data sharing is nevertheless a positive takeaway, given that these data also feed into continental and global genomic monitoring networks. Overall, the continental average delay from specimen collection to sequencing submission is 87 days, with 10 countries having an average turnaround time of less than 60 days and Botswana of less than 30 days (fig. S8).

Most importantly, in the context of optimizing genomic surveillance, we found that the route taken to sequencing affects the speed of data generation. Of the three frameworks we investigated, local sequencing has statistically faster sequencing turnaround times (median of 51 days), followed by sequencing within regional sequencing networks in Africa (median of 93 days) and finally outsourced sequencing to countries outside Africa (median of 113 days) (Fig. 2E). This finding strongly supports the investments in local genomic surveillance to generate timely and regular data for local and regional decision-making. Finally, we show that it is beneficial in several ways for countries to undertake genomic surveillance through several sequencing laboratories rather than by centralizing efforts. For instance, we estimate strong correlations between the numbers of sequencing laboratories per country and the total number of genomes produced by that country (Pearson correlation, 0.75), the total number of epiweeks for which sequencing data was produced (Pearson correlation, 0.81), and, importantly, sequencing turnaround time (Pearson correlation, -0.37) (Fig. 2F).

With the increase in sequencing capacity on the continent, a decrease in the time taken to detect new variants was observed. For example, the Beta variant was identified in December 2020 in South Africa (4), but sampling and molecular clock analyses suggest that the variant originated in September 2020. This 3-month lag in detection means that a new variant, like Beta, has ample time to spread over a large geographic region before its detection. However, by the end of 2021, the time to detect a new variant was substantially improved. Phylogenetic and molecular clock analyses suggest that the Omicron variant originated around 9 October 2021 (95% highest posterior density: 30 September to 20 October 2021), and the variant was described on 23 November 2021 (3). Thus, Omicron was detected within ~5 weeks

from origin compared with the Beta variant (~16 weeks) and the Alpha variant, which was detected in the United Kingdom (~10 weeks). More importantly, the time from sequence deposition to the WHO declaring the new variant a VOC was substantially shortened to 72 hours for the Omicron variant.

To interpret insights from the described genomic surveillance in Africa, it is important to understand the context of epidemiological reporting and sampling strategies used for sequencing on the continent (table S3). Most countries provided daily reports of newly recorded cases, whereas a few provided weekly and monthly reports. For most countries, surveillance was mainly focused on the major cities, suggesting potential cryptic circulation in rural areas. We find that at the onset of the pandemic, surveillance was focused on identification of imported cases from incoming travelers or local residents returning from various countries. As community transmissions began to emerge, the focus shifted toward regular surveillance and outbreak investigations. Together, these three strategies account for the vast majority of samples generated on the continent and analyzed here. As the pandemic progressed and vaccines were made available, some countries on the continent began to explore other sampling strategies such as reinfections, environmental samples such as wastewater samples, and vaccine breakthrough cases to gain new insights into the evolutionary dynamics of SARS-CoV-2. The utility of sequencing for viral evolution tracking and VOC detection in the way described above is obviously also dependent on sampling proportions, especially within sampling for regular surveillance.

The speed of SARS-CoV-2 evolution has complicated sequencing efforts. Common methods of RNA sequencing include reverse transcription followed by double-stranded DNA amplification using sequence-specific primer sets (39). Ongoing SARS-CoV-2 evolution has necessitated the continual evaluation and updating of these primer sets to ensure their sustained utility during genomic surveillance efforts. Here, we examined the current set of genomes to determine aspects of the sequencing process that might be improved in the future. Many of the primer sets that were used were designed using viral sequences from the start of the pandemic and may require updating to keep pace with evolution. Indeed, the ARTIC primer sets are now in version 4.1 (40). The Entebbe primer set was designed mid-2020, well into the first year of the epidemic, and used an algorithm and design that accommodates evolution (41).

The effects of viral evolution on sequencing patterns can be seen with low median unspecified nucleotide (N) values (a consequence of primer dropout or low coverage at that site) that were observed for the first 12 months of the epidemic, with an increase from October 2020

(Fig. 3A). Additional challenges appear (as indicated by increasing median N values) as the virus further evolved into the Delta and Omicron lineages from January 2021 onward (Fig. 3A). By examining the role of sequencing technology, it appears that the two major technologies used (Illumina and ONT) have similar gap profiles (as measured by mean N count per genome), whereas Ion Torrent, MGI, and Sanger show a reduced mean N count per genome (Fig. 3B). Likely factors for this pattern are the primers used in sequencing, with primer choice playing a key role in the quantity of gaps (Fig. 3C). The mean N count per genome varied with viral lineage (Fig. 3D). There was a modest difference in mean N count per genome across the lineages. Lineages that returned no classification with Pangolin (“none”) showed the highest mean N count, suggesting that high mean N count per genome was probably the basis for failed classification. The more recent lineages, Delta (e.g., AY.39, AY.75) and Omicron (BA.1.1, BA.2), also showed higher mean N count per genome, consistent with virus evolution impairing primer function. This pattern is further explored in fig. S9, where the position of gaps shows an enrichment in the genome regions after position 19,000, with frequent gaps disrupting the spike coding region.

Phylogenetic insights into the rise and spread of VOCs in Africa

During the first wave of infections in 2020 in Africa, as was the case globally, most corresponding genomes were classified as PANGO B.1 ($n = 2456$) or B.1.1 viruses ($n = 1329$). Toward the end of 2020, more-distinct viral lineages started to appear. Of these, the most important ones that affected the African continent are B.1.525 ($n = 797$), B.1.1.318 ($n = 398$) (42), B.1.1.418 ($n = 395$), A.23.1 ($n = 358$) (15, 29, 31, 33), C.1 ($n = 446$) (29), C.1.2 ($n = 300$) (31), C.36 ($n = 305$) (30, 43), B.1.1.54 ($n = 287$) (15, 29, 31, 33), B.1.416 ($n = 272$), B.1.177 ($n = 203$), B.1.620 ($n = 138$), and B.1.160 ($n = 61$) (32) (fig. S10, A and B). Our discrete state phylogeographic inference from phylogenetic reconstruction of non-VOC African sequences and an equal number of external references revealed that African countries were primarily seeded by multiple introductions of viral lineages from abroad (mainly Europe) at the beginning of the pandemic. The observed pattern of non-VOC viral lineage movement then consistently shifted toward more intercontinental exchanges (fig. S10C). Mapping out the spatial routes of dissemination shows that various countries in all subregions of the continent acted as sources of these viral lineages at one point or another (fig. S10D). Although uneven testing rates and proportions of samples sequenced on the continent may have influenced these inferences (discussed later), the results presented here are in line with the fact that these most predominant

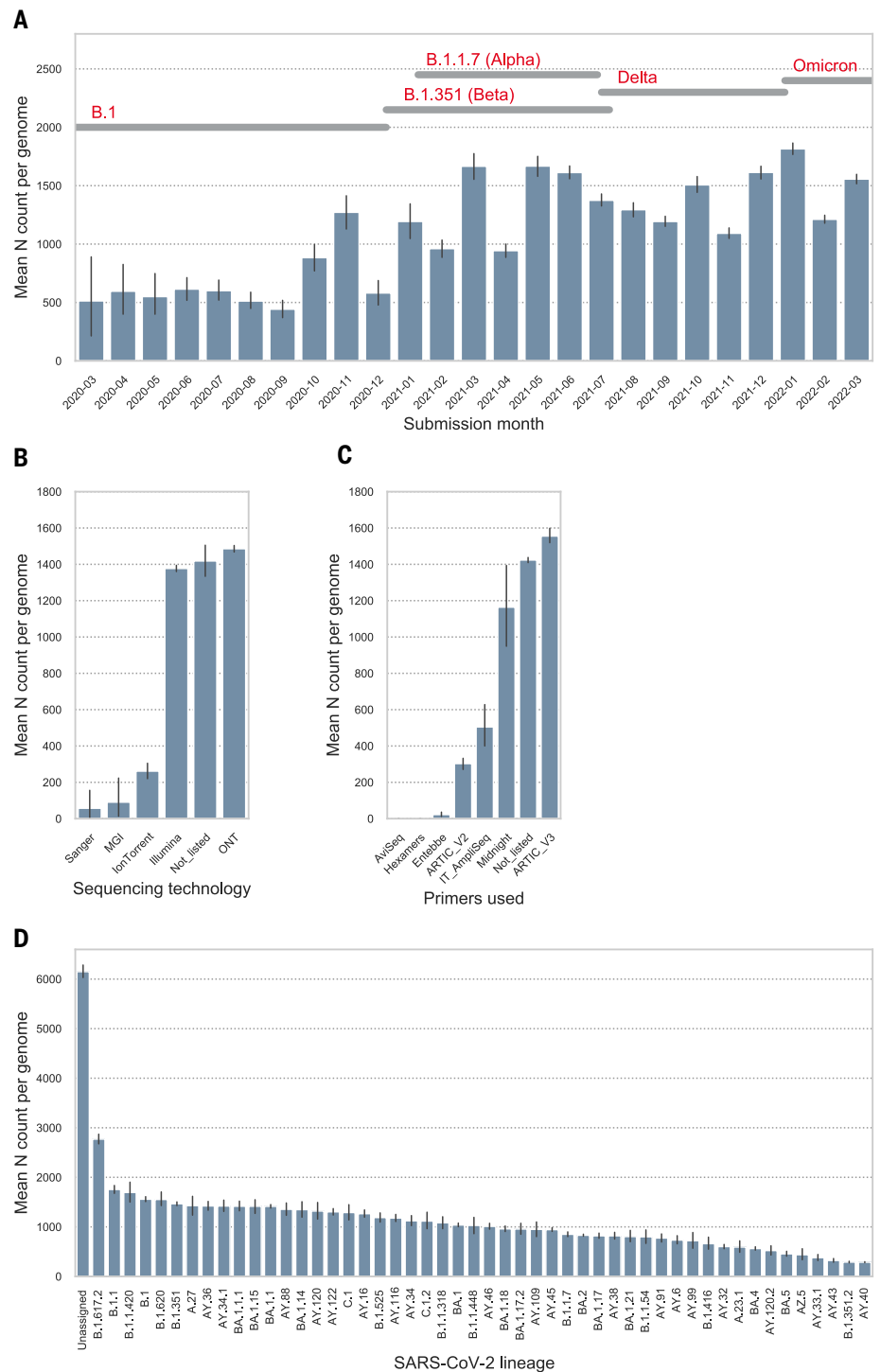


Fig. 3. Genome gap analysis. (A) The mean N count per genome by month of submission to GISAID. The time periods corresponding to the detection of important SARS-CoV-2 lineages are indicated at the top of the figure. (B) Illustration of the mean N count per genome stratified by sequencing technology. (C) The mean N count per genome stratified by the sequencing primers sets used. (D) Mean N count per genome by lineage. The mean N data were stratified by SARS-CoV-2 lineages to investigate the lineage-specific frequency of genome gaps, an indirect measure of primer mismatch. All lineages that were present at least 100 times in the genome data are presented. For (A) to (D), error bars indicate 95% confidence intervals.

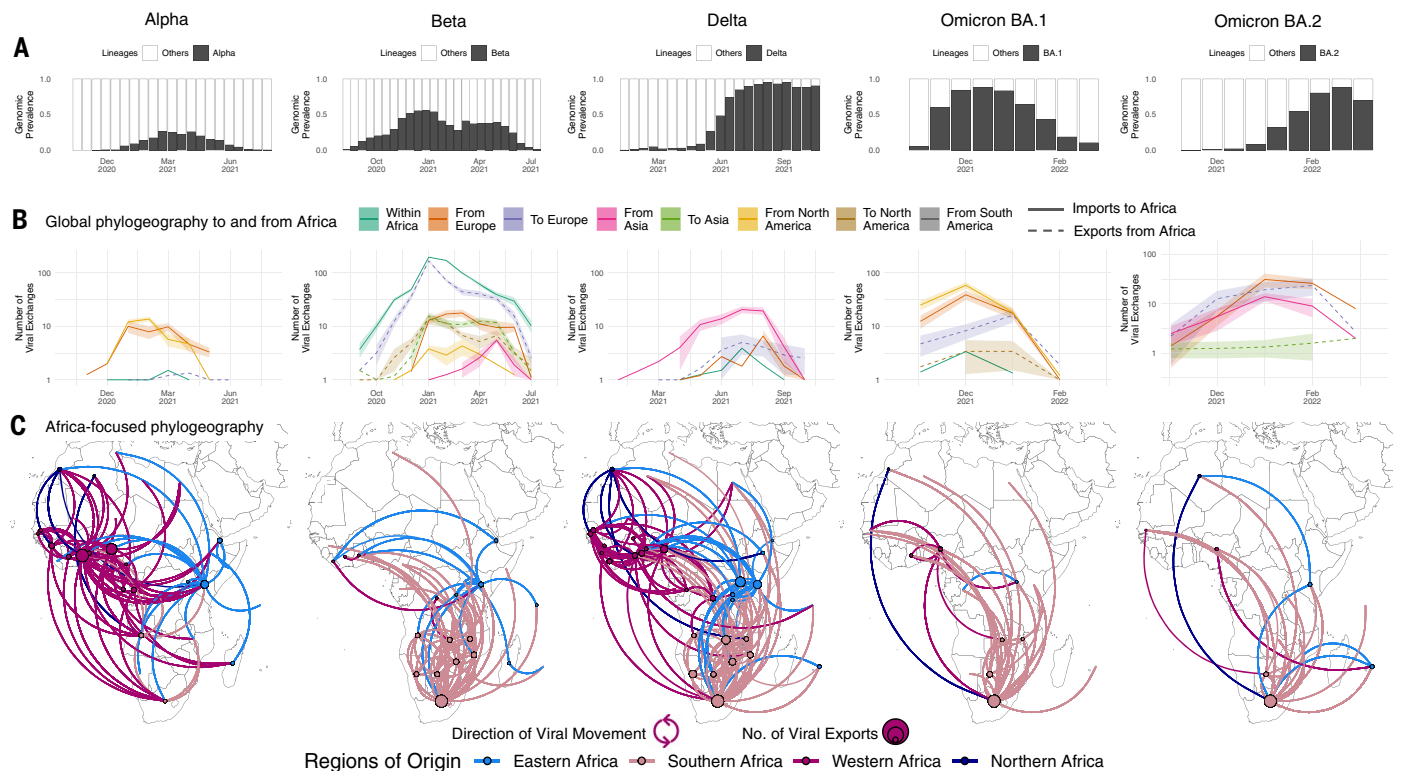


Fig. 4. Inferred viral dissemination patterns of VOCs within Africa. (A) Genomic prevalence of VOCs Alpha, Beta, Delta, and Omicron in Africa over time. (B) Inferred viral exchange patterns to, from, and within the continent of Africa for the four VOCs (Omicron as BA.1 and BA.2) based on case-sensitive phylogeographic inference. Introductions and viral transitions within Africa are shown as solid lines, and exports from Africa are shown as dotted lines; the lines are colored by continent.

The shaded areas around the lines represent the uncertainty of this analysis from 10 replicates (\pm SD). (C) Dissemination patterns of the VOCs within Africa obtained from inferred ancestral-state reconstructions performed on Africa-enriched datasets, annotated and colored by region in Africa. The countries of origin of viral exchange routes are also shown with dots, and the curves go from country of origin to destination country in a counterclockwise direction.

non-VOC lineages in Africa, except B.1.177, emerged and circulated widely in different subregions (Fig. 1).

Similar to the pandemic globally, VOCs became increasingly important in Africa toward the end of 2020. The Alpha, Beta, Delta, and Omicron variants demonstrate many similarities as well as differences in the way that they spread on the continent. For all these VOCs, we observe large regional monophyletic transmission clusters in each of their phylogenetic reconstructions in Africa (fig. S11). This suggests an important extent of continental dissemination within Africa. Alpha and Beta were epidemiologically important in distinct regions of the continent, with Alpha primarily circulating in West Africa, North Africa, and most of Central Africa; Beta circulating in southern and most of East Africa; and both only substantially cocirculating in a few countries such as Angola, Kenya, Comoros, Burundi, and Ghana (Fig. 1 and fig. S12). However, we may not have enough resolution in the geospatial data to know whether and to what extent they were truly cocirculating throughout these countries or whether there were regional outbreaks of Alpha and Beta within these countries. In Kenya, for example, Beta was detected

more frequently in coastal regions and Alpha more frequently inland (26, 44). By contrast, the Delta and Omicron variants sequentially dominated most infections on the entire continent shortly after their emergence (Fig. 4A and fig. S12).

The Alpha variant was first identified in December 2020 in the United Kingdom and has since spread globally. In Africa, Alpha was detected in 43 countries, with evidence of community transmission based on phylogenetic clustering in many countries, including Ghana, Nigeria, Kenya, Gabon, and Angola (fig. S11). Discrete state maximum likelihood reconstruction from a globally case-sensitive genomic subsampling inferred at least 80 introductions [95% confidence interval (CI): 78 to 82] into Africa, with the bulk of imports attributed to the United States (>47%) and the United Kingdom (>25%) (Fig. 4B). Only 1% of imports into any particular African country were attributed to another African nation. Phylogeographic reconstruction enriched in African sequences revealed that of those, >85% of the intercontinental Alpha exchanges in Africa originated from West African countries (Fig. 4C). This occurred in spite of initial importations of the Alpha variant from Europe into all regions of

the continent (fig. S13B) but is in line with Alpha having dominated circulation mostly in West Africa (fig. S12). In countries where Alpha was introduced but did not grow and cause an expansion of cases, this can be explained by competition with the already established Beta variant, which simultaneously circulated. The characteristics of multiple introductions of Alpha into Africa and between African countries is similar to the spread of Alpha that has been documented in the United Kingdom, Scotland, and Ireland (45–47).

The second VOC, Beta, was identified in December 2020 in South Africa (4). However, sampling and molecular clock analyses suggest that the variant originated around September 2020 (fig. S11). At the end of 2020 and beginning of 2021, Beta was driving a second wave of infection in South Africa and quickly spread to other countries within the region. The concurrent introductions and spread of Alpha and other variants (Eta, A.23.1) in other regions of the continent may have reduced the Beta variant's initial growth, limiting its spread largely to southern Africa and, to a lesser extent, the East Africa region. Beta spread to at least 114 countries globally, including 37 countries and territories in Africa. For this

variant, viral circulation and geographical exchanges occurred predominantly within the continent. Indeed, phylogeographic reconstruction from a globally case-sensitive sampling revealed that of the 810 (95% CI: 803 to 818) inferred introductions of the Beta variant into African countries, only 110 (95% CI: 105 to 115; 13%) were attributed to sources outside the continent (fig. S13C), whereas more than half of the introductions were attributed to South Africa (63%) (Fig. 4C). This is in line with expectations because the variant originated in South Africa. Beyond southern Africa, most of the introductions back into the continent were attributed to France and other European Union countries into the French overseas territories, Mayotte and Reunion, and other Francophone African countries. Africa-focused phylogeographic analysis revealed a similar spatial pattern that showed southern countries as substantial sources of the variant, followed in small numbers by countries in East Africa (Fig. 4C).

The fourth VOC observed was Delta (13), which rose to prominence in April 2021 in India, where it fueled an explosive second wave. Since its emergence, Delta has been detected in >170 countries, including 37 African countries and territories (fig. S11). Our global case-sensitive subsampled analysis infers at least 100 (95% CI: 93 to 106) introductions of the Delta variant into Africa, with the bulk attributed to India (~72%), mainland Europe (~8%), the United Kingdom (~5%), and the United States (~2.5%). Viral introductions of Delta also occurred from one African country to others in 7% of inferred introductions. From our Africa-focused phylogeographic inferences, we infer that unlike Alpha and Beta, viral dissemination of Delta within Africa was not restricted to or dominated by any particular region but rather spread across the entire continent (Fig. 4C). After introductions from Asia in the middle of 2021, Delta rapidly replaced the other circulating variants (Fig. 4A). For example, in southern African countries, the Delta variant rapidly displaced Beta and, by June 2021, was circulating at very high (>90%) frequencies (48).

The latest VOC, Omicron, was identified and characterized in November 2021 in southern Africa (3). At the time of writing, the variant had been detected and caused waves of infections in >160 countries, including 39 African countries and two overseas territories (fig. S11). Because of the genetic distance between them and their sequential (rather than simultaneous) epidemic expansion globally, phylogenies were reconstructed separately for Omicron BA.1 and BA.2. Our discrete ancestral-state reconstruction from a global case-sensitive sampling for Omicron BA.1 infers at least 55 (95% CI: 47 to 62) viral exports of BA.1 out of various African countries, of which 31 (95% CI: 25 to 36) were toward Europe and 8 (95% CI:

6 to 10) were toward North America (Fig. 4B). After explosive expansion of Omicron around the world, we inferred even more reintroductions of the variant back into Africa, at least 69 (95% CI: 60 to 78) from Europe and 102 (95% CI: 92 to 112) from North America (Fig. 4B). From our Africa-focused phylogeographic reconstructions, we determine that, as with Delta, routes of dissemination of this variant involved all regions of the continent spatially (Fig. 4C). Yet ~75% of all BA.1 viral movement volume in Africa happened between southern African countries, likely because of rapid epidemic expansion in the region soon after its detection (3). Omicron BA.2's reach in Africa was limited at the time of writing, with only 3260 sequences from 19 countries attributed to BA.2 on GISAID (date of access: 31 March 2022) (15% of all Omicron sequences from Africa). Our discrete ancestral-state reconstruction from a global case-sensitive sampling for Omicron BA.2 infers at least 68 (95% CI: 53 to 84) viral exports out of African countries, of which most were toward Europe (~88%) (Fig. 4B). We also infer at least 99 (95% CI: 87 to 109) separate introduction or reintroduction events of BA.2 back into African countries, of which ~65% are from Europe and ~30% from Asia, primarily from India (Fig. 4B). This is consistent with India having experienced one of the earliest large BA.2 waves globally. In the context of global incidence of BA.2, this case-sensitive phylogeographic analysis revealed that only 0.01% of viral movements of this lineage globally happened from one African country to another. Our Africa-focused analysis inferred a similar pattern of BA.2 spatial diffusion within African to that of BA.1 (Fig. 4C). However, given that this accounted for such a small percentage of global BA.2 movements, BA.2 diffusion from one African country to another is unlikely to have had a substantial impact on epidemiological expansion, compared with introductions from Asia, Europe, or North America.

Globally, dissemination of the SARS-CoV-2 virus throughout the pandemic was intricately linked with human mobility patterns (49–53). To determine the validity of the VOC movement patterns that we infer into and within the Africa continent in this study, we compared viral import and export events to and from South Africa with travel to the country. In December 2020, the United Kingdom accounted for the fifth-highest number of passengers entering South Africa, whereas other countries with the top-nine sources of travelers were all neighboring countries in southern Africa (fig. S14A). Considering that incidence of the Alpha variant was not meaningful in the region, this supports our inference of the United Kingdom contributing 60% of Alpha introductions to South Africa (fig. S15A). In March 2021, the United States, Germany, the United Kingdom, and India were among the top-12 sources of

travelers to South Africa after eight African countries (fig. S14B). During this time of Delta dissemination globally, we infer that ~90% of introductions of Delta into South Africa originated in the United Kingdom, the United States, and India (fig. S15B). At the end of 2021, most introductions or reintroductions of Omicron to the country came from the United Kingdom, the United States, or Botswana, corresponding to locations of both high Omicron incidence at the time and high numbers of passengers to South Africa (figs. S14C and S15C). These travel patterns also fit the findings that ~89, ~70, and ~75% of Beta, Delta, and Omicron exports, respectively, from South Africa to other African countries were directed to locations in southern Africa (figs. S14, D and E, and S15, D and E).

Discussion, limitations, and conclusions

By April 2020, a total of 20 African countries were able to sequence the virus within their own borders. This was largely made possible by other preexisting sequencing efforts on the continent that were focused on other human pathogens (e.g., HIV, tuberculosis, Ebola, and H1N1). However, these efforts were quickly limited by global supply chain issues, and, in many countries, sequencing efforts substantially slowed down or stopped toward the end of 2020. To facilitate more sequencing on the continent over the course of the past year (April 2021 to March 2022), the Africa CDC and partners invested heavily to support genomic surveillance on the continent. This included the transfer of 24 new sequencing platforms (including MinIon, GridIon, MiSeq, and NextSeq), the distribution of reagents and flow cells to support the sequencing of 100,000 positive samples, the training of >230 students and technicians in wet laboratory and bioinformatic techniques, and additional grants to support 10 regional sequencing hubs. This investment has started bearing fruit and should be intensified as the virus continues to evolve, requiring the adaptation of methodologies locally on the continent to keep pace with the emergence of variants. The continued development of sequencing protocols in Africa is of crucial importance (41, 54, 55) given the number of variants and lineages that emerged in, and were introduced to, the continent. In North Africa, the SARS-CoV-2 pandemic was caused by waves of infections that were similar to those seen in Europe (first wave attributed to B.1 descendants, second wave to Alpha, third wave to Delta, and fourth wave to Omicron); in southern Africa, the pattern was similar but with a Beta wave instead of an Alpha one. In East Africa, the pandemic was more complex, involving both Alpha and Beta as well as its own lineage A.23.1 before the arrival of Delta and Omicron. Central Africa experienced epidemic patterns that sometimes mirrored those of East Africa and other times those of southern Africa. In West Africa, Eta

made a considerable contribution to both a second wave (together with Alpha) and a third wave (together with Delta). The factors that resulted in these regional differences are not clear but could be due to differences in human mobility, founder effects, competition between lineages, or the immunity induced by earlier waves in a region.

Public health benefits of such broadly inclusive genomic surveillance are manifold. The most prominent insight from this expanded genomic surveillance in Africa has been an early warning capacity for the world after the detection of new lineages and variants, most recently relevant in the detection of Omicron BA.1, BA.2, BA.3, BA.4, and BA.5 subvariants (3, 4, 34). Furthermore, the reporting of local SARS-CoV-2 sequences made the epidemic more immediate to the Ministries of Health from the reporting African countries. It became clear early on that the viral evolution is global and that the transmission of the virus is extremely rapid, which guided mitigation strategies. The generation and availability of local sequences also validated local diagnostics and allowed investigators to determine whether nucleic acid-based diagnostics that were in use could still detect local variants. The detection of SARS-CoV-2 in returning travelers and truck drivers indicated routes that the virus might be using to enter a country and guided early efforts to slow virus entry and gain time to establish vaccination plans. Later, the difficulty of stopping the virus at borders combined with data showing that the variants were already in community circulation allowed public health officials to focus efforts and limited resources on vaccination rather than on border controls. The detection and reporting of the more-recent lineages with enhanced transmission (i.e., Omicron) and the ability to bypass existing immunity is important information and an early alert to public health officials globally that the epidemic is still proceeding. As the pandemic progresses in an evolving global context, we provide evidence that with each new variant, transmission dynamics are changing and the use of sequencing with phylogenetics could potentially alter decisions of public health measures. For example, the demonstrated shift away from regional dynamics of Alpha and Beta toward more global patterns with Delta and Omicron can provide insights to public health officials as they anticipate epidemic developments locally. With Omicron, it became clear that although the variant expanded first in Africa, the continent ultimately had a minimal role in global dissemination and that continental expansion beyond southern Africa was most influenced by external introductions, in contrast to the Beta variant. All of these public health benefits to sequencing SARS-CoV-2 are primarily amplified, as we show in this study,

if the sequencing can be conducted locally within a country, which strongly supports the continued investment into pathogen sequencing on the continent.

Despite the recent successful expansion of genomics surveillance in Africa, additional work is necessary. Even with investments from the Africa CDC–Africa Pathogen Genomics Initiative and other investments, there are still 16 countries with no sequencing capacity within their own borders. The only option for these countries is to send samples to continental sequencing hubs or to centers outside of the continent, which increases turnaround times and limits the utility of genomic surveillance for public health decision-making. Secondly, not all countries are willing to share data openly in a timely fashion for fear of being subject to travel bans or restrictions that could bring substantial economic harm. Such hesitancy has obvious potential ramifications for the future of genomic surveillance on the continent. Furthermore, with the expansion of sequencing on the continent, there is a growing need for more bioinformatics support and knowledge to allow investigators to analyze and report their data in a reasonable time frame that makes it useful for a public health response. It is also clear that the SARS-CoV-2 sequencing primers are not a static development and may require updating as the virus evolves. A number of research groups have been addressing the SARS-CoV-2 sequencing primer questions. Issues of gaps in the genomes due to missing amplicons have been discussed (56, 57). The ARTIC primer set has gone through a number of revisions to accommodate virus evolution (39, 40). Additional longer amplicon methods have been published (58–60), including methods to use a subset of ARTIC primers (61).

The patterns we describe here are of course limited to reported cases and apply to both the phylogeographic as well as the epidemiology inferences. As such, the results need to be interpreted with these limitations in mind. Our primary phylogeographic inference relied on a sampling strategy that considered all high-quality African sequences and an equal number of external references. Though this strategy has the advantage of placing all African sequences in a phylogenetic context, it introduces a bias when applied to discrete ancestral-state reconstruction because more internal nodes are inferred to be from Africa. To address this, we performed an even sampling of global cases, based on reported case counts through time, to compare against our oversampled inference. The even-sampling approach has the benefit that the discrete ancestral-state reconstruction is not biased by uneven sampling. After comparing the two, there are obvious differences, most notably that the number of inferred introductions into Africa is proportional to sampling proportions (fig. S16) because we no

longer consider all African sequences but rather just a small subset against a global sample. However, inferences from the two approaches correspond well with one another. For example, considering Alpha, we still observed that the vast majority of introductions into Africa originated from Western Europe. Patterns of dissemination within Africa are more robustly comparable between the two, for instance, that countries in West Africa were the biggest source of Alpha within the continent. High concordance between the two inference methods was also observed for other VOCs for dispersal routes within Africa, which gives us confidence in the inferred patterns we observe here. Although we represent an inference based on oversampling and case-sensitive sampling, it is, at present, not possible to explore how undersampling affects the phylogeographic reconstruction because of uneven testing rates. Additionally, the robustness of the phylogeographic inference can also be affected by the underlying methodology that is used. Broad consensus would favor the use of Bayesian methods for phylogeographic reconstruction, which is often considered to be the “gold standard” in the field. The main drawbacks of Bayesian methods are that they can only be applied to a relatively small number of sequences at a time (<1000) and they are extremely computationally and time intensive. Given the explosion of sequence data over the past 2 years, the scientific community will have to adapt or put forth new analytical methods to fully capitalize on the global sequencing efforts for SARS-CoV-2.

Despite our best attempts to consider and minimize genomic sampling bias, the accuracy of the resulting phylogenetic inferences is limited by the available epidemiological and genomic data, leading to unaccounted biases in the estimates of viral movements. This includes limited testing and subsequent sequencing in many African countries. Although the percentage of reported cases sequenced in African countries (0.01 to 10%, mean = 1.27%) is not far from global figures (0.01 to 16%, mean = 1.31%), testing rates and infection-to-detection ratios in Africa were some of the lowest globally (38, 62). Together with estimates of excess mortality being as much as 20-fold greater than the reported numbers in African countries (63), these are strong indications of undetected and underreported epidemic sizes in Africa, leading to undersampling of genomic data (62) and thus underestimates of viral exchange inferences in our study. Some countries with no publicly available SARS-CoV-2 sequences are, by definition, completely missing in our inference. This in turn means that inferred routes of viral transmission within Africa could be missing important intermediate locations, although this is potentially true around the world. Nevertheless, we believe that the viral movement inferences that we discuss in this

study provide a likely qualitative description of the patterns of SARS-CoV-2 migration into, out of, and within Africa.

Finally, we should also mention uneven sequencing and reporting standards across the different laboratories on the continent—and globally, for that matter. Different groups use different measures for what constitutes a high-quality sequence (e.g., 70 versus 80% sequence coverage) or use different sequencing depth coverage. This lack of global standardization complicates the direct comparison of sequences that may have been submitted to GISAID using different criteria, further biasing any inference. Given the sheer size of SARS-CoV-2 sequencing, with ~10 million whole-genome sequences shared on the GISAID database (date of access: 31 March 2022), there is an urgent need for global standards with regard to sequence quality and associated metadata.

Africa needs to continue expanding genomic sequencing technologies on the continent in conjunction with diagnostic capabilities. This holds true not just for SARS-CoV-2 but also for other emerging or reemerging pathogens on the continent. For example, in February 2022, the WHO announced the reemergence of wild polio in Africa, and sporadic influenza H1N1, measles, and Ebola outbreaks continue to occur on the continent. The Africa CDC has estimated that more than 100 pathogen outbreaks are reported across the continent every year. Beyond the current pandemic, continued investment in diagnostic and sequencing capacity for these pathogens could serve the public health of the continent well into the 21st century.

Methods and methods

Ethics statement

This project relied on sequence data and associated metadata that are publicly shared by the GISAID data repository and adhere to the terms and conditions laid out by GISAID (16). The African samples processed in this study were obtained anonymously from material exceeding the routine diagnosis of SARS-CoV-2 in African public and private health laboratories. Individual institutional review board references or material transfer agreements (MTAs) for countries are as follows: Angola (MTA - CON8260); Botswana—genomic surveillance in Botswana was approved by the Health Research and Development Committee (protocol HPDME 13/18/1); Egypt—surveillance in Egypt was approved by the Research Ethics Committee of the National Research Centre (Egypt) (protocol number 14 155, dated 22 March 2020); Kenya—samples were collected under the Ministry of Health protocols as part of the national COVID-19 public health response, and the whole-genome sequencing study protocol was reviewed and approved by the Scientific and Ethics Review Committee (SERU) at Kenya Medical Research Institute (KEMRI), Nairobi,

Kenya (SERU protocol #4035); Nigeria (NHREC/01/01/2007), Mali—study of the sequence of SARS-CoV-2 isolates in Mali, Letter of Ethical Committee (NO-2020/201/CE/FMPOS/FAPH of 09/17/2020); Mozambique (MTA - CON7800); Malawi (MTA - CON8265); South Africa—the use of South African samples for sequencing and genomic surveillance was approved by University of KwaZulu-Natal Biomedical Research Ethics Committee (ref. BREC/00001510/2020), the University of the Witwatersrand Human Research Ethics Committee (HREC) (ref. M180832), Stellenbosch University HREC (ref. N20/04/008_COVID-19), the University of the Free State Research Ethics Committee (ref. UFS-HSD2020/1860/2710), and the University of Cape Town HREC (ref. 383/2020); Tunisia—for sequences derived from sampling in Tunisia, all patients provided their informed consent to use their samples for sequencing of the viral genomes, and the ethical agreement was provided to the research project ADAGE (PRFCOVID19GP2) by the Committee of Protection of Persons (Tunisian Ministry of Health) under the reference CPP SUD N 0265/2020; Uganda—the use of samples and sequences from Uganda was approved by the Uganda Virus Research Institute, Research and Ethics Committee UVRI-REC Federalwide Assurance (FWA) no. 00001354, study reference GC/127/20/04/771, and by the Uganda National Council for Science and Technology, reference number HS936ES; and Zimbabwe (MTA - CON8271).

Epidemiological and genomic data dynamics

We analyzed trends in daily numbers of cases of SARS-CoV-2 in Africa up to 31 March 2022 from publicly released data provided by the Our World in Data repository for the continent of Africa (<https://github.com/owid/covid-19-data/tree/master/public/data>) as a whole and for individual countries (2). To provide a comparable view of epidemiological dynamics over time in various countries, the variable under primary consideration for Fig. 1 was “new cases per million (smoothed).” To calculate the genomic sampling proportion and frequency for each country for Fig. 2, the total number of recorded cases as of 31 March 2022 was considered, as well as the total length of time for which each country had recorded cases of SARS-CoV-2.

Genomic metadata was downloaded for all African entries on GISAID for the same time period (date of access: 31 March 2022). From this, information extracted from all entries for this study included the date of sampling, country of sampling, viral lineage and clade, originating laboratory, sequencing laboratory, and date of submission to the GISAID database. The geographical locations of the originating and sequencing laboratories were manually curated. Sequences originating and sequenced in the same country were defined as locally sequenced, irrespective of specific laboratory or

finer location. Sequences originating in one African country and sequenced in another were defined as sequenced within regional sequencing networks. Sequences sequenced in a location not within Africa were labeled as sequenced outside Africa. Sequencing turnaround time was defined as the number of days that had elapsed from specimen collection to sequence submission to GISAID. Sequencing technology information for all African entries was also downloaded from GISAID on 31 March 2022.

Primer choice and sequencing outcomes

All SARS-CoV-2 genomes from African countries were retrieved from GISAID (16) for submission dates from 1 December 2019 to 31 March 2022, yielding 100,470 entries. Associated metadata for the entries were also retrieved, including collection date, submission date, country, viral strain, and sequencing technology. Data on the primers used for the sequencing were requested from investigators and yielded primer data for 13,973 of the entries (~13%). The total N (bases with low sequence depth) per genome were counted, the results of which were then used for genome quality analysis and visualization. Gap locations in the genomes were mapped and visualized with respect to the original Wuhan strain (64).

Phylogenetic investigation

All African sequences on the GISAID sequence database (16) were downloaded on 31 March 2022 ($n = 100,470$). Of these, Alpha accounted for 3851 sequences, Beta accounted for 14,548 sequences, Delta accounted for 35,027 sequences, Omicron accounted for 21,708 sequences, and 25,336 sequences were classified as non-VOCs. Before any phylogenetic inference, we performed some quality assessment on the sequences to exclude incomplete or problematic sequences as well as sequences lacking complete metadata. Briefly, all African sequences were passed through the NextClade analysis pipeline (65) to identify and exclude (i) sequences missing >10% of the SARS-CoV-2 genome, (ii) sequences that deviate by >70 nucleotides from the Wuhan reference strain, (iii) sequences with >10 ambiguous bases, (iv) clustered mutations, and (v) sequences flagged with private mutations by NextClade. Additionally, Omicron variants were screened for traces of viral recombination with RDP5.23 (66) using default settings and a p value of ≤ 0.05 as evidence of recombination. A large number of sequences were removed ($n = 57,421$), with incomplete sequences (<90% genome coverage) being the biggest contributor. This produced a final African dataset of 43,049 high-quality African sequences. Because of the sheer size of the dataset, we opted to perform independent phylogenetic inferences on the main VOCs (Alpha, Beta, Delta, and Omicron BA.1 and BA.2) that have spread on the

African continent, as well as a separate inference for all non-VOC SARS-CoV-2 sequences.

To evaluate the spread of the virus on the African continent, we aligned the African datasets against a large number of globally representative sequences from around the world. Because of the oversampling of some variants or lineages, we performed a random down sampling while retaining the oldest two known variants from each country. Reference sequences were respectively aligned with their African counterparts independently with NextAlign (65). Each of the alignments was then used to infer maximum likelihood (ML) tree topologies in FastTree v 2.0 (67) using the general time reversible model of nucleotide substitution and a total of 100 bootstrap replicates (68). The resulting ML tree topologies were first inspected in TempEst (69) to identify any sequences that deviate more than 0.0001 from the residual mean. After the removal of potential outliers in R with the ape package (70), the resulting ML trees were then transformed into time-calibrated phylogenies in TreeTime (71) by applying a rate of 8×10^{-4} substitutions per site per year (72) to transform the branches into units of calendar time. Time-calibrated trees were then visualized, along with associated metadata, in R using ggtree (73) and other packages.

We performed a basic viral dispersal analysis for each of the VOCs (excluding Gamma) as well as for the non-VOC dataset. Briefly, a migration model was fitted to each of the time-calibrated tree topologies in TreeTime, mapping the country location of sampled sequences to the external tips of the trees. The migration model of TreeTime also infers the most likely location for internal nodes in the trees. Using a custom python script, we could then count the number of state changes by iterating over each phylogeny from the root to the external tips. We count state changes when an internal node transitions from one country to a different country in the resulting child node or tip(s). The timing of transition events is then recorded, which serves as the estimated import or export event. To infer some confidence around these estimates, we performed 10 replicates for each of the datasets by random selection from the 100 bootstrap trees. Because of the high uncertainty in the inferred locations for deep internal nodes in the trees, we truncated state changes to the earliest date of sampling in each dataset. All data analytics were performed using custom python and R scripts, and the results were visualized using the ggplot libraries (74). Such phylogeographic methods are always subject to uneven sampling through time (i.e., over the course of the pandemic) and through space (by sampling location). To address this, we have performed a case-sensitive analysis to investigate the effects of oversampling African locations on the inferred number of viral introductions. Furthermore, in a previous analysis (15), we

performed a sensitivity analysis to address some of these issues and found no substantial variations in estimates.

Case-sensitive phylogeographic inference

To address the potential oversampling of African sequences relative to global reference in the above-mentioned analyses, we performed another phylogeographic inference on subsamples based on global case counts to try to eliminate oversampling bias in our inference. To this end, we considered all high-quality sequences for each of the VOCs (Alpha, Beta, Delta, and Omicron BA.1 and BA.2) globally over the same sampling period (until 31 March 2022). We used subsampler (<https://github.com/andersonbrito/subsampler>) to generate subsamples for each variant based on globally reported cases. In short, subsampler uses a case-count matrix of daily cases, along with the fasta sequences and GISAID associated metadata, to sample a user-defined number of sequences. For each VOC and for BA.1 and BA.2, we performed 10 samplings using different number seeds to sample datasets of ~20,000. Once again, sampled sequences were screened for viral recombination as described above and sequences with signs of recombination were removed. Sub-sampler has the added advantage that it disregards poor quality sequences (e.g., <90% coverage) and sequences with missing metadata (e.g., exact date of sampling). Each dataset was then subjected to the same analytical pipeline as mentioned above to infer the viral transitions between Africa and the rest of the world.

Regional and country-specific NextStrain builds

To investigate more-granular changes in lineage dynamics within a specific country or region in Africa, we used the NextStrain pipeline (<https://github.com/nextstrain/ncov>) to generate the regional and country-specific builds for African countries (75). First, all sequence data and metadata were retrieved from the GISAID sequence database and filtered for Africa based on the “region” tab for inclusion in regional and country-specific African builds. For country-specific builds, ~4000 sequences from a given country were randomly selected and analyzed against ~1000 randomly selected sequences from the Africa “nextregions” records that do not match the focal country of interest. For regional (e.g., West Africa) builds, ~4000 sequences from the focal region were selected at random and analyzed against ~1000 randomly selected sequences from the Africa “nextregions” records that do not match the focal region of interest. The methodological pipeline for NextStrain is well documented and performs all analyses within one workflow, including filtering of sequences, alignment, tree inference, molecular clock, and ancestral-state reconstruction. For more information,

please visit <https://docs.nextstrain.org/en/latest/index.html>.

All regional and country-specific builds are regularly updated to keep track of the evolving pandemic on the continent. All builds are publicly available under the links provided in tables S1 and S2 as well as on the NextStrain web page (<https://nextstrain.org/sars-cov-2/#datasets>).

REFERENCES AND NOTES

- Q. Li et al., Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020). doi: [10.1056/NEJMoa2001316](https://doi.org/10.1056/NEJMoa2001316); PMID: 31995857
- J. Hasell et al., A cross-country database of COVID-19 testing. *Sci. Data* **7**, 345 (2020). doi: [10.1038/s41597-020-00688-8](https://doi.org/10.1038/s41597-020-00688-8); PMID: 33032556
- R. Viana et al., Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679–686 (2022). doi: [10.1038/s41586-022-04411-y](https://doi.org/10.1038/s41586-022-04411-y); PMID: 35042229
- H. Tegally et al., Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021). doi: [10.1038/s41586-021-03402-9](https://doi.org/10.1038/s41586-021-03402-9); PMID: 33690265
- D. P. Martin et al., The emergence and ongoing convergent evolution of the SARS-CoV-2 N501Y lineages. *Cell* **184**, 5189–5200.e7 (2021). doi: [10.1016/j.cell.2021.09.003](https://doi.org/10.1016/j.cell.2021.09.003); PMID: 34537136
- F. Campbell et al., Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill.* **26**, (2021). doi: [10.2807/1560-7917.ES.2021.26.24.2100509](https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509); PMID: 34142653
- B. Korber et al., Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827.e19 (2020). doi: [10.1016/j.cell.2020.06.043](https://doi.org/10.1016/j.cell.2020.06.043); PMID: 32679768
- E. Hacısuleyman et al., Vaccine breakthrough infections with SARS-CoV-2 variants. *N. Engl. J. Med.* **384**, 2212–2218 (2021). doi: [10.1056/NEJMoa2105000](https://doi.org/10.1056/NEJMoa2105000); PMID: 33882219
- D. Planas et al., Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280 (2021). doi: [10.1038/s41586-021-03777-9](https://doi.org/10.1038/s41586-021-03777-9); PMID: 34237773
- S. Yue et al., Sensitivity of SARS-CoV-2 variants to neutralization by convalescent sera and a VH3-30 monoclonal antibody. *Front. Immunol.* **12**, 751584 (2021). doi: [10.3389/fimmu.2021.751584](https://doi.org/10.3389/fimmu.2021.751584); PMID: 34630430
- S. Cele et al., Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. *Nature* **593**, 142–146 (2021). doi: [10.1038/s41586-021-03471-w](https://doi.org/10.1038/s41586-021-03471-w); PMID: 33780970
- B. Meng et al., Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep.* **35**, 109292 (2021). doi: [10.1016/j.celrep.2021.109292](https://doi.org/10.1016/j.celrep.2021.109292); PMID: 34166617
- P. Mlcochova et al., SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature* **599**, 114–119 (2021). doi: [10.1038/s41586-021-03944-y](https://doi.org/10.1038/s41586-021-03944-y); PMID: 34488225
- N. R. Faria et al., Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021). doi: [10.1126/science.abb2644](https://doi.org/10.1126/science.abb2644); PMID: 33853970
- E. Wilkinson et al., A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science* **374**, 423–431 (2021). doi: [10.1126/science.abb4336](https://doi.org/10.1126/science.abb4336); PMID: 34672751
- Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, 30494 (2017). doi: [10.2807/1560-7917.ES.2017.22.13.30494](https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494); PMID: 28382917
- C. Kuiken, B. Korber, R. W. Shafer, HIV sequence databases. *AIDS Rev.* **5**, 52–61 (2003). PMID: 12875108
- D. L. Bugembe et al., Main routes of entry and genomic diversity of SARS-CoV-2, Uganda. *Emerg. Infect. Dis.* **26**, 2411–2415 (2020). doi: [10.3201/eid2610.202575](https://doi.org/10.3201/eid2610.202575); PMID: 32614767
- T. Mashe et al., Genomic epidemiology and the role of international and regional travel in the SARS-CoV-2 epidemic in Zimbabwe: A retrospective study of routinely collected surveillance data. *Lancet Glob. Health* **9**, e1658–e1666 (2021). doi: [10.1016/S2214-109X\(21\)00434-4](https://doi.org/10.1016/S2214-109X(21)00434-4); PMID: 34695371
- A. Chouikha et al., Molecular epidemiology of SARS-CoV-2 in Tunisia (North Africa) through several successive waves of COVID-19. *Viruses* **14**, 624 (2022). doi: [10.3390/v14030624](https://doi.org/10.3390/v14030624); PMID: 35337031

21. F. Ntouni *et al.*, Genomic surveillance of SARS-CoV-2 in the Republic of Congo. *Int. J. Infect. Dis.* **105**, 735–738 (2021). doi: [10.1016/j.ijid.2021.03.036](https://doi.org/10.1016/j.ijid.2021.03.036); pmid: [33737129](https://pubmed.ncbi.nlm.nih.gov/33737129/)
22. Y. Butera *et al.*, Genomic sequencing of SARS-CoV-2 in Rwanda: Evolution and regional dynamics. medRxiv 2021.04.02.21254839 [Preprint] (2021); <https://doi.org/10.1101/2021.04.02.21254839>.
23. C. N. Agoti *et al.*, Detection of SARS-CoV-2 variant 501Y.V2 in Comoros Islands in January 2021. *Wellcome Open Res.* **6**, 192 (2021). doi: [10.12688/wellcomeopenres.16889.1](https://doi.org/10.12688/wellcomeopenres.16889.1); pmid: [35071798](https://pubmed.ncbi.nlm.nih.gov/35071798/)
24. J. M. Morobe *et al.*, Genomic Epidemiology of SARS-CoV-2 in Seychelles, 2020–2021. *Viruses* **14**, 1318 (2022). doi: [10.3390/v14061318](https://doi.org/10.3390/v14061318); pmid: [35746789](https://pubmed.ncbi.nlm.nih.gov/35746789/)
25. C. M. Morang'a *et al.*, Genetic diversity of SARS-CoV-2 infections in Ghana from 2020–2021. *Nat. Commun.* **13**, 2494 (2022). doi: [10.1038/s41467-022-30219-5](https://doi.org/10.1038/s41467-022-30219-5); pmid: [35523782](https://pubmed.ncbi.nlm.nih.gov/35523782/)
26. C. N. Agoti *et al.*, Transmission networks of SARS-CoV-2 in Coastal Kenya during the first two waves: A retrospective genomic study. *eLife* **11**, e71703 (2022). doi: [10.7554/eLife.71703](https://doi.org/10.7554/eLife.71703); pmid: [35699426](https://pubmed.ncbi.nlm.nih.gov/35699426/)
27. S. P. C. Brand *et al.*, COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science* **374**, 989–994 (2021). doi: [10.1126/science.abk0414](https://doi.org/10.1126/science.abk0414); pmid: [34618602](https://pubmed.ncbi.nlm.nih.gov/34618602/)
28. G. Githinji *et al.*, Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. *Nat. Commun.* **12**, 4809 (2021). doi: [10.1038/s41467-021-25137-x](https://doi.org/10.1038/s41467-021-25137-x); pmid: [34376689](https://pubmed.ncbi.nlm.nih.gov/34376689/)
29. H. Tegally *et al.*, Sixteen novel lineages of SARS-CoV-2 in South Africa. *Nat. Med.* **27**, 440–446 (2021). doi: [10.1038/s41591-021-02155-3](https://doi.org/10.1038/s41591-021-02155-3); pmid: [33531709](https://pubmed.ncbi.nlm.nih.gov/33531709/)
30. W. H. Roshdy *et al.*, SARS-CoV-2 Genetic diversity and lineage dynamics in Egypt. medRxiv 2022.01.05.22268646 [Preprint] (2022); <https://doi.org/10.1101/2022.01.05.22268646>
31. C. Scheepers *et al.*, Emergence and phenotypic characterization of the global SARS-CoV-2 C.1.2 lineage. *Nat. Commun.* **13**, 1976 (2022). doi: [10.1038/s41467-022-29579-9](https://doi.org/10.1038/s41467-022-29579-9); pmid: [35396511](https://pubmed.ncbi.nlm.nih.gov/35396511/)
32. G. Dudas *et al.*, Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions. *Nat. Commun.* **12**, 5769 (2021). doi: [10.1038/s41467-021-26055-8](https://doi.org/10.1038/s41467-021-26055-8); pmid: [34599175](https://pubmed.ncbi.nlm.nih.gov/34599175/)
33. D. L. Bugembe *et al.*, Emergence and spread of a SARS-CoV-2 lineage A variant (A.23.1) with altered spike protein in Uganda. *Nat. Microbiol.* **6**, 1094–1101 (2021). doi: [10.1038/s41564-021-00933-9](https://doi.org/10.1038/s41564-021-00933-9); pmid: [34163035](https://pubmed.ncbi.nlm.nih.gov/34163035/)
34. H. Tegally *et al.*, Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat. Med.* (2022). doi: [10.1038/s41591-022-01911-2](https://doi.org/10.1038/s41591-022-01911-2); pmid: [35760080](https://pubmed.ncbi.nlm.nih.gov/35760080/)
35. S. A. Madhi *et al.*, Population immunity and Covid-19 severity with omicron variant in South Africa. *N. Engl. J. Med.* **386**, 1314–1326 (2022). doi: [10.1056/NEJMoa2119658](https://doi.org/10.1056/NEJMoa2119658); pmid: [35196424](https://pubmed.ncbi.nlm.nih.gov/35196424/)
36. N. Wolter *et al.*, Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: A data linkage study. *Lancet* **399**, 437–446 (2022). doi: [10.1016/S0140-6736\(22\)00017-4](https://doi.org/10.1016/S0140-6736(22)00017-4); pmid: [35065011](https://pubmed.ncbi.nlm.nih.gov/35065011/)
37. H. C. Lewis *et al.*, SARS-CoV-2 infection in Africa: A systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *BMJ Glob. Health* **7**, e008793 (2022). doi: [10.1136/bmjgh-2022-008793](https://doi.org/10.1136/bmjgh-2022-008793); pmid: [35998978](https://pubmed.ncbi.nlm.nih.gov/35998978/)
38. R. M. Barber *et al.*, Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: A statistical analysis. *Lancet* **399**, 2351–2380 (2022). doi: [10.1016/S0140-6736\(22\)00484-6](https://doi.org/10.1016/S0140-6736(22)00484-6); pmid: [35405084](https://pubmed.ncbi.nlm.nih.gov/35405084/)
39. J. Quick, nCoV-2019 sequencing protocol v3 (LoCost) (2020).
40. J. R. Tyson, P. James *et al.*, Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv 2020.09.04.283077 [Preprint] (2020); <https://doi.org/10.1101/2020.09.04.283077>.
41. M. Cotten, D. Lule Bugembe, P. Kaleebu, M. V T Phan, Alternate primers for whole-genome SARS-CoV-2 sequencing. *Virus Evol.* **7**, veab006 (2021). doi: [10.1093/ve/veab006](https://doi.org/10.1093/ve/veab006); pmid: [33841912](https://pubmed.ncbi.nlm.nih.gov/33841912/)
42. H. Tegally *et al.*, A novel and expanding SARS-CoV-2 variant, B.1.1.318, dominates infections in Mauritius. medRxiv 2021.06.16.21259017 [Preprint] (2021); <https://doi.org/10.1101/2021.06.16.21259017>.
43. A. N. Zekri *et al.*, Characterization of the SARS-CoV-2 genomes in Egypt in first and second waves of infection. *Sci. Rep.* **11**, 21632 (2021). doi: [10.1038/s41598-021-99014-4](https://doi.org/10.1038/s41598-021-99014-4); pmid: [34732835](https://pubmed.ncbi.nlm.nih.gov/34732835/)
44. C. Nasimiyu *et al.*, Imported SARS-CoV-2 variants of concern drove spread of infections across Kenya during the second year of the pandemic. *COVID* **2**, 586–598 (2022). doi: [10.3390/covid2050044](https://doi.org/10.3390/covid2050044); pmid: [35262086](https://pubmed.ncbi.nlm.nih.gov/35262086/)
45. M. U. G. Kraemer *et al.*, Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **373**, 889–895 (2021). doi: [10.1126/science.abj0113](https://doi.org/10.1126/science.abj0113); pmid: [34301854](https://pubmed.ncbi.nlm.nih.gov/34301854/)
46. S. J. Lycett *et al.*, COVID-19 Genomics UK (COG-UK) Consortium, Epidemic waves of COVID-19 in Scotland: a genomic perspective on the impact of the introduction and relaxation of lockdown on SARS-CoV-2. medRxiv 2021.01.08.20248677 [Preprint] (2021); <https://doi.org/10.1101/2021.01.08.20248677>.
47. P. W. G. Mallon *et al.*, Whole-genome sequencing of SARS-CoV-2 in the Republic of Ireland during waves 1 and 2 of the pandemic. medRxiv 2021.02.09.21251402 [Preprint] (2021); <https://doi.org/10.1101/2021.02.09.21251402>.
48. H. Tegally *et al.*, Rapid replacement of the Beta variant by the Delta variant in South Africa. medRxiv 2021.09.23.21264018 [Preprint] (2021); <https://doi.org/10.1101/2021.09.23.21264018>.
49. S. Chang *et al.*, Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021). doi: [10.1038/s41586-020-2923-3](https://doi.org/10.1038/s41586-020-2923-3); pmid: [33171481](https://pubmed.ncbi.nlm.nih.gov/33171481/)
50. M. Chinazzi *et al.*, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020). doi: [10.1126/science.aba9757](https://doi.org/10.1126/science.aba9757); pmid: [32144116](https://pubmed.ncbi.nlm.nih.gov/32144116/)
51. M. U. G. Kraemer *et al.*, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020). doi: [10.1126/science.abb4218](https://doi.org/10.1126/science.abb4218); pmid: [32213647](https://pubmed.ncbi.nlm.nih.gov/32213647/)
52. P. Nouvellet *et al.*, Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **12**, 1090 (2021). doi: [10.1038/s41467-021-21358-2](https://doi.org/10.1038/s41467-021-21358-2); pmid: [33597546](https://pubmed.ncbi.nlm.nih.gov/33597546/)
53. C. Xiong, S. Hu, M. Yang, W. Luo, L. Zhang, Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 27087–27089 (2020). doi: [10.1073/pnas.2010836117](https://doi.org/10.1073/pnas.2010836117); pmid: [33060300](https://pubmed.ncbi.nlm.nih.gov/33060300/)
54. S. Pillay *et al.*, Whole genome sequencing of SARS-CoV-2: Adapting Illumina protocols for quick and accurate outbreak investigation during a pandemic. *Genes* **11**, 949 (2020). doi: [10.3390/genes11080949](https://doi.org/10.3390/genes11080949); pmid: [32824573](https://pubmed.ncbi.nlm.nih.gov/32824573/)
55. L. Singh *et al.*, Targeted Sanger sequencing to recover key mutations in SARS-CoV-2 variant genome assemblies produced by next-generation sequencing. *Microb. Genom.* **8**, (2022). doi: [10.1099/mgen.0.000774](https://doi.org/10.1099/mgen.0.000774); pmid: [35294336](https://pubmed.ncbi.nlm.nih.gov/35294336/)
56. A. J. Page *et al.*, Large-scale sequencing of SARS-CoV-2 genomes from one region allows detailed epidemiology and enables local outbreak management. *Microb. Genom.* **7**, (2021). doi: [10.1099/mgen.0.000589](https://doi.org/10.1099/mgen.0.000589); pmid: [34184982](https://pubmed.ncbi.nlm.nih.gov/34184982/)
57. N. De Maio, C. Walker, R. Borges, L. Weilgung, G. Slodkowitz, N. Goldman, Issues with SARS-CoV-2 sequencing data. *Virological.org* (2020); <https://virological.org/t/issues-with-sars-cov-2-sequencing-data/4743>.
58. N. E. Freed, M. Vlková, M. B. Faisal, O. K. Silander, Rapid and inexpensive whole-genome sequencing of SARS-CoV-2 using 1200 bp tiled amplicons and Oxford Nanopore Rapid Barcoding. *Biol. Methods Protoc.* **5**, bpaa014 (2020). doi: [10.1093/biomethods/bpaa014](https://doi.org/10.1093/biomethods/bpaa014); pmid: [33029559](https://pubmed.ncbi.nlm.nih.gov/33029559/)
59. J.-S. Eden *et al.*, An emergent clade of SARS-CoV-2 linked to returned travellers from Iran. *Virus Evol.* **6**, veaa027 (2020). doi: [10.1093/ve/veaa027](https://doi.org/10.1093/ve/veaa027); pmid: [32296544](https://pubmed.ncbi.nlm.nih.gov/32296544/)
60. A. S. Gonzalez-Reiche *et al.*, Introductions and early spread of SARS-CoV-2 in the New York City area. *Science* **369**, 297–301 (2020). doi: [10.1126/science.abc1917](https://doi.org/10.1126/science.abc1917); pmid: [32471856](https://pubmed.ncbi.nlm.nih.gov/32471856/)
61. K. Itokawa, T. Sekizuka, M. Hashino, R. Tanaka, M. Kuroda, Disentangling primer interactions improves SARS-CoV-2 genome sequencing by multiplex tiling PCR. *PLOS ONE* **15**, e0239403 (2020). doi: [10.1371/journal.pone.0239403](https://doi.org/10.1371/journal.pone.0239403); pmid: [32946527](https://pubmed.ncbi.nlm.nih.gov/32946527/)
62. A. X. Han *et al.*, Low testing rates limit the ability of genomic surveillance programs to monitor SARS-CoV-2 variants: a mathematical modelling study. medRxiv 2022.05.20.22275319 [Preprint] (2022); <https://doi.org/10.1101/2022.05.20.22275319>.
63. H. Wang *et al.*, Estimating excess mortality due to the COVID-19 pandemic: A systematic analysis of COVID-19-related mortality, 2020–21. *Lancet* **399**, 1513–1536 (2022). doi: [10.1016/S0140-6736\(21\)02796-3](https://doi.org/10.1016/S0140-6736(21)02796-3); pmid: [35279232](https://pubmed.ncbi.nlm.nih.gov/35279232/)
64. F. Wu *et al.*, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020). doi: [10.1038/s41586-020-2008-3](https://doi.org/10.1038/s41586-020-2008-3); pmid: [32015508](https://pubmed.ncbi.nlm.nih.gov/32015508/)
65. I. Aksamentov, C. Roemer, E. Hodcroft, R. Neher, Nextclade: Clade assignment, mutation calling and quality control for viral genomes. *J. Open Source Softw.* **6**, 3773 (2021). doi: [10.21105/joss.03773](https://doi.org/10.21105/joss.03773)
66. D. P. Martin *et al.*, RDP5: A computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evol.* **7**, veaa087 (2020). doi: [10.1093/ve/veaa087](https://doi.org/10.1093/ve/veaa087); pmid: [33936774](https://pubmed.ncbi.nlm.nih.gov/33936774/)
67. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010). doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490); pmid: [20224823](https://pubmed.ncbi.nlm.nih.gov/20224823/)
68. J. Felsenstein, Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783–791 (1985). doi: [10.1111/j.1558-5646.1985.tb00420.x](https://doi.org/10.1111/j.1558-5646.1985.tb00420.x); pmid: [28561359](https://pubmed.ncbi.nlm.nih.gov/28561359/)
69. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterogenous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016). doi: [10.1093/ve/vew007](https://doi.org/10.1093/ve/vew007); pmid: [27774300](https://pubmed.ncbi.nlm.nih.gov/27774300/)
70. A.-A. Popescu, K. T. Huber, E. Paradis, ape 3.0: New tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* **28**, 1536–1537 (2012). doi: [10.1093/bioinformatics/bts184](https://doi.org/10.1093/bioinformatics/bts184); pmid: [22495750](https://pubmed.ncbi.nlm.nih.gov/22495750/)
71. P. Sagulenko, V. Puller, R. A. Neher, TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vev042 (2018). doi: [10.1093/ve/vev042](https://doi.org/10.1093/ve/vev042); pmid: [29340210](https://pubmed.ncbi.nlm.nih.gov/29340210/)
72. S. Wang *et al.*, Molecular evolutionary characteristics of SARS-CoV-2 emerging in the United States. *J. Med. Virol.* **94**, 310–317 (2022). doi: [10.1002/jmv.27331](https://doi.org/10.1002/jmv.27331); pmid: [34506640](https://pubmed.ncbi.nlm.nih.gov/34506640/)
73. G. Yu, Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinformatics* **69**, e96 (2020). doi: [10.1002/cpb1.96](https://doi.org/10.1002/cpb1.96); pmid: [32162851](https://pubmed.ncbi.nlm.nih.gov/32162851/)
74. H. Wickham, ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011). doi: [10.1002/wics.147](https://doi.org/10.1002/wics.147)
75. J. Hadfield *et al.*, Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018). doi: [10.1093/bioinformatics/bty407](https://doi.org/10.1093/bioinformatics/bty407); pmid: [29790939](https://pubmed.ncbi.nlm.nih.gov/29790939/)
76. S. E. James, CERIKRISP/SARS-CoV-2-epidemic-in-Africa: Expanding Africa SARS-CoV-2 sequencing capacity in a fast evolving pandemic analysis. Zenodo (2022); <https://doi.org/10.5281/zenodo.7006806>.
77. T. Ward, A. Johnsen, Understanding an evolving pandemic: An analysis of the clinical time delay distributions of COVID-19 in the United Kingdom. *PLOS ONE* **16**, e0257978 (2021). doi: [10.1371/journal.pone.0257978](https://doi.org/10.1371/journal.pone.0257978); pmid: [34669712](https://pubmed.ncbi.nlm.nih.gov/34669712/)

ACKNOWLEDGMENTS

First and foremost, we acknowledge authors in institutions in Africa and beyond who have made invaluable contributions toward specimen collection and sequencing to produce and share, via GISAID, SARS-CoV-2 genomic data. We also acknowledge the authors from the originating and submitting laboratories worldwide who generated and shared SARS-CoV-2 sequence data, via GISAID, from other regions in the world, which was used to contextualize the African genomic data. A full list of GISAID sequence IDs used in the current study is available in table S4. **Funding:** Sequencing efforts in the African Union Member States were supported by the Africa Centers for Disease Control (Africa CDC)–Africa Pathogen Genomics Initiative (Africa PGI) and the World Health Organization Regional Office for Africa (WHO AFRO) through the transfer of laboratory infrastructure, the provision of reagents, and training. The Africa PGI is supported by the African Union, US Centers for Disease Control and Prevention (CDC), Bill & Melinda Gates Foundation, Illumina Inc., Oxford Nanopore Technologies, and other partners. In addition, all Institut Pasteur organizations and CERMES in Niger are part of the PEPAIR COVID-19–Africa project, which is funded by the French Ministry for European and Foreign Affairs. The KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP) and Centre for Epidemic Response and Innovation (CERI) are supported in part by grants from WHO, the Rockefeller Foundation (HTH 017), the Abbott Pandemic Defense Coalition (APDC), the US National Institutes of Health (NIH) (U01 AI151698) for the United World Antivirus Research Network (UWARN) and the INFORM Africa project through the Institute of Human Virology, Nigeria (IHVN) (U54 TW012041), H3BioNet Africa (grant no. 2020 HTH 062), the World Bank (TF0B8412), the South African Department of Science and Innovation (SA DSI), and the South African Medical Research Council (SAMRC) under the BRICS JAF #2020/049. The International Livestock Research Institute (ILRI) is also supported by the Ministry for Economic Cooperation and Federal Development of Germany (BMZ). Work conducted at the African Centre of Excellence for Genomics of Infectious Diseases (ACEGID) is made possible by support provided to ACEGID by a cohort of generous donors through TED's Audacious Project, including the ELMA Foundation, MacKenzie Scott, the Skoll Foundation, and Open

Philanthropy. Work at ACEGI) was also partly supported by grants from the National Institute of Allergy and Infectious Diseases (NIAID) (<https://www.niaid.nih.gov>), NIH-H3Africa (<https://h3africa.org>) (U01HG007480 and U54HG007480), the World Bank (projects ACE-019 and ACE-IMPACT), the Rockefeller Foundation (grant #2021 HTH), the Africa CDC through the African Society of Laboratory Medicine (ASLM) (grant #INV018978), the Wellcome Trust (project 216619/Z/19/Z), and the Science for Africa Foundation. Sequencing efforts at the National Institute for Communicable Diseases (NICD) were also supported by a conditional grant from the South African National Department of Health as part of the emergency COVID-19 response; a cooperative agreement between the NICD of the National Health Laboratory Service (NHLS) and the CDC (FAIN# U01P001048 and NU51P000930); the South African Medical Research Council (SAMRC) (project number 96838); the ASLM and the Bill & Melinda Gates Foundation (grant number INV-018978); the UK Foreign, Commonwealth and Development Office and the Wellcome Trust (grant no. 221003/Z/20/Z); and the UK Department of Health and Social Care and were managed by the Fleming Fund and performed under the auspices of the SEQAFRICA project. The NICD also acknowledges support from Hyrax Biosciences for the use of their Exatype platform. This was made possible through funding from the South African Medical Research Council, the Department of Science and Innovation, as well as support from the Health Equity Initiative at Amazon Web Services. Funding for sequencing efforts in Angola were supported through Projecto Bongola (N.º 11/ MESCTI/PDCT/2020) and Orçamento Geral do Estado Instituto Nacional de Investigação de Saúde (OGE INIS) (2020/2021). Botswana's sequencing efforts, which were led by the Botswana Harvard AIDS Institute Partnership, were supported by the Foundation for Innovative New Diagnostics (FINDdx), Bill & Melinda Gates Foundation H3ABioNet (U41HG006941), Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), and Fogarty International Center (grant no. 5D43TW009610). H3ABioNet is an initiative of the Human Health and Heredity in Africa Consortium (H3Africa) program of the African Academy of Science (AAS) and the US Department of Health and Human Services (HHS), NIH, and NIAID (5K24AI131928-04; 5K24AI131924-04); SANTHE is a DELTAS Africa Initiative (grant no. DEL-15-006). The DELTAS Africa Initiative is an independent funding scheme of the AAS's Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPADAgency) with funding from the Wellcome Trust (grant #107752/Z/15/Z) and the UK government. From Brazil, J.S.X. was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brazil (CAPES)–Finance Code 001. Sequencing efforts from Côte d'Ivoire were funded by the Robert Koch Institute and the German Federal Ministry of Education and Research (BMBF). Sequencing efforts in the Democratic Republic of the Congo were funded by the Bill & Melinda Gates Foundation under grant INV-018030 awarded to C.B.P. and further supported by funding from the Africa CDC through ASLM for Accelerating SARS-CoV-2 Genomic Surveillance in Africa, the CDC, US Army Medical Research Institute of Infectious Diseases (USAMRIID), Institut de Recherche pour le Développement (IRD)/Montpellier, University of California–Los Angeles (UCLA), and SACIDS FIND. Efforts from Egypt were funded by the Egyptian Ministry of Health, the Egyptian Academy for Scientific Research and Technology (ASRT) JESOR project #3046 (Center for Genome and Microbiome Research), the Cairo University anti-COVID-19 fund, and the Science and Technology Development Fund (STDF), project ID 41907. The sequencing effort in Equatorial Guinea was supported by a public-private partnership, the Bioko Island Malaria Elimination Project, which is composed of the government of Equatorial Guinea Ministries of Mines and Hydrocarbons, and Health and Social Welfare, Marathon EG Production Limited, Noble Energy, Atlantic Methanol Production Company, and EG LNG. Analysis for the Gabon strains was supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan International Cooperation Agency (JICA), and Japan Agency for Medical Research and Development (AMED) (grant number JP21jm0110013) and a grant from AMED (grant number JP21wm0225003). The Centre Interdisciplinaires de Recherches Médicales de Franceville (CIRMF) (Gabon) is funded by the Gabonese Government and TOTAL Energy Inc. CIRMF is a member of the Central Africa Network on Tuberculosis, HIV/AIDS and Malaria (CANTAM), which is supported by the European and Developing Countries Clinical Trials Partnership (EDCTP). The work at the West African Centre for Cell Biology of Infectious Pathogens (WACCBIP) (Ghana) was funded by a grant from the Rockefeller Foundation (2021 HTH 006), an IRD grant (ARIACOV), an African Research Universities Alliance (ARUA) Vaccine Development Hubs

grant with funds from Open Society Foundation, National Institute of Health Research (NIHR) (17.63.91) grants using UK aid from the UK Government for a global health research group for genomic surveillance of malaria in West Africa (Wellcome Sanger Institute, UK), and a World Bank African Centers of Excellence Impact grant (WACCBIP-NCDs: Awandare). In addition to the funding sources from ILRI, Kenya Medical Research Institute (KEMRI) (Kenyan) contributions to sequencing efforts were supported in part by the National Institute for Health Research (NIHR) (project references 17/63/82 and 16/136/33) using UK aid from the UK government to support global health research; the UK Foreign, Commonwealth and Development Office (FCDO) and the Wellcome Trust (grant no. 220985/Z/20/Z); and the Kenya Medical Research Institute (grant no. KEMRI/COV/SPE/012). Contributions from Lesotho were supported by the Africa CDC, ALSM, and South Africa NICD. Liberian efforts were funded by the Africa CDC through a subaward from the Bill & Melinda Gates Foundation, and efforts from Madagascar were funded by the French Ministry for Europe and Foreign Affairs through the REPAIR COVID-19–Africa project coordinated by the Pasteur International Network association. Sequencing from Malawi was supported by the Wellcome Trust. Contributions from Mali were supported by Fogarty International Center and NIAID sections of the NIH under Leidos-15X051, award numbers U2RTW010673 for the West African Center of Excellence for Global Health Bioinformatics Research Training and U19AI089696 and U19AI29387 for the West Africa International Center of Excellence for Malaria Research. Funding for surveillance, sampling, and testing in Madagascar was provided by the WHO, the CDC (grant no. U5/PO00812-05), the United States Agency for International Development (USAID) (Cooperation Agreement 72068719CA00001), and the Office of the Assistant Secretary for Preparedness and Response in the HHS (grant no. IDSEP190051-01-0200). Funding for sequencing was provided by the Bill & Melinda Gates Foundation (GCE/ID OPP1211841), Chan Zuckerberg Biohub, and the Innovative Genomics Institute at UC Berkeley. Mozambique acknowledges support from the Mozambican Ministry of Health and the President's Emergency Plan for AIDS Relief (PEPFAR) through the CDC under the terms of grant nos. GH002021 and GH001944 and from the Bill & Melinda Gates Foundation (#OPP1214435). Namibian efforts were supported by Africa CDC through a subaward from the Bill & Melinda Gates Foundation. Efforts from Niger were supported by the French Ministry for Europe and Foreign Affairs through the REPAIR COVID-19–Africa project coordinated by the Pasteur International Network association. In addition to the funding support for ACEGI already listed, Nigeria's contributions were made possible by support from Flu Lab and a cohort of donors through the Audacious Project, a collaborative funding initiative housed at TED, including the ELMA Foundation, MacKenzie Scott, the Skoll Foundation, and Open Philanthropy. COVID-19 genomic surveillance at the Centre for Human Virology and Genomics, Nigerian Institute of Medical Research, is supported by the government of Nigeria special funding for COVID-19 to the Nigerian Institute of Medical Research, Lagos, Nigeria. It is also supported by funding from the AIDS Healthcare Foundation (AHF) Global Public Health Institute (GESIT Study). Efforts from the Republic of the Congo were supported by the European and Developing Countries Clinical Trials Partnership (EDCTP) IDs PANDORA and CANTAM and the German Academic Exchange Service (DAAD) ID PACE-UP and DAAD project ID 5759234. Rwanda's contributions were made possible by funding from the African Network for Improved Diagnostics, Epidemiology and Management of Common Infectious Agents (ANDEMIA), which was granted by the German Federal Ministry of Education and Research (BMBF grants 01KA1606, 01KA2021, and 01KA2110B) and the NIHR Global Health Research program (16/136/33) using UK aid from the UK Government. In addition to the South African institutions listed above, the University of Cape Town's work was supported by the Wellcome Trust (grant no. 203135/Z/16/Z), EDCTP RADIATES (RIA2020EF-3030), the South African Department of Science and Innovation (SA DSI), and SAMRC; Stellenbosch University's contributions were supported by SAMRC; and the University of Pretoria's contributions were funded by the G7 Global Health Fund and a BMBF ANDEMIA grant. Funding from the Fleming Fund supported sequencing in Sudan. The Ministry of Higher Education and Scientific Research of Tunisia provided funding for sequencing from Tunisia. The Uganda Virus Research Institute (UVRI) (Uganda) acknowledge support from the Wellcome Trust and FCDO – Wellcome Epidemic Preparedness – Coronavirus (AFRIC019, grant agreement number 220977/Z/20/Z), the MRC (MC_UU_1201412), and the UK Medical Research Council (MRC/UKRI) and FCDO (DIASEQCO, grant agreement number MC_PC_20010). Research at the FredHutch Institute, which supported bioinformatics analyses of sequences in the present study, was supported by the Bill & Melinda Gates foundation (#INV-018979). Research support from Broad Institute colleagues was 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, Open Philanthropy, the Howard Hughes Medical Institute, and NIH (U01AI151812 and U54HG007480) (P.C.S.). Work from Quadram Institute Bioscience was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent projects BBS/E/F/000PR10348, BBS/E/F/000PR10349, BBS/E/F/000PR10351, and BBS/E/F/000PR10352 and by the Quadram Institute Bioscience BBSRC-funded Core Capability Grant (project number BB/CCG1860/1). Sequences generated in Zambia through PATH were funded by the Bill & Melinda Gates Foundation and Africa CDC. The content and findings reported herein are the sole deduction, view, and responsibility of the researcher(s) and do not reflect the official position and sentiments of the funding agencies. **Author contributions:** Conceptualization: H.Te., E.S., C.Ba., S.K.T., T.d.O., R.L., E.W.; Methodology: H.Te., J.E.S., M.C., B.Te., G.M., D.P.M., A.W.L., D.A.R., L.M.K., G.G., T.d.O., R.L., E.W.; Genomic data generation: H.Te., J.E.S., M.C., M.Moi., B.Te., G.M., D.P.M., A.W.L., A.D., D.G.A., M.M.D., A.Si., A.N.Z., A.S.G., D.A.S., A.K.Sa., A.O., A.Sow, A.O.M., A.K.Se., A.G.A., A.L., A.S.K., A.E.A., A.A.J., A.Fo., A.O.O., A.A.A., A.J., A.Kan., A.Mo., A.R., A.A., A.Kaz., A.Ba., A.Chr., A.J.T., A.Ca., A.K.K., A.Ao., A.Sou., A.A., A.Na., A.V.G., A.Nk., A.J.P., A.Y., A.V., A.N.H., A.Cho., A.Ir., A.Ma., A.L.B., A.I.s., A.A.Sy., A.G., A.Fe., A.E.S., B.Ma., B.L.S., B.S.O., B.B., B.D., B.L.H., B.T.s., B.L., B.Mv., B.N., B.T.M., B.A.K., B.K., B.A., B.P., B.Mc., C.Br., C.W., C.N., C.A., C.B.P., C.S., C.G.A., C.N.A., C.M.M., C.L., C.K.O., C.I., C.N.M., C.P., C.G., C.E.O., C.D.R., C.M.M., C.E., D.B., D.J.B., D.M., D.P., D.B., D.J., D.S., D.T., D.S.A., D.G., D.S.G., D.O.O., D.M., D.W.W., E.F., E.K.L., E.Si., E.M.O., E.N.N., E.O.A., E.O., E.Sh., E.Ba., E.B.A., E.A.Ah., E.L., E.Mu., E.F., E.Be., E.S.-L., E.A.An., F.L., F.M.T., F.W., F.A., F.T.T., F.D., F.V.A., F.T., F.O., F.N., F.M.M., F.E.R., F.A.D., F.I., G.K.M., G.T., G.L.K., G.O.A., G.U.v.z., G.A.A., G.S., G.P.M., H.C.R., H.E.O., H.O., H.A., H.K., H.N., H.Tr., H.A.A.K., H.E., H.G., H.M., H.K., I.Sm., I.B.O., I.M.A., I.O., I.B.B., I.A.M., I.S.s., I.W., I.S.K., J.W.A.H., J.A., J.S., J.C.M., J.M.T., J.H., J.G.S., J.Gi., J.Mu., J.N., J.N.U., J.N.B., J.Y., J.Mo., J.K., J.D.S., J.H., J.K.O., J.M.M., J.O.G., J.T.K., J.C.O., J.S.X., J.Gy., J.F.W., J.H.B., J.N., J.E., J.N., J.M.N., J.N., J.U.O., J.C.A., J.J.L., J.J.H.M., J.O., K.J.S., K.V., K.T.A., K.A.T., K.S.C., K.S.M., K.D., K.G.M., K.O.D., L.F., L.S., L.M.K., L.B., L.d.O.M., L.C., L.O., L.D.O., L.L.D., L.I.O., L.T., M.Mi., M.R.M., M.Mas., M.E., M.Mai., M.I.M., M.Ke., M.D., M.Mom., M.d.L.L.M., M.V., M.F.P., M.F., M.M.N., M.Mar., M.D., M.W.M., M.G.M., M.O., M.R.W., M.Y.T., M.O.A., M.Ab., M.A.B., M.G.S., M.K.A., M.M.M., M.Ka., M.S., M.B.M., M.Mw., M.A.I., M.V.P., N.Abi., N.R., N.Abr., N.I.s., N.E., N.M.T., N.D., N.Ma., N.H., N.B.S., N.M.F., N.S.A., N.B., N.Mu., N.G., N.W., N.Si., N.N., N.A.A., N.T., N.Mbh., N.H.R., N.Ig., N.Mba., O.C.K., O.S., O.Fe., O.M.A., O.Te., O.A.O., O.Fak., O.E.O., O.-E.O., O.Fay., P.S., P.O., P.C., P.N., P.S., P.E.O., P.Ar., P.K.Q., P.O.O., P.B., P.D., P.A.B., P.K.M., P.K., P.Ab., R.E., R.J., R.A.C., R.G.E., R.A., R.N., R.O.P., R.G., R.A.K., R.M.N.D., R.A.A., R.A.C., S.Gar., S.Ma., S.Bo., S.S., S.I.M., S.F., S.Mh., S.H., S.K.K., S.Me., S.T., S.H.A., S.W.M., S.D., S.M.-M., S.A., S.S.A., S.M.A., S.E., S.Mo., S.L., S.G.A., S.J., S.F.A., S.O.G., S.Gr., S.L., S.Pr., S.Ou., S.v.W., S.F.S., S.K., S.A., S.R., S.Pi., S.N., S.Be., S.L.B., S.v.d.W., T.Ma., T.Mo., T.L., T.P.V., T.S., T.G.M., T.K.B., U.J.A., U.C., U.R., U.E.G., V.E., V.N., V.G., W.H.R., W.A.K., W.A.K., W.P., W.T.C., Y.A.A., Y.R., Y.Be., Y.N., Y.Bu., Z.R.d.L., A.E.O., A.v.G., G.G., M.Moe., O.To., P.C.S., A.A.Sa., S.O.O., Y.K.T., S.K.T., T.d.O., C.H., R.L., J.N., E.W.; Data analysis: H.Te., J.E.S., M.C., M.Moi., B.Te., G.M., D.P.M., A.W.L., A.I.E., N.Sa., E.D., G.S.K., S.v.G., G.G., T.d.O., R.L., E.W.; Funding acquisition: A.E.O., A.v.G., G.G., M.Moe., O.To., A.A.Sa., S.O.O., Y.K.T., S.K.T., T.d.O., C.H.; Project administration: G.M., A.D., D.G.A., M.M.D., A.C., D.W.W., H.O., S.W.M., A.E.O., A.v.G., G.G., M.Moe., O.To., P.C.S., A.A.Sa., S.O.O., Y.K.T., S.K.T., T.d.O., C.H., R.L., J.N., E.W.; Supervision: A.E.O., A.v.G., G.G., M.Moe., O.To., P.C.S., A.A.Sa., S.O.O., Y.K.T., S.K.T., T.d.O., C.H., R.L., J.N., E.W.; Writing–original draft: H.Te., J.E.S., M.C., G.M., D.P.M., C.Ba., S.K.T., T.d.O., R.L., E.W.; Writing–review and editing: H.Te., J.E.S., M.C., M.Moi., B.Te., G.M., D.P.M., C.Ba., A.W.L., A.D., D.G.A., M.M.D., A.Si., A.N.Z., A.S.G., A.K.Sa., A.O., A.Sow, A.O.M., A.K.Se., A.I.E., A.L., A.S.K., A.E.A., A.A.J., A.Fo., A.O.O., A.A.A., A.J., A.Kan., A.Mo., A.R., A.A., A.Kaz., A.Ba., A.Chr., A.J.T., A.Ca., A.K.K., A.Ko., A.Bo., A.Sou., A.A., A.V.G., A.J.P., A.Y., A.V., A.N.H., A.Cho., A.Ir., A.Ma., A.L.B., A.I.s., A.A.Sy., A.G., A.Fe., A.E.S., B.Ma., B.L.S., B.S.O., B.B., B.D., B.L.H., B.T.s., B.L., B.Mv., B.N., B.T.M., B.A.K., B.K., B.A., B.P., B.Mc., C.Br., C.W., C.N., C.A., C.B.P., C.S., C.G.A., C.N.A., C.M.M., C.L., C.K.O., C.I., C.N.M., C.P., C.G., C.E.O., C.D.R., C.M.M., C.E., D.B., D.J.B., D.M., D.P., D.B., D.J., D.S., D.T., D.S.A., D.G., D.S.G., D.O.O., D.M., D.W.W., E.F., E.K.L., E.Si., E.M.O., E.N.N., E.O.A., E.O., E.Sh., E.Ba., E.A.An., E.Ma., F.L., F.M.T., F.W., F.A., F.T.T., F.D., F.V.A., F.T., F.O., F.N., F.M.M., F.E.R., F.A.D., F.I., G.K.M., G.T., G.L.K., G.O.A.,

G.U.v.Z., G.A.A., G.S.K., G.S., G.P.M., H.C.R., H.E.O., H.O., H.A., H.K., H.N., H.Tr., H.A.A.K., H.E., H.G., H.M., H.K., I.Sm., I.B.O., I.M.A., I.O., I.B.B., I.Ss., I.W., I.S.K., J.W.A.H., J.A., J.S., J.C.M., J.M.T., J.H., J.G.S., J.Gi., J.Mu., J.N.U., J.N.B., J.Y., J.Mo., J.K., J.D.S., J.H., J.K.O., J.M.M., J.O.G., J.T.K., J.C.O., J.S.X., J.Gy., J.H.B., J.N., J.E., J.N., J.M.N., J.N., J.U.O., J.C.A., J.J.L., J.O., K.J.S., K.V., K.T.A., K.A.T., K.S.C., K.S.M., K.D., K.G.M., K.O.D., L.F., L.S., L.B., L.d.O.M., L.C., L.O., L.L.D., L.I.O., M.M.L., M.R., M.Mas., M.E., M.Mai., M.I.M., M.K.E., M.D., M.Mom., M.d.L.L.M., M.V., M.F.P., M.F., M.M.N., M.Mar., M.D., M.W.M., M.G.M., M.O., M.R.W., M.Y.T., M.O.A., M.Aab., M.A.B., M.G.S., M.K.K., M.M.M., M.Ka., M.S., M.B.M., M.Mw., M.V.P., N.Abi., N.R., N.S., N.M.T., M.S., S.L.B., S.v.d.W., T.Ma., T.Mo., N.M.F., N.Sa., N.B., N.Mu., N.G., N.W., N.Si., N.N., N.A.A., N.T., N.Mb., N.H.R., N.M., N.Mba, O.C.K., O.S., O.Fc., O.M.A., O.Te., O.A.O., O.Fak., O.E.O., O.Fay., P.S., P.O., P.C., P.N., P.S., P.E.O., P.Ar., P.K.Q., P.O.O., P.B., P.D., P.A.B., P.K.M., P.K., P.Ab., R.E., R.J., R.K.A., R.G.E., R.A., R.N., R.O.P., R.G., R.A.K., R.A.A., R.A.C., S.Gar., S.Ma., S.S., S.I.M., S.F., S.Mh., S.H., S.K.A., S.Me., S.T., S.H.A., S.W.M., S.D., S.M.-M., S.A., S.S.A., S.M.A., S.E., S.Mo., S.L., S.Gas., S.J., S.F.A., S.Og., S.Gr., S.L., S.P., S.Ou., S.v.W., S.F.S., S.K., S.A., S.R., S.Pi., S.N., S.Be., S.L.B., S.v.d.W., T.Ma., T.Mo., T.L., T.P.V., T.S., T.G.M., T.B., U.J.A., U.C., U.R., U.E.G., V.E., V.N., V.G., W.H.R., W.A.K., W.K.A., W.P., W.T.C., Y.A.A., Y.R., Y.Be., Y.N., Y.Bu., Z.R.L.H., A.E.O., A.v.G., G.G., M.Moe., O.Yo., P.C.S., A.A.Sa., S.O.O., Y.K.T., S.K.T., T.d.O., C.H., R.L., J.N., E.W. **Competing interests:** With the exception of P.S., who is a co-founder of and consultant to Sherlock Biosciences and a Board Member of Danaher Corporation and who holds equity in the companies, the authors have no conflicts of interest to declare. **Data and materials availability:** All of the SARS-CoV-2 whole-genome sequences that were analyzed in the present study are all publicly available on the GISAID sequence database. We gratefully acknowledge the authors from the originating laboratories and the submitting laboratories, who generated and shared via GISAID genetic sequence data on which this research is based. A full list of the African sequences as well as global references are presented and acknowledged in table S4 and in our github repository (<https://github.com/CERI-KRISP/SARS-CoV-2-epidemic-in-Africa>) (76). The repositories also contain all of the metadata, raw and time-scaled ML tree topologies, and annotated tree topologies, as well as the data analysis and visualization scripts used here, which will allow for the independent reproduction of results. Furthermore, the repositories also contain all institutional review board references and material transfer agreements. Please refer to the ethics statement in the methods section for more details. **License information:** This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

Authors and their affiliations

Huriyiah Tegally^{1,2†}, James E. San^{1,2}, Matthew Cotten^{3,4}, Monika Moir⁵, Bryan Tegomoh⁶, Gerald Mboowa⁷, Darren P. Martin^{8,9}, Cheryl Baxter^{1,10}, Arnold W. Lambisia¹¹, Amadou Diallo¹², Daniel G. Amoako¹³, Moussa M. Diagne¹², Abay Sisay^{15,16}, Abdel-Rahman N. Zekri¹⁷, Abdou Salam Gueye¹⁸, Abdoul K. Sangare¹⁹, Abdoul-Salam Ouedraogo^{20,204,205}, Abdourahmane Sow²¹, Abdoulaye M. Musa^{22,23,24}, Abdul K. Sesay²⁵, Abe G. Abias²⁶, Adem I. Elzagheid²⁷, Adamou Lagare²⁸, Adedotun-Sulaiman Kemi²⁹, Aden Elmi Abar^{72,73}, Adeniji A. Johnson^{31,32}, Adeola Fowotade^{33,34}, Adeyemi O. Oluwapelumi^{35,36}, Adrienne A. Amuri^{37,38}, Agnes Juru³⁹, Ahmed Kandeil⁴⁰, Ahmed Mostafa⁴⁰, Ahmed Rebai⁴¹, Ahmed Sayed⁴², Kazeem Akano^{43,44}, Aladjie Balde^{45,46}, Alan Christoffels^{47,47}, Alexander J. Trotter⁴⁸, Allan Campbell⁴⁹, Alpha K. Keita^{50,51}, Amadou Kone⁵², Amal Bouzid^{41,53}, Amal Souissi⁴¹, Ambrose Agweyu⁵⁴, Anel Naguib⁵⁴, Ana V. Gutierrez⁴⁸, Anatole Nkeshimana⁵⁵, Andrew J. Page⁴⁸, Anges Yadouleton⁵⁶, Anika Vinze⁵⁷, Anise N. Hapji⁴³, Anissa Choukha^{58,59}, Arash Iranzadeh^{8,9}, Arisha Maharaj⁶⁰, Armel L. Batchi-Bouyou^{60,61}, Arshad Ismail¹³, Augustina A. Sylverken^{62,63}, Augustine Goba^{64,65}, Ayodele Femi^{64,65}, Ayotunde E. Sijuwola⁴³, Baba Marycelin^{66,67}, Babatunde L. Salako^{29,32}, Bamidele S. Oderinde⁶⁸, Bankole Bolajoko⁴³, Bassirou Diarra⁶², Belinda L. Herring⁴⁸, Benjamin Tsota⁶¹, Bernard Lekana-Douki^{68,69}, Bernard Mvula⁷⁰, Berthe-Marie Njanpop-Lafourcade⁶⁸, Blessing T. Marondera⁷¹,

Bouh Abdi Khaireh^{72,73}, Bourema Kouriba¹⁹, Bright Adu⁷⁴, Brigitte Pool⁷⁵, Bronwyn McInnis¹⁷, Cara Brook^{76,77}, Carolyn Williamson^{9,10,78}, Cassien Nduwimana⁵⁵, Catherine Ansombe^{79,80}, Catherine B. Pratt⁸¹, Cathrine Scheepers^{13,82}, Chantal G. Akoua-Koffi^{83,84}, Charles N. Agoti^{11,85}, Chastel M. Mapangyu^{60,86}, Cheikh Loucoubar¹², Chika K. Onwuamah⁸⁷, Chikwe Ihekweazu⁸⁸, Christian N. Malaka⁸⁹, Christophe Peyrefitte¹², Grace C. Chukwa^{43,44}, Chukwuma E. Omoruyi^{33,34}, Clotilde D. Rafai⁹⁰, Collins M. Morang⁹¹, Cyril Eramah⁹², Daniel B. Lule³, Daniel J. Bridges⁹³, Daniel Mukadi-Bamuleka³⁷, Danny Park⁵⁷, David A. Rasmussen^{94,95}, David Baker⁴⁸, David J. Nokes^{11,96}, Deogratius Ssemwanga^{3,97}, Derek Thiabialu⁶², Dominic S. Y. Amuzu⁹¹, Dominique Goedhals⁹⁸, Donald S. Grant^{64,65,99}, Donwilliams O. Ornuoyo¹¹, Dorcas Maruapula¹⁰⁰, Dorcas W. Wanjohi⁷, Ebenezer Foster-Nyarko⁴⁸, Eddy K. Lusamaki^{37,38,51}, Edgar Simulundu¹⁰¹, Edidah M. Ong'era¹¹, Edith N. Ngabana^{37,38}, Edward O. Abworo¹⁰², Edward Otieno¹¹, Edwin Shumba⁷¹, Edwine Barasa¹¹, El Bara Ahmed^{103,104}, Elhadi A. Ahmed²³, Emmanuel Lokilo³⁷, Enatha Mukantwari¹⁰⁵, Eromon Philomena⁴³, Essia Belarbi¹⁰⁶, Etienne Simon-Lorier¹⁰⁷, Etile A. Anoh⁸³, Eusebio Manueh¹⁰⁸, Fabian Leendertz¹⁰⁶, Fahn M. Taweh¹⁰⁹, Fares Wasfi⁵⁸, Fatma Abdelmoula^{41,110}, Faustinos T. Takawira³⁹, Fawzi Derrari¹¹¹, Fehintola V. Ajogbasle⁴³, Florette Ntouri^{112,113}, Folarin Onikepe^{43,44}, Francine Tourni^{60,114}, Francisca M. Muyembe^{37,38}, Frank E. Z. Ragonzingba¹¹⁵, Fred A. Drabiti^{116,117}, Fred-Akintunwa Iyanu⁴³, Gabriel K. Mbunzu³⁸, Gaetan Thilliez⁴⁸, Gemma L. Kay⁴⁸, George O. Akpede⁹², Gert U. von Zyl^{118,119}, Gordon A. Awandare⁹¹, Grace S. Kpelu^{120,121}, Grit Schubert¹⁰⁶, Gugu P. Maphalala¹²², Hafaliana C. Ranaivosoa⁷³, Hannah E. Omuonakwe¹²³, Harris Onyewara⁷, Haruka Abe¹²⁴, Hela Karray¹²⁵, Hellen Nansumba¹²⁶, Henda Triki⁵⁸, Herve Alberic Adje Kadjo¹²⁷, Hesham Elghazaly¹²⁸, Hlanzi Gumbo³⁹, Hota Mathieu¹²⁹, Hugo Kavunga-Membo³⁷, Ibtihel Smet⁴¹, Idowu B. Olawoye⁴³, Ifedayo M. O. Adetifa^{58,130}, Ikponmwosa Odiya⁹², Ilhem Boutiba-Ben Boubaker^{131,132}, Muhammad I. Ahmed⁴³, Isaac Ssewanyana¹²⁶, Isatta Wurie¹³³, Iyaloo S. Konstantinus¹³⁴, Jacqueline Wembo Afiwa Halatoko¹³⁵, James Ayei²⁶, Janaki Sonoo³⁶, Jean-Claude C. Makangara^{37,38}, Jean-Jacques M. Tamfun^{37,38}, Jean-Michel Heraud¹²⁷, Jeffrey G. Shaffer¹³⁷, Jennifer Giandhari⁷, Jennifer Musyoki¹¹, Jerome Nkurunziza¹³⁸, Jessica N. Uwanibe⁴³, Jinal N. Bhiman^{131,133}, Jiro Yasuda¹²⁴, Joana Morais^{139,140}, Jocelyn Kiconco⁹⁷, John D. Sand^{64,65}, John Huddleston¹⁴¹, John K. Odomo⁷⁴, John M. Morobe¹¹, John O. Gyapong¹²⁰, Judith T. Kanyiwa³, Johnson C. Okolie⁴³, Joicymara S. Xavier^{142,143}, Jones Gyamfi¹²⁰, Joseph F. Wamala¹⁴⁴, Joseph H. K. Bonney⁷⁴, Joseph Nyandoro¹⁴⁵, Josie Everatt¹³, Joweria Nakasegu⁹⁷, Joyce N. Ngoi⁹¹, Joyce Namulondo⁹⁷, Judith U. Oguzie^{63,44}, Julia C. Andeko⁶⁸, Julius J. Lutwama³, Juma J. H. Mogga¹⁴⁴, Justin O'Grady⁴⁸, Katherine J. Siddle⁵⁷, Kathleen Victor¹⁴⁶, Kayode T. Adeyemi^{43,44}, Kefentse A. Tumedi¹⁴⁷, Kevin S. Carvalho⁴⁸, Khadija Said Mohammed¹¹, Koussay Dellagi¹⁴⁶, Kunda G. Musonda¹⁴⁹, Kwabena O. Duedu^{120,121}, Lamia Fki-Berrajah¹²⁵, Lavanya Singh², Lenora M. Keple^{94,95}, Leon Biscornet⁷⁵, Leonardo de Oliveira Martins⁴⁸, Lucious Chabuka¹⁵⁰, Luicer Olubayo⁸, Lung Deng Ojok²⁶, Lui Lojok Deng²⁶, Lynette I. Ochola-Oyier¹¹, Lynn Tyers⁷, Madisa Mine¹⁵¹, Magalutcheeme Ramuth¹⁵⁶, Maha Mastouri^{152,153}, Mahmoud ElHefnawi¹⁵⁴, Maimouna Mbanne¹², Maithshwarelo I. Matsheka¹⁴⁷, Malebogo Keabonye¹⁵⁵, Mamadou Diop¹², Mambu Momoh^{64,65,156}, Maria da Luz Lima Mendonca¹⁴⁸, Marietjie Venter¹⁵⁷, Marietou F. Paye⁵⁷, Martin Faye²⁷, Martin M. Nyaga¹⁵⁸, Mathabo Mareka¹⁵⁹, Matoke-Muhia Damaris¹⁶⁰, Maureen W. Mburu¹¹, Maximilian G. Mpina^{161,162,163}, Michael Owusu¹⁶⁴, Michael R. Wiley^{81,165}, Mirabeau Y. Tatfeng¹⁶⁶, Mitoha Ondo'o Ayekaba¹⁶², Mohamed Abouelhoda^{167,168}, Mohamed Amine Belouaf¹¹¹, Mohamed G. Seadway^{169,170}, Mohamed K. Khalifa⁷¹, Mooko Marehabile Matobo¹⁵⁹, Mouhamed Kane¹², Mounero Salou¹⁷², Mphaphi B. Mbulawa¹⁵⁵, Mulenga Mwenda⁹³, Mushal Alam¹⁷³, My V. T. Phan³, Nabil Abid^{52,174}, Nadine Rujen^{175,176}, Nadir Abuzaied¹⁷⁷, Nalia Ismael¹⁷⁸, Nancy Elguindy⁵⁴, Ndeye Marieme Top¹², Ndongo Dial¹², Nedio Mabunda¹⁷⁸, Nei-yuan Hsiao^{9,78}, Nelson Boricó Silicho¹⁶², Ngiambululu M. Francisco¹³⁹, Ngonda Saasa¹⁷⁹, Nicholas Bbosja³, Nickson Murunga¹¹, Nikita Gumedde⁴⁸, Nicole Wolter^{13,113}, Nikita Sitharam⁷, Nnaemeka Nododo⁸⁸, Nnennaya A. Ajayi¹⁸⁰, Noel Tordo¹⁸¹,

Nokuzola Mbhele⁹, Norosoa H. Razanajatovo⁷⁷, Nosamiefan Iguosadolo⁴³, Nwando Mba⁸⁸, Ojide C. Kingsley¹⁸², Okogbenin Sylvanus⁹², Olajidi Femi¹⁸³, Olubusuyi M. Adewumi^{31,32}, Olumade Testimony^{43,44}, Oluosola A. Ogunsanya⁴³, Oluwatosin Fakayode¹⁸⁴, Onwe E. Ogah¹⁸⁵, Ope-Ewe Oludayo⁴³, Ousmane Faye¹², Pamela Smith-Lawrence¹⁵⁵, Pascale Ondoa⁷¹, Patrice Combe¹⁸⁶, Patricia Nabisub^{187,188}, Patrick Semanda²⁶, Paul E. Oluniji⁴³, Paulo Arnaldo¹⁷⁸, Peter Koo Quashie⁹¹, Peter O. Okohere^{92,189}, Philip Bejon¹¹, Philippe Dussart⁷⁷, Phillip A. Bester¹⁹⁰, Placide K. Mbala^{37,38}, Pontiano Kaleebu^{3,97}, Priscilla Abechi^{43,44}, Rabeh El-Shesheny^{40,191}, Rageema Joseph⁹, Ramy Karam Aziz^{192,193}, René G. Essomba^{194,195}, Reuben Ayivor-Djanie^{91,201,21}, Richard Njoom¹⁹⁶, Richard O. Phillips⁶³, Richmond Gorman⁶³, Robert A. Kingsley⁴⁸, Rosa Maria D. E. S. A. Neto Rodrigues^{197,198}, Rosemary A. Adu²⁹, Rosina A. A. Carr^{120,121}, Saba Gargouri¹²⁵, Saber Masmodi⁴¹, Sacha Bootsma¹⁴⁴, Safietou Sankhe¹², Sahra Isse Mohamed¹⁹⁹, Saibu Femi⁴³, Salma Mhalla^{152,200}, Salome Hosch^{161,201}, Samar Kamal Kassim¹²⁸, Samar Metha⁵⁷, Sameh Trabelsi²⁰², Sara Hassan Agwa¹²⁸, Sarah Wambui Mwangi⁷, Seydou Dombia⁵², Sheila Makiala-Mandanda^{37,38}, Sheriane Aryeetej⁶³, Shymaa S. Ahmed⁵⁴, Side Mohamed Ahmed¹⁰³, Siham Elhamoumi⁵⁷, Sikhulle Moyo^{100,203}, Silvia Lutucuta¹³⁹, Simani Gaisetsi^{100,203}, Simbirie Jalloh^{64,65}, Soa Fry Andriamandimby⁷⁷, Sobajo Ongotupe⁴³, Solène Grayo¹⁸¹, Sonia Lekana-Douki⁶⁸, Sophie Prosele⁴⁸, Soumeiya Ouangraoua^{204,205}, Stephanie van Wyk¹, Stephen F. Schaffner⁵⁷, Stephen Kanzyere^{187,188}, Steve Ahuka-Mundeki^{37,38}, Steven Rudder⁴⁸, Sureshnee Pillay², Susan Nabadda¹⁶², Sylvie Behillil²⁰⁶, Sylvie L. Budiaki¹⁵⁹, Sylvie van der Werf²⁰⁶, Tappumani Mashe^{39,207}, Thabo Mohale¹³, Thanh Le-Viet⁴⁸, Thirumalaisamy P. Velavan^{114,208}, Tobias Schindler^{161,162,201}, Tongai G. Mponga¹¹⁸, Trevor Bedford^{141,209}, Ugochukwu J. Anyanwu¹², Ugwu Chinedu^{43,44}, Upasana Ramphal^{210,210}, Uwem E. George⁴³, Vincent Enou²⁰⁶, Vishvanath Nene¹⁰², Vivianne Gorova^{211,212}, Wael H. Roshyd⁵⁴, Wasim Abdul Karim¹, William K. Ampofo²¹³, Wolfgang Preiser^{118,119}, Wondulul T. Choga^{100,214}, Yahaya Ali Ahmed¹⁸, Yajna Ramphal¹, Yaw Bediako^{91,215}, Yesheene Naidoo², Yvan Butera^{175,216,217}, Zaydah R. de Laurent¹¹, Africa Pathogen Genomics Initiative (Africa PGI)[†], Ahmed E. O. Ouma²⁷, Anne von Gottberg^{13,113}, George Githinji^{11,218}, Matshidiso Moeti¹⁸, Oyewale Tomori⁴³, Pardis C. Sabeti⁵⁷, Amadou A. Sall¹², Samuel O. Oyola¹⁰², Yeneke K. Tebeje²⁷, Sofonias K. Tessema⁷, Tulio de Oliveira^{1,210,219*}, Christian Happ^{43,44}, Richard Lessells², John Nkengasong⁷, Eduan Wilkinson^{2,4†}

¹Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa.

²KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa. ³MRC/UUVRI and LSHTM Uganda Research Unit, Entebbe, Uganda. ⁴MRC-University of Glasgow Centre for Virus Research, Glasgow, UK. ⁵The Biotechnology Centre of the University of Yaoundé I, Yaoundé, Cameroon. ⁶CDC Foundation, Atlanta, Georgia, Nebraska Department of Health and Human Services, Lincoln, NE, USA. ⁷Institute of Pathogen Genomics, Africa Centres for Disease Control and Prevention (Africa CDC), Addis Ababa, Ethiopia. ⁸Institute of Infectious Diseases and Molecular Medicine, Department of Integrative Biomedical Sciences, Computational Biology Division, University of Cape Town, Cape Town, South Africa. ⁹Division of Medical Virology, Wellcome Centre for Infectious Diseases in Africa, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa. ¹⁰Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa. ¹¹KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya. ¹²Institut Pasteur de Dakar, Dakar, Senegal. ¹³National Institute for Communicable Diseases (NICD) of the National Health Laboratory Service (NHL), Johannesburg, South Africa. ¹⁴School of Health Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, KwaZulu-Natal, South Africa. ¹⁵Department of Medical Laboratory Sciences, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia. ¹⁶Department of Microbial, Cellular and Molecular Biology, College of Natural and Computational Sciences, Addis Ababa University, Addis

- Ababa, Ethiopia. ¹⁷Cancer Biology Department, Virology and Immunology Unit, National Cancer Institute, Cairo University, Cairo, Egypt. ¹⁸World Health Organization, Africa Region, Brazzaville, Republic of the Congo. ¹⁹Centre d'Infectiologie Charles Mérieux-Mali (CICM-Mali), Bamako, Mali. ²⁰Bacteriology and Virology Department Souro Sanou University Hospital, Bobo-Dioulasso, Burkina Faso. ²¹West African Health Organisation, Bobo-Dioulasso, Burkina Faso. ²²Faculty of Medicine and Health Sciences, Kassala University, Kassala City, Sudan. ²³Department of Microbiology, Faculty of Medical Laboratory Sciences, University of Gezira, Gezira, Sudan. ²⁴General Administration of Laboratories and Blood Banks, Ministry of Health, Kassala State, Sudan. ²⁵MRC Unit The Gambia at LSHTM, Fajara, Gambia. ²⁶National Public Health Laboratory, Ministry of Health, Juba, Republic of South Sudan. ²⁷Libyan Biotechnology Research Center, Tripoli, Libya. ²⁸Center for Medical and Sanitary Research (CERMES), Niamey, Niger. ²⁹The Nigerian Institute of Medical Research, Yaba, Lagos, Nigeria. ³⁰Laboratoire de la Caisse Nationale de Sécurité Sociale, Djibouti, Republic of Djibouti. ³¹Department of Virology, College of Medicine, University of Ibadan, Ibadan, Nigeria. ³²Infectious Disease Institute, College of Medicine, University of Ibadan, Ibadan, Nigeria. ³³Medical Microbiology and Parasitology Department, College of Medicine, University of Ibadan, Ibadan, Nigeria. ³⁴Biorepository Clinical Virology Laboratory, College of Medicine, University of Ibadan, Ibadan, Nigeria. ³⁵Department of Medical Microbiology and Parasitology, Faculty of Basic Clinical Sciences, College of Health Sciences, University of Ilorin, Ilorin, Kwara State, Nigeria. ³⁶The Pirbright Institute, Woking, UK. ³⁷Pathogen Sequencing Lab, Institut National de Recherche Biomédicale (INRB), Kinshasa, the Democratic Republic of the Congo. ³⁸Université de Kinshasa (UNIKIN), Kinshasa, the Democratic Republic of the Congo. ³⁹National Microbiology Reference Laboratory, Harare, Zimbabwe. ⁴⁰Center of Scientific Excellence for Influenza Viruses, National Research Centre (NRC), Cairo, Egypt. ⁴¹Laboratory of Molecular and Cellular Screening Processes, Centre of Biotechnology of Sfax, University of Sfax, Sfax, Tunisia. ⁴²Genomics and Epigenomics Program, Research Department CCEH57357, Cairo, Egypt. ⁴³African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Ede, Osun State, Nigeria. ⁴⁴Department of Biological Sciences, Faculty of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria. ⁴⁵Laboratório de Biologia Molecular Jean Piaget, Bissau, Guinea-Bissau. ⁴⁶University Jean Piaget in Guinea-Bissau, Bissau, Guinea-Bissau. ⁴⁷SAMRC Bioinformatics Unit, SA Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa. ⁴⁸Quadram Institute Bioscience, Norwich, UK. ⁴⁹Central Public Health Reference Laboratories, Freetown, Sierra Leone. ⁵⁰Centre de Recherche et de Formation en Infectiologie de Guinée (CERFIG), Université de Conakry, Conakry, Guinée. ⁵¹TransVIHMI, Institut de Recherche pour le Développement, Institut National de la Santé et de la Recherche Médicale (INSERM), Montpellier University, 34090, Montpellier, France. ⁵²University Clinical Research Center (UCRC), University of Sciences, Techniques and Technology of Bamako, Bamako, Mali. ⁵³Sharjah Institute for Medical Research, College of Medicine, University of Sharjah, Sharjah, United Arab Emirates. ⁵⁴Central Public Health Laboratories (CPHL), Cairo, Egypt. ⁵⁵National Institute of Public Health, Bujumbura, Burundi. ⁵⁶Laboratoire des Fièvres Hémorragiques Virales du Bénin, Cotonou, Benin. ⁵⁷Infectious Disease and Microbiome Program, Broad Institute of Harvard and MIT, Cambridge, MA, USA. ⁵⁸Laboratory of Clinical Virology, WHO Reference Laboratory for Poliomyelitis and Measles in the Eastern Mediterranean Region, Pasteur Institute of Tunis, University Tunis El Manar (UTM), Tunis 1002, Tunisia. ⁵⁹Research Laboratory "Virus, Vectors and Hosts: One Health Approach and Technological Innovation for a Better Health", LR20IPT02, Pasteur Institute, Tunis 1002, Tunisia. ⁶⁰Fondation Congolaise pour la Recherche Médicale, Brazzaville, Republic of the Congo. ⁶¹Marien Ngouabi, Brazzaville, Republic of the Congo. ⁶²Kwame Nkrumah University of Science and Technology, Department of Theoretical and Applied Biology, Kumasi, Ghana. ⁶³Kumasi Centre for Collaborative Research in Tropical Medicine, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. ⁶⁴Viral Haemorrhagic Fever Laboratory, Kenema Government Hospital, Kenema, Sierra Leone. ⁶⁵Ministry of Health and Sanitation, Freetown, Sierra Leone. ⁶⁶Department of Immunology, University of Maiduguri Teaching Hospital, P.M.B. 1414, Maiduguri, Nigeria. ⁶⁷Department of Medical Laboratory Science, College of Medical Sciences, University of Maiduguri, P.M.B. 1069, Maiduguri, Borno State, Nigeria. ⁶⁸Centre Interdisciplinaires de Recherches Médicales de Franceville (CIRMF), Franceville, Gabon. ⁶⁹Département de Parasitologie-Mycologie Université des Sciences de la Santé (USS), Libreville, Gabon. ⁷⁰National HIV Reference Laboratory, Community Health Sciences Unit, Ministry of Health, Lilongwe, Malawi. ⁷¹African Society for Laboratory Medicine, Addis Ababa, Ethiopia. ⁷²National Medical and Molecular Biology Laboratory, Ministry of Health, Djibouti, Republic of Djibouti. ⁷³Africa CDC, Rapid Responder, Team Djibouti, Djibouti, Djibouti. ⁷⁴Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Ghana. ⁷⁵Seychelles Public Health Laboratory, Public Health Authority, Ministry of Health Seychelles, Victoria, Seychelles. ⁷⁶Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA. ⁷⁷Virology Unit, Institut Pasteur de Madagascar, Antananarivo, Madagascar. ⁷⁸National Health Laboratory Service (NHLS), Cape Town, South Africa. ⁷⁹Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre, Malawi. ⁸⁰Liverpool School of Tropical Medicine, Liverpool, UK. ⁸¹University of Nebraska Medical Center (UNMC), Omaha, NE, USA. ⁸²SAMRC Antibody Immunity Research Unit, School of Pathology, University of the Witwatersrand, Johannesburg, South Africa. ⁸³CHU de Bouaké, Laboratoire/Unité de Diagnostic des Virus des Fièvres Hémorragiques et Virus Émergents, Bouaké, Côte d'Ivoire. ⁸⁴UFR Sciences Médicales, Université Alassane Ouattara, Bouaké, Côte d'Ivoire. ⁸⁵School of Public Health, Pwani University, Kilifi, Kenya. ⁸⁶Faculty of Science and Techniques, University Marien Ngouabi, Brazzaville, Republic of the Congo. ⁸⁷Centre for Human Virology and Genomics, Nigerian Institute of Medical Research, Yaba, Lagos, Nigeria. ⁸⁸Nigeria Centre for Disease Control and Prevention, Abuja, Nigeria. ⁸⁹Laboratoire des Arbovirus, Fièvres Hémorragiques virales, Virus Émergents et Zoonoses, Institut Pasteur de Bangui, Bangui, Central African Republic. ⁹⁰Le Laboratoire National de Biologie Clinique et de Santé Publique (LNBCSP), Bangui, Central African Republic. ⁹¹West African Centre for Cell Biology of Infectious Pathogens (WACCBIIP), College of Basic and Applied Sciences, University of Ghana, Accra, Ghana. ⁹²Institute of Lassa Fever Research and Control, Irrua Specialist Teaching Hospital, Irrua, Nigeria. ⁹³PATH, Lusaka, Zambia. ⁹⁴Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC, USA. ⁹⁵Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA. ⁹⁶School of Life Sciences and Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research (SBIDER), University of Warwick, Coventry, UK. ⁹⁷Uganda Virus Research Institute, Entebbe, Uganda. ⁹⁸PathCare Vermaak, Pretoria, South Africa and Division of Virology, University of the Free State, Bloemfontein, South Africa. ⁹⁹College of Medicine and Allied Health Sciences, University of Sierra Leone, Freetown, Sierra Leone. ¹⁰⁰Botswana Harvard AIDS Institute Partnership and Botswana Harvard HIV Reference Laboratory, Gaborone, Botswana. ¹⁰¹Macha Research Trust, Choma, Zambia. ¹⁰²International Livestock Research Institute (ILRI), Nairobi, Kenya. ¹⁰³INRSP, Nouakchott, Mauritania. ¹⁰⁴Faculté de Médecine de Nouakchott, Nouakchott, Mauritanie. ¹⁰⁵Rwanda National Reference Laboratory, Kigali, Rwanda. ¹⁰⁶Robert Koch-Institute, Berlin, Germany. ¹⁰⁷G5 Evolutionary Genomics of RNA Viruses, Institut Pasteur, Paris, France. ¹⁰⁸Direcção Nacional da Saúde Pública, Ministério da Saúde, Luanda, Angola. ¹⁰⁹National Public Health Reference Laboratory—National Public Health Institute of Liberia, Monrovia, Liberia. ¹¹⁰Faculty of Pharmacy of Monastir, Monastir, Tunisia. ¹¹¹National Influenza Centre, Institut Pasteur d'Algérie, Algiers, Algeria. ¹¹²Department of Virology, National Health Laboratory Service (NHLS), Charlotte Maxeke Johannesburg Academic Hospital, Johannesburg, South Africa. ¹¹³School of Pathology, Faculty of Health Science, University of the Witwatersrand, Johannesburg, South Africa. ¹¹⁴Institute of Tropical Medicine, Universitätsklinikum Tübingen, Tübingen, Germany. ¹¹⁵Ministère de Santé Publique et de la Solidarité Nationale, Ndjamena, Chad. ¹¹⁶WHO Int Comoros, Moroni, Union of Comoros. ¹¹⁷World Health Organization, Africa Region, Brazzaville, Republic of the Congo. ¹¹⁸Division of Medical Virology, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Cape Town, South Africa. ¹¹⁹National Health Laboratory Service (NHLS), Tygerberg, Cape Town, South Africa. ¹²⁰UHAS COVID-19 Testing and Research Centre, University of Health and Allied Sciences, Ho, Ghana. ¹²¹Department of Biomedical Sciences, University of Health and Allied Sciences, PMB 31, Ho, Ghana. ¹²²Ministry of Health, COVID-19 Testing Laboratory, Mbabane, Kingdom of Eswatini. ¹²³Satellite Molecular Laboratory, Rivers State University Teaching Hospital, Port Harcourt, Nigeria. ¹²⁴Department of Emerging Infectious Diseases, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan. ¹²⁵CHU Habib Bourguiba, Laboratory of Microbiology, Faculty of Medicine of Sfax, University of Sfax, Sfax, Tunisia. ¹²⁶Central Public Health Laboratories (CPHL), Kampala, Uganda. ¹²⁷Institut Pasteur de Côte d'Ivoire, Département des Virus Épidémiques, Abidjan, Côte d'Ivoire. ¹²⁸Faculty of Medicine Ain Shams Research Institute (MASRI), Ain Shams University, Cairo, Egypt. ¹²⁹Doctoral School of Technical and Environmental Sciences, Department of Biology and Human Health, N'Djamena, Chad. ¹³⁰Department of Infectious Diseases Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. ¹³¹Charles Nicolle Hospital, Laboratory of Microbiology, National Influenza Center, Tunis, Tunisia. ¹³²University of Tunis El Manar, Faculty of Medicine of Tunis, Research Laboratory LR99ES09, Tunis, Tunisia. ¹³³College of Medicine and Allied Health Science, University of Sierra Leone, Freetown, Sierra Leone. ¹³⁴Namibia Institute of Pathology, Windhoek, Namibia. ¹³⁵National Institute of Hygiene, Lomé, Togo. ¹³⁶Virology/Molecular Biology Department, Central Health Laboratory, Victoria Hospital, Ministry of Health and Wellness, Port Louis, Mauritius. ¹³⁷Department of Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA. ¹³⁸WHO Burundi, Gitega, Burundi. ¹³⁹Grupo de Investigação Microbiana e Imunológica, Instituto Nacional de Investigação em Saúde (National Institute for Health Research), Luanda, Angola. ¹⁴⁰Departamento de Bioquímica, Faculdade de Medicina, Universidade Agostinho Neto, Luanda, Angola. ¹⁴¹Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA. ¹⁴²Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. ¹⁴³Institute of Agricultural Sciences, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Unaí, Brazil. ¹⁴⁴WHO South Sudan, Juba, South Sudan. ¹⁴⁵Faculty of Medicine, University of Burundi, Bujumbura, Burundi. ¹⁴⁶Pasteur Network, Institut Pasteur, Paris, France. ¹⁴⁷Botswana Institute for Technology Research and Innovation, Gaborone, Botswana. ¹⁴⁸Instituto Nacional de Saúde Pública, Praia, Cape Verde. ¹⁴⁹Zambia National Public Health Institute, Lusaka, Zambia. ¹⁵⁰Public Health Institute of Malawi, Lilongwe, Malawi. ¹⁵¹National Health Laboratory, Gaborone, Botswana. ¹⁵²Laboratory of Transmissible Diseases and Biologically Active Substances (LR99ES27), Faculty of Pharmacy, University of Monastir, Monastir, Tunisia. ¹⁵³Laboratory of Microbiology, University Hospital of Monastir, Monastir, Tunisia. ¹⁵⁴Biomedical Informatics and Chemoinformatics Group, Informatics and Systems Department, National Research Centre, Cairo, Egypt. ¹⁵⁵Ministry of Health and Wellness, Gaborone, Botswana. ¹⁵⁶Eastern Technical University of Sierra Leone, Kenema, Sierra Leone. ¹⁵⁷Zoonotic Arbo and Respiratory Virus Program, Centre for Viral Zoonoses, Department of Medical Virology, University of Pretoria, Pretoria, South Africa. ¹⁵⁸Next Generation Sequencing Unit and Division of Virology, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa. ¹⁵⁹National Reference Laboratory Lesotho, Maseru, Lesotho. ¹⁶⁰Centre for Biotechnology Research and Development, Kenya Medical Research Institute, Nairobi, Kenya. ¹⁶¹Swiss Tropical and Public Health Institute, Basel, Switzerland. ¹⁶²Laboratório de Investigações de Baney, Baney, Equatorial Guinea. ¹⁶³Ifakara Health Institute, Ifakara, Tanzania. ¹⁶⁴Department of Medical Diagnostics, Kumasi Centre for Collaborative Research in Tropical Medicine, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. ¹⁶⁵PraesensBio, Lincoln, NE, USA. ¹⁶⁶Department of Medical Laboratory Science, Niger Delta University, Bayelsa State, Nigeria. ¹⁶⁷Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Cairo, Egypt. ¹⁶⁸King Faisal Specialist Hospital and Research Center, Riyadh, Kingdom of Saudi Arabia. ¹⁶⁹Biological Prevention Department, Ministry of Defence, Cairo, Egypt. ¹⁷⁰Faculty of Science, Fayoum University, Fayoum, Egypt. ¹⁷¹Molecular Pathology Lab, Children's Cancer Hospital, Cairo, Egypt. ¹⁷²Laboratoire Biolim FSS/Université de Lomé, Lomé, Togo. ¹⁷³Department

of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, Abu Dhabi, United Arab Emirates. ¹⁷⁴High Institute of Biotechnology of Monastir, University of Monastir, Rue Taher Haddad 5000, Monastir, Tunisia. ¹⁷⁵Rwanda National Joint Task Force COVID-19, Rwanda Biomedical Centre, Ministry of Health, Kigali, Rwanda. ¹⁷⁶School of Health Sciences, College of Medicine and Health Sciences, University of Rwanda, Kigali, Rwanda. ¹⁷⁷Department of Microbiology, Faculty of Medical Laboratory Sciences, Omdurman Islamic University, Sudan. ¹⁷⁸Instituto Nacional de Saúde (INS), Marracuene, Mozambique. ¹⁷⁹Department of Disease Control, School of Veterinary Medicine, University of Zambia, Lusaka, Zambia. ¹⁸⁰Internal Medicine Department, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ¹⁸¹Institut Pasteur de Guinée, Conakry, Guinea. ¹⁸²Virology Laboratory, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ¹⁸³Department of Epidemiology and Community Health, Faculty of Clinical Sciences, College of Health Sciences, University of Ilorin, Ilorin, Kwara State, Nigeria. ¹⁸⁴Department of Public Health, Ministry of Health, Ilorin, Kwara State, Nigeria. ¹⁸⁵Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ¹⁸⁶Mayotte Hospital Center, Mayotte, France. ¹⁸⁷The African Center of Excellence in Bioinformatics and Data-Intensive Sciences, The Infectious Diseases Institute, Kampala, Uganda. ¹⁸⁸Immunology and Molecular Biology, Makerere University, Kampala, Uganda. ¹⁸⁹Department of Medicine, Faculty of Clinical Sciences, College of Medicine, Ambrose Alli University, Ekpoma, Edo State, Nigeria. ¹⁹⁰Division of Virology, National Health Laboratory Service and University of the Free State,

Bloemfontein, South Africa. ¹⁹¹Infectious Hazards Preparedness, World Health Organization, Eastern Mediterranean Regional Office, Cairo, Egypt. ¹⁹²Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo, Egypt. ¹⁹³Microbiology and Immunology Research Program, Children's Cancer Hospital Egypt, Cairo, Egypt. ¹⁹⁴National Public Health Laboratory, Ministry of Public Health of Cameroon, Yaoundé, Cameroon. ¹⁹⁵Faculty of Medicine and Biomedical Sciences, University of Yaoundé, Yaoundé, Cameroon. ¹⁹⁶Virology Service, Centre Pasteur of Cameroon, Yaounde, Cameroon. ¹⁹⁷Coordenadora da rede do Diagnóstico Tuberculose/HIV/COVID-19 na Instituição - Laboratório Nacional de Referência da Tuberculose em São Tomé e Príncipe, São Tomé, São Tomé and Príncipe. ¹⁹⁸Ponto focal para Melhoria da qualidade dos Laboratórios (SLIPTA) ao nível de São Tomé e Príncipe, São Tomé, São Tomé and Príncipe. ¹⁹⁹National Public Health Reference Laboratory (NPHRL), Mogadishu, Somalia. ²⁰⁰Faculty of Medicine of Monastir, University of Monastir, Monastir, Tunisia. ²⁰¹University of Basel, Basel, Switzerland. ²⁰²Clinical and Experimental Pharmacology Lab, LR16SP02, National Center of Pharmacovigilance, University of Tunis El Manar, Tunis, Tunisia. ²⁰³Harvard T.H. Chan School of Public Health, Boston, MA, USA. ²⁰⁴Centre MURAZ, Ouagadougou, Burkina Faso. ²⁰⁵National Institute of Public Health of Burkina Faso (INSP/BF), Ouagadougou, Burkina Faso. ²⁰⁶National Reference Center for Respiratory Viruses, Molecular Genetics of RNA Viruses, UMR 3569 CNRS, Université Paris Cité, Institut Pasteur, Paris, France. ²⁰⁷World Health Organization, Harare, Zimbabwe. ²⁰⁸Vietnamese-German Center for Medical Research, Hanoi, Vietnam. ²⁰⁹Howard Hughes Medical Institute, Fred

Hutchinson Cancer Center, Seattle, WA, USA. ²¹⁰Sub-Saharan African Network For TB/HIV Research Excellence (SANTHE), Durban, South Africa. ²¹¹World Health Organization, WHO Lesotho, Maseru, Lesotho. ²¹²Med24 Medical Centre, Ruwa, Zimbabwe. ²¹³Department of Virology, Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Ghana. ²¹⁴Division of Human Genetics, Department of Pathology, University of Cape Town, Cape Town, South Africa. ²¹⁵Yemaachi Biotech, Accra, Ghana. ²¹⁶Center for Human Genetics, College of Medicine and Health Sciences, University of Rwanda, Kigali, Rwanda. ²¹⁷Laboratory of Human Genetics, GIGA Research Institute, Liège, Belgium. ²¹⁸Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya. ²¹⁹Department of Global Health, University of Washington, Seattle, WA, USA. ***Corresponding author. Email: tulio@sun.ac.za (T.d.O.); ewilkinson@sun.ac.za (E.W.)**

†These authors contributed equally to this work.

‡Africa PGI collaborators are listed in the supplementary materials.

SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abq5358](https://doi.org/10.1126/science.abq5358)

Africa PGI Collaborator List

Figs. S1 to S16

Tables S1 to S4

Reference (77)

MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

Submitted 14 April 2022; accepted 12 September 2022
10.1126/science.abq5358

The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance

Houriya Tegally, James E. San, Matthew Cotten, Monika Moir, Bryan Tegomoh, Gerald Mboowa, Darren P. Martin, Cheryl Baxter, Arnold W. Lambisia, Amadou Diallo, Daniel G. Amoako, Moussa M. Diagne, Abay Sisay, Abdel-Rahman N. Zekri, Abdou Salam Gueye, Abdoul K. Sangare, Abdoul-Salam Ouedraogo, Abdourahmane Sow, Abdualmoniem O. Musa, Abdul K. Sesay, Abe G. Abias, Adam I. Elzagheid, Adamou Lagare, Adedotun-Sulaiman Kemi, Aden Elmi Abar, Adeniji A. Johnson, Adeola Fowotade, Adeyemi O. Oluwapelumi, Adrienne A. Amuri, Agnes Juru, Ahmed Kandeil, Ahmed Mostafa, Ahmed Rebai, Ahmed Sayed, Akano Kazeem, Aladje Balde, Alan Christoffels, Alexander J. Trotter, Allan Campbell, Alpha K. Keita, Amadou Kone, Amal Bouzid, Amal Souissi, Ambrose Agweyu, Amel Naguib, Ana V. Gutierrez, Anatole Nkeshimana, Andrew J. Page, Anges Yadouleton, Anika Vinze, Anise N. Happi, Anissa Chouikha, Arash Iranzadeh, Arisha Maharaj, Armel L. Batchi-Bouyou, Arshad Ismail, Augustina A. Sylverken, Augustine Goba, Ayoade Femi, Ayotunde E. Sijuwola, Baba Marycelin, Babatunde L. Salako, Bamidele S. Oderinde, Bankole Bolajoko, Bassirou Diarra, Belinda L. Herring, Benjamin Tsofa, Bernard Lekana-Douki, Bernard Mvula, Berthe-Marie Njanpop-Lafourcade, Blessing T. Marondera, Bouh Abdi Khaireh, Bourema Kouriba, Bright Adu, Brigitte Pool, Bronwyn McInnis, Cara Brook, Carolyn Williamson, Cassien Nduwimana, Catherine Anscombe, Catherine B. Pratt, Cathrine Scheepers, Chantal G. Akoua-Koffi, Charles N. Agoti, Chastel M. Mapanguy, Cheikh Loucoubar, Chika K. Onwuamah, Chikwe Ihekweazu, Christian N. Malaka, Christophe Peyrefitte, Chukwa Grace, Chukwuma E. Omoruyi, Clotaire D. Rafa, Collins M. Moranga, Cyril Erameh, Daniel B. Lule, Daniel J. Bridges, Daniel Mukadi-Bamuleka, Danny Park, David A. Rasmussen, David Baker, David J. Nokes, Deogratius Ssemwanga, Derek Tshiabula, Dominic S. Y. Amuzu, Dominique Goedhals, Donald S. Grant, Donwilliams O. Omuoyo, Dorcas Maruapula, Dorcas W. Wanjohi, Ebenezer Foster-Nyarko, Eddy K. Lusamaki, Edgar Simulundu, Edidah M. Ongera, Edith N. Ngabana, Edward O. Abworo, Edward Otieno, Edwin Shumba, Edwine Barasa, El Bara Ahmed, Elhadi A. Ahmed, Emmanuel Lokilo, Enatha Mukantwari, Eromon Philomena, Essia Belarbi, Etienne Simon-Loriere, Etil A. Anoh, Eusebio Manuel, Fabian Leendertz, Fahn M. Taweh, Fares Wasfi, Fatma Abdelmoula, Faustinos T. Takawira, Fawzi Derrar, Fehintola V. Ajogbasile, Florette Treurnicht, Folarin Onikepe, Francine Ntoumi, Francisca M. Muyembe, Frank E. Z. Ragomzingba, Fred A. Dratibi, Fred-Akintunwa Iyanu, Gabriel K. Mbunso, Gaetan Thilliez, Gemma L. Kay, George O. Akpede, Gert U. van Zyl, Gordon A. Awandare, Grace S. Kpeli, Grit Schubert, Gugu P. Maphalala, Hafaliana C. Ranaivoson, Hannah E. Omunakwe, Harris Onywere, Haruka Abe, Hela Karray, Hellen Nansumba, Henda Triki, Herve Albric Adje Kadio, Hesham Elgahzaly, Hlanai Gumbo, Hota Mathieu, Hugo Kavunga-Membo, Ibtihel Smeti, Idowu B. Olawoye, Ifedayo M. O. Adetifa, Ikponmwoosa Odia, Ilhem Boutiba Ben Boubaker, Iluoreh Ahmed Muhammad, Isaac Ssewanyana, Isatta Wurie, Iyaloo S. Konstantinus, Jacqueline Wemboo Afiwa Halatoko, James Ayei, Janaki Sonoo, Jean-Claude C. Makangara, Jean-Jacques M. Tamfum, Jean-Michel Heraud, Jeffrey G. Shaffer, Jennifer Giandhari, Jennifer Musyoki, Jerome Nkurunziza, Jessica N. Uwanibe, Jinal N. Bhiman, Jiro Yasuda, Joana Morais, Jocelyn Kiconco, John D. Sandi, John Huddlestone, John K. Odoom, John M. Morobe, John O. Gyapong, John T. Kayiwa, Johnson C. Okolie, Joicymara S. Xavier, Jones Gyamfi, Joseph F. Wamala, Joseph H. K. Bonney, Joseph Nyandwi, Josie Everatt, Joweria Nakaseegu, Joyce M. Ngoi, Joyce Namulondo, Judith U. Oguzie, Julia C. Andeko, Julius J. Lutwama, Juma J. H. Mogga, Justin OGrady, Katherine J. Siddle, Kathleen Victoir, Kayode T. Adeyemi, Kefentse A. Tumed, Kevin S. Carvalho, Khadija Said Mohammed, Koussay Dellagi, Kunda G. Musonda, Kwabena O. Duedu, Lamia Fki-Berrajah, Lavanya Singh, Lenora M. Kepler, Leon Biscornet, Leonardo de Oliveira Martins, Lucious Chabuka, Luicer Olubayo, Lul Deng Ojok, Lul Lojok Deng, Lynette I. Ochola-Oyier, Lynn Tyers, Madisa Mine, Magalutcheeme Ramuth, Maha Mastouri, Mahmoud ElHefnawi, Maimouna Mbanne, Maitshwarelo I. Matsheka, Malebogo Kebabonye, Mamadou Diop, Mambu Momoh, Maria da Luz Lima Mendona, Marietjie Venter, Marietou F. Paye, Martin Faye, Martin M. Nyaga, Mathabo Mareka, Matoke-Muhia Damaris, Maureen W. Mburu, Maximilian G. Mpina, Michael Owusu, Michael R. Wiley, Mirabeau Y. Tatteng, Mitoha Ondoo Ayekaba, Mohamed Abouelhoda, Mohamed Amine Beloufa, Mohamed G. Seadawy, Mohamed K. Khalifa, Mooko Marethabile Matobo, Mouhamed Kane, Mounerou Salou, Mphaphi B. Mbulawa, Mulenga Mwenda, Mushal Allam, My V. T. Phan, Nabil Abid, Nadine Rujeni, Nadir Abuzaid, Nalia Ismael, Nancy Elguindy, Ndeye Marieme Top, Ndongo Dia, Ndio Mabunda, Nei-yuan Hsiao, Nelson Boric Silochi, Ngiambudulu M. Francisco, Ngonda Saasa, Nicholas Bbosa, Nickson Murunga, Nicksy Gumede, Nicole Wolter, Nikita Sitharam, Nnaemeka Ndodo, Nnennaya A. Ajayi, Nol Tordo, Nokuzola Mbhele, Norosoa H. Razanajatovo, Nosamiefan Iguosadol, Nwando Mba, Ojide C. Kingsley, Okogbenin Sylvanus, Oladiji Femi, Olubusuyi M. Adewumi, Olumade Testimony, Olusola A. Ogunsanya, Oluwatosin Fakayode, Onwe E. Ogah, Ope-Ewe Oludayo, Ousmane Faye, Pamela Smith-Lawrence, Pascale Ondo, Patrice Combe, Patricia Nabisubi, Patrick Semanda, Paul E. Oluniyi, Paulo Arnaldo, Peter Kojo Quashie, Peter O. Okokhere, Philip Bejon, Philippe Dussart, Phillip A. Bester, Placide K. Mbala, Pontiano Kaleebu, Priscilla Abechi, Rabeh El-Shesheny, Rageema Joseph, Ramy Karam Aziz, Ren G. Essomba, Reuben Ayivor-Djanie, Richard Njouom, Richard O. Phillips, Richmond Gorman,

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Robert A. Kingsley, Rosa Maria D. E. S. A. Neto Rodrigues, Rosemary A. Audu, Rosina A. A. Carr, Saba Gargouri, Saber Masmoudi, Sacha Bootsma, Safietou Sankhe, Sahra Isse Mohamed, Saibu Femi, Salma Mhalla, Salome Hosch, Samar Kamal Kassim, Samar Metha, Sameh Trabelsi, Sara Hassan Agwa, Sarah Wambui Mwangi, Seydou Doumbia, Sheila Kialala-Mandanda, Sherihane Aryeetey, Shymaa S. Ahmed, Side Mohamed Ahmed, Siham Elhamoumi, Sikhulile Moyo, Silvia Lutucuta, Simani Gaseitsiwe, Simbirie Jalloh, Soa Fy Andriamandimby, Sobajo Oguntope, Solne Grayo, Sonia Lekana-Douki, Sophie Prosolek, Soumeya Ouangraoua, Stephanie van Wyk, Stephen F. Schaffner, Stephen Kanyerezi, Steve Ahuka-Mundeke, Steven Rudder, Sureshnee Pillay, Susan Nabadda, Sylvie Behillil, Sylvie L. Budiaki, Sylvie van der Werf, Tapfumane Mashe, Thabo Mohale, Thanh Le-Viet, Thirumalaisamy P. Velavan, Tobias Schindler, Tongai G. Maponga, Trevor Bedford, Ugochukwu J. Anyaneji, Ugwu Chinedu, Upasana Ramphal, Uwem E. George, Vincent Enouf, Vishvanath Nene, Vivianne Gorova, Wael H. Roshdy, Wasim Abdul Karim, William K. Ampofo, Wolfgang Preiser, Wonderful T. Choga, Yahaya Ali Ahmed, Yajna Ramphal, Yaw Bediako, Yeshnee Naidoo, Yvan Butera, Zaydah R. de Laurent, Africa Pathogen Genomics Initiative (Africa PGI), Ahmed E. O. Ouma, Anne von Gottberg, George Githinji, Matshidiso Moeti, Oyewale Tomori, Pardis C. Sabeti, Amadou A. Sall, Samuel O. Oyola, Yenew K. Tebeje, Sofonias K. Tessema, Tulio de Oliveira, Christian Happi, Richard Lessells, John Nkengasong, Eduan Wilkinson, Aaron L. Shibe, Abasi Ene Obong, Abayomi Fadeyi, Abbad Anas, Abd Elazeez Shabaan, Abd Monaem Adel, Abd Moniem Ain Shoka, W. Abdelhamid, Abdelilah Laraqui, Abdelkader Laatiris, Abdelkrim Meziane Bellefquih, Abdellah Faouzi, F. Abdelmoulah, Abdelmunim Essabbar, Abderrahmane Bimouhen, Abderraouf Hilali, I. Abdo, Abdou Padane, Abdoul Karim Sangar, Abdoul Karim Soumah, Abdoulaye Djimde, Abdoulaye Toure, Abdoulie Kante, Abdulla Bashein, Abdullah Salama, Abe Lojuan, Abebe Genetu Bayih, Abel Abera Negash, Abel Lissom, N. Abid, Abba A. Konou, Abo Shama, O. Abosedo, S. Abouelnaga, Abraham Ali, Abraham Kwabena Anang, Abraham Tesfaye, Adam K. Khan, Adamu Tayachew, Adane Mihret, Adba Alfatih AlEmam, Adede Hawi, A. Adesegun, Adey Feleke Desta, Adib Ghassan, Adjaratou Traor, B. A. Adjiratu Aissatou, Adodo Sadi, Adrian Egli, Adriano Mendes, Adugna Abera, Abdul Cand, Afaf Alaoui, Afonso Pedro, B. Agbodzi, A.M. Ageez, Ahidjo Ayoub, Ahlam Alarif, Ahmed E Kayed, Ahmed El-Taweel, Ahmed Elsayed, Ahmed F. Gad, Ahmed Fakhakh, Ahmed Kandeil, M. Ahmed, Ahmed Mostafa, O. S. Ahmed, Ahmed Reggad, Ahmed Taha, Ah Yong Vida, Aicha Bensalem, Aida Sivo, Aissam Hachid, F. Ajili, F.V. Ajogbasile, Akim A. Adegnika, Akoele Siliadin, Akwii Patience Natasha, Aladje Balde, Alan Lemtudo, Alaoui Sanaa-amine, A. M. Alaruusi, Alassane Ouro-Medeli, Albert Nyunja, Alberto Rizzo, Alemseged Abdissa, Alemu Tike Debela, Alessandro Mancon, Alessandro Marcello, Alexander Goredema, Alexander Greninger, Alexis Ndjolo, Alexis Niyomwungere, Alfredo Mari, Alfredo Mayor, M.A. Ali, Ali Zumla, Alia Ben Kahla, Alia Grad, Alice Kabanda, Alie Tia, Alimou Camara, Alimuddin Zumla, Alle Baba Dieng, Almoustapha I. Maiga, Amadou Alpha Sall, Amadou Daou, Amal Naguib, Amal Souiri, Amal Zouaki, Amalou Ghita, Amandine Mveang-Nzoghe, Amariane Kon, Amariane M. M. Kon, Ambroise Ahouidi, Amel Benyahia, Amel Naguib, K. E. Amer, Ameyo Dorkenoo, Amina Barkat, Aminata Dia, Aminata Mbaye, Aminata Mboup, Aminata Sileymane Thiam, Amine Idriss Lahlou, Amira Suliaman Wadi, Amivi Ehlan, Amujal Marion, Amuri Aziza, Amy Strydom, Anass Abbad, Anatole Nkeshimana, Andargachew Mulu, Anderrahmane Maaroufi, Andrea E. Luquette, Andreas Shiningavamwe, Andres Moreira-Soto, Andrew Azman, Andrew J. Bennett, Andrew Tarupiwa, Anga Latifa, Ange Badjo, Angel Angelov, Angela Brisebarre, Angela M. Detweiler, Angoune Ndong, Anja Werno, Anna Julienne Selbe NDiaye, Anna-Lena Sander, B. Annajr, Annamaria D'Aprile, Anne van der Linden, Annemiek van der Eijk, Annette Erhart, Anou M. Somboro, Anoumou Dagnran, Anthony Ahumibe, Anthony Lévassieur, Antje van der Linden, Antoine Dara, Anu Jegede-Williams, M. Aouni, A. Arjarquah, Arlene Uwituze, Arlo Upton, Armel Poda, Arsne Som, Arsne Zongo, Arsenia Massinga, K.M. Asare, Ashaba Fred Katabazi, Asma Ferjani, Assane Dieng, Astou Gaye-Gaye, N. Atiga, Atsbeha G. Weldemariam, Augustina Arjaquah, A. Auld, Awa Ba-Diallo, Awatef ElMoussi, Awunyo Sena, Aya Mohamed, Ayman Farhaly, B. Ayo-Ale, F. Ayoade, Ayola Akim Adegnika, Ayong More, Ayorinde Babatunde James, Azami Nawfel, Azaria Diergaardt, Azuka Patrick Okwuraiwe, Babafemi O. Taiwo, S. B. Baboo, B. S. Bahadoor, A.A. Bahnassy, Bakary Sanyang, U. Bakry, Bamba Fatoumata Tour, Bamidele Iwalokun, R. Banda, S. Bane, Bankole Johnson, Barada Cisse, Barbra Murwira, Bas Oude Munnink, Armel L. Batchi-Bouyou, R. Batra, Beatrice Dhaala, Belayachi Lamiae, A.B. Belkhir, I. Ben Ayed, Ben Morton, Ben Moussa, Ben Wulf, Bndicte Ndeboko, Benhida Rachid, Benjamiin B. Lindsey, Benjamin H. Foulkes, Benjamin Hounkpatin, Benjamin Selekon, M. Bensaid, Bernard Mpairwe, Bernard Ssentalo Bagaya, Bert Vanmechelen, Bertrand Lell, Beth Mutai, Bethlehem Adnew, Beuty Makamure, B. Bighignoli, Birahim Piere Ndiaye, Bishwo N. Adhikari, S.M. Bitrou, Blaise MBoringong Akenji, Bode Shobayo, Boitumelo Zuze, Bonifacio Manguire Nlavo, Bouchra Belfquih, Bouchra Boujemla, Bouna Yatassaye, Bouzidi Aymane, Brian Andika, Bright K. Yemi, Bronwyn Kleinhans, Bruna Galvao, Bubacar Delgado Pinto Embal, Bulelani Manene, Butoyi Pascal, Camille Capel, Campbell Anu, Carine Tchibozo, Carla Madeira, Carlos Cortes, Carniel Elisabeth, Carol Kifude, Carolle Yanique Tayimetha, Catherine E. Arnold, Catherine Okoi, Cecilia Waruhiu, Celestin Godwe, Celestina Obiekea, Celine Nkenfou, L. B. Chabuka, Chakib Nejjari, H. Chambaro, D. Chanda, K. Changula, Charifa Drissi Touzani, Charles Kayuki, Charles Nyagupe, Charles Ssuuna, Charoute Hicham, Chaselynn M. Watters, H. Chawech, Cheikh Sokhna, Chenaoui

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Mohamed, Chiamaka Nwuba, C. Chilufya, Chimaobi Chukwu, Chinyere Anyika, P.J. Chipimo, S. Chitanga, M. Chitenje, M. J. Chiwaula, Chouati Taha, Chris Mansell, Christelle Butel, Christian A. Devaux, Christian Drosten, Christian Ranaivoson, Christian Utpatel, Christophe Malabat, A. Chtourou, G. Chukwu, Claudia Daubenberger, Clement G. Kakai, Clement Masakwe, Clotaire Donatien Rafai, Collins Chenwi, Collins M. Misita, Collins Muli, Corine H. GeurtsvanKessel, Corinne Maufrais, Coulbaly Mbegnan, M. Crawford, Cristina M. Tato, Cyrus Yiaba, N. D'Amore, Dabiri Damlari, Dalia Ramadan, D. Damena, Daniel Ouso, Daniela Pisanelli, Danilo Licastro, David Hammer, David Nieuwenhuijse, David Patrick Kateete, David Wilkinson, Davy Leger Mouangala, Dawit Hailu Alemayehu, R.M. Dawood, R. De Sanctis, Deborah N. A. Mettle, Demba Koita, Dennis Kenyi Lodiongo, Dennis Laryea, Dennis N. Wandera, Dereje Leta, Dersi Noureddine, Desire Takou, Dessalegn Abeje Tefera, Dhvani Batra, Dia Ndongo, A. Diab, Diabou Diagne, Diane A. Mabika, Dian Bamourou, Dianke Samat, B. Diarra, Didier Raoult, Dieudonne Mutangana, Dineo Emang Tshiamo. Gape Nyepetsi, Diosdado Odjama Nseng Ada, Djimabi Salah, Donatella Cedola, Doris Harding, Dorothy Yeboah-Manu, Dossou Ange, Dowbiss Meta Djomsi, Dragana Drinkovic, C. Draper, Dumbuya Foday Sahr, Ebenezer Odewale, Eclou Sedjro, Edang-Minko Armand, Edgar Kigozi, Edith Koskei, Edith Nkwembe, Edmilson F. de Oliveira Filho, Edmira Maria da Costa, Edward Kiritu, Efreem S. Lim, Egon A. Ozer, B. Egyir, Ehab Abdelkader Abuelenein, Eitel Mpoudi Ngole, El Aliani Aissam, El Ansari Fatima Zahra, El Hamouchi Adil, El Oualid Abdelmjid, H. El-Shaqnqery, M. El-Zayat, Elamin Abualas, Elannaz Hicham, A. Elargoubi, Eleni Kidane, Elham Rizgalla, M.M. Elhoseny, F.W. Elhosiény, Elisabeth Carniel, Elizabeth Nyakarungu, S.M. Elkhateeb, Elmostafa Benaissa, Elmostafa El Fahime, M. Elouanass, M.H. Elsisy, Elvyre Mbongo-Nkama, Elwaleed M. Elamin Sara A.I Latif, Emame Edu, Emanuele Orsini, Emilio Skarwan, Emily K. Stefanov, Emmanuel Nasinghe, Emmanuel Nepolo, Emmanuel Ogunbayo, Emmanuelle Munger, Emmanuelle Permal, Emna Gaies, H. Ennibi, Ennibi Khalid, Erasmus Kotey, Erasmus Smit, Eric A. Lelo, Eric Adu, Eric Delaporte, Eric Katagirya, Eric Kezakarayagwa, Eric Muthanje, Erica Luis Maria Magalhes, Ernest Asiedu, Esemu Livo, Esperance Umumararungu, Esther Chitechi, Esther Omuseni, Etoundi Mballa Alain, Evelyn Bonney Quansah, Evelyn Y. Bonney, M.H. Ezzelarab, Faatu Cassama, Fabien Roch Niama, Fabio Arena, G. Faggioni, E. Fahsbender, Faith Sigei, Faouzi Abdellah, Farah Jouali, H. Farawyla, Farida Hilali, Fatima Chgouri, Fatima El Falaki, Fatima Zahra El ansari, Fatma Ebied, Fatna Bssaibis, Fattah Al Onifade, Faye Ousmane, Faye Ahmed Khardine, Faye Khardine, A.V. Fedorov, Fekadu Alemu, Fekak Jamal, Fengming Sun, Fetouma Doudou, A.E. Fikry, S. Fillo, Fiorenza Bracchitta, Firmin Kabor, L. Fki-berrajah, Flora Donati, Florence Fenollar, Foday Sahr Samuel Sorie, O. Folarino, Francine Folorunsho, Francisca Kabatesi, Francisco Muyembe-Mawete, Francois Malagon, Franklin Kiemd, Franklyn Egbe Asiedu-Bekoe, Frdrick Nkongho, Frederick Lemoine, Tei-Maya, R. H. Freitas, T. Fuh-Neba, I. Gaaloul, Gabriel Kabamba, Gadissa Gutema, Gaies Emna, Galadima Gadzama, Galal Mahmoud, Garba Ouangole, S. Gargouri, R. Garry, Gary McAuliffe, Gathii Kimita, Gdon Prince Manouana, Gelanew Tesfaye, George Awinda, George B. Kyei, George Michuki, Georgelin Nguema Ondo, Gerhard van Rooyen, Gert Marais, Getachew Abichu, Getachew Tesfaye Beyene, Getachew Tollera, Getnet Hailu, Ghamaz Hamza, Ghazi Kayali, Ghizlane El Amin, Gibson Mhlanga, Gilbert Kibet, Gildas Hounkanrin, F. Giordani, Gizaw Teka, Godfrey Masete, R. Goldstone, C. Gomaa, Gomaa Mokhtar, Gora Lo, B. I. Gosnell, A. Gottberg, Grace Angong Belournou, Grace Oni, Grace Vincent, Gregory Wani, Guedi Ali Barreh, Guillermo Garcia, Guohong Deng, N. P. Guseva, Guy Paterne Malonga Mbembo, Guy Stéphane Padzys, S. Gwayi, Habiba Ben Romdhane, Habiba Naija, Haby Diallo, A. Hadad, M.M. Hafez, Hafsia Ladhari, Hagar Elshora, Hailu Dadi, Hajjaji Mohammed, Hakima Kabbaj, Hala Hafez, A. Halafawy, Halidou Tinto, Halimatou Diop Ndiaye, Hamadi Assane, T.N. Hamdani, M.S. Hamdy, Hamidah Namagembe, M. Hammad, A. Hammami, Hamza Gharmaz, Hana Sofia Andersson, Hanae Dakka, Hanen El Jebari, Hany K. Soliman, Harmak Houda, R. Harvey, Hasnae Benkirane, Hassan Aguenau, Hassan Ihazmad, H. Hassan, W. Hassan, Hatem Ahmed, Heitzer Sogodogo, Helena Seth-Smith, Helisoa Razafimanjato, Hellen Koka, Hemlali Mouhssine, Henry Mwebesa, Hermes Perez, Herve Christian Paho Tchoudjin, Hicham Elannaz, Hicham Oumzil, Hilda Opoku Frempong, Hinson Fidelia, Hoda Ezz Elarab, Hong Xie, Houda Benrahma, Housna Arrouchi, Housseem Guedouar, Hu Luo, Hubert Bassne, Huiqing Si, Ian Goodfellow, Ibrahima Guindo, Ibrahima Halilou, Idil Salah Abdillahi, Idriss-Amine Lahlou, Idrissa Diawara, Ifunanya Egoh, F.A. Ige, A.A. Iknane, Ikram Bnouyahia, Ikram Omar Osman, Ilhem Boutiba Ben Boubaker, Iman Abdillahi Hassan, Iman Foda, Imane Smyej, Imen Kacem, Imen M dini, Imen Mkada, Inacio Mandomando, Iaki Comas, Ines M dini, L. Ingls, Innocent Mudau, Ireoluwa Yinka JOEL, Irina Chestakova, Irving Cancino, Isaac Phiri, Ismael Pierrick Mikelet Boussoukou, A. Ismail, Israel Osei-Wusu, Issaka Maman, Ivan Barilar, Ivy A. Asante, Izuwayo Gerard, Jaafar Heikel, Jacob Sououpgui, Jacques Marx, D. Jalal, Jalal Nourilil, Jalil El Atar, Jalila Ben Khelil, Jalila Rahoui, Jamal Fekkak, James Mutisya, James Ussher, Jan Felix Exler, Jane MaCauley, Janet Majanja, N. Jannoo, Jarra Manneh, Jasmin Scharnberg, Jasmin Schlotterbeck, Jasper Chimedza, Jawad Bouzid, Jean B. Niyibigira, Jean Claude Djontu, Jean Maritz, Jean-Claude Makangara Cigolo, Jean-louis Monemou, Jeanne dArc Umuringa, Jeremy Delerce, Jerome Ndaruhutse, Jerome Nkurunziza, Jesse Addo Asamoah, Jill Sherwood, Jing Wang, Joan Marti-Carerras, Joe K. Mutungi, Joe Mutungi, Joel Fleury Djoba Siaway, Joel Koivogui, Joep de Ligt, John Njuguna, John Rumunu, John Tembo, John Waitumbi, Johnson A. Adeniji, Jonathan Rigby, Jorn Hellemans, Joseph Fokam, Joseph L. DeRisi,

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Joseph Makhema, Joseph Mugisha, Joseph Ojonugwa Shaibu, Joseph Oliver-Commey, Josephine Bwogi, Josh Freeman, Josiah Ayoola Isong, Josphat Nyataya, Joy Ayoola, Joyce Appiah-Kubi, Judd F. Hultquist, Jude Gedeon, Judith Sokei, Julia Howard, Julia Schneider, Julian Campbell, Juliet Elvy, A.B. Jumaa, Justin Lee, Justin Lessler, Justin O'Grady, Kaba Kourouma, Kais Ghedira, K. Kalantar, Kamela Mahlakwane, S. Kamoun, Kangwa Mulonga, P.C. Kapata, F. Kapaya, M. Kapin'a, H. Karray Hakim, W. Kasambara, Kasmi Yassine, Kassahun Tesfaye, Kathleen Subramoney, A.D. Katyshev, N. Kayeyi, Kayla Barnes, Kayla Delaney, E.V. Kazorina, Kedumetse Seru, S. Keita, Keith Durkin, Keith R Jerome, Keke K. Ren, Kena Swanson, Kenneth K. Maeka, Keren Okyerebea Attiku, Kevin Sanders, Kevine Zang Ella, Keyru Tuki, Khabab A. Elhag, Khadim Gueye, Khaled Amer, Khalid Ennibi Mostafa, N. Kharat, Z. Khumalo, Kilian Stoecker, L. Kim, Kimberly A. Bishop-Lilly, J. Kimotho, Kitane Driss Lahlou, Kofi Bonney, Kokou Tegueni, Kolawole Wasiu Wahab, E.V. Kolomoets, Komal Jain, Kominist Asmamaw, Komlan Kossi, Kondwani Jambo, Kouadio T. Yao, Kouriba Drr, Kra Ouffou, Y.M. Krasnov, Krishna Kumar Kandaswamy, Kristian Andersen, A.A. Kritsky, Kumbelembe David, Kutlo Macheke, V.V. Kutyrev, S. Kwenda, Kwitaka Maluzi, Kwok Lee, Kyle A. Long, Kyra Grantz, Lacy M. Simons, Laetitia Serrano, Lagos State Government, Laila Elsayy, Laila Sbabou, Lallepak Lamboni, LaRinda A. Holland, Lasata Shrestha, Lassana Sangar, Latifa Anga, Lauren Jelly, Laurien Hoornaert, Le Thi Kieu Linh, Legodile Kooepile, Leigh-Anne MC Intyre, Lon Mutesa, Leona Okoli, Lopold Ouedraogo, Kuate-lere Lesego, D. Leta, A. Letaief, G. Liboro, Lilian Kanjau, L. Lin, Linda Boatemaa, Linda Houhamdi, Lipkin W. Ian, F. Lista, M.M. Liwewe, Lloyd Mulenga, Logan J. Voegtly, Loide Shipingana, Loris Micelli, Lorreta Kwasah, Loubna Allam, Louise Lefrancois, Loukman Salma, Lucas N. Amenga-Etego, Ludivine Brechar, Ludovic Mewono, Luis A. Estrella, Lusya Mhuulu, Lwanga Newton, M. Ulrich, M.T. Mogotsi, Maaroufi Abderrahmane, W. Mad, W.Y. Madi, Madlen Stange, S. Magdeldin, Maher Kharrat, B. Mahlangu, Mahmoud Shehata, N. Mahrous, A. Maida, Makhtar Camara, T. Makori, K. Malama, Malena S. Bestehorn-Willmann, I. Malolo, Mamadou Beye Cheikh Ibrahima Lo, Mamadou Bhoeye Keita, Mamadou Saliou Bah, Mamadou Saliou Sow, Mamoudou Harouna Djingarey, Mamoudou Maiga, Manal Hamdy Elsaid, Manal Hamdy Zahran, Mandiou Diakite, Manel Ben Sassi, Mangombi Pambou, N. Manickchund, Manoli Torres Puente, S. S. Manraj, T. Mansour, Maowia M. Mukhtar, Marcel Tongo, Marchoudi Nabila, Margaret Mills, Maria Artesi, Maria Pia Patrizio, Maria Rita Gismondo, Maria Rosaria Lipsi, Mariam Kehinde SULAIMAN, Mariama Kujabi, Marie Amougou, Marie Claire Okomo, Marie Madeleine Chabert-Consen, Marie-Astrid Vernet, Marie-Pierre Hayette, Mariem Gdoura, Marijke Reynders, Marion Barbet, Marion Koopmans, Marjan Boter, Mark Siedner, Markos Abebe, Markus H. Antwerpen, Marouane Melloul, Martin Maidadi Foudi, Martine Peeters, Marvin Hsiao, Mary DeAlmeida, Mary Lalemi, Mary-Ann Davies, K. Masahiro, Masse Sambou, M. Mathabo, Mathew D. Parker, Mathias C. Walter, H. Mathur, Matt Blakiston, Matt Storey, Matthew Bates, Matthew Rogers, Matthias Pauthner, Maud Vanpeene, Maurizio Margaglione, Max Bloomfield, May Abdelfattah, May Sherif Soliman, Mbengu Fall, K. Mdallose, Meei-Li Huang, S. Mehta, Mlanie Albert, Melchior A. Jol Assi, Mline Bizard, Merabet Mouad, Meriem Laamarti, Messanh Douffan, S. Mhalla, Michael Addidle, Michael Marks, Michael Nagel, Michael V. Deschenes, Micheala Davids, Michelle Balm, Michelle Lin, Michelle Tan, A. Mihrete, Mikhail Olayinka BUHARI, Milanca Agostinho C, Mildred Adusei-Poku, Milkah Mwangi, Mina Kamel, J. Miranda, Mireille Prince-David, Miriam Eshun, Misaki Wayengera, Mitai Mishra, Mjid Eloualid, Mly Abdelaziz Elaoui, A. Mnguni, F. Mnyameni, Mogomotsi Matshaba, T. Mohale, Mohamed Abdel-Salam Elgohary, Mohamed Ahmed Ali, Mohamed Ben Moussa, Mohamed Chenaoui, Mohamed El Sayes, Mohamed Elhadidi, Mohamed Gomaa Seadawy, Mohamed Hassan Abdoelraheem, Mohamed Hassany, Mohamed Houmed Aboubaker, K.S. Mohamed, Mohamed Kamal, Mohamed Rhajaoui, Mohamed Seadawy, Mohamed Shamel, Mohamed Shemis, K.S. Mohammed, Mohammed Walid Chemao Elfihri, Mohcine Bennani Mechita, Mokhtar Gomaa, Molalegne Bitew, M. Momoh, Mona O.A. Alkarim, Monemo Pacome, Monilade Akinola, A. Monte, G. Monuir, M. Monze, M. Mooko, A.N. Morales, Moreira-Soto Andres, Moriba Povogui, Mosepele Mosepele, Moses Chilufya, Moses Joloba, Moses Luutu, Mostafa Elouennass, Mostfa Elhoseiny, Mostfa Elnakib, Mostfa Yakout, Mouhcine Gardoul, Mouhssine Hemlali, A. Mouity Matoumba, Mouna Ben Sassi, Mouna Safer, Mounneem Essabbar, Moustapha Mbow, Moustapha Nzamba Maloum, Moustapha Sakho, Moyinoluwa Odugbemi, P. Mtshali, B. Mubemba, Muchaneta Mugabe, M. Mufinda, Muhammad Faisal, Muinah Adenike Fowora, Mukantwari Enatha, W. Muleya, Muntaser elTayeb Ibrahim, F. Mupeta, Murebwayire Clarisse, Mushal Ali, Mushal Allam, Mustapha Mouallif, S.N. Muuo, W. Mwangomba, Myriam Seffar, T. E. Mzumara, N. N'dilimabaka, Nabila Soara, A. Nabli, Nadia El Mrimar, Nadia Rodrigues, Nadia Siteo, Nafisatou Leye, A. Naguib, K. S. Nalubamba, Nancy M. El Guindy, Nandi Siegfried, Nardjes Hihhi, Narjis Amar, E.A. Naryshkina, Nathalia Endjala, Nathalia Garus-Oas, Nathan Kapata, Ndack Ndiaye, Ndahafa Frans, N.T. Ndam, Ndye Coumba Tour Kane, Ndiaye Ndack, Ndiene Fainguem, Ndodo Nnaemeka, Ndumbu Pentikainen, L. Ndwiga, Nedio Mabunda, N. Neff, A.A. Negash, Nejla Stambouli, Z. Neto, A.M. Ngonga Dikongo, W. Ngosa, Ngozi Mirabel Otuonye, F. S. Niatou-Singa, Nicaise T. Ndam, Nicholas Feasey, Nicholas Mwikwabe, J. Nicod, Nicole Vidal, Nikki Freed, Nischay Mishra, Nissaf Ben Alaya, Nitin Savaliya, Noah Baker, No Patrick Mbondoukwe, Nokukhanya Mdallose, Nonso Nduka, Noura M Abo Shama, Nourlil Jalal, Nsubuga Gideon, N. Ntuli, Nuro Abilio, Nyam Itse Yusuf, Oby Wayoro, M. Ochwoto, Ofonime Ebong, L. Ofori-Boadu, Okomo Assoumou Marie Claire, Ola Elroby, Olabisi Olabisi, Olajumoke Ojo, Popoola, L. Olfert, Olin

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensees American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Silander, Olufemi Obafemi, Olufemi Samuel Amoo, Olukunke Oluwasemowo, Olusola Anuoluwapo Akanbi, Oluwakemi Laguda-Akingba, Oluwatimilehin Adewumi, Omar Askander, Omar Elahmer, S. Omar, S. Omilabu, Omnia Kutkat, Omoare Adesuyi, Omondi Francis Carey, Onalethata Lesetedi, E. Ongera, Ontlametse T. Bareng, C.K. Onwuamah, O. Ope-Ewe, Osama Mansour, Oscar Kanjerwa, F. Oteng, Otmane Touzani, Oumaima Ait Si Mohammed, Oumy Diop, Ouna Ouadghiri, Ousseynou Gueye, C. Owusu-Nyantakyi, A. Oyefolu, Oyeronke Ayansola, P Nthiga, S. Palomba, L. Panja, Papa Alassane Diaw, D. Park, Pasacaline Manga, H. Patel, Patience Motshosi, M. Pato, Patrick Amoth, Patrick Descheemaeker, Patrick Mavingui, Patrick Tuyisenge, M. Pattoo, Paul Dobi, Paul Liberator, Paulin N. Essone, Paulina Joozinho da Costa Jarra Manneh, Pauline Yacine Sene, Paulo A. Carralero R.R. Paixo, Pavitra Roychoudhury, Peace O. Uche, Pei Zhou, Penda Malhado Diallo, A. Pereira, Petas Akogbeto, Peter Bauer, Peter T. Skidmore, Petra Raimond, Phasha-Muchemenye Mmatshopho, Philip Ashton, C. Philip, Philip El-Duah, Philip M. Soglo, Philip Wonder Phiri, Phillip Wagner, Philippe Colson, Philippe Dussart, Philippe Lavrard Meyer, Phillip Ashton, Phionah Tushabe, Pierre-Edouard Fournier, Piet Maes, Popova A. Yu, Portia Manangazira, Praise Adewumi, Qi Yang, Quaneeta Mohktar, E.B. Quansah, P. Quashie, Rabia Quedraogo, Rachel Maqsood, Rachid Githii, Rachid Abi, Rachid Benhida, Rachid El Jaoudi, Rahaman A. Mentag, Rahma Ahmed, Raiva Algeriani, Rajiha Simbi, Ramalia Chabi Abubeker, Ramon Nari, Ramy Lorenzo-Redondo, Galal, A. Raouf, Raoul Saizonou, Raphael Lumembe, Ravena Mubichi, Regina Z. Cer, Reham Dawood, Reham Kassab, Rehema Liyai, A. Rehn, Rei Jos Pereira, Reina Sikkema, Rfaki Abderrazak, Riad Mounir Armanious, Riadh Daghfous, Riadh Gouider, Richard Adegbola, Richard Lino Loro Lako, Richard Molenkamp, Richard Webby, Richmond Yeboah, S. Rick, Rida Tagajdid, Rine Zeh Nfor, Rivalyn Nakoune Yandoko, Robert Newton, Robert Rutayisire, Rodney S. Daniels, Rodrigue Bikangu, Rodrigue K. Kohoun, Rodrigue Kamga, Rodrigue Mintsu Nguema, Roger Shapiro, J. Rogers, Rogers Kamulegeya, Rokaia Laamrti, Romo Aim Laclong Lontchi, Ronald Kiiza, Ronald M. Galiwango, Rosella De Nittis, W.H. Roshdy, Rotimi Myrabelle Avome Houechehou, Roua Ben Othman, Rui Inndi, M.A. Saad, Saad Amzazi, Sabin Nsanzimana, Sada Diallo, Y.A. Sadji, Safae El Mazouri, Safae Elkochri, Safae Ghouleme, Safiatou Karidioula, Safietou Sankhe, Sahr Gevao, P.F. Sahr, J.O. Saibu, Sadou Ouedraogo, S. Saiid, Sainabou Laye Ndure, Sakoba Keita, D. Salah, Salah Eldin Hussein, H. Salah, A.A. Saleh, M. Saleh, Salifou Sourakatou, Sally Roberts, Salma Abid, Salma Sayed, M. Salou, O.B. Salu, Sam Lissauer, Sam O'neilla Oye Bingono, Samba Ndiour, Sameira M. Fageer, Samia Abdou Girgis, M. Samir, O. Samir, Samira Benkeroum, Samira F. Ibrahim, Samira M. Fageer, Samira Zoa Assoumou, Samirah Saiid, Samoel Ashimosi Khamadi, Samson Konongoi Limbaso, Samuel Armoo, Samuel Kirimunda, Samuel Sorie, Samwel Lifumo Symekher, Samwel Owaka, Sana Ferjani, Sanaa Alaoui-Amine, Sander Anna-Lena, Sankhe Safietou, Santiago Jimnez-Serrano, Santigie Kamara, Sara Chammam, Sara Mahmoud, Sarah Jefferies, Sarah Rubin, Sarah Stanley, Saraswathi Sathes, Saro Abdella, Sarra Chamman, Savannah Mwesigwa, H. Sawa, M.G. Seadawy, Sean Ellis, Sbastien Bontems, Sefetogi Ramaologa, Sekesai Zinyowera, Selassie Kumordjie, Sen Claudine Henriette Ngomtcho, Sena Awunyo, Serge Alain Sadeuh-Mba, Serigne Saliou Niane, Seth Okeyo, Settimia Altamura, Seyram Bless Agbenyo, Shah Mohamed Bakhsh, Shahin Lockman, Shahinaz.A. Bedri, Shaimaa Soliman, L. Shalaby, Shannon Wilson, Sharmini Muttaiyah, Sharon Abimbola, Sharon Hsu, S.A. Shcherbakova, Shebbar Osiany, Shereen Shawky, Sherine Helmy, A.P. Shevtsova, Shimaa Moustafa, Shirlee Wohl, Shirley Johane, Sidonie A.M. Kagnissode, Silvanus Mukunzi Opanda, Sim Mayaphi, Simo Tchuda Bit, Simeone Dal Monego, Simon Peter Ruhweza, Simone Eckstein, Sindayiheba Reuben, N. Sinyange, Sivaramakrishna Rachakonda, Soa Fy Andriam, Sofia Viegas, Sokhna Sogodogo, Sola Ndongo, Solomon Ajibaye, Sombo Langat, Somda Soro Georgina Fwoloshi, Sonal P. Charlene, Sondes Henson, Sonia V. Haddad, Bedi, E.A. Sosedova, Souad Kartti, Souissi Amira, Souleymane Mboup, Soundl Mat, Sounkalo Dao, Sourakatou Salifou, Srinivas Reddy Pallerla, Stefan Niemann, Steffen Borrmann, Stephen Asimwe, Stephen M. Egan, Stephen Ochola, Stephen Wanok, Steven J. Reynolds, Subomi Olorunnimbe, Sudhir Bunga, C. Sujeewon, Sunday Babatunde, Susan Engelbrecht, Susan Morpeth, Susan Taylor, Susann Handrick, Swaibu Gatere, Sylvie Melingui, S.L. Symeker, Syntyche Devatchagni, A.G. Taha, Taha Chouati, Taha Maatoug, Tahar Bajjou, A. Takada, K.A. Taloa, Tamrayehu Seyoum, M. Tan, Tania Stander, Tanja Niemann, Tarek Aanniz, Tarek Refaat Elnagdy Raafat Zaher, Tatenda Takawira, C. Tato, D.A. Tefera, Tei-Maya Fred, Tesfaye Gelanew, Tesfaye Rufael, Thanh Le Viet, Thela Tefelo, Thrse Kagon, Thibaut Armel Cherif Gnimadi, K. Thiongo, Thomas Briese, Van L. Thomas, Thongbotho Mphoyakgosi, Thushan I de Silva, Tim Roloff, Timothy Blackmore, Tiziana Rollo, Tom Lutalo, Tomade A.M. Ibrahim, T. B. Tombolomako, Tomiwa Adepetun, C. Tomkins-Tinch, Tony Wawina-Bokalanga, Tope Sobajo, Touil Nadia, Touria Essayagh, Toy Nwako, Traor, H. Triki, R. Tsiry, Tsiry Randriambolanantsoa, Tumisang Madisa, M. Turki, C.A. Ugwu, Uyi Emokpae, Valeria Delli Carri, Valeria Micheli, G. Van Rooyen, M. Vanaerschot, Vanessa Magnussen, Vanessa Mohr, Vani Sathyendran, Veronica Playle, R. Viana, U. Vickos, Victor Max Corman, Victor Mukonka, Victor Ofula, Vida Ahyong, Vincent Appiah, Vincent Bours, Violette V. M'cormack, Virginia Hope, Vivi Hue-Trang Lieu, Vololoniaina Raharinosy, Wadegu Meshack, N. Wadonda, J. Wadula, Wael Ali, Waidi Sule, Wallace D Bulimo, Warren Cyrus Yiaba, C.N. Waruhui, Wasfi Fares, R. Webby, A.G. Weldemariam, Wendy Karen Jo, R. Woelfel, S. Wolh, I. Wurie, Xavier Crespin, Xiaoyun Ren, Xiuhua Wang, Ya.M. Krasnov, Yacine Amet

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).



Dia, Yacine DIA Seni Ndiaye, Yacob Mohamed Yusuf, Yacouba Sawadogo, Yahya Yadouleton, Yakob Gebregziabher Maidane, Yao Tsegay, Yasser Layibo, Yassine El Hady, Yassmin Sekhsokh, Yawo A. Moatasim, Sadj, C. Yeboah, Youbi Mohammed, Yousif Rabih Makki, Youssef Akhoud, Yuri Ushijima, Yusuf Jimoh, Yvette Badou, J. Matoke D. Zablou, J.O. Zablou, Zaineb Hamzaoui, Zakia Regragui, Zara Wuduri, Zein Souma, Zeinab Ali Waberi, Zeinab S. Imam, Zekiba Tarnagda, Zemmouri Faouzia, J. Zhang, Zhenghui Li, Zimmerman Maiga, Zohour Kasmy, Minko O. Zong, A. Zorgani, Zouheir Yassine, Zoukaneirii Issa, P. Zulu, and Zuzheng Xiang

Science, **378** (6615), eabq5358.

DOI: 10.1126/science.abq5358

Surveillance across Africa

The past 2 years, during which waves of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants swept the globe, have starkly highlighted health disparities across nations. Tegally *et al.* show how the coordinated efforts of talented African scientists have in a short time made great contributions to pandemic surveillance and data gathering. Their efforts and initiatives have provided early warning that has likely benefited wealthier countries more than their own. Genomic surveillance identified the emergence of the highly transmissible Beta and Omicron variants and now the appearance of Omicron sublineages in Africa. However, it is imperative that technology transfer for diagnostics and vaccines, as well the logistic wherewithal to produce and deploy them, match the data-gathering effort. —CA

View the article online

<https://www.science.org/doi/10.1126/science.abq5358>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

3. Improved monitoring of parasites that evade diagnosis, treatment or vaccination

This chapter contains the following publications:

Hosch *et al.* **Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment.** *Malaria Journal* 2022

Yoboue *et al.* **Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests.** *PLOS Neglected Tropical Diseases* 2022

Bechtold *et al.* **Development and evaluation of PlasmoPod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance.** *PLOS Global Public Health* 2023

Hosch *et al.* **PHARE: A bioinformatics pipeline for compositional profiling of multiclonal *Plasmodium falciparum* infections from long-read Nanopore sequencing data.** Under review by *Journal of Antimicrobial Chemotherapy*

3.1. Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment

Published in Malaria Journal, 2022

RESEARCH

Open Access



Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment

Salome Hosch^{1,2}, Charlene Aya Yoboue^{1,2,7}, Olivier Tresor Donfack³, Etienne A. Guirou^{1,2}, Jean-Pierre Dangy^{1,2}, Maxmillian Mpina^{1,4,5}, Elizabeth Nyakurungu^{4,5}, Koranan Blöchliger^{1,2}, Carlos A. Guerra³, Wonder P. Phiri³, Mitoha Ondo'ó Ayekaba⁶, Guillermo A. García³, Marcel Tanner^{1,2}, Claudia Daubenberger^{1,2*†} and Tobias Schindler^{1,2,5*†} 

Abstract

Background: Surveillance programmes often use malaria rapid diagnostic tests (RDTs) to determine the proportion of the population carrying parasites in their peripheral blood to assess the malaria transmission intensity. Despite an increasing number of reports on false-negative and false-positive RDT results, there is a lack of systematic quality control activities for RDTs deployed in malaria surveillance programmes.

Methods: The diagnostic performance of field-deployed RDTs used for malaria surveys was assessed by retrospective molecular analysis of the blood retained on the tests.

Results: Of the 2865 RDTs that were collected in 2018 on Bioko Island and analysed in this study, 4.7% had a false-negative result. These false-negative RDTs were associated with low parasite density infections. In 16.6% of analysed samples, masked *pfhrp2* and *pfhrp3* gene deletions were identified, in which at least one *Plasmodium falciparum* strain carried a gene deletion. Among all positive RDTs analysed, 28.4% were tested negative by qPCR and therefore considered to be false-positive. Analysing the questionnaire data collected from the participants, this high proportion of false-positive RDTs could be explained by *P. falciparum* histidine rich protein 2 (PfHRP2) antigen persistence after recent malaria treatment.

Conclusion: Malaria surveillance depending solely on RDTs needs well-integrated quality control procedures to assess the extent and impact of reduced sensitivity and specificity of RDTs on malaria control programmes.

Keywords: Molecular malaria surveillance, False-positive malaria rapid diagnostic test, PfHRP2 persistence, *pfhrp2* gene deletion

Background

According to the World Health Organization (WHO), more than 409,000 malaria-related deaths were reported in 2019, most of them in children below the age of 5 years [1]. The majority of malaria infections (94%) and malaria-related deaths (95%) occurred in sub-Saharan Africa (SSA) [2], where *Plasmodium falciparum* is the dominant

*Correspondence: claudia.daubenberger@swisstph.ch;
tobias.schindler@swisstph.ch

†Claudia Daubenberger and Tobias Schindler contributed equally to this work

¹ Swiss Tropical and Public Health Institute, Basel, Switzerland
Full list of author information is available at the end of the article



© The Author(s) 2022. **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.

malaria parasite [1]. The test-treat-track strategy advised by WHO is one of the backbones of current malaria control and elimination programmes [3]. This strategy entails every suspected malaria case be tested, every confirmed case be treated, and the disease be tracked through surveillance systems [4]. Testing relies heavily on rapid diagnostic tests (RDTs), exemplified by the more than 348 million RDTs distributed globally in 2019 [1]. In SSA, RDTs have almost completely replaced light microscopy for malaria diagnosis, accounting for an estimated 75% of all malaria tests conducted in 2017 [5]. RDTs are point-of-care tests that detect circulating antigens, such as the *P. falciparum*-specific histidine rich protein 2 (PfHRP2) or histidine rich protein 3 (PfHRP3), as well as the pan-*Plasmodium* spp. enzymes, lactate dehydrogenase (pLDH) or aldolase [6]. More than 90% of RDTs currently in use target the PfHRP2 antigen because of its higher sensitivity compared to non-PfHRP2 antigens [7]. PfHRP2-based RDTs used for the diagnosis of febrile patients that suffer from malaria infection are highly sensitive and specific [8]. RDTs are often used by national malaria surveillance programmes. However, when individuals are asymptomatic with low parasite densities, RDTs often fail to detect the parasites due to low antigen concentrations [9, 10].

A recent study showed that false-negative RDTs (FN-RDT) are more common in lower malaria transmission settings, younger subjects and in urban areas in SSA [11]. Reduced diagnostic performance of RDTs has also been attributed to genetic diversity of the *pfhrp2* gene [12], differences in expression levels of PfHRP2 antigen in parasite field strains [13], or deletion of *pfhrp2* and *pfhrp3* genes in isolates [14]. *Pfhrp2* gene deletions appear to be common and therefore are relevant as they might be a threat to malaria control programmes based on monitoring of malaria prevalence through RDTs [15, 16].

Less attention has been given to the specificity of malaria RDTs used in malaria surveys that potentially result in false-positive results. False-positive RDTs (FP-RDT) have been associated with high levels of circulating rheumatoid factor [17–19] or acute typhoid fever [20]. There is evidence of FP-RDTs in patients infected with *Schistosoma mekongi* [21] or human African trypanosomiasis [22]. FP-RDTs are also caused by persisting antigen circulation in peripheral blood after successful *P. falciparum* drug treatment. A meta-analysis revealed that around half of the PfHRP2-detecting RDTs remain positive 15 days (95% CI 5–32) post *P. falciparum* treatment, which is 13 days longer than RDTs based on the pLDH antigen [23]. The latter study also reported a higher persistent RDT positivity among individuals treated with artemisinin combination therapy (ACT) than those treated with other anti-malarial drugs. Since RDTs are

instrumental to malaria surveillance programmes, their diagnostic performance should be systematically monitored over time using sensitive and highly specific methods detecting *Plasmodium* spp. molecular markers. Described here is an approach for quality control of field-deployed RDTs by retrospective molecular analysis of the parasite DNA retained on them using RT-qPCR.

Methods

The 2018 malaria indicator survey conducted on Bioko Island as a biobank of RDTs for molecular malaria surveillance

A malaria indicator survey (MIS) has been conducted annually since 2004 on the island of Bioko, Equatorial Guinea, to evaluate the impact of malaria control interventions [24]. The survey uses a standard questionnaire developed by the Roll Back Malaria initiative to gather information on selected households and their occupants. The 2018 Bioko Island MIS covered 4774 households with 20,012 permanent residents, among whom 13,505 persons consented to storage and molecular analysis of their RDT. Briefly, consenting individuals living in surveyed households are tested for malaria and malaria-related anaemia. Malaria testing was done with the CareStart™ Malaria HRP2/pLDH (Pf/PAN) combo test (ACCESS BIO, NJ, USA). PfHRP2-positive RDTs were recorded as *P. falciparum*, pLDH-positive RDTs as *Plasmodium* spp. and RDT-positive for both antigens as mixed. The haemoglobin level in peripheral blood was measured during the MIS using a battery-operated portable HemoCue system (HemoCue AB, Ängelholm, Sweden). The anaemia status (mild, moderate, severe) was categorized based on definitions published by WHO [25] stratified by age, gender and pregnancy status. Households were assigned scores based on the type of assets and amenities they own to derive a surrogate of their socio-economic status (SES), using principal component analysis (PCA). After ranking all households based on their score, they were divided into five equal categories (quintiles), each with approximately 20% of the households. The first quintile corresponded to the lowest wealth index and the fifth to the highest wealth index. The household wealth index categories were also assigned to permanent household members.

Detection and quantification of *Plasmodium* spp. nucleic acids extracted from RDTs

A previously published dataset generated with the Extraction of Nucleic Acids from RDTs (ENAR) protocol developed by the authors was extended for this study [26]. Briefly, RDTs were barcoded, stored at room temperature and shipped to Basel, Switzerland, for nucleic acid (NA) extraction and detection. This approach simplifies small

volume blood collection, transport and storage logistics, and allows linking outcomes of molecular-based detection of parasite-derived NA with the demographic and socio-economic information collected from each corresponding MIS participant at high throughput.

All 2865 samples were initially screened with the PlasQ RT-qPCR assay [27]. In this RT-qPCR assay, the high copy number *P. falciparum*-specific varATS region [28] and the pan-*Plasmodium* 18S rDNA gene were targeted [29, 30]. Samples with cycle of quantification (Cq) value < 45 in two replicates of either of the two targets, varATS or 18S rDNA, were considered positive for active blood-stage malaria infection. *Plasmodium falciparum* parasites were quantified based on their Cq value for varATS [26]. In addition, only samples with Cq value < 35 for amplification of the internal control gene, the human *rnasep* gene were included, to demonstrate that the NA extracted from the RDTs is sufficient for reliable molecular analysis of malaria parasites. Non-*falciparum* malaria species identification of samples positive for the pan-*Plasmodium* target 18S rDNA was performed with a multiplex RT-qPCR assay based on species-specific 18S rDNA sequences as described previously [26].

Quality control and categorization of RDT outcomes

A RDT was considered positive if a healthcare worker recorded a positive signal for the PfHRP2, pLDH or both targets during the MIS. Among these positive RDTs, a true-positive RDT (TP-RDT) result was defined as a RDT with detectable *Plasmodium* spp. NA (two replicates with varATS and/or 18S rDNA Cq < 45 and human *rnasep* Cq < 35). A FP-RDT result was defined as positively read and recorded RDT in the field but with a negative outcome for *Plasmodium* spp. NA based on PlasQ RT-qPCR in the presence of human *rnasep* Cq < 35. Negative RDTs were classified as being read as negative by a healthcare worker during the MIS and recorded in the database. A true-negative RDT (TN-RDT) result was defined as a RDT whose negative result collected in the field was confirmed by the PlasQ RT-qPCR. A FN-RDT result was defined as negatively read by a healthcare worker in the field with a positive PlasQ RT-qPCR result based on two replicate amplifications with varATS and/or 18S rDNA Cq < 45 and the human *rnasep* Cq < 35.

qHRP2/3-del assay for detection of *pfhrp2* and *pfhrp3* deletions

The previously published qHRP2/3-del assay that simultaneously amplifies the *pfhrp2* and *pfhrp3* genes together with the internal control gene *pfrnr2e2* was adapted to accommodate for the lower input of NA [31]. Briefly, the probe for the internal control gene *pfrnr2e2* was labelled with fluorescein (FAM) instead of Cy5 to improve its

detectability. Additionally, the final concentration of all primers was increased from 0.3 μ M to 0.45 μ M. Concentrations of 0.15 μ M were used for the *pfrnr2e2* probe, and 0.225 μ M for the *pfhrp2* and *pfhrp3* probes each. All samples were run in triplicates and the number of amplification cycles was increased from 45 to 50. Every 96-well qPCR plate contained control DNA extracted from a known *pfhrp2*-deleted *P. falciparum* strain (Dd2), a *pfhrp3*-deleted *P. falciparum* strain (HB3), and a *P. falciparum* strain without *pfhrp2* and *pfhrp3* gene deletions (NF54) as well as a non-template control (NTC). Successful amplification was defined as a mean Cq < 40 for *pfrnr2e2* calculated from at least two replicates for each sample. The qHRP2/3-del assay only was run with NA extracted from RDTs that had displayed a Cq < 35 for the *varATS* target in the PlasQ RT-qPCR.

Pfrnr2e2, *pfhrp2* and *pfhrp3* are all single-copy genes and they show comparable performances in the multiplex qPCR assay [31]. One approach to detect *P. falciparum* strains with *pfhrp2* and/or *pfhrp3* gene deletions in mixed *P. falciparum* strain infections (herein defined as masked gene deletions), is to calculate the difference in Cq values obtained between *pfhrp2* or *pfhrp3* and *pfrnr2e2* amplifications (Δ Cq values). This is done by subtracting the Cq value obtained during the amplification of *pfrnr2e2* from the Cq value of *pfhrp2* or *pfhrp3*, respectively. Combining all runs that were conducted, the mean Δ Cq for *pfhrp2* in controls (NF54 and HB3) was 0.00 (SD \pm 0.52) and for *pfhrp3* the mean Δ Cq in controls (NF54 and Dd2) was 1.19 (SD \pm 0.83). For *pfhrp2* the Δ Cq cut-off value of 2.0 determined by Schindler et al. [31] was used to identify masked gene deletions. For *pfhrp3* a Δ Cq cut-off value of 4.0 was chosen to identify masked gene deletions due to the shift in mean Δ Cq in the controls.

Genotyping of *Plasmodium falciparum* *pfmsp1* and *pfmsp2* genes

Genotyping with *pfmsp1* and *pfmsp2* was performed following published procedures using nested PCR [32]. The first two PCR reactions amplify conserved sequences within the polymorphic regions of *pfmsp1* and *pfmsp2*, respectively. The second, nested PCR targets allele-specific sequences in five separate reactions. Samples were run in 20- μ L total volume with 1 \times Hot Firepol Master Mix (Solys BioDyne, Estonia), 0.25 μ M of forward and reverse primers and 2- μ L template DNA. The cycling conditions for the first PCR were 95 $^{\circ}$ C for 12 min, 25 cycles of 95 $^{\circ}$ C for 30 s, 58 $^{\circ}$ C for 1 min and 72 $^{\circ}$ C for 2 min and 72 $^{\circ}$ C for 10 min. For the second PCR, the cycling conditions for the three allele-specific *pfmsp1* primer pairs were 95 $^{\circ}$ C for 12 min, 35 cycles of 95 $^{\circ}$ C for 30 s, 56 $^{\circ}$ C for 40 s and 72 $^{\circ}$ C for 40 s and 72 $^{\circ}$ C for 10 min. For the two *pfmsp2* allele-specific reactions the conditions were: 95 $^{\circ}$ C for 12 min, 35 cycles of

95 °C for 30 s, 58 °C for 40 s and 72 °C for 40 s and 72 °C for 10 min. Presence and size of PCR products was determined and documented visually on a 1% agarose gel with a 100 bp DNA ladder.

Genotyping of *Plasmodium malariae* circumsporozoite protein (pmcsp)

The *pmcsp* gene was amplified by semi-nested PCR for all samples with a positive signal for *Plasmodium malariae* in the non-*falciparum* malaria species identification assay [26]. The first PCR was run with 3 µL of DNA template in a reaction volume of 20 µL. The reaction mix contained 1 × Hot Firepol Master Mix and 0.25 µM of each of the primers *csp_OF* [33] and *csp-R* [34]. The conditions for the first PCR were: 95 °C for 12 min; 35 cycles of 95 °C for 15 s, 53 °C for 30 s and 65 °C for 90 s and final elongation at 65 °C for 10 min. The second, semi-nested PCR used 1.5 µL of the product from the first reaction in a total volume of 15 µL. The reaction mix contained 1 × Hot Firepol Master Mix and 0.33 µM of the primers *csp_IF* [33] and *csp-R*. The conditions for the second PCR were: 95 °C for 12 min; 35 cycles of 95 °C for 15 s, 52 °C for 30 s and 62 °C for 90 s and final elongation at 62 °C for 10 min. The PCR product was sent to Microsynth (Microsynth AG, Switzerland) for bidirectional sanger sequencing. The 15 sequences of *P. malariae* circumsporozoite protein from Bioko Island have been deposited into GenBank under the Accession Numbers MW963324–MW963338.

Data analysis and statistics

The generated (RT)-qPCR data was initially analysed with the CFX Maestro Software (Bio-Rad Laboratories, CA, USA). Thresholds for each fluorescence channel were set manually and Cq values were then uploaded to the ELIMU-MDx platform for data storage and analysis [23]. Sequence analysis was performed using Geneious Prime 2019.1.1 (<https://www.geneious.com>). Statistical analysis and data visualization was performed using the R statistical language (version 4.0.3) based on packages *data.table*, *dplyr*, *epiDisplay*, *epitools*, *ggplot2*, *ggpubr*, *ggridges*, *gridExtra*, *lme4*, *readxl*, *reshape2*, *scales*, *stringr*, *tidyr*, *tidyverse*. Wilcoxon rank sum test was used for numeric values. Fisher's exact test (two-sided) was used for contingency tables. A generalized linear mixed-effects model with fixed and random effects was used for calculation of odds ratios and their confidence intervals.

Results

Integration of molecular diagnostic methods into the national malaria control programme to assess the performance of malaria RDTs

In total, 2865 RDTs (21.2%) collected during the 2018 MIS were included in this study. The median age of

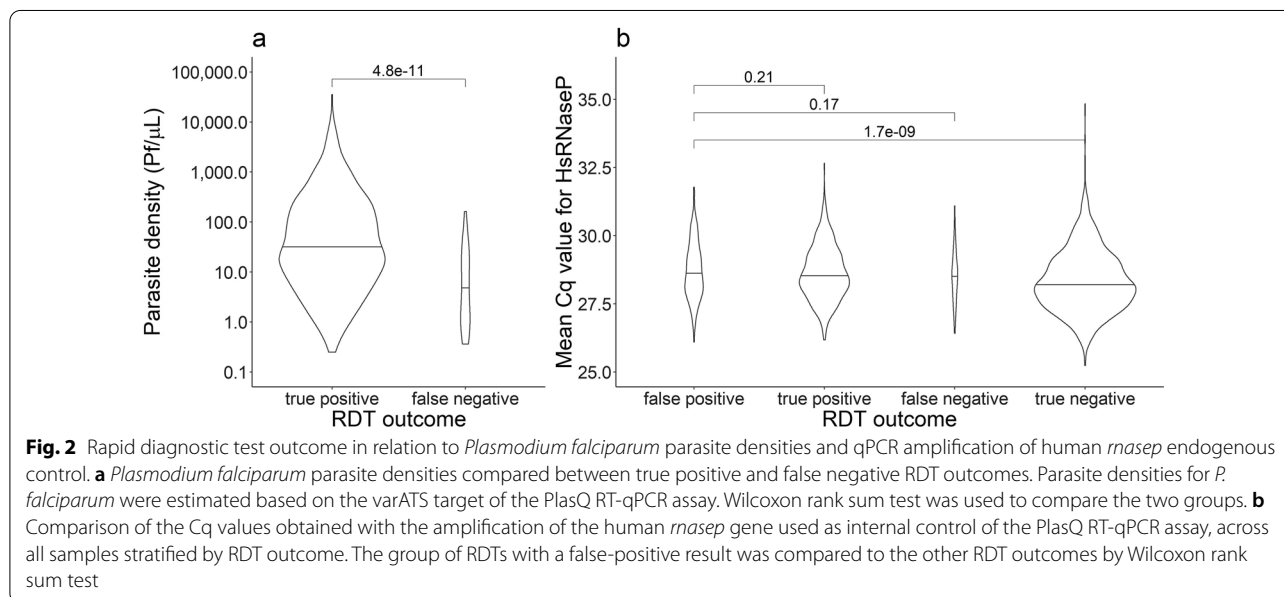
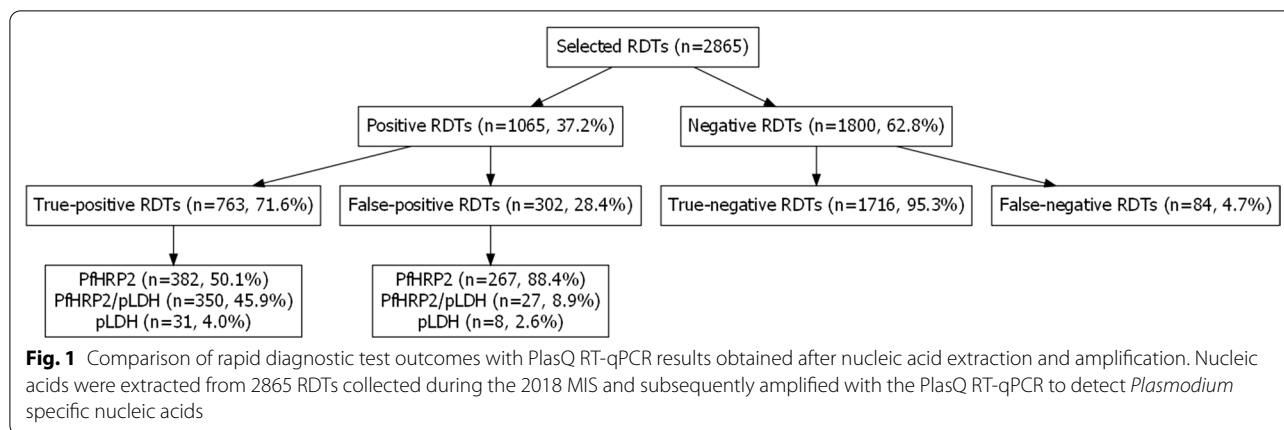
volunteers included in this sample collection was 22 years (interquartile range 9 to 38 years), female participants were over-represented (58.2%), and 97.8% of the participants were asymptomatic, non-febrile individuals. Of the 507 (17.7%) participants that reported to have been sick in the 2 weeks preceding the survey, 81.5% (413/507) had had fever. Other common symptoms were headache in 34.1% (173/507), followed by articular and bone pain in 21.3% (108/507), pallor and weakness in 13.0% (66/507), vomiting in 11.9% (60/507), and shaking chills in 7.5% (38/507). Fever was accompanied by other symptoms in 87.2% (360/413) of those who reported to have had fever. More than two-thirds of the RDTs were collected in the urban areas of the capital city Malabo on Bioko Island.

Following NA extraction, a PlasQ RT-qPCR result was generated for 1800 malaria negative and 1065 positive RDTs, as recorded in the MIS database. By comparison between PlasQ RT-qPCR results with the RDT results collected in the field, RDTs were grouped into four categories, namely true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN), respectively (Fig. 1). The PlasQ RT-qPCR was used as a gold standard to evaluate the performance of the RDT, and this resulted in an overall sensitivity of 90.0% and specificity of 85.0% of field-deployed RDTs.

When stratified by type of antigen, RDTs classified as FP-RDTs were predominantly those that detected the PfHRP2 antigen (88.4%); whereas 8.9 and 2.6% of the FP-RDTs were those that detected both PfHRP2 plus pLDH antigens or the pLDH antigen only, respectively. Around half of RDTs classified as TP-RDTs were those that detected the PfHRP2 antigen only (50.1%), followed by those that detected PfHRP2 plus pLDH antigens (45.9%) and lastly, those that detected the pLDH antigen only (4.0%).

Low parasite density infections are likely to cause FP-RDT results in the field

The ENAR approach used in this study detects 10–100 times lower asexual blood stage parasite densities than the PfHRP2-based RDT [26]. The data confirms that a clear association exists between FN-RDT, TP-RDT and *P. falciparum* parasite densities assessed by the PlasQ RT-qPCR outcome. TP-RDT had higher geometric mean parasite densities (35.0 *P. falciparum*/µL, IQR: 7.2–166.0) compared to FN-RDTs (4.6 *P. falciparum*/µL, IQR: 1.1–20.0) (Fig. 2a, Wilcoxon rank sum test, $p < 0.001$). Although *P. falciparum* was the most common (93.8%) *Plasmodium* spp. species among RT-qPCR positive RDTs, *P. malariae* (4.0%) and *Plasmodium ovale* spp. (1.1%) were also identified. Co-infections between *Plasmodium* spp. species were included in these prevalence calculations. In 3.2% (27/847) of *Plasmodium*



spp.-positive samples, no species could be assigned, possibly due to low parasite density and the generally lower sensitivity of the species-specific qPCR assays. No *Plasmodium vivax* and *Plasmodium knowlesi* parasite NAs were detected. The central repeat region of the *P. malariae* circumsporozoite protein (*pmcsp*) was amplified by PCR and Sanger sequenced to reconfirm the presence of *P. malariae* derived NA (Additional file 1: Fig. S1b). Nucleotide sequences were unique among all the 15 *P. malariae* PCR fragments sequenced and also the number of NAAG and NDAG repeats varied between these isolates indicating high diversity of the local *P. malariae* population. *Plasmodium malariae* was found among 6.6% of FN-RDTs compared to 3.8% among TP-RDTs. Similarly, *P. ovale* spp. was more prevalent in FN-RDTs (2.6%) than in TP-RDTs (0.9%).

To exclude the possibility that FP-RDTs are the consequence of failed amplification related to the degradation of NA retained on the RDTs, an additional analysis was carried out. During the PlasQ RT-qPCR, the human *rnasep* gene was used as an internal control to monitor the amount of NA extracted from each RDT. On average, the human *rnasep* was amplified with a Cq value of 28.5 (SD ± 1.0). There was no significant difference in the Cq values of the human *rnasep* gene amplification among RDTs, which were categorized as FP (28.6, SD ± 1.0), TP (28.5, SD ± 1.0), or FN (28.6, SD ± 1.0). TN-RDTs had a significantly lower median Cq value (28.2, SD ± 1.1) (Fig. 2b). These results indicate that the lack of detectable *P. falciparum* NA in the blood retained on FP-RDTs is not related to poor NA extraction performance or a failure in detecting NAs.

FN-RDT results are not associated with parasites carrying *pfhrp2* and *pfhrp3* gene deletions

Plasmodium falciparum strains were genotyped to identify strains with *pfhrp2* and/or *pfhrp3* gene deletions. The number of samples available was limited based on the combination of low parasite density infections and the limited amount of blood retained on RDTs as a source of NA. The single copy gene *pfmr2e2*, serving as the internal control of the qHRP2/3-del assay, was amplified with $Cq < 40$ in 184/406 (45.3%) samples. To avoid false reporting of *pfhrp2* and/or *pfhrp3* gene deletions, the analysis was restricted to samples that had an additional successful amplification in either *pfmsp1* (32/47, 68.1%) or *pfmsp2* (31/47, 66.0%). No amplification in *pfmsp1* or *pfmsp2* was observed in 23.4% (11/47) of samples. Based on the available data from the 27 samples with successful *pfmsp1* and *pfmsp2* genotyping (Additional file 1: Fig. S1a), polyclonal infections consisting of two or more distinct *P. falciparum* clones were found in 63.0% (17/27) of samples. Association between parasite density and amplification of each of the three distinct reference genes (*pfmr2e2*, *pfmsp1* or *pfmsp2*) is shown in Fig. 3a–c. At least two out of three reference genes were amplified in 36 samples, which were then included in the analysis of the *pfhrp2* and *pfhrp3* deletion status. No evidence for parasites carrying a *pfhrp2* gene deletion was found in these 36 samples, but 4 out of 36 samples (11.1%) were likely to carry *pfhrp3* gene deletions. All 4 samples with *pfhrp3* deletion were recorded as positive for PfHRP2 by RDT.

The qHRP2/3-del assay was used to identify *pfhrp2* and/or *pfhrp3* gene deletions in polyclonal *P. falciparum* infections by calculating the ΔCq values as the difference of Cq values between *pfhrp2* and *pfhrp3* gene amplification and the *pfmr2e2* internal control. Figure 3d shows the distribution of samples with their respective ΔCq values for *pfhrp2* and *pfhrp3*. Of the 36 samples included, 2 samples (5.6%) had increased ΔCq values for both genes, 2 samples (5.6%) only for the *pfhrp2* gene and 2 samples (5.6%) only for the *pfhrp3* gene, respectively. Importantly, all 36 samples, which were screened for *pfhrp2* and *pfhrp3* gene deletions, were positive for PfHRP2 by RDT. Three out of 6 samples with increased ΔCq values for *pfhrp2* and/or *pfhrp3* were successfully genotyped with *pfmsp1* and *pfmsp2*. Two genotypes were found in one sample with increased ΔCq value for *pfhrp2* and *pfhrp3* each and a single genotype in one sample with increased ΔCq value for *pfhrp3*.

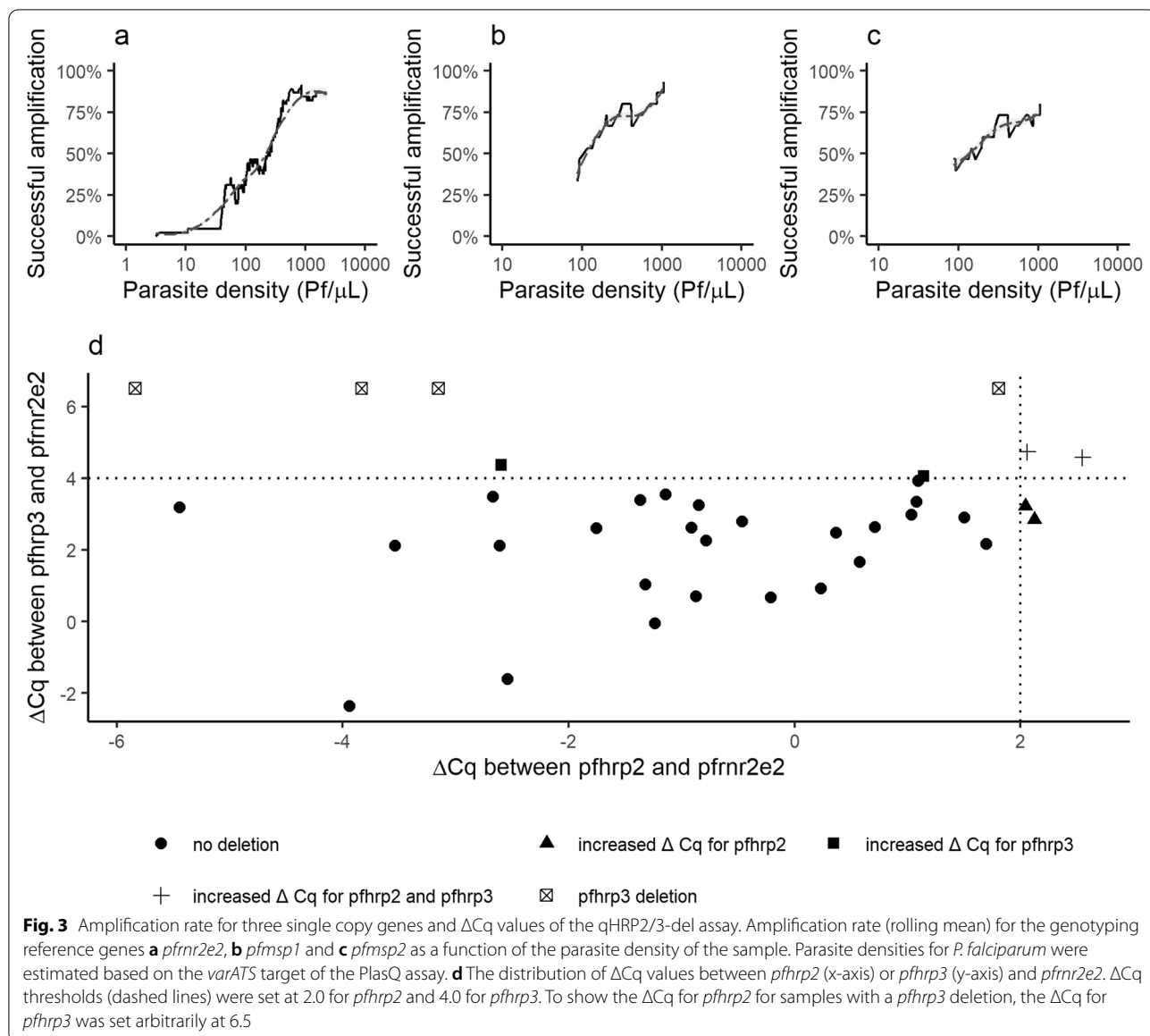
FP-RDT results are associated with recent use of anti-malarial drugs

The rate of FP-RDTs differed across age, level of anaemia, geographical location of residence, and the SES

(Additional file 1: Fig. S2). Interestingly, no study participant with a FP-RDT had a fever (> 37.5 °C) at the time of survey, while 1.6% (12/754) of those with TP-RDTs were recorded with fever. Eight variables collected during the MIS were used to identify risk factors associated with FP-RDTs through multivariate logistic regression analysis in which the outcome of the test was set as the outcome variable (Additional file 1: Table S1). FP-RDTs ($n = 297$) were compared to TP-RDTs ($n = 754$). Because sample collection was clustered within communities, community affiliation was introduced as a random effect to the model. The MIS included 299 communities, of which 201 (67.2%) were represented in the dataset. The median number of samples from a community was 3. Survey participants belonging to higher socio-economic classes (aOR 1.51 $p = 0.01$) had increased odds of having a FP-RDT. Participants who were reported to have been treated with an anti-malarial drug 2 weeks preceding the survey had more than four times the odds of a FP-RDT result than a TP-RDT (aOR 4.52, $p < 0.001$). Noteworthy, 46.6% (136/292) of the participants who had received an anti-malarial treatment in the 2 weeks preceding the survey did recall what drug they had been treated with. The majority of MIS participants (80.9%, 110/136), who reported to have received recent anti-malarial treatment, mentioned that they had received artemisinin derivatives or ACT. Due to the small number of MIS participants treated with non-ACT anti-malarials, the variety of anti-malarials used within this group, and the fact that this information is self-reported, it was decided not to include any further analysis, including a breakdown into individual drugs. In contrast, moderate to severe anaemia reduced the odds of having a FP-RDT (aOR 0.60, $p = 0.02$). Those who reside in the rural Bioko Sur Province had also decreased odds of having a FP-RDT (aOR 0.44, $p = 0.01$). Age, gender, bed-net use, and reported sickness in the 2 weeks preceding the survey were not significantly associated with FP-RDTs (Fig. 4).

The impact of asymptomatic malaria infections on anaemia status might be underestimated by FP-RDT results

It was hypothesized that high rates of FP-RDTs are likely to result in underestimating the impact of asymptomatic malaria infections on the anaemia status. Among malaria-infected children aged < 5 years, the prevalence of anaemia was 67.7% if malaria status was assessed by RDT. Proportion of anaemic children with FP-RDT result (48.9%) is similar to children with TN-RDT result (41.4%) ($p = 0.85$, Fisher exact test), whereas children with a TP-RDT result are more likely to suffer from anaemia (78.3%) ($p = 0.0005$, Fisher exact test) (Fig. 5). This significant effect is even more pronounced among

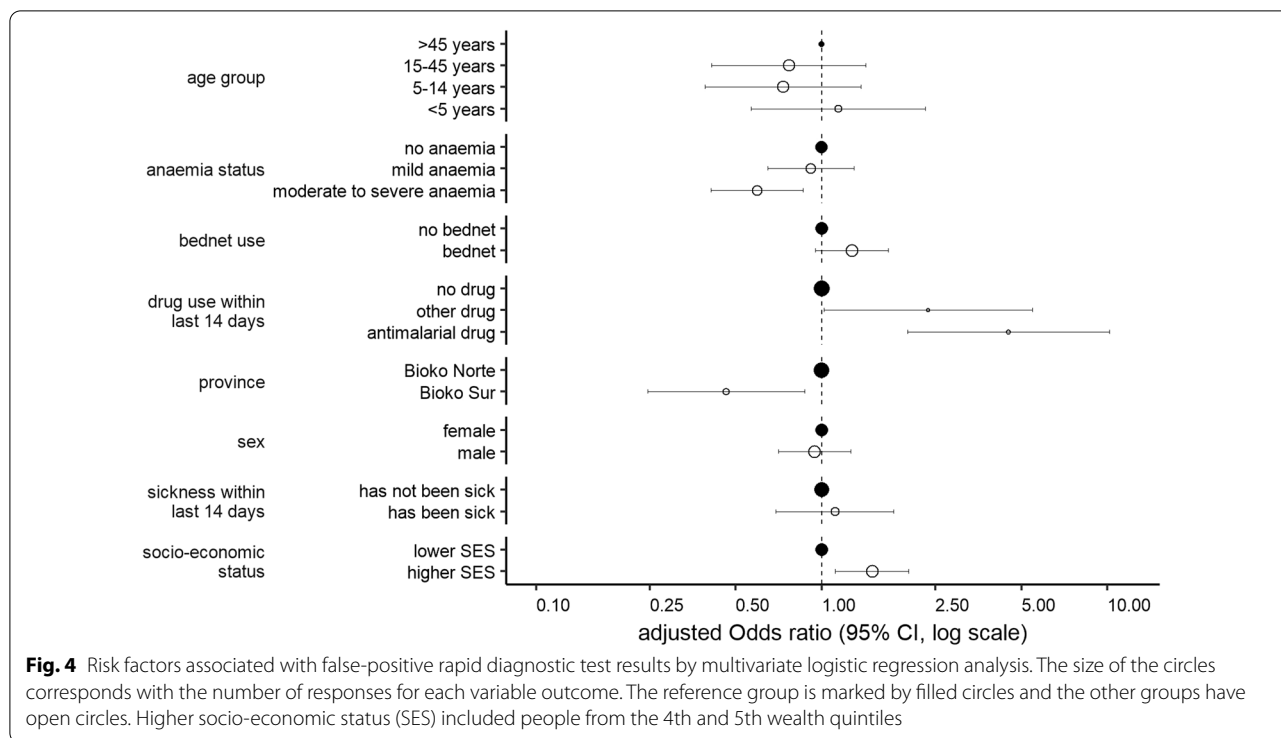


children < 5 years with moderate and severe anaemia if compared to mild anaemia. Removing all FP-RDTs in this association between malaria infection status and anaemia levels in children < 5 years reveals that the association between asymptomatic malaria with moderate or severe anaemia might be even stronger. In older children and adults, the impact of FP-RDTs on assessing the anaemia status is negligible.

Discussion

Malaria control programmes rely on continuous and systematic collection of surveillance data for decision making and resource allocation [35]. A critical measure

that closely reflects malaria transmission intensity is the parasite rate, which is the proportion of the population found to carry parasites in their blood [36]. RDTs, more specifically PfHRP2-based RDTs, are the most widely used test to measure parasite rates in endemic countries and are a cornerstone of malaria control. However, diagnostic performance issues of PfHRP2-based RDTs were identified to be particularly related to limited specificity. Therefore, malaria surveillance depending solely on RDTs might profit from well-integrated quality control procedures assessing the potential impact of reduced sensitivity and specificity of the RDT used. Presented in this report is an efficient approach to assess the performance of field-deployed RDTs used for



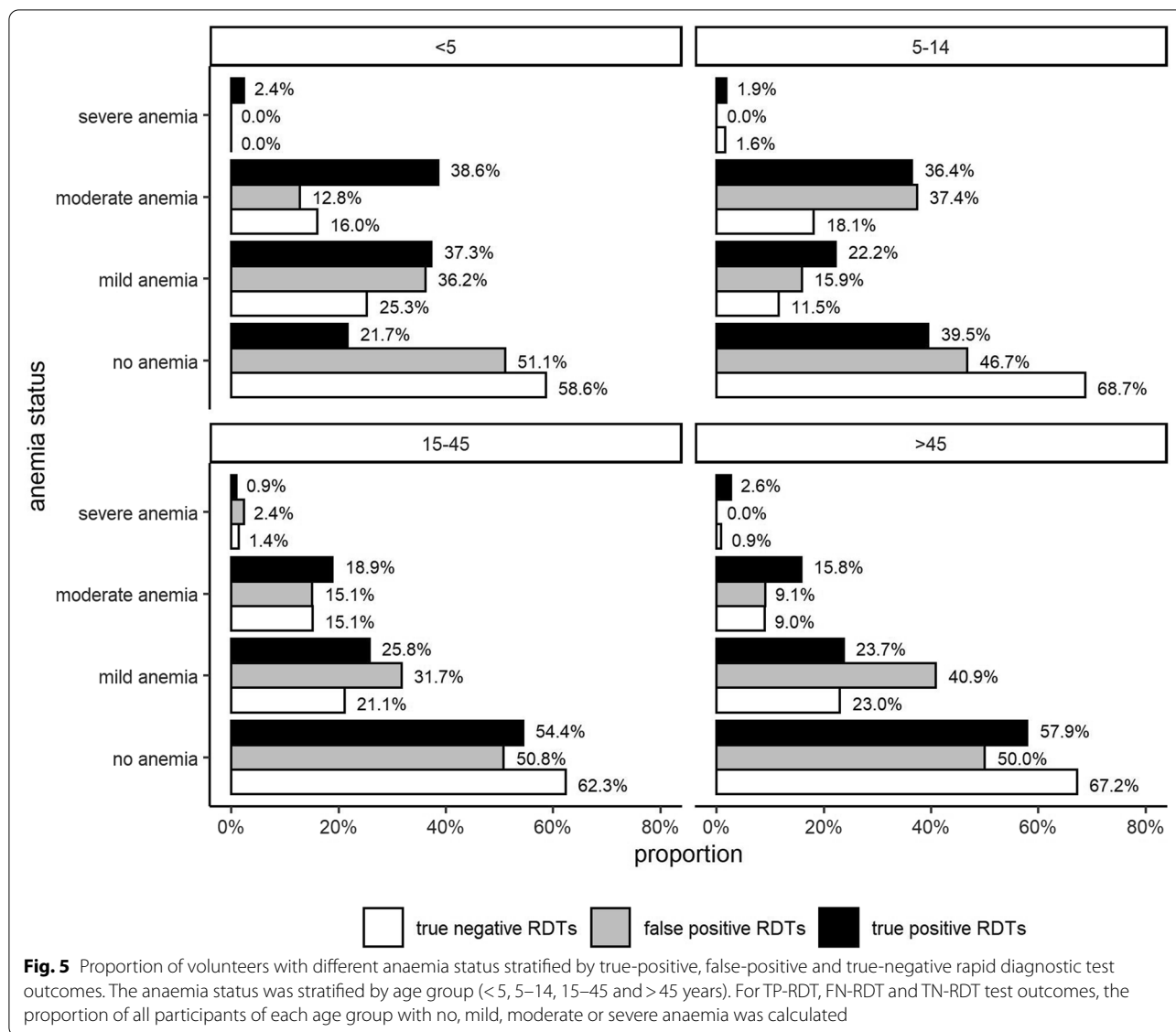
malaria surveillance based on NA extraction from the RDTs followed by qPCR analyses.

Plasmodium spp. NA was found in 4.7% (84/1800) of the negative RDTs and were classified as FN-RDTs. The low proportion of FN-RDTs can be explained by the low parasite densities in these asymptomatic individuals (geometric mean of 5.4 *P. falciparum*/μL) and the low amount of blood (one drop corresponds to approximately 5 μL) used as starting material for the molecular analysis. This is a certainly one of the major limitations of the approach. In a previous study conducted among asymptomatic blood donors in Malabo, PfHRP2-based RDTs showed a sensitivity of only 23.1% and more than 75% of infections had densities below 100 *P. falciparum*/μL [27]. Therefore, the true proportion of FN-RDTs in a high prevalence setting such as Bioko Island is likely to be higher than reported here.

Plasmodium falciparum isolates were identified with potential *pfhrp3* deletions but not a single isolate with a confirmed *pfhrp2* deletion. Given the overall high frequency of polyclonal *P. falciparum* infections in this setting (63.0% by *pfmsp1*/*pfmsp2* genotyping), it was assumed that if *P. falciparum*-carrying *pfhrp2* deletions exist, then they would be most likely masked by co-infecting *P. falciparum* isolates without *pfhrp2* gene deletions. Of all the samples included for final analysis, 11.1% had an increased ΔCq value for *pfhrp2* and 11.1% for *pfhrp3* amplification, indicating for the first time that

there are likely *P. falciparum* strains circulating on Bioko Island carrying deletions in their *pfhrp2* and/or *pfhrp3* genes. So far, one report described *P. falciparum* strains carrying *pfhrp2* and *pfhrp3* deletions in blood samples collected on the continental region of Equatorial Guinea [37]. Since travel activity between Bioko Island and the mainland of Equatorial Guinea is high, it can be assumed that parasite strains are exchanged frequently between these locations [38]. Most importantly, blood samples with *P. falciparum* clones indicative of masked *pfhrp2* and *pfhrp3* gene deletions were recorded as PfHRP2 positive by RDT. Likely, the co-circulating *P. falciparum* clones compensate for the lack of PfHRP2 expression resulting in RDT-positive testing. In 462 clinical samples from different African countries, 7.4% (34/462) samples carried a *pfhrp2* deletion and 10.6% (49/462) a *pfhrp3* deletion, while masked *pfhrp2* and *pfhrp3* deletions were found in 3.0 and 3.2% of samples, respectively [39].

The data support the notion that in settings where polyclonal *P. falciparum* infections are common assays with the ability to identify masked *pfhrp2* and/or *pfhrp3* gene deletions should be used [40]. Importantly, to avoid false reporting of *pfhrp2* and/or *pfhrp3* gene deletions, a robust and multi-layered approach was used by which only samples with a pre-defined parasite density, successful amplification of the assays' internal control, and additional, independent amplification of either *pfmsp1* or *pfmsp2* genes were included into the analysis.



In this study, a significant proportion of FP-RDTs were discovered. The findings are not unique to Bioko Island. In a study conducted in Tanzania, 22% of malaria-positive RDTs were negative by molecular analysis for *P. falciparum* [41]. A study performed in Guinea-Bissau reported 26% FP-RDTs [42], and in Western Kenya, approximately one-third of positive RDTs were negative by molecular detection methods for *P. falciparum* [43]. With introduction of a novel RDTs labelled as ‘ultra-sensitive’, detecting lower concentrations of the PfHRP2 antigen, the problem of FP-RDT results is expected to become greater, as already shown in a recent study [44].

The wrong positivity of RDTs based on PfHRP2 detection could be associated with recent use of anti-malarial drugs confirming previous reports [23, 45–48]. It has

been well established that anti-malarial treatment leads to FP-RDT results because the PfHRP2 antigen persists in the blood days to weeks after parasite clearance [23, 45–48].

In addition, an association was found between FP-RDTs and potential access to anti-malarial drugs based on higher SES and on living in urban parts of the Island.

The impact of FP-RDTs differs greatly depending on the setting in which RDTs are deployed. In clinical settings, FP-RDTs might be less common, but the consequences are serious since wrong prescription of anti-malarials might increase risk of overlooking other life-threatening diseases causing fever [49]. In cases where RDTs are used for epidemiological surveys, a high proportion of FP-RDTs due to PfHRP2 antigen persistence might lead to

an overestimation of malaria prevalence, particularly in populations with good access to anti-malarial treatment. Using RDT only as test for malaria infection status might underestimate the negative consequences of asymptomatic malaria infections on haemoglobin levels, particularly in children < 5 years of age [50].

The benefits and the challenges that come with large-scale deployment of molecular techniques for malaria surveillance in malaria-endemic regions have been discussed [51]. Alternative and non-molecular approaches such as automated malaria diagnosis using haematology analysers [52] should be further evaluated for malaria surveillance purposes. The ongoing COVID-19 pandemic has raised the awareness of the value of introducing novel methods as surveillance tools in the public health systems in Africa [53]. Building on this experience will potentially accelerate efforts to integrate sensitive and specific tools for continuous, large-scale surveillance of malaria in control programmes.

Conclusion

Malaria surveillance programmes based on RDT assessments of malaria prevalence should be strengthened by the integration of molecular epidemiological data in the same setting. These data will serve as an early warning system for (i) spread of *P. falciparum* strains evading widely used diagnostic tests; (ii) understanding overuse of malaria drugs; (iii) help with identifying fever-causing diseases beyond malaria; and, (iv) help to clarify the burden of asymptomatic malaria as a cause of severe to moderate anaemia, particularly in children < 5 years.

Abbreviations

RDT: Rapid diagnostic test; PfHRP2/3: Histidine rich protein 2/3; ENAR: Extraction of nucleic acids from RDTs; MIS: Malaria indicator survey; RT-qPCR: Reverse transcription quantitative polymerase chain reaction; Cq: Quantification cycle; sSA: Sub-Saharan Africa; ACT: Artemisinin combination therapy; SES: Socio-economic status; NA: Nucleic acid; WHO: World Health Organization; PCA: Principal component analysis; FAM: Fluorescein; NTC: Non-template control.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12936-022-04043-7>.

Additional file 1: Figure S1. Genetic diversity of *Plasmodium falciparum* and *Plasmodium malariae* length polymorphic genes. **Figure S2.** False-positive rapid diagnostic tests as a proportion of all positive rapid diagnostic tests. **Table S1.** Multivariable logistic regression analysis of risk factors associated with false-positive rapid diagnostic tests.

Acknowledgements

The authors would like to thank all MIS participants for their contribution and the BIMCP staff for their commitment and support during sample collection. We would like to thank Christin Gump, Christian Scheurer and Sergio Wittlin from the Swiss TPH Malaria Drug Discovery Group for their help with cultivating PfNF54, PfDD2 and PfHB3 parasites, whose DNA was used as controls for

the qHRP2/3-del assay. We would like to acknowledge Amanda Ross for her support and guidance with the statistical analysis used in this manuscript.

Authors' contributions

Conceptualization: SH, CD, TS. Data curation and validation: SH, TS, OTD. Formal analysis and visualization: SH. Funding acquisition: CD, MT, GAG, WPP. Investigation: OTD, GAG, WPP, MOA, CAG. Methodology: SH, CAY, EAG, JPD, KB. Resources: MM, EN, OTD, GAG, WPP, CAG. Project administration and supervision: CD, TS. Writing original draft: SH, TS, CD. All authors read and approved the final manuscript.

Funding

This study was funded by a public-private partnership, the Bioko Island Malaria Elimination Project (BIMEP), composed of the Government of Equatorial Guinea, Marathon EG Production Limited, Noble Energy, and Atlantic Methanol Production Company. Etienne A. Guirou and Charlene Aya Yoboue are recipients of Swiss Government Excellence Scholarships (Number 2016.1250 and 2017.0748, respectively) granted by the State Secretariat for Education, Research and Innovation.

The funding sources had no role in the study design, the collection, analysis, and interpretation of data, as well as in writing this manuscript and in the decision to submit the paper for publication.

Availability of data and materials

All data needed to evaluate the conclusions in the paper are present in the manuscript or the Additional files. Further information will be made available to interested researchers.

Declarations

Ethics approval and consent to participate

The Ministry of Health and Social Welfare of Equatorial Guinea and the Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref. No. LSHTM: 5556) approved the 2018 malaria indicator survey. Written informed consent was obtained from all adults and from parents or guardians of children who agreed to participate. Only samples for which an additional consent for molecular analysis was obtained were included in this study. We confirm that all experiments were performed in accordance with relevant national and international guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Swiss Tropical and Public Health Institute, Basel, Switzerland. ²University of Basel, Basel, Switzerland. ³Medical Care Development International, Malabo, Equatorial Guinea. ⁴Ifakara Health Institute, Bagamoyo, United Republic of Tanzania. ⁵Laboratorio de Investigaciones de Bane, Bane, Equatorial Guinea. ⁶Ministry of Health and Social Welfare, Malabo, Equatorial Guinea. ⁷Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire.

Received: 21 June 2021 Accepted: 7 January 2022

Published online: 24 January 2022

References

1. WHO. World malaria report. 20 years of global progress and challenges. Geneva: World Health Organization; 2020. p. 2020.
2. Bosco AB, Nankabirwa JI, Yeka A, Nsobyia S, Gresty K, Anderson K, et al. Limitations of rapid diagnostic tests in malaria surveys in areas with varied transmission intensity in Uganda 2017–2019: implications for selection and use of HRP2 RDTs. *PLoS ONE*. 2021;15: e0244457.
3. Tediosi F, Lengeler C, Castro M, Shretta R, Levin C, Wells T, et al. Malaria control. In: Holmes KK, Bertozzi S, Bloom BR, Jha P, Gelband H, DeMaria LM, Horton S, editors., et al., Major infectious diseases. 3rd ed. Washington

- (DC): The International Bank for Reconstruction and Development/The World Bank; 2017. p. 1–27.
4. WHO. T3: Test. Treat. Track. Scaling up diagnostic testing, treatment and surveillance for malaria. Geneva: World Health Organization; 2012.
 5. WHO. World malaria report 2018. Geneva: World Health Organization; 2018.
 6. Moody A. Rapid diagnostic tests for malaria parasites. *Clin Microbiol Rev.* 2002;15:66–78.
 7. WHO. Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 8 (2016–2018). Geneva: World Health Organization; 2018.
 8. Hofmann NE, Antunes Moniz C, Holzschuh A, Keitel K, Boillat-Blanco N, Kagoro F, et al. Diagnostic performance of conventional and ultrasensitive rapid diagnostic tests for malaria in febrile outpatients in Tanzania. *J Infect Dis.* 2019;219:1490–8.
 9. Moutacho JC, Goldring JPD. Malaria rapid diagnostic tests: challenges and prospects. *J Med Microbiol.* 2013;62:1491–505.
 10. Bousema T, Okell L, Felger I, Drakeley C. Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nat Rev Microbiol.* 2014;12:833–40.
 11. Watson OJ, Sumner KM, Janko M, Goel V, Winskill P, Slater HC, et al. False-negative malaria rapid diagnostic test results and their impact on community-based malaria surveys in sub-Saharan Africa. *BMJ Global Health.* 2019;4: e001582.
 12. Baker J, Ho M-F, Pelecanos A, Gatton M, Chen N, Abdullah S, et al. Global sequence variation in the histidine-rich proteins 2 and 3 of *Plasmodium falciparum*: implications for the performance of malaria rapid diagnostic tests. *Malar J.* 2010;9:129.
 13. Baker J, Gatton ML, Peters J, Ho MF, McCarthy JS, Cheng Q. Transcription and expression of *Plasmodium falciparum* histidine-rich proteins in different stages and strains: implications for rapid diagnostic tests. *PLoS ONE.* 2011;6: e22593.
 14. Gamboa D, Ho M-F, Bendezu J, Torres K, Chiodini PL, Barnwell JW, et al. A large proportion of *P. falciparum* isolates in the Amazon region of Peru lack pfhpr2 and pfhpr3: implications for malaria rapid diagnostic tests. *PLoS ONE.* 2010;5: e8091.
 15. Gendrot M, Fawaz R, Dormoi J, Madamet M, Pradines B. Genetic diversity and deletion of *Plasmodium falciparum* histidine-rich protein 2 and 3: a threat to diagnosis of *P. falciparum* malaria. *Clin Microbiol Infect.* 2019;25:580–5.
 16. Verma AK, Bharti PK, Das A. HRP-2 deletion: a hole in the ship of malaria elimination. *Lancet Infect Dis.* 2018;18:826–7.
 17. Lee J-H, Jang JW, Cho CH, Kim JY, Han ET, Yun SG, Let, et al. False-positive results for rapid diagnostic tests for malaria in patients with rheumatoid factor. *J Clin Microbiol.* 2014;52:3784–7.
 18. Iqbal J, Sher A, Rab A. *Plasmodium falciparum* histidine-rich protein 2-based immunocapture diagnostic assay for malaria: cross-reactivity with rheumatoid factors. *J Clin Microbiol.* 2000;38:1184–6.
 19. Grobusch MP, Alpermann U, Schwenke S, Jelinek T, Warhurst DC. False-positive rapid tests for malaria in patients with rheumatoid factor. *Lancet.* 1999;353:297.
 20. Meatherall B, Preston K, Pillai DR. False positive malaria rapid diagnostic test in returning traveler with typhoid fever. *BMC Infect Dis.* 2014;14:377.
 21. Leshem E, Keller N, Guthman D, Grossman T, Solomon M, Marva E, et al. False-positive *Plasmodium falciparum* histidine-rich protein 2 immunocapture assay results for acute schistosomiasis caused by *Schistosoma mekongi*. *J Clin Microbiol.* 2011;49:2331–2.
 22. Gillet P, Mumba Ngoyi D, Lukuka A, Kande V, Atua B, van Griensven J, et al. False positivity of non-targeted infections in malaria rapid diagnostic tests: the case of human African trypanosomiasis. *PLoS Negl Trop Dis.* 2013;7: e2180.
 23. Dalrymple U, Arambepola R, Gething PW, Cameron E. How long do rapid diagnostic tests remain positive after anti-malarial treatment? *Malar J.* 2018;17:228.
 24. Cook J, Hergott D, Phiri W, Rivas MR, Bradley J, Segura L, et al. Trends in parasite prevalence following 13 years of malaria interventions on Bioko island, Equatorial Guinea: 2004–2016. *Malar J.* 2018;17:62.
 25. WHO. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System. Geneva, World Health Organization, 2011 (WHO/NMH/NHD/MNM/11.1). <http://www.who.int/vmnis/indicators/haemoglobin.pdf>. Accessed 13 Jan 2022.
 26. Guirou EA, Schindler T, Hosch S, Donfack OT, Yoboue CA, Krähenbühl S, et al. Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Sci Rep.* 2020;10:12305.
 27. Schindler T, Robaina T, Sax J, Bieri JR, Mpina M, Gondwe L, et al. Molecular monitoring of the diversity of human pathogenic malaria species in blood donations on Bioko Island, Equatorial Guinea. *Malar J.* 2019;18:9.
 28. 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:e1001788.
 29. Kamau E, Alemayehu S, Feghali KC, Komisar J, Regules J, Cowden J, et al. Measurement of parasitological data by quantitative real-time PCR from controlled human malaria infection trials at the Walter Reed Army Institute of Research. *Malar J.* 2014;13:288.
 30. Kamau E, Alemayehu S, Feghali KC, Saunders D, Ockenhouse CF. Multiplex qPCR for detection and absolute quantification of malaria. *PLoS ONE.* 2013;8:e71539.
 31. Schindler T, Deal AC, Fink M, Guirou E, Moser KA, Mwakasungula SM, et al. A multiplex qPCR approach for detection of pfhpr2 and pfhpr3 gene deletions in multiple strain infections of *Plasmodium falciparum*. *Sci Rep.* 2019;9:13107.
 32. Snounou G, Zhu X, Siripoon N, Jarra W, Thaithong S, Brown KN, et al. Biased distribution of msp1 and msp2 allelic variants in *Plasmodium falciparum* populations in Thailand. *Trans R Soc Trop Med Hyg.* 1999;93:369–74.
 33. Saralamba N, Mayxay M, Newton PN, Smithuis F, Nosten F, Archasukhan L, et al. Genetic polymorphisms in the circumsporozoite protein of *Plasmodium malariae* show a geographical bias. *Malar J.* 2018;17:269.
 34. Tahar R, Ringwald P, Basco LK. Heterogeneity in the circumsporozoite protein gene of *Plasmodium malariae* isolates from sub-Saharan Africa. *Mol Biochem Parasitol.* 1998;92:71–8.
 35. WHO. Malaria surveillance, monitoring and evaluation: a reference manual. Geneva: World Health Organization; 2020.
 36. Smith DL, Guerra CA, Snow RW, Hay SI. Standardizing estimates of the *Plasmodium falciparum* parasite rate. *Malar J.* 2007;6:131.
 37. Berzosa P, González V, Taravillo L, Mayor A, Romay-Barja M, García L, et al. First evidence of the deletion in the pfhpr2 and pfhpr3 genes in *Plasmodium falciparum* from Equatorial Guinea. *Malar J.* 2020;9:99.
 38. Guerra CA, Kang SY, Citron DT, Hergott DEB, Perry M, Smith J, et al. Human mobility patterns and malaria importation on Bioko Island. *Nat Commun.* 2019;10:2332.
 39. Grignard L, Nolder D, Sepúlveda N, Berhane A, Mihreteab S, Kaaya R, et al. A novel multiplex qPCR assay for detection of *Plasmodium falciparum* with histidine-rich protein 2 and 3 (pfhpr2 and pfhpr3) deletions in polyclonal infections. *EBioMedicine.* 2020;55:102757.
 40. Agaba BB, Yeka A, Nsoby S, Arinaitwe E, Nankabirwa J, Opigo J, et al. Systematic review of the status of pfhpr2 and pfhpr3 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.* 2019;18:355.
 41. Ishengoma DS, Lwitiho S, Madebe RA, Nyagonde N, Persson O, Vestergaard LS, et al. Using rapid diagnostic tests as source of malaria parasite DNA for molecular analyses in the era of declining malaria prevalence. *Malar J.* 2011;10:6.
 42. Nag S, Ursing J, Rodrigues A, Crespo M, Krogsgaard C, Lund O, et al. Proof of concept: used malaria rapid diagnostic tests applied for parallel sequencing for surveillance of molecular markers of anti-malarial resistance in Bissau, Guinea-Bissau during 2014–2017. *Malar J.* 2019;8:252.
 43. Robinson A, Busula AO, Muwanguzi JK, Powers SJ, Masiga DK, Bousema T, et al. Molecular quantification of *Plasmodium* parasite density from the blood retained in used RDTs. *Sci Rep.* 2019;9:5107.
 44. Acquah FK, Donu D, Obboh EK, Bredu D, Mawuli B, Amponsah JA, et al. Diagnostic performance of an ultrasensitive HRP2-based malaria rapid diagnostic test kit used in surveys of afebrile people living in Southern Ghana. *Malar J.* 2021;20:125.
 45. Iqbal J, Siddique A, Jameel M, Hira PR. Persistent histidine-rich protein 2, parasite lactate dehydrogenase, and panmalarial antigen reactivity after clearance of *Plasmodium falciparum* mono-infection. *J Clin Microbiol.* 2004;42:4237–41.

46. Markwalter CF, Gibson LE, Mudenda L, Kimmel DW, Mbambara S, Thuma PE, et al. Characterization of *Plasmodium* lactate dehydrogenase and histidine-rich protein 2 clearance patterns via rapid on-bead detection from a single dried blood spot. *Am J Trop Med Hyg.* 2018;98:1389–96.
47. Chinkhumba J, Skarbinski J, Chilima B, Campbell C, Ewing V, San Joaquin M, et al. Comparative field performance and adherence to test results of four malaria rapid diagnostic tests among febrile patients more than five years of age in Blantyre, Malawi. *Malar J.* 2010;9:209.
48. Houzé S, Boly MD, Le Bras J, Deloron P, Faucher J-F. Pf HRP2 and Pf LDH antigen detection for monitoring the efficacy of artemisinin-based combination therapy (ACT) in the treatment of uncomplicated falciparum malaria. *Malar J.* 2009;8:211.
49. Sansom C. Overprescribing of antimalarials. *Lancet Infect Dis.* 2009;9:596.
50. Lufungulo Bahati Y, Delanghe J, Bisimwa Balaluka G, Sadiki Kishabongo A, Philippe J. Asymptomatic submicroscopic *Plasmodium* infection is highly prevalent and is associated with anemia in children younger than 5 years in South Kivu/Democratic Republic of Congo. *Am J Trop Med Hyg.* 2020;102:1048–55.
51. Nsanzabana C. Strengthening surveillance systems for malaria elimination by integrating molecular and genomic data. *Trop Med Infect Dis.* 2019;4:139.
52. Pillay E, Khodajji S, Bezuidenhout BC, Litshie M, Coetzer TL. Evaluation of automated malaria diagnosis using the Sysmex XN-30 analyser in a clinical setting. *Malar J.* 2019;18:15.
53. Ondoa P, Kebede Y, Loembe MM, Bhiman JN, Tessema SK, Sow A, et al. COVID-19 testing in Africa: lessons learnt. *Lancet Microbe.* 2020;1:e103–4.

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



3.2. Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests


Published in PLOS Neglected Tropical Diseases, 2022

RESEARCH ARTICLE

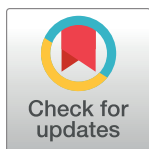
Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests

Charlene Aya Yoboue ^{1,2,3}, Salome Hosch ^{1,2}, Olivier Tresor Donfack⁴, Etienne A. Guirou ^{1,2}, Bonifacio Manguire Nlavo⁵, Mitoha Ondo'o Ayekaba⁶, Carlos Guerra ⁴, Wonder P. Phiri⁴, Guillermo A. Garcia⁴, Tobias Schindler ^{1,2} ^{*}, Claudia A. Daubenger ^{1,2} ^{*}

1 Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland, **2** University of Basel, Basel, Switzerland, **3** Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire, **4** Medical Care Development International, Malabo, Equatorial Guinea, **5** Marathon EG Production Ltd, Malabo, Equatorial Guinea, **6** Ministry of Health and Social Welfare, Malabo, Equatorial Guinea

 These authors contributed equally to this work.

* tobias.schindler@swisstph.ch (TS); claudia.daubenger@swisstph.ch (CAD)



OPEN ACCESS

Citation: Yoboue CA, Hosch S, Donfack OT, Guirou EA, Nlavo BM, Ayekaba MO, et al. (2022) Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests. PLoS Negl Trop Dis 16(1): e0009798. <https://doi.org/10.1371/journal.pntd.0009798>

Editor: Richard Stewart Bradbury, Federation University Australia, AUSTRALIA

Received: September 9, 2021

Accepted: January 10, 2022

Published: January 31, 2022

Copyright: © 2022 Yoboue et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: C.A.Y. and E.G. are recipients of Swiss Government Excellence Scholarships (Numbers 2017.0748 and 2016.1250, respectively) granted by the State Secretariat for Education, Research and Innovation. This study was partially funded by

Abstract

Background

Regular and comprehensive epidemiological surveys of the filarial nematodes *Mansonella perstans* and *Loa loa* in children, adolescents and adults living across Bioko Island, Equatorial Guinea are lacking. We aimed to demonstrate that blood retained on malaria rapid diagnostic tests, commonly deployed for malaria surveys, could be used as a source of nucleic acids for molecular based detection of *M. perstans* and *L. loa*. We wanted to determine the positivity rate and distribution of filarial nematodes across different age groups and geographical areas as well as to understand level of co-infections with malaria in an asymptomatic population.

Methodology

M. perstans, *L. loa* and *Plasmodium* spp. parasites were monitored by qPCR in a cross-sectional study using DNA extracted from a subset malaria rapid diagnostic tests (mRDTs) collected during the annual malaria indicator survey conducted on Bioko Island in 2018.

Principal findings

We identified DNA specific for the two filarial nematodes investigated among 8.2% (263) of the 3214 RDTs screened. Positivity rates of *M. perstans* and *L. loa* were 6.6% and 1.5%,

a public–private partnership, the Bioko Island Malaria Elimination Program (BIMEP), consisting of the Government of Equatorial Guinea, Marathon EG Production Limited, Noble Energy, and Atlantic Methanol Production Company. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

respectively. *M. perstans* infection were more prominent in male (10.5%) compared to female (3.9%) survey participants. *M. perstans* parasite density and positivity rate was higher among older people and the population living in rural areas. The socio-economic status of participants strongly influenced the infection rate with people belonging to the lowest socio-economic quintile more than 3 and 5 times more likely to be *L. loa* and *M. perstans* infected, respectively. No increased risk of being co-infected with *Plasmodium* spp. parasites was observed among the different age groups.

Conclusions/Significance

We found otherwise asymptomatic individuals were infected with *M. perstans* and *L. loa*. Our study demonstrates that employing mRDTs probed with blood for malaria testing represents a promising, future tool to preserve and ship NAs at room temperature to laboratories for molecular, high-throughput diagnosis and genotyping of blood-dwelling nematode filarial infections. Using this approach, asymptomatic populations can be reached and surveyed for infectious diseases beyond malaria.

Author summary

Mansonella perstans and *Loa loa* are filarial nematodes that infect millions of people living in less developed areas, predominantly in sub-Saharan Africa. Both parasites are neglected among other filarial nematodes because both are regarded as causing mainly asymptomatic infections. The aim of this study was to explore the feasibility of using malaria rapid diagnostic tests (mRDTs) deployed during malaria surveys as a convenient sampling strategy for molecular surveillance of blood-dwelling filarial nematode infections. Our findings demonstrate the potential of mRDTs as a source of parasite DNA beyond malaria, providing an opportunity to expand current knowledge on the distribution and populations mostly affected by *M. perstans* and *L. loa* infections to Equatorial Guinea, located in Central-West Africa.

Introduction

Human filariases are vector-borne infectious diseases that encompass Mansonellosis and loiasis [1]. Mansonellosis is caused by three main nematode species, *M. perstans*, *M. streptocerca* and *M. ozzardi* [2]. Recently, an additional species has been described in Gabon named *Mansonella* sp "DEUX" [3]. Mansonellosis is one of the most neglected tropical diseases despite the fact that in large parts of sub-Saharan Africa, as well as in Latin America an estimated 100 million people are infected [2,4]. The life cycles of *Mansonella* spp. generally alternate between an insect vector and humans who are the final hosts. The insect vectors transmitting *M. perstans* belong to the genus *Culicoides* [2]. When feeding on an infected human, female vectors pick-up microfilariae (mf) circulating in the blood. The mf penetrate the insect's gut and undergo developmental stages in the thoracic flight muscle, migrate to the head and proboscis from where *M. perstans* is transferred to humans during the next feeding round [2]. The third-stage infective larvae (L3) actively penetrate the human skin before migrating and maturing into adult worms that can be found in serous body cavities, mainly the peritoneal cavity [2]. Adult male and female worms mate and female worms begin to produce unsheathed mf circulating

in peripheral blood. Unsheathed mf of *M. perstans* are detected by microscopic examination in thick or thin blood films stained with Giemsa in blood samples taken at any time of the day [5,6]. Little is known about the clinical outcome of chronic *M. perstans* infections in endemic populations, and as for other filarial infections most infections seem to be clinically silent [2,5]. Clinical symptoms attributed to *M. perstans* infections include eosinophilia, angioedema, arthralgia, fever, headache, pruritus, skin eruption, serositis, neurologic manifestations, ocular or palpebral pruritus, visual impairment and chest pain [1].

At least 10 million people are infected with *Loa loa* in endemic countries of Central, and Western Sub-Saharan Africa [7]. *L. loa* larvae are transmitted to humans during blood feeding of an infected vector fly belonging to the genus *Chrysops* spp. [1]. The sheathed adult worms live freely in the subcutaneous tissues where they produce thousands of sheathed mf daily, usually with a peak between 10 AM and 4 PM [8]. Diagnosis of *L. loa* can be challenging since adult worms can be present without detectable mf in blood [8].

In areas where different filarial nematodes are co-endemic, misclassification of mf by microscopy can be problematic [8,9]. Therefore, molecular methods have been developed to improve filarial nematode detection [10] and qPCR-based molecular techniques have shown higher sensitivity to detect low parasite density infections, and to discriminate between different filarial nematode species [8]. Large-scale implementation of molecular diagnostic methods for neglected tropical diseases has been regarded as a challenge in the public health systems of low and middle-income countries based on cost, human resource requirement and complexity of supply chain management [11].

The aim of this work was to demonstrate that blood retained on malaria rapid diagnostic tests (mRDTs) is a source of nucleic acids for molecular based detection of *M. perstans* and *L. loa* in otherwise asymptomatic individuals. In doing so, we described the association of filarial infections with demographic and geographic factors and assessed the level of multi-parasitism of these nematodes with the highly endemic malaria parasites on Bioko Island. We also used Cq values as a proxy of filarial parasite density and compared this measurement against demographic and geographic characteristics of the investigated population as well as the time of the day of RDT sample collection.

Material and methods

Ethics statement

The MIS in 2018 was approved by the Ministry of Health and Social Welfare of Equatorial Guinea and the Ethics Committee of the London School of Hygiene & Tropical Medicine (Ref. No. LSHTM: 5556). Consent was sought from eligible respondents before the administration of the questionnaire. A signed authorization was requested from the parents or legal guardians of children, and adult participants to use their samples for further analyses. Laboratory experiments were performed in accordance with relevant guidelines and regulations.

Study area

Bioko Island is located on the West African continent shelf, precisely in the Gulf of Guinea and separated from Cameroon by no more than 32 kilometres of shallow ocean. With its total land surface of approximately 2000 km², Bioko forms part of the nation of Equatorial Guinea and is administratively divided into four districts: Malabo, the capital of Equatorial Guinea, and Baney both located in the northern part of the Island; and Luba and Riaba both located in the Southern part. The Island has an estimated population size of 270'000 people, with the majority (85%) living in Malabo [12]. Bioko has a typical equatorial climate, with high temperatures, high humidity, and heavy rainfall [13].

Malaria has historically been hyper-endemic in Bioko Island with a prevalence of 45% among 2–14 year old children before the launch of the Malaria control program [14,15]. The Bioko Island Malaria Control Program—implemented since 2003—has successfully reduced malaria prevalence and malaria related morbidity and mortality [14,16]. Malaria indicator surveys (MIS) have been conducted annually since 2004 within sentinel sites across the island to enable monitoring and evaluation of this programme [14,17].

Study design

The 2018, the MIS on Bioko Island was conducted on a representative sample using the Malaria Pf/PAN (HRP2/pLDH) Ag Combo rapid diagnostic test (ACCESS BIO, NJ, USA). Consent to store used mRDTs for further molecular analyses was obtained from 13,505 survey participants and unique barcode labels were affixed to the used mRDTs that were shipped to the Swiss Tropical and Public Health Institute for further analyses. Each mRDT barcode was linked to a household unique identification code [18]. This allowed a detailed geographic allocation of the mRDT results. The age distribution of MIS 2018 participants differed slightly between the four districts. Malabo district with median age of 17 years, (IQR 7–30) and is characterized by a slightly younger MIS 2018 population than Riaba (17 years, IQR 7–40), Luba (19 years, IQR 8–45) or Baney (16 years, IQR 7–32). To derive parasite positivity rate estimates, all individuals found positive and all tested in the sample were aggregated at a 1x1 km² grid. Coinfections were estimated at the same spatial resolution, for comparison.

A subset of the mRDT were selected for molecular analysis of the blood retained on the mRDT based on the mRDT test outcome for malaria. Out of the 1376 malaria positive mRDTs identified during the 2018 MIS, we analysed the nucleic acids from 1065 mRDTs (77.4%). In addition, other malaria negative mRDT were selected for the molecular analysis.

Extraction of nucleic acids from used mRDTs

We reused the extracted nucleic acids (NAs) from a study which was published recently by our group [19]. Briefly, the entire and uncut nitrocellulose strip in the used mRDT was carefully removed and incubated in a lysing buffer at 60°C for two hours. After several washing steps, NAs were eluted in a final volume of 75 µL and stored at -20°C. The extracted NAs were amplified and detected by reverse transcription quantitative real time PCR (RT-qPCR) to identify and quantify *Plasmodium* spp. parasites [19]. For the study presented here, we extracted and analysed NAs from additional mRDTs collected during the same MIS and extend the approach by using the extracted NAs to detect the filarial nematodes, *M. perstans* and *L. loa*, by qPCR using well described marker genes [20,21]. We have calculated the median number of days between blood collection onto the mRDT during the MIS 2018 and extraction of NA in our laboratory in Basel to be 253 days (IQR 138–352 days).

Loa loa and *Mansonella perstans* detection by a multiplex qPCR assay

A multiplex qPCR assay, herein referred to as llmp-qPCR, was developed and performed to detect *L. loa* and *M. perstans* DNA. In a multiplex qPCR reaction, the *L. loa* specific qPCR assay amplifies a 62 base pair (bp) fragment from the hypothetical protein LLMF72 [20] and the *M. perstans* specific qPCR assay is based on amplification of a 187 bp fragment of the 18S ribosomal RNA gene [21]. The human rnasep gene (RNaseP) served as an amplification internal control to monitor the successful extraction and amplification of DNA [22]. All reactions were run in duplicates on 96-well plates. Molecular biology grade H₂O was run as a non-template control, and a mix of *M. perstans* and *L. loa* DNA served as a positive control for each run. Each reaction contained 2 µL of template DNA and 8 µL master mix consisting of 1 x

Luna Universal Probe qPCR Master Mix (New England Biolabs, Ipswich), 0.4 μ M of *M. perstans* forward primer, 0.4 μ M of *M. perstans* reverse primer, 0.25 μ M of the Yakima Yellow-labelled *M. perstans* probe, 0.8 μ M of *L. loa* forward primer, 0.8 μ M of *L. loa* reverse primer, 0.4 μ M of FAM-labelled *L. loa*-specific minor groove binder (MGB) probe, 0.4 μ M HsRNaseP forward primer, 0.4 μ M of HsRNaseP-reverse primer and 0.4 μ M of Cy5-labelled HsRNaseP probe. Using the Bio-Rad CFX96 Real-Time PCR System (Bio-Rad Laboratories, California, USA), amplification program was 1min at 95°C, followed by 50 cycles of 15s at 95°C and 45s at 55°C. Samples were considered positive if the quantification cycle (Cq) value was <50.

Co-infection of *Mansonella perstans* and *Loa loa* with *Plasmodium* spp.

The PlasQ is a multiplex RT-qPCR assay for *Plasmodium* spp. and *P. falciparum* detection and quantification that has been developed by our group and described previously [19,22]. This qPCR assay consisted of amplification of two targets combined in a multiplex assay, namely the Pan-*Plasmodium* 18S rDNA sequence (Psp18S) and the *P. falciparum*-specific acidic terminal sequence of the var genes (PfvATs). The human RNaseP (HsRNaseP) gene served as an internal control to assess the quality of DNA extraction and qPCR amplification. The PlasQ was performed on NA extracted from mRDT and samples with Cq value <45 of either of the two targets, PfvATs or Psp18S, were considered positive for *Plasmodium* spp. Then, results obtained were linked to llmp-qPCR results obtained from the same aliquot of NA extracted from identical mRDT to assess co-infection status between *Plasmodium* spp., *L. loa* and *M. perstans*. Coinfections were estimated at the same 1x1 km² grid, for comparison.

Sanger sequencing analysis of *Mansonella perstans* and *Loa loa*

The ribosomal internal transcribed spacer 1 region was amplified with a set of primers that bind universally to all filarial species and are designed to highlight interspecific differences [23]. The PCR products were 484 base pairs (bp) for *M. perstans* and 457 bp for *L. loa*. PCR products of 10 and 23 samples tested positive by llmp-qPCR for *L. loa* and *M. perstans*, respectively, were sequenced from both ends by Sanger Sequencing (Microsynth AG, Balgach, Switzerland) to confirm specificity of the qPCR assays. Samples covering a large and representative range of different Cq values were selected for sanger sequencing of the ribosomal internal transcribed spacer 1 region (S4 Fig). Sequence analysis was realized using Geneious Prime 2019.1.1 (<https://www.geneious.com>). A consensus sequence of all 23 *M. perstans* and 10 *L. loa* sequences of 417 bp and 390 bp length, respectively, served as a query sequence to identify all GenBank entries with >90% identity and >95% coverage using BLAST. Additionally, representative sequences for *M. streptocerca* (KR868771) and *M. ozzardi* (EU272180) deposited to GenBank were included. Geneious Prime software (version 2021.0.3) was used for the multiple sequence alignment and to generate the phylogenetic analysis using its in-build neighbour-joining (NJ) clustering method [24]. Branch lengths were estimated with the Tamura-Nei model [25] with *Onchocerca volvulus* (EU272179) as an outgroup. The resulting newick file was imported into R for final phylogeny and visualization using the ggplot2, ggtree, and treeio packages.

Data analysis

Households were assigned scores based on the type of assets and amenities they own (radio, television, sofa, fan, air-conditioner, car, etc) to derive a surrogate of Socio-Economic Status (SES), using Principal Component Analysis (PCA). Households were ranked based on their score and the distribution was further divided into five equal categories (quintiles), each with approximately 20% of the households. The first quintile corresponded to the lowest wealth

index (WI) and the fifth to the highest WI. The household WI categories were also assigned to permanent (de jure) household members. Predicted co-infection rates for *Plasmodium* spp. and *M. perstans* or *L. loa* were expressed as the product and 95% confidence interval (95% CI) of *Plasmodium* spp., *M. perstans* and *L. loa* the prevalence stratified by age group. The ELIM-U-MDx platform was used for quality control, management and analysis of qPCR data [26]. Statistical analysis and visualization of data were conducted using R version 3.5.1. Univariate analysis (Fisher's exact test and Wilcoxon-Mann-Whitney-Test, as appropriate) was used for comparison between groups. P-value < 0.05 was considered statistically significant.

Results

Sample selection and study population characteristics

A total of 4774 households, including 20'012 individuals, from across Bioko Island participated in the MIS 2018. 13'505 participants provided an additional consent for molecular analysis of the mRDT collected. To increase the probability to identify individuals with filarial nematodes and *Plasmodium* spp. co-infections, we over-sampled mRDTs from two specific sub-populations. Firstly, malaria positive mRDTs were preferentially selected and processed and secondly, for filarial nematodes infections, mRDTs from adults living in rural districts were enriched for selection and analysis. A graphical representation of the over-sampling is shown in [S1 Fig](#). Among the mRDTs selected for NAs extraction, 1065 mRDTs were malaria positive, accounting for 75.8% of all positive mRDTs identified during the 2018 MIS. Significantly higher proportions of adults and people living outside of urban Malabo were included. The subset of mRDTs which were selected to investigate the positivity rates of *M. perstans* and *L. loa* infections stratified by geographical location, age and socio-economic status is shown in [Table 1](#). Noteworthy, from each district or age group at least 10% of the collected mRDTs were included into the analysis. In summary, the majority of the samples included were collected in Malabo (64%). The proportion of mRDTs collected from women was higher compared to men. The mean age was 22 years (SD = 19.7) and participants aged <20 belonged to the most common age group (45.4%). Socio-economic status was higher in participants living in Malabo and Baney compared to the two southern districts (Luba and Riaba). The mean haemoglobin value of all participants was 12.02 g/dl (SD = 1.9) and 99.4% of people did not have fever at time of the sample collection.

Positivity rates of *L. loa* and *M. perstans* among participants of the annual malaria indicator survey

Using the llmp-qPCR assay, detecting simultaneously *M. perstans* and *L. loa* in a single qPCR reaction ([S2 Fig](#)), of the 3214 mRDTs that were tested, 8.2% (263) were positive for *M. perstans* and/or *L. loa*. The proportion of mRDTs positive for *M. perstans* was 6.6% (213) compared to 1.5% (50) for *L. loa*. [Fig 1](#) details the positivity rates of *M. perstans* and *L. loa* stratified by age (A), socio-economic status (B) and gender (C). People living in rural districts have significantly higher positivity rates for *M. perstans* than people living in the urban areas. Positivity rates in rural districts ranged from 17.1% (Luba) to 13.2% (Baney) compared to 2.1% in the urban district of Malabo. On the contrary, no significant differences in *L. loa* positivity rates were observed between rural and urban districts. *L. loa* was more prevalent in the two Southern districts, Riaba (3.9%) and Luba (2.7%), compared to the Northern districts of Malabo (1.5%) and Baney (0.7%). *M. perstans* infection rates in high-endemic rural districts increased significantly with age and the highest positivity rate was observed in participants older than 60 years ($p < 0.00001$). *L. loa* was found at higher rates in participants older than 20 years of age

Table 1. Distribution of population included by age, gender, sociodemographic status and district of residence.

Characteristics	Malabo (n = 2064)	Baney (n = 690)	Luba (n = 257)	Riaba (n = 203)	Total (n = 3214)
Gender					
Women (%)	1261 (61.1)	384 (55.7)	139 (54.1)	100 (49.3)	2086 (55.0)
Men (%)	803 (38.9)	306 (44.3)	118 (45.9)	103 (50.7)	1704 (45.0)
Age (years)					
0–19 (%)	1261 (61.1)	110 (15.9)	42 (16.4)	47 (23.2)	1460 (45.4)
20–39 (%)	576 (27.9)	379 (54.9)	58 (22.7)	46 (22.7)	1059 (33)
40–59 (%)	111 (5.4)	141 (20.4)	81 (31.6)	58 (28.6)	391 (12.2)
≥ 60 (%)	116 (5.6)	60 (8.7)	75 (29.3)	52 (25.6)	303 (9.4)
Socio-economic status (quintile)					
1 lowest	185 (9.0)	153 (22.5)	93 (36.3)	94 (46.3)	525 (16.4)
2 second lowest	348 (16.8)	132 (19.4)	66 (25.8)	53 (26.1)	599 (18.7)
3 middle	473 (22.9)	130 (19.1)	29 (11.3)	32 (15.8)	664 (20.7)
4 second highest	487 (23.6)	119 (17.5)	49 (19.1)	22 (10.8)	677 (21.1)
5 highest	571 (27.7)	145 (21.4)	19 (7.4)	2 (1.0)	737 (23.0)

<https://doi.org/10.1371/journal.pntd.0009798.t001>

in urban as well as in rural areas ($p = 0.0001$) (Fig 1A). Among children below the age of five, a positivity rate of 1.4% (4/296) for *M. perstans* and not a single infection for *L. loa* was observed. In older children and adolescents, positivity rates of 2.4% and 0.7% were estimated for *M. perstans* and *L. loa*, respectively. Infection rates were strongly influenced by the socio-economic status of the individuals (Fig 1B). People from rural district assigned to the lowest SES were three times more likely to harbour an *M. perstans* infection than people belonging to the highest SES. The same was observed for *L. loa* where the positivity rate was also higher among lowest SES compared to highest SES.

While the proportion of *L. loa* infections was comparable between male and female individuals combined for rural and urban districts ($p = 0.4$), *M. perstans* was significantly higher among male (21.1%) compared to the female (9.0%) inhabitants of rural districts ($p < 0.00001$) (Fig 1C).

Filarial nematode species identification by ribosomal internal transcribed spacer 1 region sequence analysis

Specificity of the qPCR-based species identification was confirmed by sequence analysis of the conserved ribosomal internal transcribed spacer 1 region that encompassed 390 to 450 bp depending on the filarial species (Fig 2). To our knowledge, this is the first time that a molecular marker, regularly used for filarial nematode species identification, was amplified and sequenced from DNA extracted from blood retained on mRDTs. Twenty-three samples positive for *M. perstans* and ten samples positive for *L. loa* identified by the lmp-qPCR assay were all confirmed by sequence analysis. The *M. perstans* sequences clustered closely with each other and other *M. perstans* sequences, but are distinct from *Mansonella* spp. DEUX sequences. The *L. loa* sequences generated in this study are closely related to sequences from Central- and West-Africa deposited in GenBank.

Geographical distribution of *M. perstans*, *L. loa* and *Plasmodium* spp. show distinct patterns

On Bioko Island the environmental living conditions for the population differ starkly between the urban centre in Malabo and the rural districts in Baney, Riaba and Luba. Therefore, we

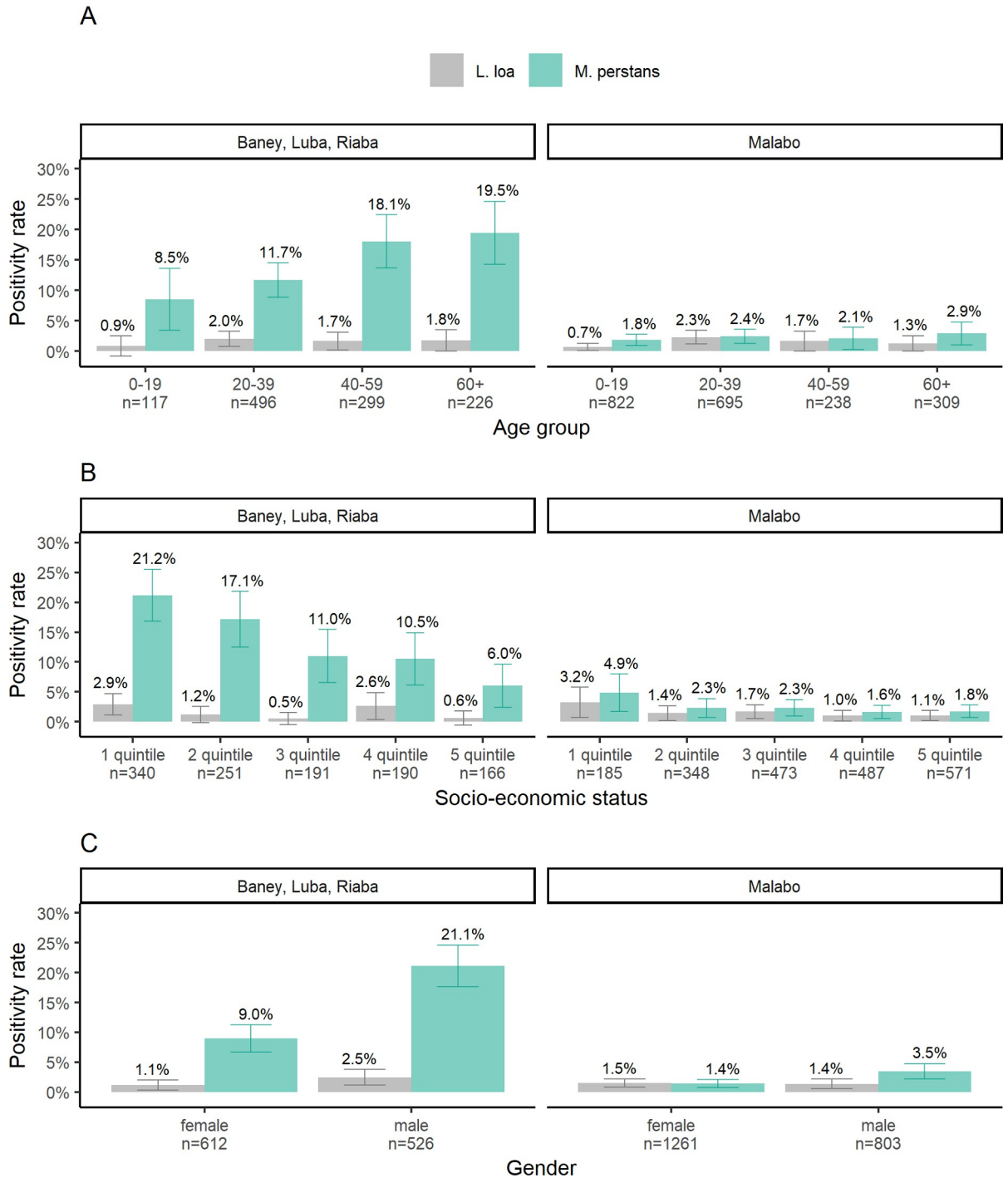


Fig 1. Positivity rates of *M. perstans* and *L. loa* by rural and urban districts. (A) across all age groups investigated, (B) grouped according to socio-economic status and (C) stratified by gender. Positivity rates and 95% confidence intervals were calculated as the proportion of llmp-qPCR positive mRDTs in all tests carried out in each group and are given on top of each bar. Data supporting the Fig 1 are detailed in S2 Table.

<https://doi.org/10.1371/journal.pntd.0009798.g001>

mapped the prevalence and geographical distribution of *P. falciparum* for all samples collected during the MIS based on mRDT positivity in Fig 3A. To investigate the level of co-infections between malaria and filarial parasites in different populations, malaria positive RDTs were

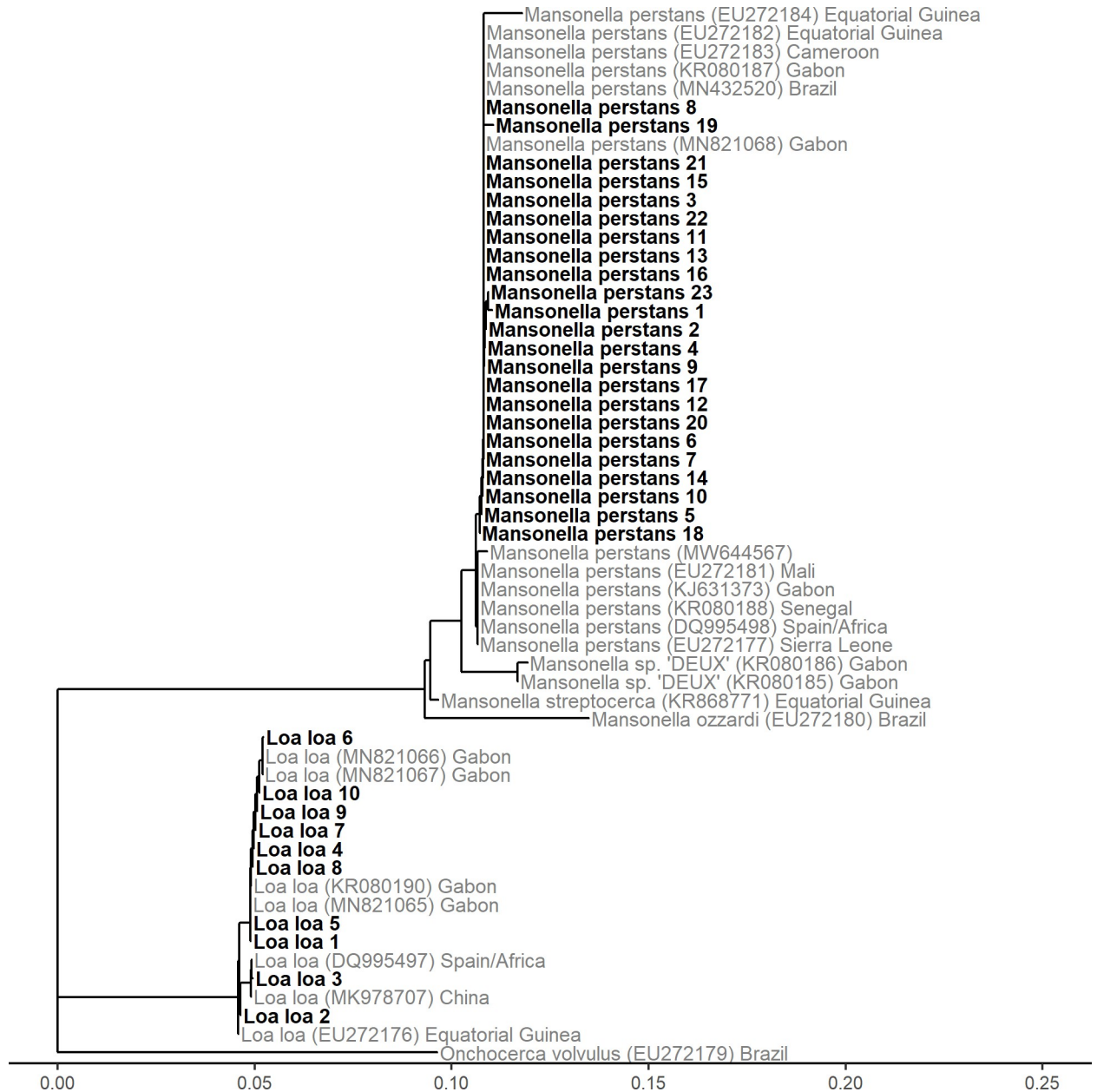


Fig 2. Phylogenetic tree of the ribosomal internal transcribed spacer 1 region of *M. perstans* and *L. loa*. PCR products were amplified from DNA extracted from mRDTs and sequenced. The scale on the x-axis corresponds to the number of substitutions per site (number of changes/sequence length).

<https://doi.org/10.1371/journal.pntd.0009798.g002>

given priority when selected for molecular analysis described here. This intentional enrichment resulted in higher positivity rates of *Plasmodium* spp. when analysed by RT-qPCR (Fig 3B). High positivity rates of *M. perstans* were found in areas along the East coast as well as in the Southern districts; while in urban areas around Malabo, *M. perstans* was found at low rates or was even absent (Fig 3C). *L. loa* positivity rates are generally low and no distinct patterns are seen. Interestingly, no *L. loa* were found in the district of Baney, where positivity rates of *M. perstans* are the highest (Fig 3D).

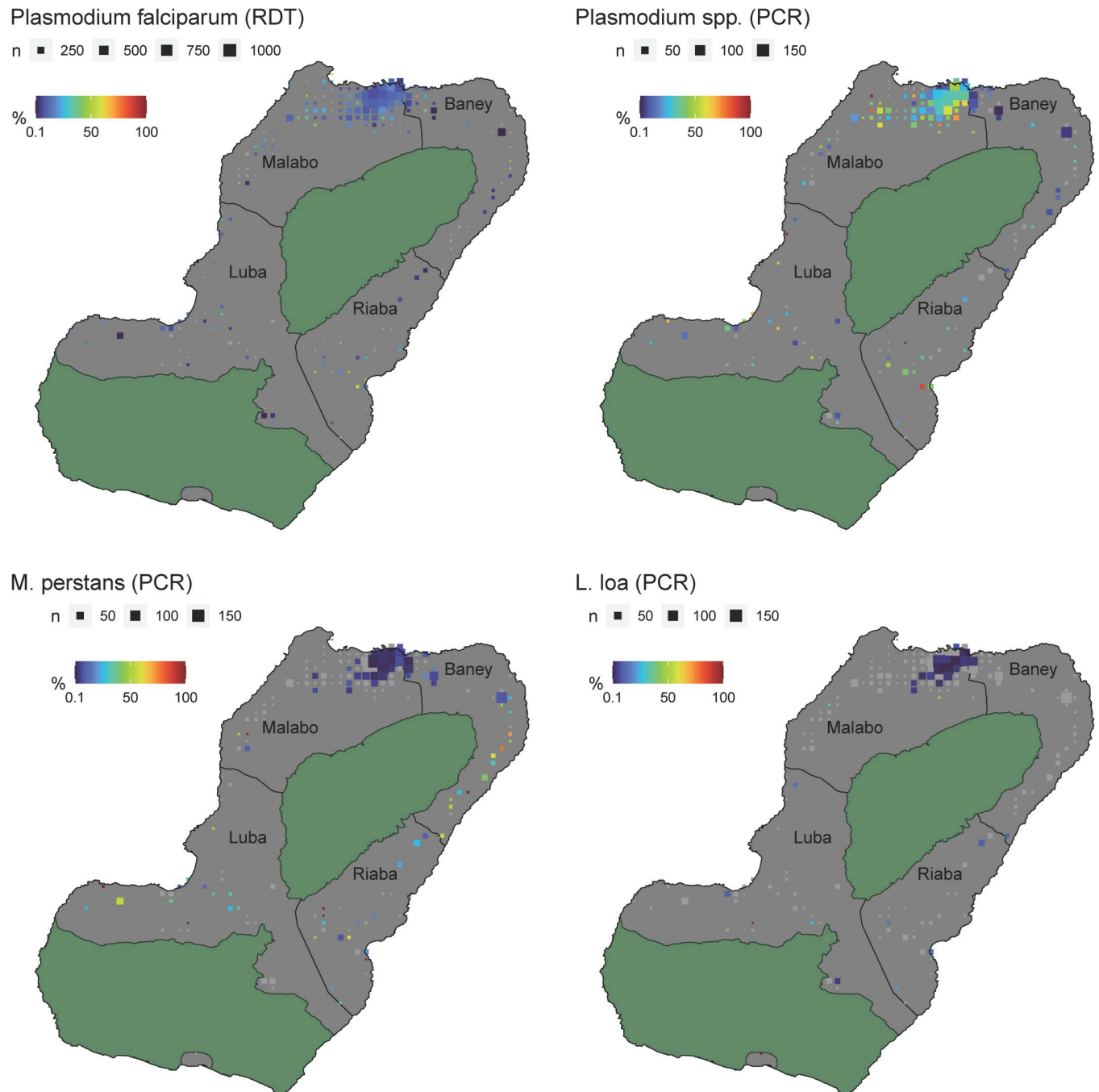


Fig 3. Positivity rate of *Plasmodium* spp., *M. perstans* and *L. loa* across Bioko Island. (A) Prevalence of PfHRP2-positive mRDTs for the entire MIS 2018 population. (B) *Plasmodium* spp. RT-qPCR positivity rate of mRDTs selected for molecular analysis. (C) *M. perstans* qPCR based positivity rate. (D) *L. loa* qPCR based positivity rate. The size of the squares represents the number of people analysed living in the corresponding 1x1 km² grid. The areas marked in green are nature reserves. Greyed out spots on the maps represent settlements from which mRDTs were collected with null positivity rate for *Plasmodium* spp., *M. perstans* and *L. loa*.

<https://doi.org/10.1371/journal.pntd.0009798.g003>

Co-infection of *M. perstans* or *L. loa* with *Plasmodium* spp. parasites

Next, we wanted to estimate the proportion of malaria positive individuals co-infected with filarial nematodes at the molecular level. From the total 3214 mRDTs extracted, we analysed

2775 mRDTs for *Plasmodium* spp., by using the PlasQ assay because of limited availability of NA. Ten out of 2775 mRDTs were positive for both, *Plasmodium* spp., and *L. loa*, 32 were positive for both, *Plasmodium* spp., and *M. perstans*, and six were positive for both *M. perstans* and *L. loa*. Triple infection of *M. perstans*, *L. loa*, and *Plasmodium* spp., was found in three samples (Fig 4A). Then, we analyzed the likelihood that predicted co-infection rates between *Plasmodium* spp. and either *M. perstans* or *L. loa* differed from observed proportions indicative of biological interaction between these infectious diseases as described previously [27]. We did not find evidence for unexpected higher or lower proportions of co-infections in any of the investigated age groups (Fig 4B). Coinfections were mapped with areas positive for either *M. perstans* (Fig 4C) or *L. loa* (Fig 4D). Infections were highlighted according to the presence of people infected with more than one of the investigated parasites. Most coinfections were observed in Malabo, the area with the highest population density.

Age is associated with variation of microfilaria levels in peripheral blood

The llmp-qPCR Cq values for both, *M. perstans* and *L. loa*, were used as approximations of parasite density and compared between infected individuals older and younger than 40 years of age. The median age of all *M. perstans* positive individuals was 40 years. Persons older than 40 years had a significantly lower *M. perstans* Cq values compared to individuals below the age of 40 years (geom. Mean of 34.0 versus 35.1, $p = 0.02$) (Fig 5A). The difference in *M. perstans* infection parasite densities associated with age is not the result of variation in blood volumes analysed or amount of DNA amplified since the Cq values of the human RNaseP gene is similar between the two groups. The same outcome can be observed in Fig 5B. For instance, at a Cq value of 33, the cumulative frequency of individuals older than 40 years of age is 33.6% compared to only 15.5% for younger individuals. Analysing the distribution of the *M. perstans* Cq values reveals a clear shift towards lower Cq values in infected individuals above the age of 40 years (Fig 5C). Combining all these findings directs towards the conclusion that older individuals have a tendency to higher *M. perstans* parasite densities. No significant differences were observed among the individuals infected with *L. loa*. Apart from age (S3A Fig), no significant difference was observed between female and male gender (S3B Fig), while the parasite density of *M. perstans* infections was higher in rural areas comprising districts of Baney, Luba and Riaba compared to the more urban district of Malabo ($p = 0.018$) (S3C Fig). Interestingly, no difference of llmp-qPCR Cq values were seen in samples collected during the morning versus afternoon (S3D Fig).

Discussion

We conducted a larger scale, cross-sectional study of samples including paediatric, adolescent, adult and elderly populations residing in urban and rural regions of Bioko Island. To the best of our knowledge, this report represents the first molecular epidemiological study of *M. perstans* and *L. loa* in Equatorial Guinea that also includes the evaluation of geographical distribution and association with socio-economic status.

Between 1978 and 2020, a total of 20 publications described Mansonellosis, Loasis, Onchocerciasis, and lymphatic filariasis in the context of Equatorial Guinea (S1 Table). Five of these publications described *L. loa* and *M. perstans* infections among Equato-Guineans living in Spain [28–31] or Singapore [32], while three were case reports of foreigners returning from Equatorial Guinea [33–35]. One cross-sectional study investigated the prevalence of *L. loa* and *M. perstans* on Bioko Island using microscopy and qPCR-based detection [36].

We have extended our mRDT-based molecular surveillance tool, originally developed for malaria, to the filarial nematodes *M. perstans* and *L. loa*. The widespread availability and use of

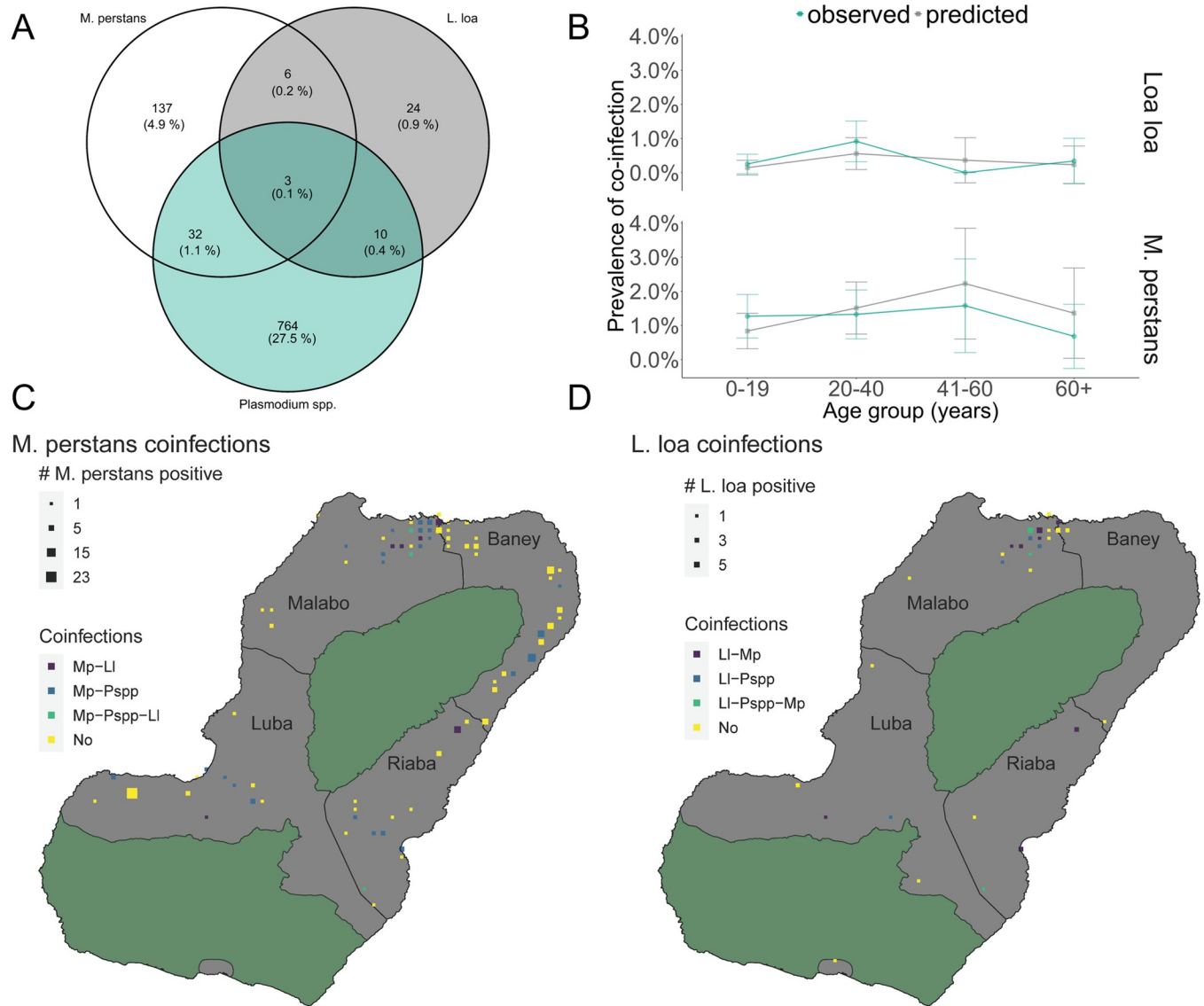


Fig 4. *Plasmodium* spp., *M. perstans* and *L. loa* multi-parasitism infections on Bioko Island. (A) Number of *M. perstans*, *L. loa* and *Plasmodium* spp. co-infections amongst 2775 individuals tested. (B) Positivity rates of *Plasmodium* spp., *M. perstans* and *L. loa* co-infections, stratified by age group. The blue lines and orange lines depict the observed and predicted co-infection rates, respectively. (C) Geographical distribution of *M. perstans* co-infections with malaria and *L. loa*. (D) Geographical distribution of *L. loa* co-infections with malaria and *M. perstans*. In (C) and (D) only 1x1 km² grids with at least one case of *M. perstans* or *L. loa* infection are presented. The areas marked in green are nature reserves.

<https://doi.org/10.1371/journal.pntd.0009798.g004>

mRDTs in malaria endemic regions that are also endemic to *M. perstans* and *L. loa*, the simplicity of mRDT collection and storage, would make this approach convenient for large-scale molecular epidemiological studies covering *Plasmodium* spp., *M. perstans* and *L. loa*. Using our extraction protocol based on mRDTs, high quality and sufficient quantities of *M. perstans* and *L. loa* specific DNA fragments were obtained as demonstrated by successful Sanger sequencing of the ribosomal internal transcribed spacer 1 region. Therefore, apart from amplifying short DNA fragments usually used for qPCR, our NA extraction method also allows to amplify larger fragments suitable for genotyping of the pathogens of interest. In future, switching to more polymorphic markers for genotyping in combination with next generation sequencing technologies might improve tracking of infections and importantly help to

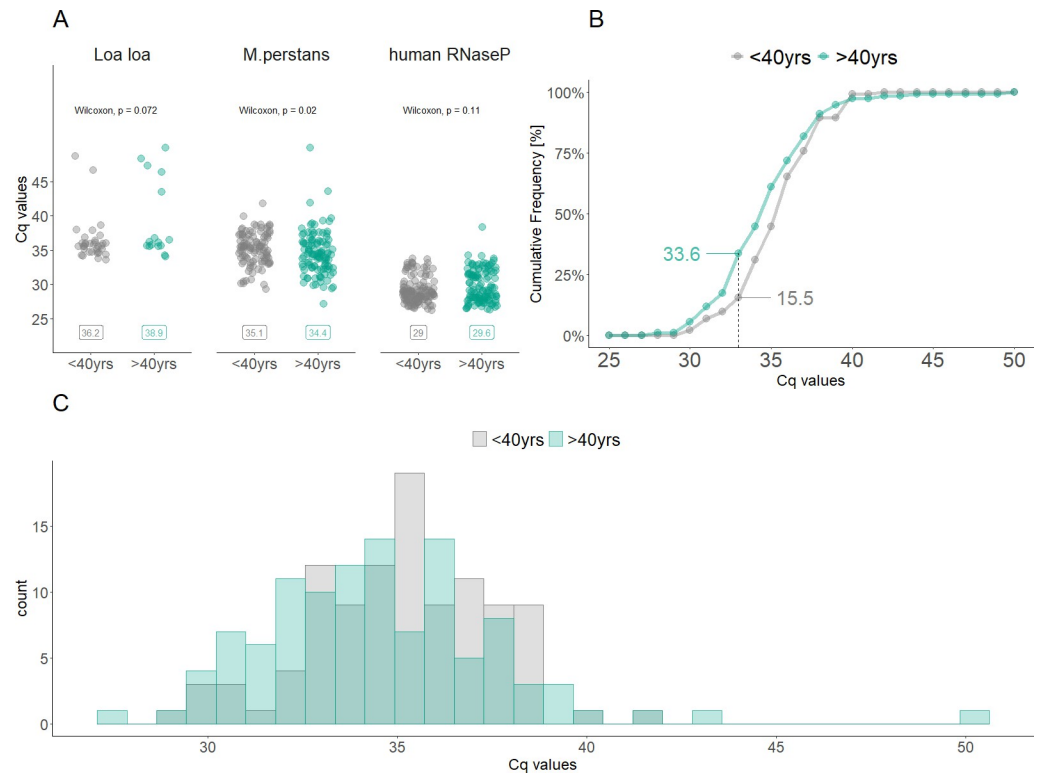


Fig 5. Comparison of Cq values obtained from *M. perstans* and *L. loa* infected individuals aged above and below 40 years. (A) Scatter plots of Cq values for *L. loa*, *M. perstans* qPCR and the corresponding Cq values for the human RNaseP gene qPCR. The geometric mean values for each group are shown. (B) Cumulative frequency of Cq values for *M. perstans* infected individuals. At a Cq value of 33 (dashed line), the cumulative frequency among individuals aged above 40 years is 33.6% compared to 15.5% among individuals aged below 40 years. (C) Histogram of the distribution of Cq values for *M. perstans* infected individuals stratified according to age.

<https://doi.org/10.1371/journal.pntd.0009798.g005>

understand if there are multiple strain infections that possibly accumulate over time in the elderly population showing the highest parasite density of *M. perstans* infections by qPCR.

The cost of our mRDT-based *M. perstans* and *L. loa* test system was estimated to be \$4 per sample, from which \$3 were spent for NA extraction and \$1 on the lmp-qPCR assay. Noteworthy, the same aliquot of extracted NA was used for genomic characterization and quantification of malaria parasites as described [19] making this approach highly cost efficient. The low cost and high scalability of our approach enables systematic monitoring of impact of public health interventions against blood borne pathogens through large scale surveillance.

Here, we found that infections with *M. perstans* (6.6%) are more prevalent than *L. loa* (1.5%) and that *M. perstans* infections can be mostly found in the older, male, population living in rural parts of Bioko Island. This finding reconfirms previous reports [37–39]. We found similar prevalence data for *M. perstans* and *L. loa* compared to a qPCR-based cross-sectional study conducted on Bioko Island in 2018. Ta and colleagues found that 8.8% and 0.7% of persons tested were infected with *M. perstans* and *L. loa*, respectively [36]. The similar proportions found in these two independent studies indicate that detection rates for both filarial nematode species are comparable in spite of different sampling methods (dried blood recovered from mRDT versus freshly collected whole blood) and blood volumes (5 μ L versus 200 μ L) used. The prevalence of *L. loa* found in both studies could be underestimated and partially explained by the fact that 70% of infected individuals do not show mf circulation in peripheral blood, with occult infection or occasional presence of adult worms under conjunctive tissue [40].

The higher parasite density, as expressed by the qPCR's Cq values, of *M. perstans* infection found in rural areas on the East coast of Bioko Island might reflect the distribution or abundance of the vector and its active transmission in those areas. The vector *Culicoides* presence is more likely associated to aquatic environments, banana and plantain stems [9] that might describe the environmental characteristics of rural areas in Bioko Island. Entomological monitoring for the presence of these vectors would be justified to improve the understanding of the geographic patterns observed and inform control interventions. Elderly people above the age of 60 were proportionally the most affected age group. The increased infection rates in combination with higher *M. perstans* parasite densities compared to younger people might be due to the cumulative effect of reinfection during their lifespan [39].

It has been shown that filarial worms including *M. perstans* and *L. loa* cause chronic infections that are associated with strong immune modulation in the human host [41,42]. These long-standing and strong immune modulatory effects particularly of *M. perstans* might negatively impact on mf clearance [37] as well as on co-infections like malaria or tuberculosis outcomes in the same host [43]. In addition, albeit not clinically overt, *M. perstans* infections might impair vaccine induced immune responses and protection by exerting strong immune modulatory effects as described for other helminth infections [44,45]. Therefore, molecular epidemiological studies using the methodology outlined here may prove critical in identifying cofounders of the protective efficacy of experimental malaria vaccine studies currently ongoing in Equatorial Guinea [46,47].

Quantitative measurements of filarial nematodes might become important in the context of development of novel drug interventions against Mansonellosis, loiasis, lymphatic filariasis and onchocerciasis in areas with high co-infections between these parasites [48]. The lack of an international standard with predefined numbers of mf of each of the filariasis causing parasites that could be used for quantitative assessment of microfilaremia based on Cq values measured is one of the tools limiting our approach.

Our study presented had some limitations. We restricted our analysis to *L. loa* and *M. perstans*, both blood-borne pathogens. *Onchocerca volvulus* and *M. streptocerca* were described on Bioko Island [36] but their mf are located in the skin and therefore are detected using primarily skin biopsies for microscopy for molecular analyses. Although Lymphatic filariasis has not been reported on Bioko Island, using mRDTs collected during daylight as a source of the blood sample would not allow exploring the presence of *Wuchereria bancrofti*. Also, we have used primers/probe combinations in our Llmp-qPCR assay that could most likely not amplify the newly identified *Mansonella* sp "DEUX, thereby potentially omitting this novel *Mansonella* species described recently [4]. We conducted a feasibility study to demonstrate that it is possible and sensible to use the mRDTs and metadata collected during an annual MIS to assess at very low additional costs the positivity rate of highly neglected nematode filarial infections for different demographic and socio-economic groups. However, a full analytical and clinical performance evaluation to determine the sensitivity and specificity of our approach based on a direct comparison with microscopy would be needed to fully understand the limitations of our molecular testing for routine surveillance of filarial nematodes in endemic regions.

Conclusion

In summary, our approach of repurposing used mRDT as source of NA provides a promising, future tool that enables a cost-effective approach to monitor the prevalence, genotypes, parasite densities and co-infections of filarial nematodes and potentially other blood borne infectious diseases. Also, PCR amplification and sequencing of DNA fragments allowing for genotyping extends the range of possible applications of using NA stored on mRDTs. This

method might be of particular interest in settings with limited access to cool chains, laboratory infrastructure and in populations not necessarily served by clinics and health posts in rural areas.

Supporting information

S1 Table. Literature review on studies carried out on filarial nematodes in Equatorial Guinea or Equato-Guineans living abroad.

(DOCX)

S2 Table. Positivity rates of *M. perstans* and *L. loa* stratified by gender, age, district and socio-economic status.

(DOCX)

S1 Fig. Selection of mRDTs used for NA extraction and molecular analysis stratified by mRDT result (A), age groups (B) and district (C).

(TIFF)

S2 Fig. Representative amplification plots for the llmp-qPCR assay. (A) Multiplex qPCR amplification of the human RNase P gene, *M. perstans* and *L. loa*. (B) Curves in purple show amplification of the RNaseP gene used as an internal extraction and qPCR amplification control. (C) Curves in yellow show the amplification and detection of *M. perstans*-specific 18S target by qPCR. (D) Curves in blue show the amplification and detection of the *L. loa*-specific LLMF72 target by qPCR.

(TIF)

S3 Fig. Variation of Llmp-qPCR Cq values for *L. loa* (white) and *M. perstans* (grey). Grouped by age (A), gender (B), urban or rural residence (C) and day time of blood collection (D).

(TIFF)

S4 Fig. Graphical representation of samples selected for the Sanger Sequencing experiment of the ribosomal internal transcribed spacer 1 region. All *M. perstans* or *L. loa* positive samples, sorted by their Cq values are shown. Samples selected for Sanger Sequencing are highlighted in red.

(TIFF)

Author Contributions

Conceptualization: Charlene Aya Yoboue, Tobias Schindler, Claudia A. Daubenberger.

Data curation: Charlene Aya Yoboue, Salome Hosch, Tobias Schindler.

Formal analysis: Charlene Aya Yoboue, Salome Hosch, Tobias Schindler.

Funding acquisition: Claudia A. Daubenberger.

Investigation: Charlene Aya Yoboue, Salome Hosch, Olivier Tresor Donfack, Tobias Schindler.

Methodology: Charlene Aya Yoboue, Salome Hosch, Etienne A. Guirou, Tobias Schindler.

Project administration: Tobias Schindler, Claudia A. Daubenberger.

Resources: Bonifacio Manguire Nlavo, Mitoha Ondo'o Ayekaba, Wonder P. Phiri, Guillermo A. Garcia, Claudia A. Daubenberger.

Software: Salome Hosch, Olivier Tresor Donfack.

Supervision: Tobias Schindler, Claudia A. Daubenberger.

Validation: Charlene Aya Yoboue, Salome Hosch, Tobias Schindler.

Visualization: Charlene Aya Yoboue, Salome Hosch, Carlos Guerra, Tobias Schindler.

Writing – original draft: Charlene Aya Yoboue, Tobias Schindler, Claudia A. Daubenberger.

Writing – review & editing: Charlene Aya Yoboue, Salome Hosch, Olivier Tresor Donfack, Etienne A. Guirou, Bonifacio Manguire Nlavo, Mitoha Ondo'o Ayekaba, Carlos Guerra, Wonder P. Phiri, Guillermo A. Garcia, Tobias Schindler.

References

1. Knopp S, Steinmann P, Hatz C, Keiser J, Utzinger J. Nematode infections: filariases. *Infectious disease clinics of North America*. 2012; 26(2):359–81. <https://doi.org/10.1016/j.idc.2012.02.005> PMID: 22632644
2. Ta-Tang TH, Crainey JL, Post RJ, Luz SL, Rubio JM. Mansonellosis: current perspectives. *Research and reports in tropical medicine*. 2018; 9:9–24. <https://doi.org/10.2147/RRTM.S125750> PMID: 30050351
3. Mourembou G, Fenollar F, Lekana-Douki JB, Ndjoiy Mbiguino A, Maghendji Nzondo S, Matsiegui PB, et al. *Mansonella*, including a Potential New Species, as Common Parasites in Children in Gabon. *PLoS neglected tropical diseases*. 2015; 9(10):e0004155. <https://doi.org/10.1371/journal.pntd.0004155> PMID: 26484866
4. Sandri TL, Kreidenweiss A, Cavallo S, Weber D, Juhas S, Rodi M, et al. Molecular epidemiology of *Mansonella* species in Gabon. *The Journal of infectious diseases*. 2020.
5. Simonsen PE, Onapa AW, Asio SM. *Mansonella perstans* filariasis in Africa. *Acta tropica*. 2011; 120 Suppl 1:S109–20. <https://doi.org/10.1016/j.actatropica.2010.01.014> PMID: 20152790
6. Mediannikov O, Ranque S. Mansonellosis, the most neglected human filariasis. *New microbes and new infections*. 2018; 26:S19–s22. <https://doi.org/10.1016/j.nmni.2018.08.016> PMID: 30402239
7. Metzger WG, Mordmuller B. *Loa loa*-does it deserve to be neglected? *The Lancet Infectious diseases*. 2014; 14(4):353–7. [https://doi.org/10.1016/S1473-3099\(13\)70263-9](https://doi.org/10.1016/S1473-3099(13)70263-9) PMID: 24332895
8. Whittaker C, Walker M, Pion SDS, Chesnais CB, Boussinesq M, Basanez MG. The Population Biology and Transmission Dynamics of *Loa loa*. *Trends in parasitology*. 2018; 34(4):335–50. <https://doi.org/10.1016/j.pt.2017.12.003> PMID: 29331268
9. Wanji S, Tayong DB, Ebai R, Opoku V, Kien CA, Ndongmo WPC, et al. Update on the biology and ecology of *Culicoides* species in the South-West region of Cameroon with implications on the transmission of *Mansonella perstans*. *Parasites & vectors*. 2019; 12(1):166. <https://doi.org/10.1186/s13071-019-3432-9> PMID: 30975194
10. Alhassan A, Li Z, Poole CB, Carlow CK. Expanding the MDx toolbox for filarial diagnosis and surveillance. *Trends in parasitology*. 2015; 31(8):391–400. <https://doi.org/10.1016/j.pt.2015.04.006> PMID: 25978936
11. Zainabadi K, Adams M, Han ZY, Lwin HW, Han KT, Ouattara A, et al. A novel method for extracting nucleic acids from dried blood spots for ultrasensitive detection of low-density *Plasmodium falciparum* and *Plasmodium vivax* infections. *Malaria journal*. 2017; 16(1):377. <https://doi.org/10.1186/s12936-017-2025-3> PMID: 28923054
12. Guerra CA, Kang SY, Citron DT, Hergott DEB, Perry M, Smith J, et al. Human mobility patterns and malaria importation on Bioko Island. *Nature communications*. 2019; 10(1):2332. <https://doi.org/10.1038/s41467-019-10339-1> PMID: 31133635
13. Herrador Z, Garcia B, Ncogo P, Perteguer MJ, Rubio JM, Rivas E, et al. Interruption of onchocerciasis transmission in Bioko Island: Accelerating the movement from control to elimination in Equatorial Guinea. *PLoS neglected tropical diseases*. 2018; 12(5):e0006471. <https://doi.org/10.1371/journal.pntd.0006471> PMID: 29723238
14. Kleinschmidt I, Schwabe C, Benavente L, Torrez M, Ridl FC, Segura JL, et al. Marked increase in child survival after four years of intensive malaria control. *The American journal of tropical medicine and hygiene*. 2009; 80(6):882–8. PMID: 19478243
15. Bradley J, Rehman AM, Schwabe C, Vargas D, Monti F, Ela C, et al. Reduced prevalence of malaria infection in children living in houses with window screening or closed eaves on Bioko Island, equatorial Guinea. *PLoS one*. 2013; 8(11):e80626. <https://doi.org/10.1371/journal.pone.0080626> PMID: 24236191

16. Cano J, Berzosa PJ, Roche J, Rubio JM, Moyano E, Guerra-Neira A, et al. Malaria vectors in the Bioko Island (Equatorial Guinea): estimation of vector dynamics and transmission intensities. *J Med Entomol*. 2004; 41(2):158–61. <https://doi.org/10.1603/0022-2585-41.2.158> PMID: 15061273
17. Cook J, Hergott D, Phiri W, Rivas MR, Bradley J, Segura L, et al. Trends in parasite prevalence following 13 years of malaria interventions on Bioko island, Equatorial Guinea: 2004–2016. *Malaria journal*. 2018; 17(1):62. <https://doi.org/10.1186/s12936-018-2213-9> PMID: 29402288
18. García GA, Hergott DEB, Phiri WP, Perry M, Smith J, Osa Nfumu JO, et al. Mapping and enumerating houses and households to support malaria control interventions on Bioko Island. *Malaria journal*. 2019; 18(1):283. <https://doi.org/10.1186/s12936-019-2920-x> PMID: 31438979
19. Guirou EA, Schindler T, Hosch S, Donfack OT, Yoboue CA, Krähenbühl S, et al. Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Scientific reports*. 2020; 10(1):12305. <https://doi.org/10.1038/s41598-020-69268-5> PMID: 32703999
20. Fink DL, Kamgno J, Nutman TB. Rapid molecular assays for specific detection and quantitation of *Loa loa* microfilaremia. *PLoS neglected tropical diseases*. 2011; 5(8):e1299. <https://doi.org/10.1371/journal.pntd.0001299> PMID: 21912716
21. Bassene H, Sambou M, Fenollar F, Clarke S, Djiba S, Mourembou G, et al. High Prevalence of *Mansonella perstans* Filariasis in Rural Senegal. *The American journal of tropical medicine and hygiene*. 2015; 93(3):601–6. <https://doi.org/10.4269/ajtmh.15-0051> PMID: 26078318
22. Schindler T, Robaina T, Sax J, Bieri JR, Mpina M, Gondwe L, et al. Molecular monitoring of the diversity of human pathogenic malaria species in blood donations on Bioko Island, Equatorial Guinea. *Malaria journal*. 2019; 18(1):9. <https://doi.org/10.1186/s12936-019-2639-8> PMID: 30646918
23. Jiménez M, González LM, Carranza C, Bailo B, Pérez-Ayala A, Muro A, et al. Detection and discrimination of *Loa loa*, *Mansonella perstans* and *Wuchereria bancrofti* by PCR-RFLP and nested-PCR of ribosomal DNA ITS1 region. *Experimental parasitology*. 2011; 127(1):282–6. <https://doi.org/10.1016/j.exppara.2010.06.019> PMID: 20599994
24. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987; 4(4):406–25. <https://doi.org/10.1093/oxfordjournals.molbev.a040454> PMID: 3447015
25. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*. 1993; 10(3):512–26. <https://doi.org/10.1093/oxfordjournals.molbev.a040023> PMID: 8336541
26. Krähenbühl S, Studer F, Guirou E, Deal A, Mächler P, Hosch S, et al. ELIMU-MDx: a web-based, open-source platform for storage, management and analysis of diagnostic qPCR data. *Biotechniques*. 2020; 68(1):22–7. <https://doi.org/10.2144/btn-2019-0064> PMID: 31588775
27. Salim N, Knopp S, Lweno O, Abdul U, Mohamed A, Schindler T, et al. Distribution and risk factors for *Plasmodium* and helminth co-infections: a cross-sectional survey among children in Bagamoyo district, coastal region of Tanzania. *PLoS neglected tropical diseases*. 2015; 9(4):e0003660. <https://doi.org/10.1371/journal.pntd.0003660> PMID: 25837022
28. Cuello MR, Cuadros EN, Claros AM, Hortelano MG, Fontelos PM, Peña MJ. [Filariasis infestation in patients emanating from endemic area. 14 cases series presentation]. *Anales de pediatria (Barcelona, Spain: 2003)*. 2009; 71(3):189–95. <https://doi.org/10.1016/j.anpedi.2009.04.022> PMID: 19640814
29. Placinta IA, Pascual CI, Chiari-Toumit C, Mata-Moret L, Sanchez-Cañizal J, Barranco-González H. Ocular loiasis affecting a child and its assessment by Anterior Segment Optical Coherence Tomography. *Rom J Ophthalmol*. 2019; 63(2):184–7. PMID: 31334399
30. Puente S, Ramírez-Olivencia G, Lago M, Subirats M, Bru F, Pérez-Blazquez E, et al. Loiasis in sub-Saharan migrants living in Spain with emphasis of cases from Equatorial Guinea. *Infectious diseases of poverty*. 2020; 9(1):16. <https://doi.org/10.1186/s40249-020-0627-4> PMID: 32029005
31. Puente S, Lago M, Subirats M, Sanz-Esteban I, Arsuaga M, Vicente B, et al. Imported *Mansonella perstans* infection in Spain. *Infectious diseases of poverty*. 2020; 9(1):105. <https://doi.org/10.1186/s40249-020-00729-9> PMID: 32703283
32. Lee LS, Paton NI. Importation of seven cases of an unusual helminthic infection into Singapore and assessment of the risk of local transmission. *Singapore medical journal*. 2004; 45(5):227–8. PMID: 15143359
33. El Haouri M, Erragragui Y, Sbai M, Alioua Z, Louzi, El Mellouki W, et al. [Cutaneous filariasis *Loa Loa*: 26 moroccan cases of importation]. *Annales de dermatologie et de venerologie*. 2001; 128(8–9):899–902. PMID: 11590341
34. Wang X, Zhang X, Zong Z. A Case of loiasis in a patient returning to China diagnosed by nested PCR using DNA extracted from tissue. *Journal of travel medicine*. 2012; 19(5):314–6. <https://doi.org/10.1111/j.1708-8305.2012.00635.x> PMID: 22943273

35. Priest DH, Nutman TB. Loiasis in US Traveler Returning from Bioko Island, Equatorial Guinea, 2016. *Emerging infectious diseases*. 2017; 23(1):160–2. <https://doi.org/10.3201/eid2301.161427> PMID: 27983940
36. Ta TH, Moya L, Nguema J, Aparicio P, Miguel-Oteo M, Cenzual G, et al. Geographical distribution and species identification of human filariasis and onchocerciasis in Bioko Island, Equatorial Guinea. *Acta tropica*. 2018; 180:12–7. <https://doi.org/10.1016/j.actatropica.2017.12.030> PMID: 29289559
37. Asio SM, Simonsen PE, Onapa AW. *Mansonella perstans* filariasis in Uganda: patterns of microfilaraemia and clinical manifestations in two endemic communities. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2009; 103(3):266–73. <https://doi.org/10.1016/j.trstmh.2008.08.007> PMID: 18809192
38. Downes BL, Jacobsen KH. A Systematic Review of the Epidemiology of Mansonelliasis. *Afr J Infect Dis*. 2010; 4(1):7–14. <https://doi.org/10.4314/ajid.v4i1.55085> PMID: 23878696
39. Debrah LB, Nausch N, Opoku VS, Owusu W, Mubarik Y, Berko DA, et al. Epidemiology of *Mansonella perstans* in the middle belt of Ghana. *Parasites & vectors*. 2017; 10(1):15. <https://doi.org/10.1186/s13071-016-1960-0> PMID: 28061905
40. Akue JP, Eyang-Assengone ER, Dieki R. *Loa loa* infection detection using biomarkers: current perspectives. *Research and reports in tropical medicine*. 2018; 9:43–8. <https://doi.org/10.2147/RRTM.S132380> PMID: 30050354
41. Ritter M, Ndongmo WPC, Njouendou AJ, Nghochuzie NN, Nchang LC, Tayong DB, et al. *Mansonella perstans* microfilaremic individuals are characterized by enhanced type 2 helper T and regulatory T and B cell subsets and dampened systemic innate and adaptive immune responses. *PLoS neglected tropical diseases*. 2018; 12(1):e0006184. <https://doi.org/10.1371/journal.pntd.0006184> PMID: 29324739
42. Ricciardi A, Nutman TB. IL-10 and Its Related Superfamily Members IL-19 and IL-24 Provide Parallel/Redundant Immune-Modulation in *Loa loa* Infection. *The Journal of infectious diseases*. 2021; 223(2):297–305. <https://doi.org/10.1093/infdis/jiaa347> PMID: 32561912
43. Metenou S, Babu S, Nutman TB. Impact of filarial infections on coincident intracellular pathogens: *Mycobacterium tuberculosis* and *Plasmodium falciparum*. *Curr Opin HIV AIDS*. 2012; 7(3):231–8. <https://doi.org/10.1097/COH.0b013e3283522c3d> PMID: 22418448
44. Stelekati E, Wherry EJ. Chronic bystander infections and immunity to unrelated antigens. *Cell Host Microbe*. 2012; 12(4):458–69. <https://doi.org/10.1016/j.chom.2012.10.001> PMID: 23084915
45. Nkurunungi G, Zirimenya L, Natukunda A, Nassuuna J, Oduru G, Ninsiima C, et al. Population differences in vaccine responses (POPVAC): scientific rationale and cross-cutting analyses for three linked, randomised controlled trials assessing the role, reversibility and mediators of immunomodulation by chronic infections in the tropics. *BMJ Open*. 2021; 11(2):e040425. <https://doi.org/10.1136/bmjopen-2020-040425> PMID: 33593767
46. Billingsley PF, Maas CD, Olotu A, Schwabe C, García GA, Rivas MR, et al. The Equatoguinean Malaria Vaccine Initiative: From the Launching of a Clinical Research Platform to Malaria Elimination Planning in Central West Africa. *The American journal of tropical medicine and hygiene*. 2020; 103(3):947–54. <https://doi.org/10.4269/ajtmh.19-0966> PMID: 32458790
47. Jongo SA, Urbano V, Church LWP, Olotu A, Manock SR, Schindler T, et al. Immunogenicity and Protective Efficacy of Radiation-Attenuated and Chemo-Attenuated PfSPZ Vaccines in Equatoguinean Adults. *The American journal of tropical medicine and hygiene*. 2021; 104(1):283–93. <https://doi.org/10.4269/ajtmh.20-0435> PMID: 33205741
48. Chesnais CB, Pion SD, Boullé C, Gardon J, Gardon-Wendel N, Fokom-Domgue J, et al. Individual risk of post-ivermectin serious adverse events in subjects infected with *Loa loa*. *EClinicalMedicine*. 2020; 28:100582. <https://doi.org/10.1016/j.eclinm.2020.100582> PMID: 33294807

3.3. Development and evaluation of Plasmopod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance

Published in PLOS Global Public Health, 2023

RESEARCH ARTICLE

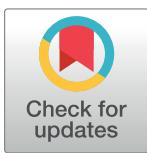
Development and evaluation of Plasmopod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance

Philippe Bechtold^{1,2*}, Philipp Wagner^{3,4}, Salome Hosch^{3,4}, Michele Gregorini^{1,2}, Wendelin J. Stark^{1,2}, Jean Chrysostome Gody⁵, Edwige Régina Kodia-Lenguetama⁶, Marilou Sonia Pagonendji⁷, Olivier Tresor Donfack⁸, Wonder P. Phiri⁸, Guillermo A. García⁸, Christian Nsanzanbana^{3,4}, Claudia A. Daubengerger^{3,4}, Tobias Schindler^{3,4}, Ulrich Vickos^{9,10}

1 Institute for Chemical and Bioengineering, ETH Zurich, Zuerich, Switzerland, **2** Diaxxo AG, Zuerich, Switzerland, **3** Swiss Tropical and Public Health Institute, Basel, Switzerland, **4** University of Basel, Basel, Switzerland, **5** Paediatric Hospital and University Complex of Bangui, Bangui, Central African Republic, **6** National Blood Transfusion Centre, Ministry of Health, Bangui, Central African Republic, **7** Laboratory of Parasitology, Institute Pasteur of Bangui, Bangui, Central African Republic, **8** MCD Global Health, Malabo, Equatorial Guinea, **9** Infectious and Tropical Diseases Unit, Department of Medicine, Amitié Hospital, Bangui, Central African Republic, **10** Microbiology and Diagnostic Immunology Unit, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy

✉ These authors contributed equally to this work.

* p.bechtold@diaxxo.com (PB); tobias.schindler@swisstph.ch (TS)



OPEN ACCESS

Citation: Bechtold P, Wagner P, Hosch S, Gregorini M, Stark WJ, Gody JC, et al. (2023) Development and evaluation of Plasmopod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance. *PLOS Glob Public Health* 3(9): e0001516. <https://doi.org/10.1371/journal.pgph.0001516>

Editor: Sarah Auburn, Menzies School of Health Research, AUSTRALIA

Received: December 30, 2022

Accepted: September 4, 2023

Published: September 27, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgph.0001516>

Copyright: © 2023 Bechtold et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data that support the findings of this study are available as a

Abstract

Malaria surveillance is hampered by the widespread use of diagnostic tests with low sensitivity. Adequate molecular malaria diagnostics are often only available in centralized laboratories. Plasmopod is a novel cartridge-based nucleic acid amplification test for rapid, sensitive, and quantitative detection of malaria parasites. Plasmopod is based on reverse-transcription quantitative polymerase chain reaction (RT-qPCR) of the highly abundant *Plasmodium* spp. 18S ribosomal RNA/DNA biomarker and is run on a portable qPCR instrument which allows diagnosis in less than 30 minutes. Our analytical performance evaluation indicates that a limit-of-detection as low as 0.02 parasites/ μ L can be achieved and no cross-reactivity with other pathogens common in malaria endemic regions was observed. In a cohort of 102 asymptomatic individuals from Bioko Island with low malaria parasite densities, Plasmopod accurately detected 83 cases, resulting in an overall detection rate of 81.4%. Notably, there was a strong correlation between the Cq values obtained from the reference RT-qPCR assay and those obtained from Plasmopod. In an independent cohort, using dried blood spots from malaria symptomatic children living in the Central African Republic, we demonstrated that Plasmopod outperforms malaria rapid diagnostic tests based on the PfHRP2 and panLDH antigens as well as thick blood smear microscopy. Our data suggest that this 30-minute sample-to-result RT-qPCR procedure is likely to achieve a diagnostic performance comparable to a standard laboratory-based RT-qPCR setup. We believe that the Plasmopod rapid NAAT could enable widespread accessibility of high-quality and cost-effective molecular malaria surveillance data through decentralization of testing and surveillance activities, especially in elimination settings.

supplementary document uploaded to the journal's website.

Funding: Funding for PB, MG and WJS was provided by the Botnar Research Centre for Child Health as part of the Fast Track Call for Acute Global Health Challenges as well as the BRIDGE programme by Swiss National Science Foundation and Innosuisse. TS is supported by the Bioko Island Malaria Elimination Project (BIMEP). BIMEP is a public private partnership funded through the Government of Equatorial Guinea, Marathon Oil, Noble Energy, SonaGas, GEPetrol, and Atlantic Methanol (AMPCO). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: We have read the journal's policy and the authors of this manuscript have the following competing interests: PB, MG, and WJS are co-inventors on a corresponding patent application around the diaxxoPCR platform. MG, TS and WJS are shareholders of the ETH spin-off company Diaxxo AG. All other authors declare no competing interests.

Introduction

Malaria is an infectious disease caused by different *Plasmodium* spp. species and is transmitted through female *Anopheles* spp. mosquitoes to humans [1]. Although significant progress on combating the spread of the disease has been achieved, more than 600'000 people still die annually [2]. Considerable improvements in the accuracy and availability of diagnostics need to be achieved to reduce the overall burden of the disease further with the aim of malaria elimination [3]. Currently employed methods comprise blood smear microscopy, antigen rapid diagnostic tests (RDT) and nucleic acid amplification tests (NAAT) [4]. Diagnosis of malaria by light microscopy using Giemsa-stained thick or thin blood smears has been the gold standard since the early 20th century. Well trained and experienced microscopists can reach a limit of detection (LOD) of 50–100 parasites/ μ L blood [5]. Expert microscopists are however chronically lacking, and the sample throughput is rather low. RDTs are based on detection of parasite antigens in blood and can provide valuable and rapid answers in remote areas without the need for extensive training. They are low-cost and widely available with 419 million WHO pre-qualified malaria RDTs sold globally in 2020 [2]. Despite the improvements in the quality of malaria RDTs through programs like the WHO product testing [6, 7], their sensitivity remains limited as they detect antigens without involving target molecule amplification. A LOD, of 100–200 parasites/ μ L for PfHRP2-based RDTs [8] and about 1000 parasites/ μ L for panLDH-based RDTs [9] renders them unsuitable for surveillance in endemic and low-transmission environments and in elimination settings [10]. The presence of *P. falciparum* strains with *pfhrp2* gene deletions poses an additional challenge for malaria surveillance as these strains cannot be detected by PfHRP2-based RDTs [11]. NAAT such as reverse transcription quantitative polymerase chain reaction (RT-qPCR) have shown a LOD of 0.05 parasites/ μ L [12], which is by a factor of 1000 more sensitive than microscopy. This approach is especially favourable in areas where a high proportion of asymptomatic malaria carriers are living, maintaining the transmission cycle of the parasite [13, 14]. However, in resource-constrained settings the use of highly sensitive RT-qPCR based malaria testing has been restricted to well-equipped centralized laboratories due to high initial investment cost, sophisticated supply chain management and shortage of trained laboratory personnel [15]. Simple, rapid, highly sensitive and reliable molecular diagnostic tools are needed more than ever in conjunction with a functional surveillance system to enable and sustain malaria elimination.

The diaxxoPCR technology is based on a portable and easy-to-use qPCR instrument, which is designed for rapid identification, quantification and genotyping of pathogens at affordable costs. Only minimal hands-on-operations are needed, and pathogens can be detected in less than 30 minutes, based on an innovative temperature control strategy that allows reaching unprecedented high heating and cooling rates ($> 13^{\circ}\text{C/s}$) during the PCR [16]. Importantly, no cold-chain during shipping, storage or usage of the reagents is needed since the RT-qPCR reactions are run in aluminium-based cartridges, which come preloaded with all reagents in dried form. The cartridges are equipped with a total of 20 wells that can be loaded according to specific requirements. Each well accommodates a single qPCR assay to measure one sample. With a single qPCR assay, the maximum number of patient samples per cartridge is 20. However, when incorporating controls or a standard curve, the number of samples per cartridge will be reduced accordingly.

For mobile testing applications, the device can be powered using a car battery and the results can be accessed directly on the device's screen or through the browser of a smartphone or laptop. The diaxxoPCR device performs RT-qPCR amplification with very small reagent and sample input volumes, rendering it highly cost efficient in addition to its unparalleled speed. The PlasmoPod offers cost advantages over standard RT-qPCR assays due to its minimal reagent volumes, with an estimated cost-per-sample of around EUR 1.5 in small-size

batches and the potential to decrease below EUR 1.0 at scale. The diaxxoPCR platform has delivered results comparable to state-of-the-art qPCR devices for SARS-CoV-2 detection and genotyping [17]. The instrument achieved excellent diagnostic performance when tested with RNA extracted from culture-derived SARS-CoV-2 Variants of Concern (VOC) lineages and clinical samples collected in Equatorial Guinea, Central-West Africa [17].

In the current study, we describe the development of a *Plasmodium* spp. cartridge for the diaxxoPCR device referred to as “PlasmoPod”. We further report on the performance of PlasmoPod as a rapid and highly sensitive NAAT-based diagnostic tool for malaria and compare it to other currently available diagnostic tests using samples collected from children and adults living in two different Central African countries.

Materials and methods

Ethics statement

The malaria indicator survey conducted on Bioko Island, Equatorial Guinea was approved by the Ministry of Health and Social Welfare of Equatorial Guinea and the Ethics Committee of the London School of Hygiene & Tropical Medicine (Ref. No. LSHTM: 5556). Written informed consent was obtained from all adults and from parents or guardians of children who agreed to participate. Only samples for which an additional consent for molecular analysis was obtained were included in this study. The study in Bangui, Central African Republic was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics and Scientific Committee from the University of Bangui (approval n°3/UB/FACSS/CSCVPER/PER) and by the Ministry of Health of the Central African Republic (approval n°0277/MSPP/CAB/DGSP/DMPM/ SMEE du 05 août 2002) as part of the communicable and endemic diseases surveillance diagnostic program. The patients were informed about the objectives of the study and nature of their participation. Then, written and signed informed consents were obtained from the participants or the parents on behalf of their children.

PlasmoPod NAAT development and analytical performance evaluation

The experiments on the diaxxoPCR platform were performed using the PlasmoPod cartridges supplied by Diaxxo AG (Zuerich, Switzerland). The 20 well cartridges contain all reagents necessary for running a RT-qPCR in preloaded and in dried form. The *Plasmodium* spp. assay used for PlasmoPod is a TaqMan probe-based qPCR assay which uses a 6-Carboxyfluorescein (6-FAM) labelled probe to enable the detection of a specific amplification product produced during PCR. Published oligonucleotide sequences and concentrations for detection of *Plasmodium* spp. parasites are used [18]. Analytical performance of PlasmoPod was evaluated with purified NAs from different relevant pathogens. Cross-reactivity was tested against DNA or RNA extracted from bacteria (*Salmonella enterica subsp. enterica* serovar Typhi), viruses (Dengue virus serotype 3, Chikungunya virus, Yellow fever virus and Zika virus) and closely related apicomplexan parasites (*Cryptosporidium parvum* and *Cryptosporidium hominis*). Additionally, the assay's specificity was evaluated using *Plasmodium* spp.-free human blood and serum samples from four different donors. The sensitivity, accuracy, and reproducibility of the PlasmoPod NAAT was assessed using DNA extracted from cultivated and synchronized ring-stage NF54 *P. falciparum* parasites. Four technical replicates from a total of 15 DNA titration steps with concentrations ranging from 500 to 0.0008 parasites/ μ L, were analysed with PlasmoPod using the following cycling parameters on the diaxxoPCR instrument: reverse transcription of 300 seconds at 53°C, initial polymerase activation for 60 seconds at 90°C and then 45 cycles of 10 seconds at 94°C and 20 seconds at 56°C. Raw data was analysed by diaxxoPCR software and Cq values were automatically assigned to the samples.

Collection and characterization of samples from the asymptomatic malaria cohort

Bio-banked samples collected during The Malaria Indicator Surveys (MIS) conducted on Bioko Island, Equatorial Guinea in 2018 and 2019 were used for this study. The MIS involved voluntary participation of permanent residents and short-term visitors. The volunteers included in the survey were classified as asymptomatic for malaria and the survey was conducted at their respective places of residence. They were tested for malaria using the CareStart Malaria HRP2/pLDH Combo RDT. The used RDTs were stored at room temperature in plastic bags with desiccants and subsequently transported to the Swiss Tropical and Public Health Institute for additional molecular analysis. Total nucleic acids were extracted by the “Extraction of Nucleic Acids from RDTs” (ENAR) protocol and [19, 20] and the pan-*Plasmodium* spp. 18S ribosomal DNA and RNA molecules were targeted [18, 21] and detected by a highly-sensitive RT-qPCR (herein referred to Psp18S RT-qPCR assay) [19]. A total of 102 samples, found positive by the Psp18S RT-qPCR assay (Quantification Cycle (Cq) values <40,) were included into the PlasmoPod evaluation study.

Collection and characterization of samples from the clinical malaria cohort

The samples from the clinical cohort were collected at the Paediatric Hospital and University Complex of Bangui (CHUPB), located in the Central African Republic (CAR). Collection took place between March 8th and 13th, 2021. The patients were children aged between 2 months and 15 years that were admitted to the emergency department with fever as their main clinical symptom. In case malaria was suspected and after obtaining informed consent from their legal guardians, whole blood samples were collected in EDTA blood collection tubes. A malaria RDT (A&B Rapid Test Malaria P.f./Pan, Luca, Italy), thick blood smear (TBS) microscopy and a complete blood count were routinely performed. An aliquot of the whole blood was prepared as dried blood spots (DBS) on filter papers. The DBS were stored at room temperature and sent to the Swiss Tropical and Public Health Institute, Basel, Switzerland for further molecular analysis.

Molecular characterization of malaria parasites identified among clinical cohort samples with reference molecular assays

A molecular reference dataset from the DBS collected in the CAR was established to be used as a gold standard against which the performance of PlasmoPod was compared to. The reference dataset included the species identification of *Plasmodium* spp. positive samples as well as the analysis of the *pfhrp2/3* deletion status and quantification of the parasite density of all *P. falciparum* positive samples. Briefly, the New Extraction Method (NEM) protocol developed by Zainabadi *et al.* was used to extract total nucleic acids (NA), including DNA and RNA, from the DBS [22]. In short, one entire DBS, which corresponds to 30–50 μ L of whole blood, was lysed at 60°C for 2 h. NAs were subsequently purified and eluted in 100 μ L elution buffer as described elsewhere [19]. The same Psp18S RT-qPCR assay targeting pan-*Plasmodium* spp. 18S ribosomal DNA and RNA molecules as for the analysis of the asymptomatic malaria cohort, was used. The Psp18S assay was analysed and samples with Cq values <40 were considered malaria positive. All samples positive for the Psp18S assay were analysed by species-specific qPCR assays as described previously [23]. All samples positive for *P. falciparum* were screened for *pfhrp2* and/or *pfhrp3* deletions using a multiplex qPCR assay detecting *pfhrp2/3* deletions [24]. Only samples with a Cq value < 35 for the internal control of the *pfhrp2/3* deletion assay were considered eligible for analysis of deletion status. All reference qPCR and RT-qPCR assays were run on a Bio-Rad CFX96 Real-Time PCR System (Bio-Rad Laboratories,

California, USA). Samples were analysed in duplicate with positive (DNA from *P. falciparum* strain NF54) and non-template controls (molecular biology grade H₂O) added to each run.

qPCR-based quantification of *P. falciparum* parasite density of clinical malaria samples

The *P. falciparum* parasite density was determined based on the amplification and detection of the *P. falciparum*-specific single copy gene ribonucleotide reductase R2_e2 (herein referred to as PfRNR2 assay) [24, 25]. Briefly, the WHO International Standard for *Plasmodium falciparum* DNA for NAAT-based assays (PfIS) [26] was used to generate a serial dilution in parasite-free whole blood, ranging from 0.01 to 100'000 parasites/μL. Thirty μL of each dilution step was put on a DBS and dried, followed by NA extraction and qPCR quantification by the PfRNR2 assay. The resulting standard curve, including the slope and y-axis intercept, was used to quantify the parasite densities in the clinical samples. Based on a cut-off value of 5000 parasites/μL, the malaria positive children were categorized into high and moderate parasite density infection groups. The parasite density cut-off was determined based on clinical study criteria for symptomatic and severe malaria [27–29]. In case the PfRNR2 assay was negative while the more sensitive Psp18S assay was positive the child was assigned to the moderate parasite density group.

PlasmoPod NAAT evaluation using the asymptomatic and clinical malaria cohort samples

Asymptomatic malaria cohort. One replicate of 4.5 μL of extracted total NAs was loaded into a well of the PlasmoPod and run on the diaxxoPCR device using its standard cycling program and data analysis as described above. Raw data was analysed by diaxxoPCR software and C_q values were automatically assigned to the samples.

Clinical malaria cohort. Rapid extraction was performed on a single DBS punch with a 3mm diameter. The DBS punch was submerged in 100 μL of a 5% Chelex (Bio-Rad, California, USA) solution and heated to 95°C for 3 min. The supernatant of the resulting solution was used directly for PlasmoPod analysis. Briefly, 4.5 μL of eluate per sample were loaded in duplicates onto the 20-well cartridge, covered with paraffin oil (Sigma-Aldrich, St. Louis, USA) and cycled for a duration of 25 minutes (for 45 cycles) in the diaxxoPCR device. Each run contained two wells with positive (DNA from *P. falciparum* strain NF54) and two wells of a non-template control (molecular biology grade H₂O) control. The cycling parameters on the diaxxoPCR were as follows: reverse transcription of 300 seconds at 50°C, initial polymerase activation for 60 seconds at 92°C and then 45 cycles of 2 seconds at 92°C and 15 seconds at 55°C. Raw data was analysed by diaxxoPCR software and C_q values were automatically assigned to the samples. Specimens with amplification with a C_q < 40 in 2/2 of replicates were considered positive.

Data analysis

Statistical analysis and data visualization was performed using the R statistical language (version 4.1.2) based on packages dplyr, epiR, ggplot2, ggpubr, gridExtra, readxl, reshape2, scales, tidyr, tidyverse, cowplot, and plyr.

Results

PlasmoPod is a cartridge-based NAAT for rapid *Plasmodium* spp. detection

The diaxxoPCR device is a novel, small-scale and standalone qPCR instrument (Fig 1A) which can be used to run and analyse ready-to-use cartridges which contains all RT-qPCR reagents

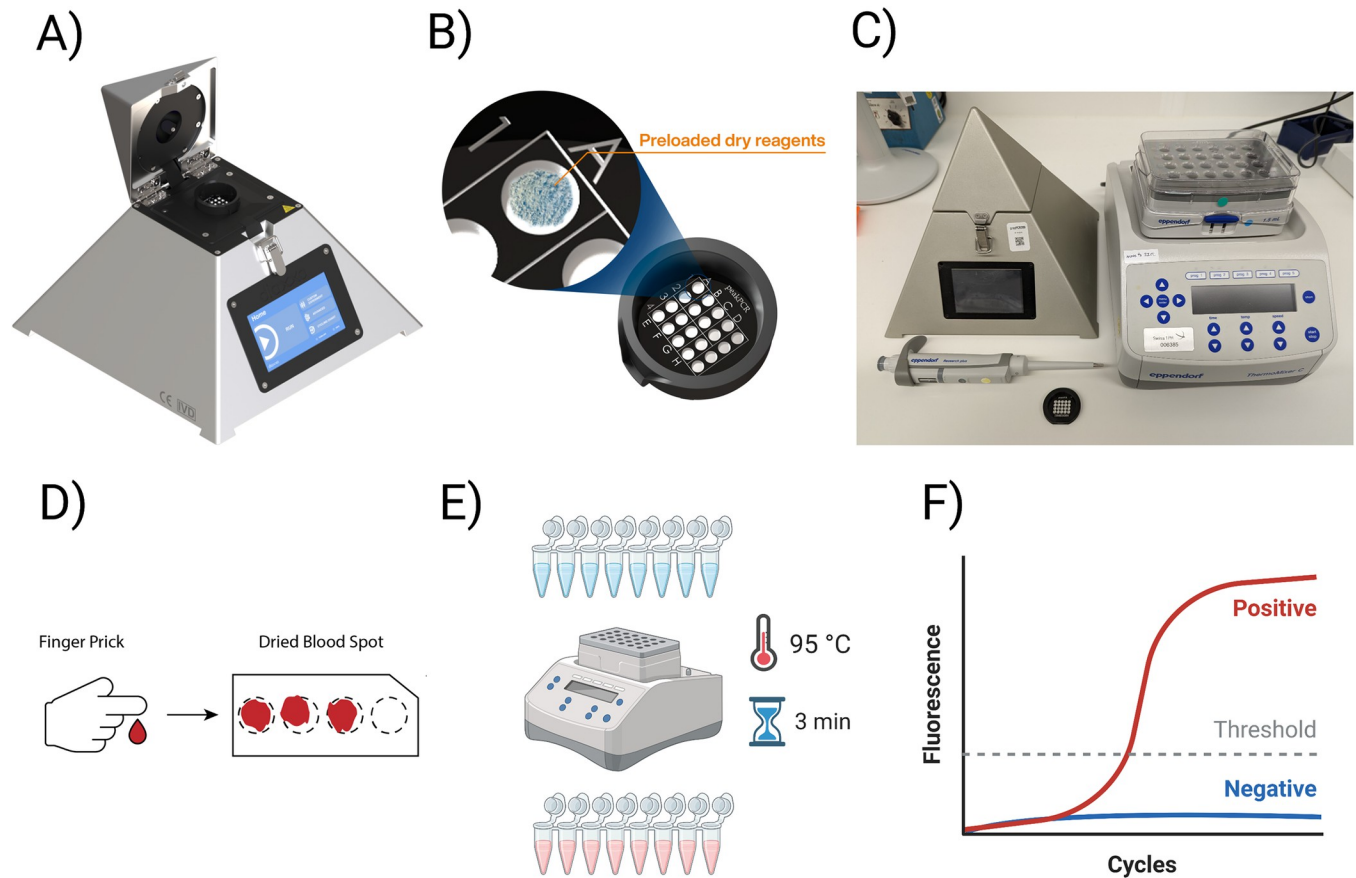


Fig 1. DiaxxoPCR and PlasmoPod setup for rapid molecular malaria detection. A) The DiaxxoPCR instrument is a pyramid-shaped stand-alone device for rapid qPCR cycling and fluorescence acquisition. (B) The PlasmoPod assay is based on a single-use cartridge which is pre-loaded with all qPCR reagents in dried form. (C) The PlasmoPod laboratory setup consisting of the diaxxoPCR device, a heatblock, a pipette and PlasmoPod cartridges. (D) Finger prick blood is sampled and stored as dried blood spots (DBS). (E) The setup for rapid NA extraction from DBS. (F) The NA are amplified and signals analysed using the diaxxoPCR device and its integrated analysis software. Images A and B are republished from diaxxo AG under a CC BY license, with permission from Dr Michele Gregorini (who is an author of this manuscript), original copyright [diaxxo AG, 2020–2022]. The granted permission to the manuscript has been uploaded. Partially created with [Biorender.com](https://www.biorender.com).

<https://doi.org/10.1371/journal.pgph.0001516.g001>

and oligonucleotidesoligo nucleotides in dried form (Fig 1B). The cartridge developed and evaluated in this study was named “PlasmoPod”. The diaxxoPCR device was selected as the core component for our minimal laboratory setup for molecular malaria diagnosis due to its compact size and robustness (Fig 1C). Unlike many other devices, the diaxxoPCR does not incorporate moving parts, making it highly suitable for mobile testing applications. In terms of dimensions, the device is comparable to a standard laboratory heating block, ensuring ease of portability and integration into various testing environments. As a first application for malaria diagnosis using diaxxoPCR, we developed a novel molecular diagnostic approach for rapid, sensitive, and quantitative detection of malaria parasites from blood sampled and stored on DBS. This approach included a rapid NA extraction procedure from DBS by submerging and boiling a single 3 mm diameter DBS punch in a Chelex solution (Fig 1D). During the 3-minute incubation step the blood preserved on the DBS is dissolved into the solution, the cells are lysed, and potential PCR inhibitors are removed by the Chelex (Fig 1E). Without any further processing 4.5 μ L of this solution are directly loaded into a well of the PlasmoPod. Using the diaxxoPCR device, the reverse transcription and a total of 45 PCR cycles are run in less than 30 minutes. The results can then be accessed through the screen of the diaxxoPCR device, a

connected smartphone or a computer (Fig 1F). In the current study we evaluated the PlasmoPod cartridge run on the diaxxoPCR rapid PCR device by comparison with a standard laboratory-based diagnostic approach for malaria based on state-of-the-art NA extraction procedure and RT-qPCR detection.

The analytical performance evaluation of PlasmoPod enables quantitative detection of *Plasmodium* spp. parasites with high specificity, reproducibility and sensitivity

The primary objective of the analytical performance evaluation was to assess the performance of the qPCR itself. Thus, we utilized purified NAs, without considering the rapid extraction procedure's potential impact on the analytical performance. The potential for cross-reactivity and unspecific amplification of PlasmoPod was tested with purified NA from a range of pathogens co-circulating in malaria endemic countries and parasite-free human-derived samples (Fig 2A). PlasmoPod measurements with bacteria (*Salmonella enterica subsp. enterica* serovar Typhi), viruses (Dengue virus serotype 3, Chikungunya virus, Yellow fever virus and Zika virus), closely related apicomplexan parasites (*Cryptosporidium parvum* and *Cryptosporidium hominis*) and *Plasmodium* spp.-free human blood resulted in delta fluorescence values (end-point minus baseline fluorescence) and maximum amplification curve slope values below the pre-defined positivity cut-off values. As comparison, results obtained with purified total NAs from culture-derived ring-stage synchronized *P. falciparum* parasites analysed at different concentrations are shown. To demonstrate the ability to also detect non-*falciparum* human pathogenic *Plasmodium* spp. species, we analysed clinical samples positive for *P. vivax*, *P. ovale* spp. and *P. malariae*. For *P. falciparum*, the two dilutions with the lowest input concentration were

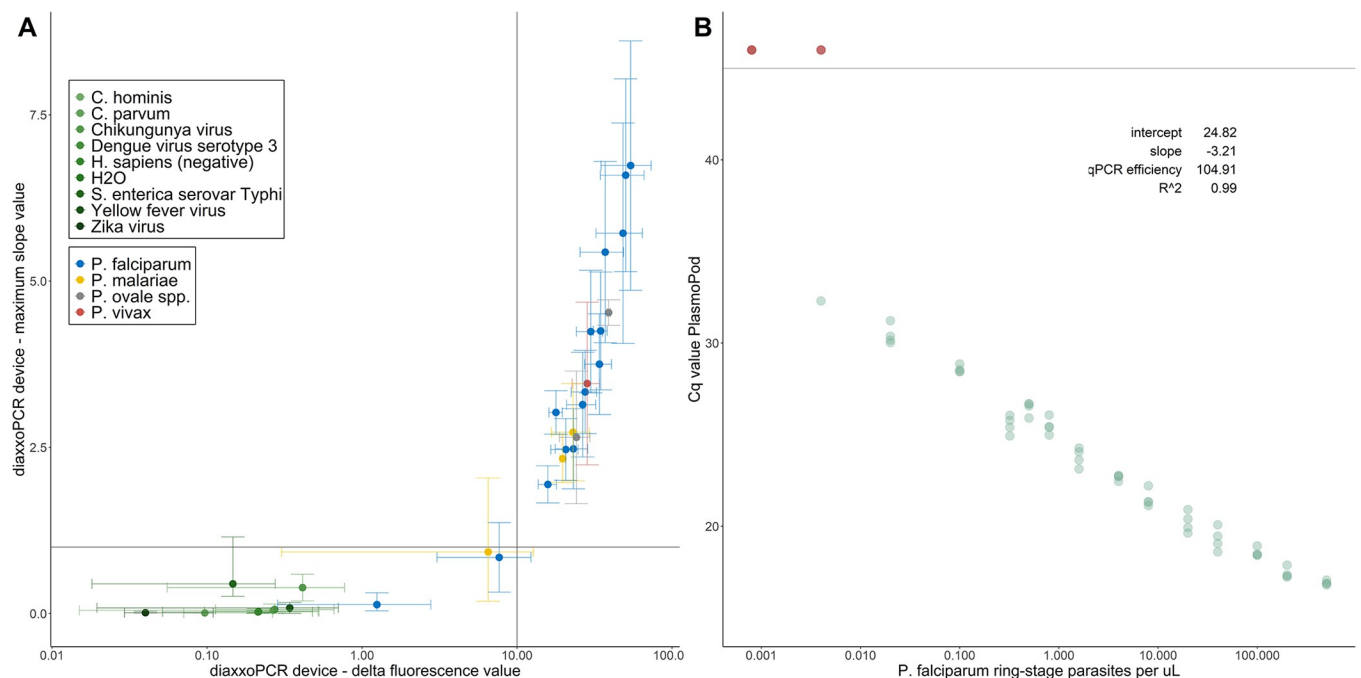


Fig 2. Analytical performance evaluation of PlasmoPod. (A) DiaxxoPCR derived delta fluorescence and maximum slope after amplification with nucleic acid from different pathogens. Lines indicate the delta fluorescence positivity cutoff (vertical) and maximum slope positivity cutoff (horizontal). (B) Cq value of PlasmoPod versus dilutions of NAs extracted from culture-derived ring-stage synchronized *P. falciparum* parasites. Sample without amplification are colored in red.

<https://doi.org/10.1371/journal.pgph.0001516.g002>

below the pre-defined positivity thresholds but still distinguishable from the non-malaria samples (Fig 2A). Using the same serial dilution of DNA extracted from culture-derived ring-stage synchronized *P. falciparum* parasites, the relationship between concentration and Cq values was established (Fig 2B). Four replicates of all concentrations were run on three different PlasmoPod cartridges. The lowest concentration which resulted in a positive signal for all four replicates was 0.02 parasites/ μ L. The relationship between Cq values and parasite concentration is described by an R^2 value of 0.99. The slope of -3.21 translates into an almost perfect qPCR efficiency of 104.91%.

Diagnostic performance of PlasmoPod evaluated with samples from asymptomatic parasite carriers using NAs extracted from archived RDTs

To assess the performance of PlasmoPod as a sensitive molecular tool for malaria surveillance, we examined extracted NA from 102 asymptomatic individuals carrying malaria parasites, who were part of the annual malaria indicator survey conducted on Bioko Island, Equatorial Guinea. For this study, only samples found positive by a previous qPCR screening were included. Among the participants, 47 tested positive for both PfHRP2 and panLDH antigens, 23 for panLDH alone, 12 for PfHRP2 alone, and 20 were negative for both antigens. NAs were extracted directly from the blood stored on the archived RDTs and analyzed using the laboratory RT-qPCR assay targeting *Plasmodium* spp. 18S rDNA/rRNA as the gold standard diagnostic test. All 102 individuals had detectable *Plasmodium* spp. NA on their archived RDTs, with Cq values ranging from 23.6 to 39.0 and a median of 33.4. PlasmoPod correctly identified 83 out of the 102 samples, yielding an overall detection rate of 81.4%. A strong correlation was observed between the Cq values obtained from the reference RT-qPCR assay on the standard laboratory platform and the Cq values derived from PlasmoPod (Fig 3A). The detection probability of PlasmoPod was dependent on the input target molecule number, as indicated by the Cq values obtained from the reference RT-qPCR assay (Fig 3B). At an ultra-high Cq value of 39.9, the estimated detection rate was 64.1% (95% CI: 42.6–85.6%), suggesting a high recall rate even at very low target molecule concentrations.

This dataset encompassed individuals of all age groups, ranging from 1 to 75 years old, which is of particular interest since older asymptomatic individuals are expected to exhibit higher natural immunity and consequently lower parasite densities. To conduct a more detailed analysis, we stratified the cohort into children (up to 15 years old) and adults (Fig 3C). Children exhibited a higher detection rate (88.4%) compared to adults (76.3%), likely due to the typically higher parasite densities observed in children. The Cq values obtained with PlasmoPod were lower in the children's group, although this difference was not statistically significant (Fig 3D).

Parasitological and clinical characteristics of clinical malaria cohort used for PlasmoPod test evaluation

Next, the performance of the PlasmoPod was evaluated by analyzing blood samples collected from febrile patients admitted to the Paediatric Hospital and University Complex of Bangui. DBS from a total of 47 children were included and an overview of the parasitological and demographic characteristics of these children are shown in Table 1. The age of the children ranged from 2 months to 15 years with 48.9% (23/47) being female. DBS collected from these children were screened for *Plasmodium* spp. NAs with a high sensitivity diagnostic RT-qPCR assay based on the parasites' 18S ribosomal DNA/RNA (Psp18S assay) using the Bio-Rad CFX96 qPCR device. Parasite density in *P. falciparum* positive children was estimated by the PfrNR2 qPCR assay and children were stratified accordingly into moderate (<5000 parasites/

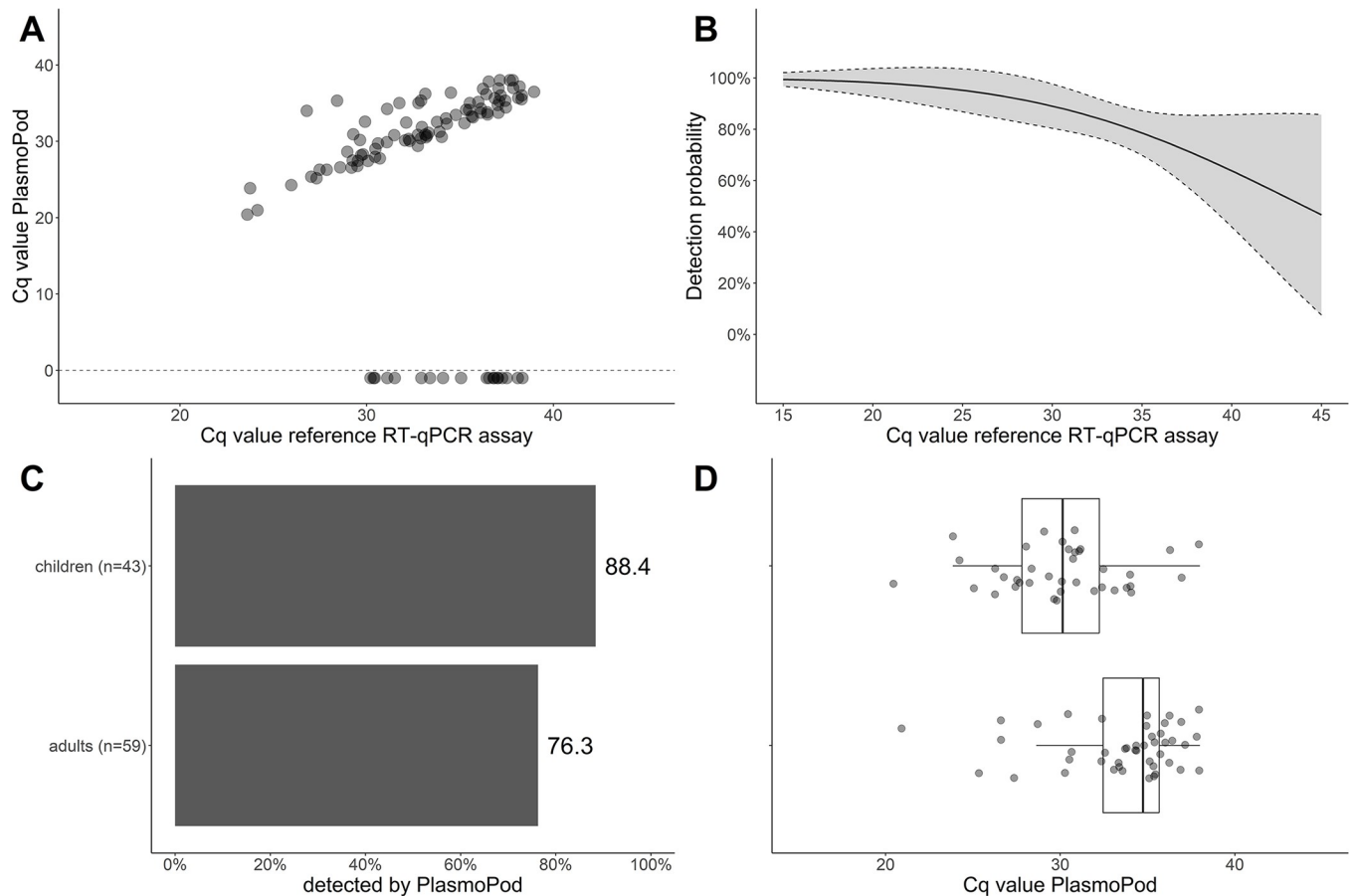


Fig 3. Analysis of asymptomatic malaria cohort. (A) Correlation of of Cq values obtained from reference RT-qPCR run on the Biorad CFX96 instrument and PlasmoPod run on diaxoPCR. Samples negative for PlasmoPod were assigned a Cq value of -1. (B) Detection probability for PlasmoPod modelled based on reference Cq values. The grey area represents the 95% confidence interval. (C) Detection rate of PlasmoPod stratified by age group. (D) Cq values stratified by age group.

<https://doi.org/10.1371/journal.pgph.0001516.g003>

µL) and high (≥ 5000 parasites/µL) parasite density groups. Out of 47 children, 16 were negative for *Plasmodium* spp. and *P. falciparum* by RT-qPCR screening, 16 had moderate *P. falciparum* parasite densities and 15 had higher *P. falciparum* parasite density infection. The children assigned to the higher parasite density group were younger compared to the children with moderate parasite densities or children without detectable parasites.

A strong correlation between parasite densities derived from TBS microscopy and qPCR was observed (Fig 4A). Interestingly, five out of the seven Psp18S qPCR-positive samples

Table 1. Parasitological and demographic characteristics of the study population selected for evaluation of PlasmoPod from Central African Republic.

Malaria stratification	Number of children	Age	Sex (% female)	Parasite density (parasites/µL)
		Median and Range		
Negative for malaria	16	5.5 years 6 month– 15 years	50.0%	0
Moderate parasite density <5000 per µL	16	3 years 10 month– 14 years	46.7%	Median: 1158 IQR: 447–2858
High parasite density >5000 per µL	15	13 months 2 month–8 years	53.3%	Median: 25'800 IQR: 12'171–47'960

<https://doi.org/10.1371/journal.pgph.0001516.t001>

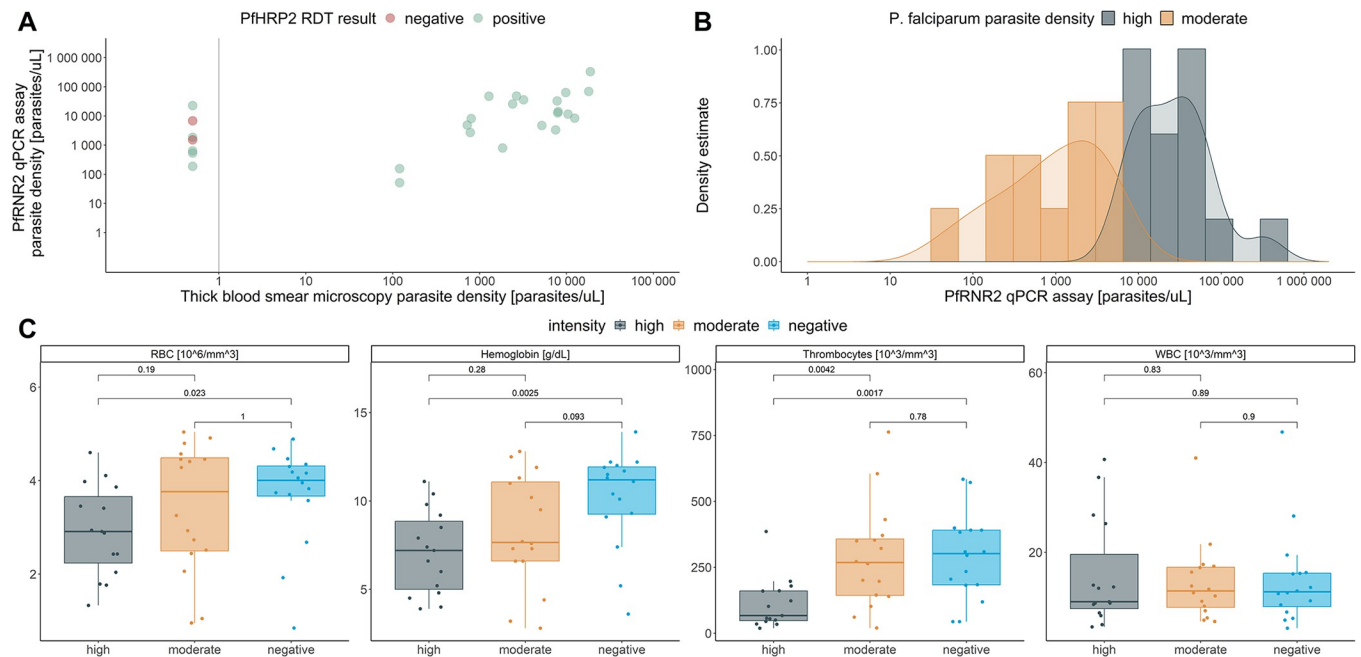


Fig 4. Clinical and parasitological characterization of study population. (A) Correlation of parasite density assessed by the PfRRN2 assay and thick blood smear microscopy. (B) Modelled distribution of parasite densities measured by PfRRN2 assay and highlighted for high- and moderate-density infections. (C) Hematological parameters, including red blood cell counts (RBC), Hemoglobin levels (Hb), Thrombocyte counts (PLT) and white blood cell counts (WBC) compared between malaria negative children and children with high- and moderate-density malaria infections.

<https://doi.org/10.1371/journal.pgph.0001516.g004>

which were negative by TBS microscopy, were positive by PfHRP2-based RDT, indicating a low diagnostic performance of the TBS microscopy in this setting. The modelled parasite density for the moderate- and high-density groups are visualized in Fig 4B. A large dynamic range, covering parasite densities from 52 to 332'983 parasites/ μ L, is included in this diagnostic test evaluation. Next, we wanted to compare the four hematological parameters, including counts on erythrocytes (RBC), leukocytes (WBC), thrombocytes (PLT) as well as the hemoglobin concentration (Hb) that had been collected from these children during blood collection stratified by malaria infection status (Fig 4C). RBC, PLT and Hb were significantly lower among the high parasite density infection group compared to the malaria negative children.

The parasitological characteristics of the malaria positive children was further analyzed by additional RT-qPCR assays in which the *Plasmodium* spp. species as well as the *pfhrp2* and *pfhrp3* deletion status was investigated. The highly sensitive Psp18S RT-qPCR assay was run as a multiplex assay combined with the internal control of the screening assay, the human *rnasep* gene (S1 Fig). Human-derived NAs were found in all 47 samples with an average Cq value of 27.8 and a standard deviation of 1.5, indicating that the nucleic acid extraction procedure worked efficiently and consistently. All 31 malaria positive children were tested positive for *P. falciparum* and no other *Plasmodium* spp. species were found (S1 Fig). The *P. falciparum* parasites identified in this cohort were analyzed for the presence of *pfhrp2* and/or *pfhrp3* gene deletions and not a single case of a *P. falciparum* strain with *pfhrp2*/*pfhrp3* gene deletion was found (S1 Fig).

PlasmoPod, coupled with a quick NA extraction procedure from DBS, enables rapid malaria diagnosis

A clinical evaluation dataset, including the well characterized malaria negative and positive samples described above, was used to compare the diagnostic performance of PlasmoPod with

malaria diagnosis based on PfHRP2/panLDH RDTs and TBS microscopy. During the clinical evaluation stage, PlasmoPod was run with NAs extracted by the rapid Chelex-based procedure on the diaxzoPCR instrument. As the gold standard for qualitative comparisons, we used the outcome of the highly sensitive Psp18S RT-qPCR assay based on amplification of NAs extracted with the NEM protocol and run on the Bio-Rad CFX96 qPCR instrument. All quantitative analysis was conducted based on the parasite densities obtained by the highly accurate PfrNR2 qPCR assay. In comparison to the gold standard the sensitivities and the specificities were calculated summarized in **Table 2**. PlasmoPod showed an overall sensitivity of 93.6%. With 83.9%, the PfHRP2-based RDT achieved the second highest sensitivity, while TBS microscopy and panLDH-RDT outcomes resulted in overall sensitivities below 70%. As expected, all diagnostic methods achieved higher sensitivity in children with high parasite densities compared to children with moderate parasite densities. Among children with moderate parasite densities, PlasmoPod missed 2/16 children resulting in a sensitivity of 87.5% for this group. Interestingly, the two false-negative children had parasite densities below the LOD of the PfrNR2 qPCR and were only detected by the highly sensitive RT-qPCR assay based on amplification of 18S ribosomal total NAs. Sensitivities among the moderate parasite density ranged from 75.0% to 43.8% for the other three diagnostic methods. Only TBS microscopy and PlasmoPod, performed with a 100% specificity. The panLDH-RDT and PfHRP2-RDT tests were wrongly positive in 1, and 2 out of 16 *Plasmodium* spp. negative children, respectively.

A strong correlation of Cq values derived from the PlasmoPod measurements obtained from the di-axzoPCR device and the Cq values obtained from two reference qPCR assays run on the Bio-Rad CFX96 instrument was observed (**Fig 5A**). The correlation of PlasmoPod with the Psp18S RT-qPCR was stronger than the correlation with the DNA-based PfrNR2 qPCR assay. There was an overall high correlation between two independent qPCR assays, run with DBS extracted NAs following standard procedures on a standard qPCR instrument like Bio-Rad CFX96 qPCR instrument, and our novel approach based on rapid extraction procedure from DBS in combination with ready-to-use PlasmoPod cartridges PlasmoPods and rapid PCR cycling. Additionally, a significant correlation between PlasmoPod Cq values and parasite densities measured by thick blood smear microscopy was observed (**Fig 5B**). In summary, the data presented is a strong indication that PlasmoPod allows for quantitative measurements of malaria parasites in DBS collected under field conditions.

Discussion

Accurate and reliable diagnostic tests are the fundamental backbone of healthcare systems. Yet, 47% of the global population has little to no access to diagnostics [30]. The global technical

Table 2. Diagnostic performance of RDT (PfHRP2/panLDH), TBS microscopy and PlasmoPod compared to the gold standard RT-qPCR assay run on the BioRad CFX96 qPCR instrument.

	Sensitivity (95% CI)			Specificity (95% CI)
	All positive children	Moderate-density infections	High-density infections	-
	-	(<5000 parasites/ μ L)	(>5000 parasites/ μ L)	-
	n = 31	n = 16	n = 15	n = 16
TBS microscopy	64.55% (20/31) (45.4% - 80.8%)	43.8% (7/16) (21.3% - 73.4%)	86.7% (13/15) (59.5% - 98.3%)	100.0% (16/16) (79.4% - 100.0%)
PfHRP2-RDT	83.9% (26/31) (66.3% - 94.6%)	75.0% (12/16) (47.6% - 92.7%)	93.3% (14/15) (68.1% - 99.8%)	87.5% (14/16) (61.7% - 98.5%)
panLDH-RDT	67.7% (21/31) (48.6% - 83.3%)	62.5% (10/16) (32.3% - 83.7%)	80.0% (12/15) (51.9% - 95.7)	93.8% (15/16) (69.8% - 99.8%)
PlasmoPod	93.6% (29/31) (78.6% - 99.2%)	87.5% (14/16) (61.7% - 98.5%)	100.0% (15/15) (78.2% - 100.0%)	100.0% (16/16) (79.4% - 100.0%)

CI = Confidence interval

<https://doi.org/10.1371/journal.pgph.0001516.t002>

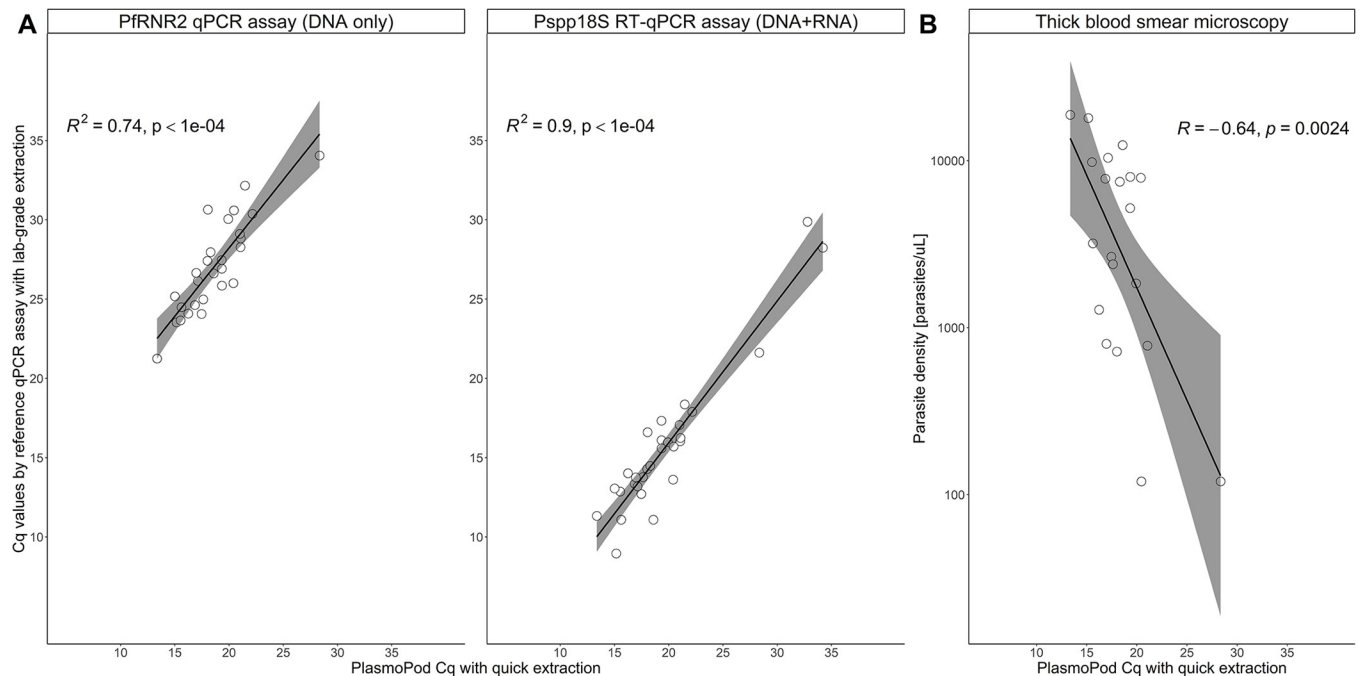


Fig 5. Quantification of *P. falciparum* parasites using PlasmoPod. (A) Correlation of Cq values derived from PlasmoPod and reference (RT)-qPCR assays based on *Plasmodium* spp. 18S ribosomal DNA and RNA (left panel) and the *P. falciparum* ribonucleotide reductase R2_e2 assays (right panel). (B) Correlation between PlasmoPod and thick blood smear microscopy for quantification of *P. falciparum* parasites. The grey color represents the 95% confidence interval.

<https://doi.org/10.1371/journal.pgph.0001516.g005>

strategy for malaria 2016–2030 sets the target of reducing global malaria incidence and mortality rates by at least 90% by 2030 [31]. One of the major pillars of the strategy ensuring access to malaria prevention, treatment and diagnosis. Expansion of diagnostic testing is required to provide timely and accurate surveillance data. This data is crucial for tracking the successes or drawbacks of malaria control and elimination efforts. Furthermore, in low transmission settings aiming for malaria elimination, the large-scale deployment of highly sensitive and specific diagnostic techniques is required to ensure the accurate diagnosis of low density asymptomatic parasite carriers [32]. Due to limited access to sensitive molecular tests for malaria surveillance and despite their low diagnostic performance, PfHRP2-based RDTs are still the most widely used diagnostic tests for malaria surveillance in endemic regions [2]. Molecular diagnostic techniques, in particular PCR-based tests, are much more accurate tools for surveillance. Incorporating NAATs in reactive case detection (RCD) [33] or monitoring activities related to mass drug administration (MDA) [34] programs provides clear advantages over antigen-based RDTs due to their higher sensitivity. However, these diagnostic tests are rarely used in malaria control programs since the infrastructure is often limited to centralized testing facilities and the costs of equipment and consumables are relatively high [35]. Traditional approaches to molecular malaria surveillance are centralized, where samples are collected and analysed at reference laboratories. This can be limit the process by logistical and financial constraints, leading to under-detection and under-reporting of malaria cases. Decentralized testing, on the other hand, involves the use of portable diagnostic devices and decentralized laboratories. These can be deployed at the point-of-care or in community settings. This approach enables more rapid and widespread testing, leading to more timely and accurate disease surveillance. In addition, decentralized testing can reduce the burden on central laboratories and improve access to testing in underserved or remote areas.

As a first step towards our goal of developing novel tools for improved and decentralized malaria surveillance, we designed, developed and extensively evaluated a rapid NAAT-based diagnostic test for *Plasmodium* spp. parasites using the portable and low-cost diaxxoPCR platform. Starting from DBS, in less than 30 minutes a diagnostic test for malaria is conducted with an analytical performance similar to sensitive, state-of-the-art laboratory-based RT-qPCR assays. Our novel approach is based on PlasmoPod cartridges which requires little hands-on time, no cold chain and almost no technical expertise as they are preloaded with all RT-qPCR reagents. In this initial study, we introduced a proof-of-concept for a cartridge-based molecular malaria test and presented validation data for our PlasmoPod platform. However, we acknowledge that further validation of this platform is necessary through direct testing conducted in endemic countries. A major limitation of our study is that the validation was conducted in a controlled laboratory setting, which does not necessarily represent the real-world conditions found in malaria-endemic testing sites. Additionally, we have identified several areas that can be improved. The current process of rapid extraction and loading of the PlasmoPods still relies on manual pre-processing of the samples. The pipetting steps involved present a challenge for widespread implementation of our approach, as it requires a certain level of expertise. Therefore, the next generation of the diaxxoPCR platform should incorporate automated nucleic acid extraction and loading of the PlasmoPods to reduce hands-on time and manual sample handling. By overcoming these challenges, PlasmoPods in combination with rapid PCR cyclers could be well-suited for deployment to satellite public health laboratories enabling decentralization of molecular malaria surveillance activities or for remote health care settings which do not have a fully equipped laboratory infrastructure.

In this initial study, we used a three-step approach to evaluate PlasmoPod for malaria diagnosis. We started with an analytical performance evaluation using well characterized samples and laboratory strains. Two different samples sets originating from two Central Africa countries, the Central African Republic and Equatorial Guinea, were used to further test the performance of PlasmoPod. Samples from malaria asymptomatic individuals from Bioko Island, Equatorial Guinea and symptomatic children from Bangui, Central African Republic, were analysed with PlasmoPod.

Our analytical performance evaluation suggests that a LOD as low as 0.02 parasites/ μ L can be achieved if NA are extracted from whole blood and no cross-reactivity with other pathogens common in malaria endemic regions was observed. This high sensitivity is achieved by using the highly abundant *Plasmodium* spp. 18S ribosomal total NA (RNA and DNA) as a biomarker for malaria infection. The parasite's 18S ribosomal NAs are present as a multicopy gene in the DNA as well as transcribed as RNA molecules.

PlasmoPod exhibited an 81.4% detection rate when analysing samples from 102 asymptomatic malaria-positive individuals on Bioko Island, emphasizing its efficacy as a sensitive molecular tool for malaria surveillance, particularly among asymptomatic cases.

The evaluation of the PlasmoPod and diaxxoPCR for malaria diagnosis among symptomatic children was performed with DBS collected from children attending the emergency department at the Paediatric Hospital and University Complex of Bangui. The results revealed that the PlasmoPod has a sensitivity of 100% if tested with children having high parasite densities and 87.5% if tested with children having a moderate-density infection. Since the gold-standard test used purified nucleic acids extracted from an entire DBS corresponding to approximately 30 μ L of blood, while the PlasmoPod approach utilized only a 3 mm DBS punch containing 1–2 μ L of blood [36], we conclude that PlasmoPod is likely to achieve a performance similar to a laboratory-based diagnostic platform for diagnosing symptomatic patients. Using a highly abundant biomarker like 18S ribosomal RNA and DNA for the RT-qPCR, the lower sample input and lack of highly pure NA due to the rapid extraction process,

can be compensated resulting in a robust approach suitable for field applications. Furthermore, it was shown that, diagnostic performance of the TBS microscopy and the PfHRP2/panLDH-RDTs lack sensitivity. Even among the 15 symptomatic children carrying parasites densities above 5'000 parasites/ μ L, two were missed by TBS microscopy, one by PfHRP2-RDT and three by the panLDH-RDT. In addition to lack of sensitivity, also the specificity of the RDTs was reduced. Two out of 16 *Plasmodium* spp. negative children were wrongly tested positive for *P. falciparum* by PfHRP2-based antigen RDT. False-positive RDTs are common in endemic regions and are likely caused by persisting PfHRP2 antigen circulation post anti-malarial treatment [37, 38]. However, a limitation of our study was that we used an RDT that had not undergone WHO prequalification.

In summary, using samples from two different independent cohorts, including asymptomatic individuals and symptomatic patients, PlasmoPod achieved sensitivities above 80% compared to a highly sensitive RT-qPCR assay. While the initial evaluation of PlasmoPod using samples from 149 individuals shows promising results, further validation studies are necessary to fully assess its performance and reliability. During this initial testing phase, PlasmoPods underwent successful evaluation using NAs extracted through various methods. These methods encompassed commercial column-based extraction kits, the ENAR protocol from archived RDTs, and chelex-based extraction from DBS samples. While the method of NA extraction influences overall sensitivity, it's noteworthy how PlasmoPod demonstrates efficacy across varying levels of NA quantity and purity.

Over the past decade, various NAATs for malaria diagnosis have been published, utilizing different technologies such as qPCR [18, 21, 39–42] and isothermal amplification [43–48]. Each of these assays has its own advantages and disadvantages concerning analytical performance, throughput, simplicity, storability, and laboratory setup requirements. In this study, we introduced the PlasmoPod as a proof-of-concept for a cartridge-based NAAT, aiming to simplify and standardize molecular malaria diagnosis and surveillance. Currently, the PlasmoPod is designed to detect conserved NA sequences present in all human pathogenic *Plasmodium* spp. species. However, future development of this platform should focus on the design and validation of cartridges specifically designed for identifying individual *Plasmodium* spp. species. Given that the diaxoPCR device is a universal qPCR platform, adapting published multiplex species-specific qPCR assays [23, 49, 50] to the PlasmoPod could be a straightforward process. Additionally, separate cartridges could be developed for the molecular characterization of *P. falciparum* isolates. Incorporating assays that enable the detection of *pfhrp2* gene deletions [24, 51, 52] or molecular markers of anti-malarial drug resistance [53, 54] could be particularly valuable in supporting decentralized malaria surveillance efforts. By expanding the capabilities of the PlasmoPod platform to include species identification and molecular characterization of drug resistance, we can enhance its utility and contribute to more effective malaria control and surveillance programs.

Conclusions

In conclusion, we have established a 30-minute sample-to-result RT-qPCR procedure that delivers results with similar diagnostic performance as state-of-the-art RT-qPCR assays for malaria diagnosis. In most malaria endemic regions, molecular malaria diagnostics are only available in centralized laboratories and inaccessible at peripheral health facilities where they are needed most. We believe that the PlasmoPod rapid NAAT can bridge this gap and will enable widespread accessibility of high-quality, sensitive and easy to handle molecular malaria testing at the individual as well as the population level allowing decentralization of testing and surveillance activities.

Supporting information

S1 Fig. Molecular analysis of *Plasmodium* spp. parasites identified in the CAR dataset.

Three different molecular assays were used to (A) screen for *Plasmodium* spp. parasites, (B) identify *Plasmodium* spp. species and (C) detect *pfhrp2/3* gene deletion. Each child is represented in a column stratified according to malaria infection status. Green colors represent negative measurements for the respective qPCR assay, while grey colors were chosen for tests which were not conducted. All tests were run on the Bio-Rad CFX96 qPCR instrument. (TIF)

S1 Data. Supplementary dataset. All data that support the findings of this study are available as a supplementary document uploaded to the journal's website. (XLSX)

Acknowledgments

The authors would like to thank the administrative and laboratory staff of the Paediatric Hospital and University Complex of Bangui (CHUPB) for their support and fruitful cooperation. The authors would also like to thank all patients and their legal guardians who participated in this study. We would also like to thank A&B Professional company located in Lucca, Italy which donated the malaria RDTs used in this study. The authors would also like to acknowledge the contribution of the technical and scientific personnel of Medical Care Development Global Health conducting the yearly malaria indicator surveys on Bioko Island.

Author Contributions

Conceptualization: Philippe Bechtold, Michele Gregorini, Tobias Schindler, Ulrich Vickos.

Data curation: Olivier Tresor Donfack, Tobias Schindler, Ulrich Vickos.

Formal analysis: Philippe Bechtold, Philipp Wagner, Tobias Schindler.

Funding acquisition: Wendelin J. Stark, Wonder P. Phiri, Guillermo A. García.

Investigation: Jean Chrysostome Gody, Marilou Sonia Pagonendji, Olivier Tresor Donfack, Ulrich Vickos.

Methodology: Philippe Bechtold, Philipp Wagner, Salome Hosch, Edwige Régina Kodia-Lenguetama, Marilou Sonia Pagonendji, Christian Nsanzanbana, Ulrich Vickos.

Project administration: Tobias Schindler.

Resources: Wendelin J. Stark, Edwige Régina Kodia-Lenguetama, Olivier Tresor Donfack, Wonder P. Phiri, Guillermo A. García, Christian Nsanzanbana, Claudia A. Daubenberger.

Supervision: Claudia A. Daubenberger, Tobias Schindler, Ulrich Vickos.

Validation: Philippe Bechtold, Michele Gregorini, Tobias Schindler.

Visualization: Philippe Bechtold, Philipp Wagner, Salome Hosch.

Writing – original draft: Philippe Bechtold, Claudia A. Daubenberger, Tobias Schindler.

Writing – review & editing: Philippe Bechtold, Claudia A. Daubenberger, Tobias Schindler.

References

1. Meibalan E, Marti M. Biology of Malaria Transmission. Cold Spring Harbor Perspectives in Medicine. 2017; 7: a025452. <https://doi.org/10.1101/cshperspect.a025452> PMID: 27836912

2. World Health Organization. World malaria report 2021. Geneva: World Health Organization; 2021. Available: <https://apps.who.int/iris/handle/10665/350147>
3. Landier J, Parker DM, Thu AM, Carrara VI, Lwin KM, Bonnington CA, et al. The role of early detection and treatment in malaria elimination. *Malaria Journal*. 2016; 15: 363. <https://doi.org/10.1186/s12936-016-1399-y> PMID: 27421656
4. Picot S, Cucherat M, Bienvenu A-L. Systematic review and meta-analysis of diagnostic accuracy of loop-mediated isothermal amplification (LAMP) methods compared with microscopy, polymerase chain reaction and rapid diagnostic tests for malaria diagnosis. *International Journal of Infectious Diseases*. 2020; 98: 408–419. <https://doi.org/10.1016/j.ijid.2020.07.009> PMID: 32659450
5. Gatton M. *Methods Manual*. World Health Organization. 2012; 1–109.
6. Cunningham J, Jones S, Gatton ML, Barnwell JW, Cheng Q, Chiodini PL, et al. A review of the WHO malaria rapid diagnostic test product testing programme (2008–2018): performance, procurement and policy. *Malaria Journal*. 2019; 18: 387. <https://doi.org/10.1186/s12936-019-3028-z> PMID: 31791354
7. Malaria rapid diagnostic test performance. Results of WHO product testing of malaria RDTs: Round 8 (2016–2018). [cited 20 May 2023]. Available: <https://www.who.int/publications-detail-redirect/9789241514965>
8. Das S, Peck RB, Barney R, Jang IK, Kahn M, Zhu M, et al. Performance of an ultra-sensitive *Plasmodium falciparum* HRP2-based rapid diagnostic test with recombinant HRP2, culture parasites, and archived whole blood samples. *Malaria Journal*. 2018; 17: 118. <https://doi.org/10.1186/s12936-018-2268-7> PMID: 29549888
9. Ochola L, Vounatsou P, Smith T, Mabaso M, Newton C. The reliability of diagnostic techniques in the diagnosis and management of malaria in the absence of a gold standard. *The Lancet Infectious Diseases*. 2006; 6: 582–588. [https://doi.org/10.1016/S1473-3099\(06\)70579-5](https://doi.org/10.1016/S1473-3099(06)70579-5) PMID: 16931409
10. Bosman A, Cunningham J, Lindblade KA, Noor A. WHO Technical Consultation on research requirements to support policy recommendations on highly sensitive malaria diagnostic tests. *World Malaria Report*. 2018; 1–34.
11. Verma AK, Bharti PK, Das A. HRP-2 deletion: a hole in the ship of malaria elimination. *Lancet Infect Dis*. 2018; 18: 826–827. [https://doi.org/10.1016/S1473-3099\(18\)30420-1](https://doi.org/10.1016/S1473-3099(18)30420-1) PMID: 30064667
12. Roth JM, Korevaar DA, Leeflang MMG, Mens PF. Molecular malaria diagnostics: A systematic review and meta-analysis. *Critical Reviews in Clinical Laboratory Sciences*. 2016; 53: 87–105. <https://doi.org/10.3109/10408363.2015.1084991> PMID: 26376713
13. Mwenda MC, Fola AA, Ciobotariu II, Mulube C, Mambwe B, Kasaro R, et al. Performance evaluation of RDT, light microscopy, and PET-PCR for detecting *Plasmodium falciparum* malaria infections in the 2018 Zambia National Malaria Indicator Survey. *Malaria Journal*. 2021; 20: 386. <https://doi.org/10.1186/s12936-021-03917-6> PMID: 34583692
14. Bousema T, Okell L, Felger I, Drakeley C. Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nat Rev Microbiol*. 2014; 12: 833–840. <https://doi.org/10.1038/nrmicro3364> PMID: 25329408
15. Amir A, Cheong F-W, De Silva JR, Lau Y-L. Diagnostic tools in childhood malaria. *Parasites & Vectors*. 2018; 11: 53. <https://doi.org/10.1186/s13071-018-2617-y> PMID: 29361963
16. Gregorini M, Mikutis G, Grass RN, Stark WJ. Small-Size Polymerase Chain Reaction Device with Improved Heat Transfer and Combined Feedforward/Feedback Control Strategy. *Ind Eng Chem Res*. 2019; 58: 9665–9674. <https://doi.org/10.1021/acs.iecr.9b01209>
17. Bechtold P, Wagner P, Hosch S, Siegrist D, Ruiz-Serrano A, Gregorini M, et al. Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable peakPCR Platform. *Anal Chem*. 2021; 93: 16350–16359. <https://doi.org/10.1021/acs.analchem.1c02368> PMID: 34852455
18. Kamau E, Tolbert LS, Kortepeter L, Pratt M, Nyakoe N, Muringo L, et al. Development of a highly sensitive genus-specific quantitative reverse transcriptase real-time PCR assay for detection and quantitation of *plasmodium* by amplifying RNA and DNA of the 18S rRNA genes. *J Clin Microbiol*. 2011; 49: 2946–2953. <https://doi.org/10.1128/JCM.00276-11> PMID: 21653767
19. Guirou EA, Schindler T, Hosch S, Donfack OT, Yoboue CA, Krähenbühl S, et al. Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Scientific Reports*. 2020; 10: 12305. <https://doi.org/10.1038/s41598-020-69268-5> PMID: 32703999
20. Yoboue CA, Hosch S, Donfack OT, Guirou EA, Nlavo BM, Ayekaba MO, et al. Characterising co-infections with *Plasmodium* spp., *Mansonella perstans* or *Loa loa* in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests. *PLOS Neglected Tropical Diseases*. 2022; 16: 1–18. <https://doi.org/10.1371/journal.pntd.0009798> PMID: 35100277

21. Kamau E, Alemayehu S, Feghali KC, Saunders D, Ockenhouse CF. Multiplex qPCR for Detection and Absolute Quantification of Malaria. Langsley G, editor. PLoS ONE. 2013; 8: e71539. <https://doi.org/10.1371/journal.pone.0071539> PMID: 24009663
22. Zainabadi K, Adams M, Han ZY, Lwin HW, Han KT, Ouattara A, et al. A novel method for extracting nucleic acids from dried blood spots for ultrasensitive detection of low-density *Plasmodium falciparum* and *Plasmodium vivax* infections. *Malaria Journal*. 2017; 16: 377. <https://doi.org/10.1186/s12936-017-2025-3> PMID: 28923054
23. Schindler T, Robaina T, Sax J, Bieri JR, Mpina M, Gondwe L, et al. Molecular monitoring of the diversity of human pathogenic malaria species in blood donations on Bioko Island, Equatorial Guinea. *Malaria Journal*. 2019; 18: 9. <https://doi.org/10.1186/s12936-019-2639-8> PMID: 30646918
24. Schindler T, Deal AC, Fink M, Guirou E, Moser KA, Mwakasungula SM, et al. A multiplex qPCR approach for detection of pfrp2 and pfrp3 gene deletions in multiple strain infections of *Plasmodium falciparum*. *Scientific Reports*. 2019; 9: 13107. <https://doi.org/10.1038/s41598-019-49389-2> PMID: 31511562
25. Munro JB, Jacob CG, Silva JC. A Novel Clade of Unique Eukaryotic Ribonucleotide Reductase R2 Subunits is Exclusive to Apicomplexan Parasites. *Journal of Molecular Evolution*. 2013; 77: 92–106. <https://doi.org/10.1007/s00239-013-9583-y> PMID: 24046025
26. Padley DJ, Heath AB, Sutherland C, Chiodini PL, Baylis SA. Establishment of the 1st World Health Organization International Standard for *Plasmodium falciparum* DNA for nucleic acid amplification technique (NAT)-based assays. *Malaria Journal*. 2008; 7: 139. <https://doi.org/10.1186/1475-2875-7-139> PMID: 18652656
27. Sirima SB, Mordmüller B, Milligan P, Ngoa UA, Kironde F, Atuguba F, et al. A phase 2b randomized, controlled trial of the efficacy of the GMZ2 malaria vaccine in African children. *Vaccine*. 2016; 34: 4536–4542. <https://doi.org/10.1016/j.vaccine.2016.07.041> PMID: 27477844
28. White MT, Verity R, Griffin JT, Asante KP, Owusu-Agyei S, Greenwood B, et al. Immunogenicity of the RTS,S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial. *The Lancet Infectious Diseases*. 2015; 15: 1450–1458. [https://doi.org/10.1016/S1473-3099\(15\)00239-X](https://doi.org/10.1016/S1473-3099(15)00239-X) PMID: 26342424
29. A Phase 3 Trial of RTS,S/AS01 Malaria Vaccine in African Infants. *New England Journal of Medicine*. 2012; 367: 2284–2295. <https://doi.org/10.1056/NEJMoa1208394> PMID: 23136909
30. Fleming KA, Horton S, Wilson ML, Atun R, DeStigter K, Flanigan J, et al. The Lancet Commission on diagnostics: transforming access to diagnostics. *The Lancet*. 2021; 398: 1997–2050. [https://doi.org/10.1016/S0140-6736\(21\)00673-5](https://doi.org/10.1016/S0140-6736(21)00673-5) PMID: 34626542
31. Global technical strategy for malaria 2016–2030, 2021 update. [cited 9 Dec 2022]. Available: <https://www.who.int/publications-detail-redirect/9789240031357>
32. McMorro ML, Aidoo M, Kachur SP. Malaria rapid diagnostic tests in elimination settings—can they find the last parasite? *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2011; 17: 1624. <https://doi.org/10.1111/j.1469-0691.2011.03639.x> PMID: 21910780
33. Stuck L, Fakih BS, Al-mafazy AH, Hofmann NE, Holzschuh A, Grossenbacher B, et al. Malaria infection prevalence and sensitivity of reactive case detection in Zanzibar. *International Journal of Infectious Diseases*. 2020; 97: 337–346. <https://doi.org/10.1016/j.ijid.2020.06.017> PMID: 32534138
34. PCR incidence of *Plasmodium falciparum* infections in cohort samples over time during a malaria MDA randomized control trial in southern province Zambia | Cochrane Library. [cited 31 May 2023]. <https://doi.org/10.1002/central/CN-01462021>
35. Kerr G, Robinson LJ, Russell TL, Macdonald J. Lessons for improved COVID-19 surveillance from the scale-up of malaria testing strategies. *Malaria Journal*. 2022; 21: 223. <https://doi.org/10.1186/s12936-022-04240-4> PMID: 35858916
36. Hewawasam E, Liu G, Jeffery DW, Gibson RA, Muhlhausler BS. Estimation of the Volume of Blood in a Small Disc Punched From a Dried Blood Spot Card. *European Journal of Lipid Science and Technology*. 2018; 120: 1700362. <https://doi.org/10.1002/ejlt.201700362>
37. Dalrymple U, Arambepola R, Gething PW, Cameron E. How long do rapid diagnostic tests remain positive after anti-malarial treatment? *Malaria Journal*. 2018; 17: 228. <https://doi.org/10.1186/s12936-018-2371-9> PMID: 29884184
38. Hosch S, Yoboue CA, Donfack OT, Guirou EA, Dangy J-P, Mpina M, et al. Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment. *Malaria Journal*. 2022; 21: 23. <https://doi.org/10.1186/s12936-022-04043-7> PMID: 35073934

39. Kamau E, Alemayehu S, Feghali KC, Juma DW, Blackstone GM, Marion WR, et al. Sample-ready multiplex qPCR assay for detection of malaria. *Malar J*. 2014; 13: 158. <https://doi.org/10.1186/1475-2875-13-158> PMID: 24767409
40. 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 Medicine*. 2015; 12: e1001788. <https://doi.org/10.1371/journal.pmed.1001788> PMID: 25734259
41. Mangold KA, Manson RU, Koay ESC, Stephens L, Regner M, Thomson RB, et al. Real-Time PCR for Detection and Identification of *Plasmodium* spp. *J Clin Microbiol*. 2005; 43: 2435–2440. <https://doi.org/10.1128/JCM.43.5.2435-2440.2005> PMID: 15872277
42. Carlier L, Baker SC, Huwe T, Yewhalaw D, Haileselassie W, Koepfli C. qPCR in a suitcase for rapid *Plasmodium falciparum* and *Plasmodium vivax* surveillance in Ethiopia. *PLOS Global Public Health*. 2022; 2: e0000454. <https://doi.org/10.1371/journal.pgph.0000454> PMID: 36962431
43. Poon LL, Wong BW, Ma EH, Chan KH, Chow LM, Abeyewickreme W, et al. Sensitive and Inexpensive Molecular Test for Falciparum Malaria: Detecting *Plasmodium falciparum* DNA Directly from Heat-Treated Blood by Loop-Mediated Isothermal Amplification. *Clinical Chemistry*. 2006; 52: 303–306. <https://doi.org/10.1373/clinchem.2005.057901> PMID: 16339303
44. Sirichaisinthop J, Buates S, Watanabe R, Han E-T, Suktawonjaroenpon W, Krasaesub S, et al. Evaluation of Loop-Mediated Isothermal Amplification (LAMP) for Malaria Diagnosis in a Field Setting. *Am J Trop Med Hyg*. 2011; 85: 594–596. <https://doi.org/10.4269/ajtmh.2011.10-0676> PMID: 21976556
45. Cook J, Aydin-Schmidt B, González IJ, Bell D, Edlund E, Nassor MH, et al. Loop-mediated isothermal amplification (LAMP) for point-of-care detection of asymptomatic low-density malaria parasite carriers in Zanzibar. *Malaria Journal*. 2015; 14: 43. <https://doi.org/10.1186/s12936-015-0573-y> PMID: 25627037
46. Kemleu S, Guelig D, Eboumbou Moukoko C, Essangui E, Diesburg S, Mouliom A, et al. A Field-Tailored Reverse Transcription Loop-Mediated Isothermal Assay for High Sensitivity Detection of *Plasmodium falciparum* Infections. *PLoS One*. 2016; 11: e0165506. <https://doi.org/10.1371/journal.pone.0165506> PMID: 27824866
47. Kersting S, Rausch V, Bier FF, von Nickisch-Rosenegk M. Rapid detection of *Plasmodium falciparum* with isothermal recombinase polymerase amplification and lateral flow analysis. *Malaria Journal*. 2014; 13: 99. <https://doi.org/10.1186/1475-2875-13-99> PMID: 24629133
48. Lalremruata A, Nguyen TT, McCall MBB, Mombongo-Goma G, Agnandji ST, Adegnikaa AA, et al. Recombinase Polymerase Amplification and Lateral Flow Assay for Ultrasensitive Detection of Low-Density *Plasmodium falciparum* Infection from Controlled Human Malaria Infection Studies and Naturally Acquired Infections. *J Clin Microbiol*. 2020; 58: e01879–19. <https://doi.org/10.1128/JCM.01879-19> PMID: 32102854
49. Shokoples SE, Ndao M, Kowalewska-Grochowska K, Yanow SK. Multiplexed Real-Time PCR Assay for Discrimination of *Plasmodium* Species with Improved Sensitivity for Mixed Infections. *J Clin Microbiol*. 2009; 47: 975–980. <https://doi.org/10.1128/JCM.01858-08> PMID: 19244467
50. Rougemont M, Van Saanen M, Sahli R, Hinrikson HP, Bille J, Jatou K. Detection of Four *Plasmodium* Species in Blood from Humans by 18S rRNA Gene Subunit-Based and Species-Specific Real-Time PCR Assays. *Journal of Clinical Microbiology*. 2004; 42: 5636–5643. <https://doi.org/10.1128/JCM.42.12.5636-5643.2004> PMID: 15583293
51. Grignard L, Nolder D, Sepúlveda N, Berhane A, Mihreteab S, Kaaya R, et al. A novel multiplex qPCR assay for detection of *Plasmodium falciparum* with histidine-rich protein 2 and 3 (pfrp2 and pfrp3) deletions in polyclonal infections. *EBioMedicine*. 2020; 55: 102757. <https://doi.org/10.1016/j.ebiom.2020.102757> PMID: 32403083
52. Vera-Arias CA, Holzschuh A, Oduma CO, Badu K, Abdul-Hakim M, Yukich J, et al. High-throughput *Plasmodium falciparum* hrp2 and hrp3 gene deletion typing by digital PCR to monitor malaria rapid diagnostic test efficacy. *Kana BD, editor. eLife*. 2022; 11: e72083. <https://doi.org/10.7554/eLife.72083> PMID: 35762586
53. Kamau E, Alemayehu S, Feghali KC, Tolbert LS, Ogutu B, Ockenhouse CF. Development of a TaqMan Allelic Discrimination assay for detection of single nucleotide polymorphisms associated with anti-malarial drug resistance. *Malar J*. 2012; 11: 23. <https://doi.org/10.1186/1475-2875-11-23> PMID: 22264294
54. Nsanzabana C, Ariey F, Beck H-P, Ding XC, Kamau E, Krishna S, et al. Molecular assays for antimalarial drug resistance surveillance: A target product profile. *PLoS One*. 2018; 13: e0204347. <https://doi.org/10.1371/journal.pone.0204347> PMID: 30235327

3.4. PHARE: A bioinformatics pipeline for compositional profiling of multiclonal *Plasmodium falciparum* infections from long-read Nanopore sequencing data

Short running title: Finding *P. falciparum* haplotypes in multiclonal infections

Salome Hosch^{1,2,†}, Philipp Wagner^{1,2,*†}, Johanna Nouria Giger^{1,2}, Nina Dubach^{1,2}, Elis Saavedra^{1,2}, Carlo Federico Perno³, Jean-Chrysostome Gody⁴, Marilou Sonia Pagonendji⁵, Carine Ngoagoni⁶, Christophe Ndoua⁷, Christian Nsanzabana^{1,2}, Ulrich Vickos³, Claudia Daubenberger^{1,2} and Tobias Schindler^{1,2}

¹ University of Basel, Petersplatz 1, 4001 Basel, Switzerland

² Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland

³ Ospedale Pediatrico Bambino Gesù, Piazza di Sant'Onofrio, 4, 00165 Roma, Italy

⁴ Pediatric University Hospital Centre of Bangui, Bangui, Central African Republic

⁵ Laboratory of Parasitology, Institut Pasteur of Bangui, Bangui, Central African Republic

⁶ Medical Entomology Unit, Institut Pasteur of Bangui, Bangui, Central African Republic

⁷ National Malaria Control Program, Ministry of Health, Bangui, Central African Republic

*Corresponding author: E-mail: philipp.wagner@swisstph.ch, Phone: +41 61 284 88 43

†These authors contributed equally

Submitted to *Journal of Antimicrobial Chemotherapy*

Synopsis

Background: The emergence of drug-resistant clones of *Plasmodium falciparum* is a major public health concern, and the ability to detect and track the spread of these clones is crucial for effective malaria control and treatment. However, in endemic settings, malaria infected people often carry multiple *P. falciparum* clones simultaneously making it likely to miss drug resistant clones using traditional molecular typing methods.

Objectives: Our goal was to develop a bioinformatics pipeline for compositional profiling in multiclonal *P. falciparum* samples, sequenced using the Oxford Nanopore Technologies MinION platform.

Methods: We developed the “Finding *P. falciparum* haplotypes with resistance mutations in polyclonal infections” (PHARE) pipeline using existing bioinformatics tools and custom scripts written in python. PHARE was validated on three control datasets containing *P. falciparum* DNA of four laboratory strains at varying mixing ratios. Additionally, the pipeline was tested on clinical samples from children admitted to a paediatric hospital in the Central African Republic.

Results: The PHARE pipeline achieved high recall and accuracy rates in all control datasets. The pipeline can be used on any gene and was tested with amplicons of the *P. falciparum* drug resistance marker genes *pf dhps*, *pf dhfr*, and *pf K13*.

Conclusions: The PHARE pipeline helps to provide a more complete picture of drug resistance in the circulating *P. falciparum* population and can help to guide treatment recommendations. PHARE is freely available under the GNU Lesser General Public License v3.0 on GitHub: <https://github.com/Fippu/PHARE>

Introduction

Malaria is one of the top three infectious diseases globally, with an estimated 247 million cases reported in 2021 and continuing efforts are needed for its control¹. Artemisinin combination therapy (ACT) is the primary treatment for *Plasmodium falciparum* (*Pf*), the deadliest malaria species. ACT consists of an artemisinin derivative combined with one of six partner drugs or drug combinations^{2,3}. Delivered together, the fast-acting artemisinin component rapidly kills most of the asexual blood stage parasites within a few days, and the longer-acting partner drug clears the residual parasite populations^{2,4}. Recent reports have confirmed the increased prevalence of *Pf* strains showing reduced clearance rate after ACT treatment indicative of partial resistance development in Africa⁵. Modelling of widespread resistance to both artemisinin and a partner drug in Africa indicated a potential scenario of 16 million additional annual malaria cases⁶ resulting in nearly 80,000 additional malaria deaths each year, in addition to economic losses of 1 billion US dollar⁷.

Adaptations of guidelines for treatment with antimalarial drugs are developed based on parasitological data determined in therapeutic efficacy studies^{8,9}. Alternatively, sensitivity to antimalarial drugs can be assessed *in vitro* with IC₅₀ studies¹⁰ and sensitivity to artemisinin by performing ring stage survival assays¹¹. However, since both *in vitro* and *in vivo* studies are expensive and labour-intensive, molecular markers, such as single nucleotide polymorphism (SNPs) and copy number variations (CNV), can be used as indicators of resistance to a particular antimalarial drug^{12,13}. These markers can be identified through rapid and relatively inexpensive molecular biology techniques such as amplicon sequencing¹⁴ and quantitative PCR-based technologies^{15,16}. Incorporating blood sampling and storage on filter papers enables the analysis of samples from remote settings for the presence of drug resistance markers.^{17,18} Using molecular markers may be more appropriate for approaches where samples are collected regularly and analysed more frequently over a longer time period to monitor potential changes in drug resistance patterns over time¹⁹. Longitudinal monitoring of antimalarial drug efficacy supported with molecular surveillance on resistance markers is essential for making data-driven decisions on combination therapy guidelines in a country or region²⁰.

Malaria infections often consist of multiple parasite clones, known as multiclonal infections. Multiclonal malaria infections are common in areas with high malaria transmission rates²¹ and can complicate the identification of molecular markers, especially if some clones are present at lower densities within a sample. Compositional profiling aims to characterize the different clones present in the infection and their relative proportions.

Advances in single molecule sequencing technologies and bioinformatics analyses have improved the ability to detect and analyse genetic variation within the parasite population, enabling the identification of molecular markers even in multiclonal infections²²⁻²⁴. Additionally, using novel long read sequencing methods such as Oxford Nanopore Technologies (ONT) MinION, full length genes can be sequenced in one read. All SNPs present in this gene can therefore be analysed as one haplotype instead of SNP sites being analysed individually. However, most bioinformatics tools for haplotype calling have been developed for Illumina short read data and are difficult to adapt to single-molecule long-read sequencing data derived from ONT sequencing. Therefore, we developed the “Finding *P. falciparum* haplotypes with resistance mutations in polyclonal infections” (PHARE) pipeline, which is designed to identify the molecular markers of drug resistance of all detectable clones in a sample, leveraging the technological advancements made available through long read sequencing.

Materials and Methods

Clinical and laboratory datasets to validate the PHARE pipeline

To develop and test our bioinformatics analysis pipeline, *Pf in vitro* culture strains Dd2, HB3, K1 and NF54 were mixed, followed by DNA extraction and ONT sequencing to generate three datasets (supplementary methods). Clinical testing and comparison with conventional methods was done with blood samples collected at the Paediatric Hospital and University Complex of Bangui (CHUPB) in the Central African Republic from children with fever between March 8th and 13th, 2021. The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics and Scientific Committee from the University of Bangui (approval n°3/UB/FACSS/CSCVPER/PER) and by the Ministry of Health of the Central African Republic (approval n°0277/MSPP/CAB/DGSP/DMPM/SMEE from 5 august 2002) as part of the communicable and endemic diseases surveillance diagnostic program. The patients were informed about the objectives of the study and nature of their participation. Then, written and signed informed consent was obtained from the parents on behalf of their children.

PCR amplification and Sanger sequencing

The antimalarial drug resistance markers *pfdhps*, *pfdhfr* and *pfk13* were amplified using previously established nested PCR protocols^{27,28} and sent to Microsynth AG (Balgach, Switzerland) for Sanger sequencing using the same primers as the nested PCR.

PCR amplification, library preparation and ONT sequencing

The full-length drug resistance genes *pfdhfr*, *pfdhps* and *pfk13* were amplified separately using published primers²⁹. Amplified DNA was purified using 0.8x volume AMPure® XP beads (Beckman Coulter) and quantified with the Qubit dsDNA HS Assay Kit (Invitrogen). For the control datasets using laboratory strains, the Native Barcoding Kit 96 (ONT SQK-NBD112.96) and R10.4 flow cell (ONT FLO-MIN112) were used. For the clinical samples, Ligation Sequencing Kit (ONT SQK-LSK109), Native Barcoding Expansion 13-24 (EXP-NBD114) and R9.4.1 flow cell (ONT FLO-MIN106D) were used. Sequencing libraries were prepared according to manufacturer's instruction. Briefly, 200 fmol amplicon DNA were end-prepped, unique barcodes were ligated, and sequencing adapters were ligated. The libraries were loaded onto the flow cell and sequenced on the MinION Mk1C.

Data analysis and availability

Sanger sequencing data was analysed in SeqScape v4.0 (Applied Biosystems). The sequences and basecalling were manually checked. Mutations were marked and a mutation report was generated.

The mutation report was then saved as a .csv file and imported into R. ONT sequencing data was extracted as raw data in FAST5 format. Basecalling was done using the “super high accuracy” basecalling in guppy (v6.4.2). To test the impact of the guppy basecalling model, the “high accuracy” and the “fast” model were also tested with the second control dataset. All sequencing data was uploaded to the NCBI Sequence Read Archive under BioProject ID PRJNA974955.

Programming of PHARE pipeline

The pipeline uses the tools Filtlong (v0.2.1) ³⁰, SeekDeep extractor (v3.0.1) ³¹, minimap2 (v2.24) ³², samtools (v1.16.1) ³³ and pysam (v0.20.0) ³⁴. The code is written in bash, R (v4.2.2) and python (v3.7.11).

Results

We developed PHARE, a novel multi-step bioinformatics pipeline to comprehensively profile the composition of genetic drug resistance markers in multiclonal samples. To fine tune parameters and validate the PHARE pipeline, we analysed a total of four datasets. The PHARE pipeline is documented and available for download and modification under the GNU Lesser General Public License v3.0 on GitHub: <https://github.com/Fippu/PHARE>

PHARE is a multi-step pipeline combining quality control (1), SNP identification (2) and graphical representation of haplotypes (3). Adjustable parameters and required reference files are listed in Table S1.

(1) The first step of the pipeline is to filter the input reads based on read quality and read length using Filtlong. Optionally, amplicons get extracted using SeekDeep extractor based on primer recognition and read length. Following this, an alignment to the provided reference sequence is done using minimap2.

(2) We developed a python-based algorithm that reduces the complexity of the data, by first defining SNPs, then discarding the remaining information of the sequenced read. This approach improves the analysis of error prone ONT reads. SNPs are found by iterating over the alignment and finding nucleotide sites, where at least two different bases are found with a frequency above a predefined threshold (SNP selection cut-off parameter, default 10%) or where the majority of bases differ from the nucleotide provided in the reference.

ONT has higher error rates in low complexity regions since the translocation speed of DNA through the nanopore is not constant and the basecaller cannot accurately determine the length of homopolymers where the electrical signal remains the same ³⁵. During alignment, minimap2 introduces gaps at the ends of these homopolymers to correct for the erroneous reads. This leads to a lower coverage and a higher error rate at these sites. Therefore, SNPs at sites below a coverage minimum of 80% are filtered out.

In an optional step, silent mutations which do not affect the amino acid sequence, are removed. When investigating phenotypic variations such as drug resistance, silent SNPs are of limited interest and can reduce the yield of the pipeline, since every read must fulfil a minimal quality score (minqual parameter, default Phred score of 15) at each SNP site. If at any SNP site of a read a gap is present or the quality score at that position is below the threshold the read is excluded from analysis. This leads to a linear decline in the number of reads passing quality control when adding SNPs at random sites (Figure S1).

The core step of the PHARE pipeline is the extraction of haplotypes, defined here as a set of genetic determinants located on a single gene, from each read. Like the detection of SNP sites, this is an iterative approach. The algorithm loops through all reads which did align to the reference and concatenates all the SNP sites to result in one haplotype per single read. In this step, reads with gaps at any of the SNP sites and reads with insufficient base quality are removed from further analysis. The output of the pipeline is a tab-separated file for each sample, which lists all present haplotypes and their respective read counts.

(3) The tab-separated files with the results of the PHARE pipeline for each sample are analysed in an R script to generate a visual output of the data. In this step, the minimum haplotype frequency filter (default 5%) is applied, to filter out low abundance haplotypes that are considered noise.

PHARE can identify SNPs with 100% accuracy and recall

To tune the parameters and test the limits of detection of the PHARE pipeline we prepared simulated multiclonal samples from ONT sequencing data of three *in vitro* culture strains with mixing ratios ranging from 1:99 to 99:1. In addition to the limits of detection, the SNP finding process was also validated on this simulated dataset with a total of 39 samples. A SNP selection cut-off of 10% was used. SNPs in the *pfdhfr*, *pfdhps* and *pfK13* genes were correctly identified in all samples, resulting in 100% accuracy and recall. Figure 1a is a graphical representation of the nucleotide distribution and SNP finding process in *pfdhps*, using the 50% HB3 and 50% K1 mixture sample: At two nucleotide positions, one of the minority nucleotides was more frequent than the SNP selection cut-off of 10% and these positions got correctly identified as SNPs. A dip in coverage below the SNP coverage minimum of 80% (Figure 1a) can be observed in the three homopolymer regions of the *pfdhps* gene starting at nucleotide positions 983, 1727 and 2190, respectively (Figure 1b). By using a minimum haplotype frequency of 5%, the minor clone could be distinguished from noise, hence no false haplotypes were identified in all three genes (Figure 1c).

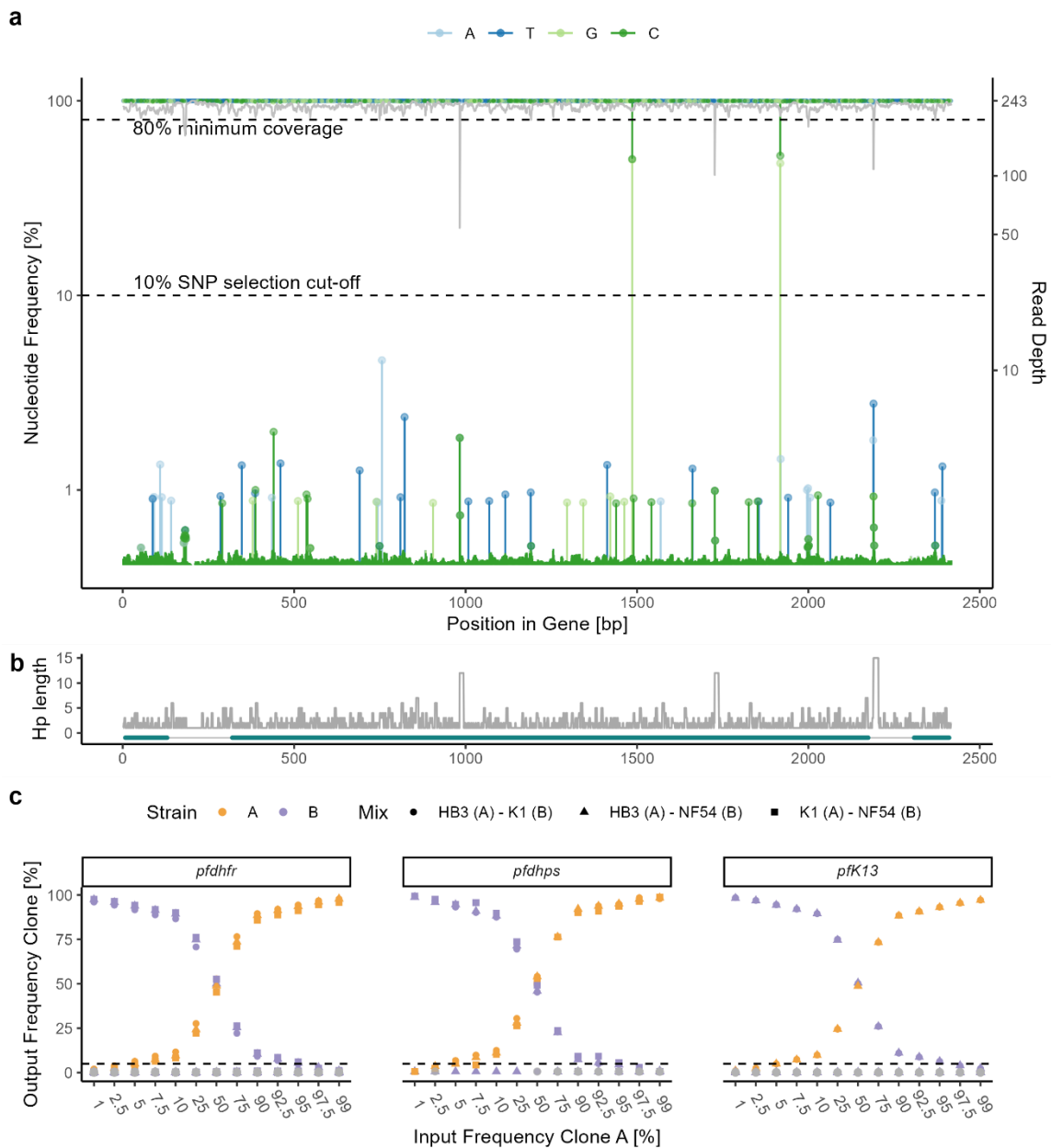


Figure 1. Simulated multiclonal infections were created by mixing reads from laboratory strains HB3, K1 and NF54 at different ratios ranging from 1% to 99%. a) SNP calling in *pfdhps* with a mixture of 50% HB3 and 50% K1 reads. The nucleotide position is on the x-axis. The read depth is shown in grey and the relative frequency of different bases in the respective colour. The lower dotted line is the SNP selection cut-off, which was set at 10%. The upper dotted line is the minimum coverage which was set at 80%. b) The homopolymer length over the whole gene based on the reference sequence is shown in grey. Below are the exons (thick dark-green lines). c) The input frequencies are compared against the output frequencies of the plot. A minimum haplotype frequency of 5% (dotted black line) was set to distinguish true haplotypes from noise (grey).

The PHARE pipeline performs better with more accurate basecalling models and can distinguish true haplotypes from sequencing errors

In a second control experiment, culture media of two *in vitro* culture strains (NF54 and Dd2) with highly similar parasitaemia were mixed at different concentrations. *Pfdhps* was amplified in three technical replicates and ONT sequenced to test the impact of the guppy basecalling model on the pipeline's output and to test the quantitative performance and limit of detection of the PHARE pipeline. The performance of the PHARE pipeline improves with more accurate basecalling (Figure 2a). Haplotypes could be recalled with a high accuracy and low variability when using the "super high accuracy" basecalling. The three replicates had very little variation (IQR 0.95%) and haplotypes of the laboratory strains 436f/613s (Dd2) and S436/A613 (NF54) were correctly identified. No false haplotypes were observed using the "super high accuracy" basecalling. Variation between replicates increased when changing the basecalling model to "high accuracy" and even more when using the "fast" basecalling model. Furthermore, SNP calling could not be reliably done using the "fast" basecalling dataset, since too much noise was masking the true SNPs and the coverage at many nucleotides was below 80% (Figure S2). Therefore, the same SNPs which were found in the "super high accuracy" data were manually set for the "fast" basecalling in this experiment. All subsequent analyses were performed using "super high accuracy" basecalling.

Over all concentrations, the percentage of reads mapping to the NF54 haplotype was lower than the frequency of NF54 in the original sample. To test for a bias in the pipeline, we compared the results generated with PHARE to mapping the full length reads directly to the *pfdhps* reference sequences of the two laboratory strains NF54 and Dd2 with minimap2. The haplotypes found by both methods have a high correlation ($R^2 > 0.99$) (Figure S3).

We determined the impact of the minqual parameter of the pipeline which sets both, the minimal quality that must be reached at the SNP sites, as well as the minimal quality of the whole read. The minimal quality is estimated by the guppy basecaller and corresponds to a Phred score Q which is defined as $Q = -10 \times \log_{10}(P)$, where P is the sequencing error probability. As an example, a Phred score of 10 corresponds to 10% error probability. The effect of the minqual parameter on the ratio of noise haplotypes compared to true haplotypes and its influence on the read depth is shown in Figure 2b. Increasing minqual led to a reduction in noise, though the benefits were decreasing. Meanwhile the read depth declined linearly with increasing minqual, with a rapid decline after minqual 20. We settled on a minqual of 15 for further analysis. We then assessed the effect of the minimum haplotype frequency which differentiates true haplotypes from false haplotypes on precision, recall and f1 score (Figure 2c). Precision reached its maximum at a minimum haplotype frequency of over 3.6%, while the

recall reached a plateau between 5.0% and 9.8%. We therefore decided on a minimum haplotype frequency of 5% for subsequent experiments.

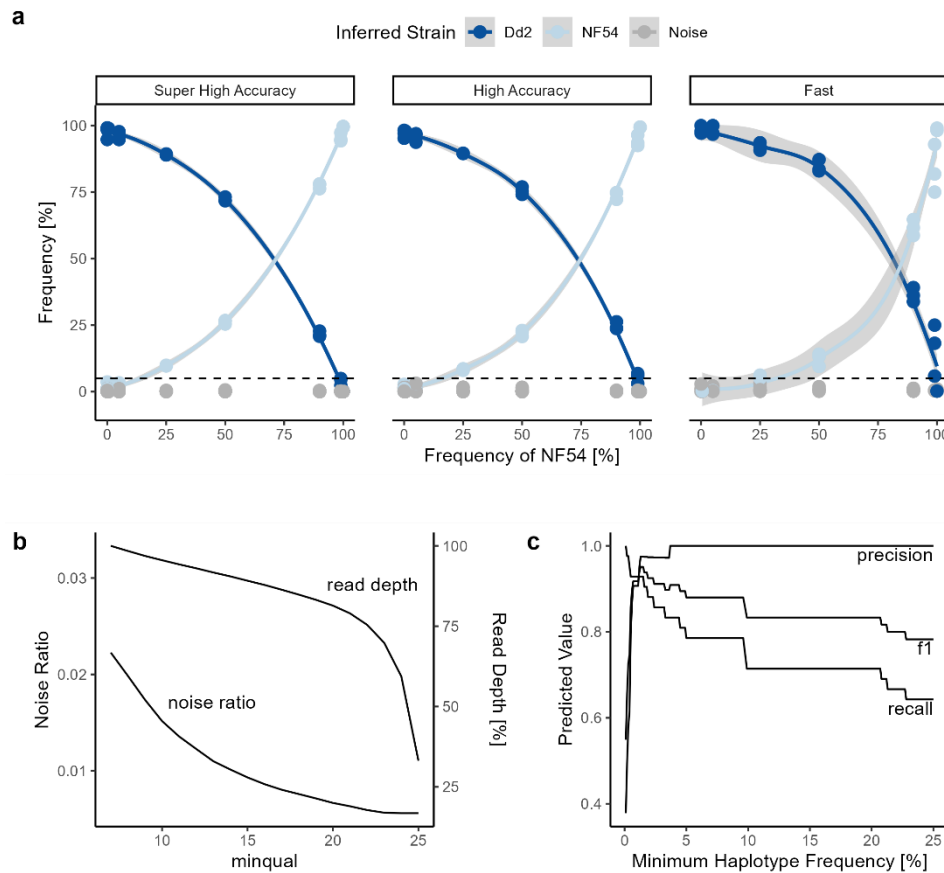


Figure 2. Data from *pfdhps* experiment with NF54 and Dd2 in vitro culture strains at different mixing ratios. a) Using a minimal quality score of 15 and three different basecalling models. The colours indicate the strain that was inferred based on the haplotypes found by the PHARE pipeline (436f/613s [fs] in Dd2 and S436/A613 [SA] in NF54). b) Read depth and noise ratio (the fraction of noise reads as part of the total amount of reads) at different minimum read qualities (Phred scores). c) Precision, recall and f1 score with the minqual parameter set to 15. Both, b) and c) use data generated by “super high accuracy” basecalling.

Compositional profiling of four haplotypes is possible with the PHARE pipeline

We tested the ability of the PHARE pipeline to detect more than two haplotypes in a single sample, using the three drug resistance marker genes *pfdhfr*, *pfdhps* and *pfk13*. DNA of the four *in vitro* culture strains NF54, Dd2, HB3 and K1 was mixed in equimolar ratios for two or four strains and amplified by PCR. Figure 3 shows the PHARE pipeline outputs for mixes of NF54+Dd2, HB3+K1 and all four culture strains combined. All expected haplotypes and no false haplotypes were detected. The estimated haplotype frequency did not always correspond to the input amount of DNA prior to PCR. For example,

in the equimolar mixture of NF54 and Dd2, the haplotype corresponding to NF54 was found with a relative frequency of 67.3% in *pfdhfr* but only 27.6% in *pfdhps*. In *pfdhps* and *pfdhfr* the four haplotypes differ by one to three SNPs leading to an amino acid exchange, while in *pfK13* only HB3 differs from the other three strains by one single SNP making a distinction not always possible.

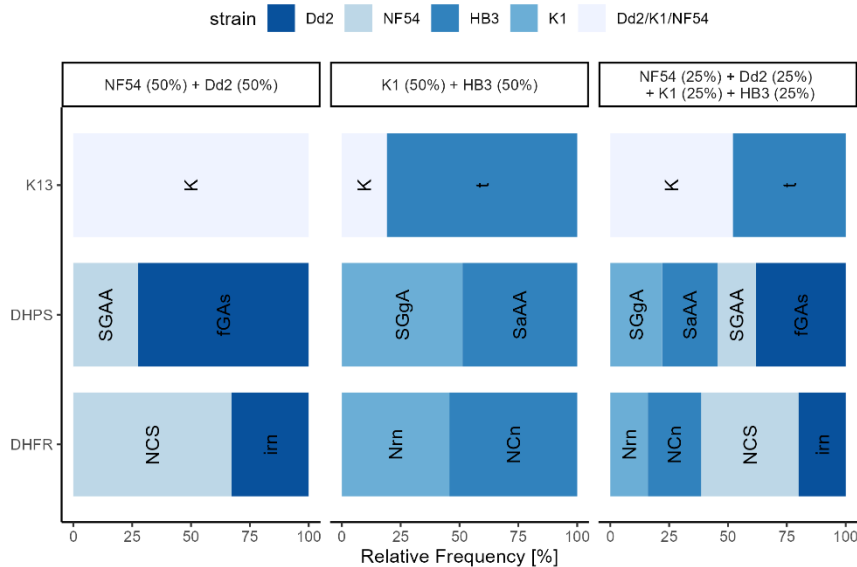


Figure 3. PHARE pipeline output for *pfdhfr*, *pfdhps* and *pfK13*. The width of the bars indicates the relative frequency of the haplotype written on the bar as determined by the PHARE pipeline. The colour indicates the laboratory strain matching the haplotype found by the pipeline. SNP sites are N51i, C59r and S108n for *pfdhfr*; S436f, G437a, A581g and A613s for *pfdhps* and K189t for *pfK13*.

The SNPs detected by the PHARE pipeline share a high agreement with conventional Sanger sequencing, while enabling the detection of complex haplotypes

Next, we used a sample set of twelve dried blood spots from the Central African Republic to understand the performance on clinical samples sequenced on ONT flow cell R9.4.1. Since older chemistry was used for sequencing, adapted parameters were used for SNP calling, namely the coverage minimum was set to 50%.

PHARE and Sanger sequencing detected the same three and five SNP sites in *pfdhfr* and *pfdhps*, respectively (Table 1). In *pfdhfr*, the same variants were detected for all 12 samples. At the G437a SNP site in *pfdhps*, PHARE and Sanger sequencing disagreed in three samples: the first sample was G437/437a with PHARE and 437a with Sanger, the second sample was G437/437a with PHARE and G437 with Sanger and the third sample was G437 with PHARE and G437/437a with Sanger. In *pfK13* both, PHARE and Sanger found no SNPs in the 849 bp region analysed by Sanger sequencing. However, outside of this region the K189n and K189t SNPs were found by PHARE.

Marker	SNP	Agreement over 12 Samples	wt	mut	mixed	
<i>pfdhfr</i>	N51i	100%	1	11	0	
	C59r	100%	0	12	0	
	S108n	100%	0	12	0	
<i>pfdhps</i>	I431v	100%	10	1	1	
	S436a	100%	2	6	4	
	G437a	PHARE	75%	2	5	5
		Sanger		2	6	4
	K540e	100%	1	1	2	
	A581g	100%	11	0	1	
<i>pfK13</i>	wild type*	100%	12	0	0	

Table 1. Comparison of drug resistance SNPs detected with Sanger and ONT sequencing for twelve clinical isolates from the Central African Republic. * Outside of the 849 bp region analysed by Sanger, the K189n and K189t SNPs were found.

To demonstrate the default output of the pipeline, the results for *pfdhps* are shown in Figure 4. Using the PHARE pipeline we were able to find six different *pfdhps* haplotypes based on five SNP sites. We found up to four haplotypes in a single sample. In *pfdhfr*, we found two different haplotypes based on N51i. Eleven of the twelve samples had the 51i mutated haplotype and one the N51 wild type haplotype. The SNPs C59r and S108n SNPs were mutated in all samples. Furthermore, all SNP sites which were found by the PHARE pipeline in *pfdhps* and *pfdhfr* are known sites of drug resistance.

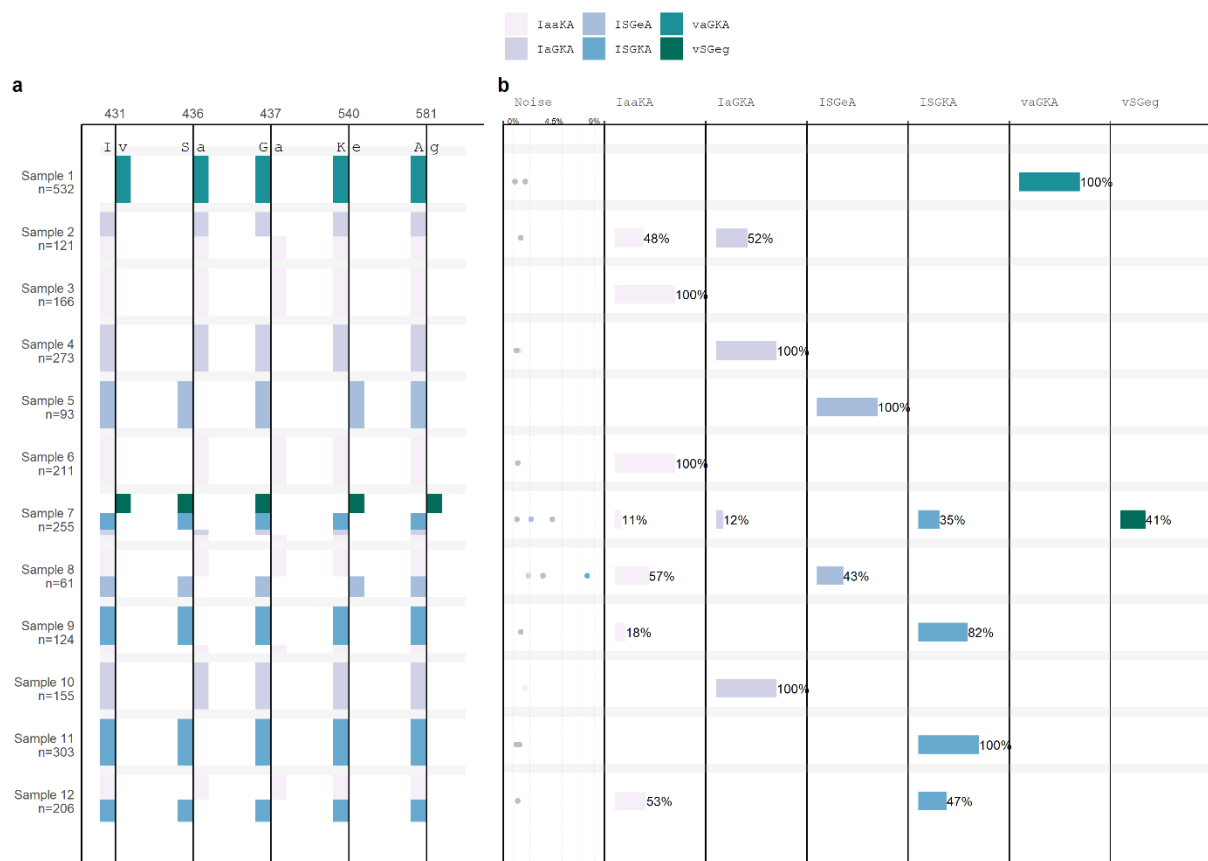


Figure 4. PHARE pipeline output for 12 clinical samples from the Central African Republic. a) SNP sites and the relative frequencies of mutated and wild type SNPs for each sample. b) Relative frequencies of each haplotype per sample, as well as the noise level. Every point in the noise graph represents a hypothetical haplotype that was filtered out because it is below the minimum haplotype frequency.

In *pfK13*, three different haplotypes were found, with one to three haplotypes per sample based on K189 which was mutated to either threonine or asparagine. As opposed to *pfdhps* and *pfdhfr*, the SNP found by the PHARE pipeline in *pfK13* in this clinical sample set has not been reported to be associated with drug resistance.

Discussion

Here we present the development of the PHARE pipeline, a novel method for compositional profiling of haplotypes in multiclonal *Pf* positive samples. Our pipeline has shown excellent performance in both control and clinical datasets, detecting up to four clones in a sample and detecting minor clones with an occurrence of as low as 5%. With upcoming technical improvements of ONT, the sensitivity of the pipeline is expected to increase further. Against popular misconception, it is possible to analyse SNPs in ONT data. Numerous studies have shown the feasibility of using Nanopore data for SNP analysis^{36,37} and multiple tools have been developed for this purpose, such as Longshot³⁸, NanoSNP³⁹ and Clair⁴⁰.

While qualitative performance of the PHARE pipeline was excellent, the estimated haplotype frequency did not always correspond to the input amount of DNA prior to PCR. Bias during PCR is a possible cause and could be reduced with a lower number of PCR cycles as well as sufficient template DNA⁴¹. Furthermore, while several publications state that ONT sequencing does not suffer from a GC content bias in terms of reads sequenced^{35,41-43}, others did find biases resulting from GC content and read length⁴⁴. Further investigation into the quantitative performance of the PHARE pipeline is needed.

Differentiating real, well-described haplotypes from false haplotypes that occur due to sequencing errors is a trade-off between avoiding false positives and being sufficiently sensitive. To avoid false positives and reach a specificity of 100%, we used a minimum haplotype frequency of 5%, discarding haplotypes with a prevalence below. This limits the sensitivity and prohibits detecting minority clones with frequencies below 5%. Depending on the research question, a lower value could be used, increasing the sensitivity. Read depth did not have a strong influence on the relative frequency of haplotypes reported by the pipeline when comparing replicates. However, the required sequencing depth depends on the number of variable sites of the genetic marker. Furthermore, the beginning and end of ONT reads tend to be of lower quality than the middle section^{35,43}. Therefore, it is important to design amplicons long enough to accommodate for this, preferably with primer binding sites outside of the gene of interest.

It is crucial to differentiate between samples containing one clone with a combination of drug resistance SNPs and samples with the same SNPs spread over multiple clones, because for some antimalarial resistance genes, drug resistance increases exponentially with the addition of further SNPs to the same gene⁴⁵. If the SNP sites responsible for generating malaria drug resistance are far apart on a single gene, it is necessary to sequence and analyse the full-length gene using long read sequencing. This is where ONT sequencing combined with the PHARE pipeline offers significant

benefits over other methods generating shorter reads such as Illumina < 600 bp⁴⁶ and Sanger < 900 bp⁴⁷.

Analysing more samples and conducting further studies will provide a deeper and more refined understanding of its ability in compositional profiling of multiclonal *Pf* infections. We believe that the PHARE pipeline has important implications for tracking the spread of resistance against malaria drugs, especially, since drug resistance markers are monitored in minority clones and across the whole gene without prior knowledge of SNP sites.

Acknowledgments

The authors would like to thank the administrative and laboratory staff of the Paediatric University Hospital Centre of Bangui, Bangui for their support and fruitful cooperation. The authors would also like to thank all patients who participated in this study.

Additionally, from the Swiss Tropical and Public Health Institute, we would like to thank Carla Beuret, Anina Schnoz and Maura Concu, for their help with analysing the clinical samples from the Central African Republic and Christin Gump, Christian Scheurer and Sergio Wittlin for their help with cultivating PfNF54, PfDd2, PfHB3 and PfK1 parasites.

Funding

This study was funded by a public–private partnership, the Bioko Island Malaria Elimination Project, financially supported by the Government of Equatorial Guinea, Marathon EG Production Limited, Noble Energy, and Atlantic Methanol Production Company.

Transparency declarations

None to declare.

References

1. Geneva: World Health Organization. World malaria report 2022. <https://www.who.int/publications/i/item/9789240064898> (2022).
2. WHO. *Guidelines for the treatment of malaria*. (World Health Organization, 2015).
3. WHO. The use of artesunate-pyronaridine for the treatment of uncomplicated malaria. *Who/Htm/Gmp/2019.13* 1–2 (2019).
4. Sinclair, D., Zani, B., Donegan, S., *et al.* Artemisinin-based combination therapy for treating uncomplicated malaria. *Cochrane Database Syst. Rev.* (2009) doi:10.1002/14651858.CD007483.PUB2/MEDIA/CDSR/CD007483/IMAGE_N/NCD007483-CMP-010-01.PNG.
5. Uwimana, A., Umulisa, N., Venkatesan, M., *et al.* Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *Lancet Infect. Dis.* **21**, 1120–1128 (2021).
6. Slater, H. C., Griffin, J. T., Ghani, A. C., *et al.* Assessing the potential impact of artemisinin and partner drug resistance in sub-Saharan Africa. *Malar. J.* **15**, 1–11 (2016).
7. Geneva: World Health Organization. *Strategy to respond to antimalarial drug resistance in Africa*. *Who/Ucn/Gmp/2022.04* vol. 04 (2022).
8. WHO. METHODS FOR SURVEILLANCE OF ANTIMALARIAL DRUG EFFICACY WHO Library Cataloguing-in-Publication Data : Methods for surveillance of antimalarial drug efficacy. (2009).
9. Geneva: World Health Organization. Tools for monitoring antimalarial drug efficacy. *Global Malaria Programme* <https://www.who.int/teams/global-malaria-programme/case-management/drug-efficacy-and-resistance/tools-for-monitoring-antimalarial-drug-efficacy>.
10. Smilkstein, M., Sriwilaijaroen, N., Kelly, J. X., *et al.* Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. *Antimicrob. Agents Chemother.* **48**, 1803–1806 (2004).
11. Witkowski, B., Menard, D., Amaratunga, C., *et al.* Ring-stage Survival Assays (RSA) to evaluate the in-vitro and ex-vivo susceptibility of Plasmodium falciparum to artemisinins. at <https://www.wwarn.org/sites/default/files/INV10-Standard-Operating-Procedure-Ring-Stage-Survival-Assays.pdf> (2013).
12. Plowe, C. V., Roper, C., Barnwell, J. W., *et al.* World Antimalarial Resistance Network (WARN) III: Molecular markers for drug resistant malaria. *Malar. J.* **6**, 1–10 (2007).
13. Picot, S., Olliaro, P., De Monbrison, F., *et al.* A systematic review and meta-analysis of evidence for correlation between molecular markers of parasite resistance and treatment outcome in falciparum malaria. *Malar. J.* **8**, 1–15 (2009).
14. Jones, S., Kay, K., Hodel, E. M., *et al.* Should deep-sequenced amplicons become the new gold standard for analyzing malaria drug clinical trials? *Antimicrob. Agents Chemother.* **65**, (2021).
15. Pholwat, S., Liu, J., Stroup, S., *et al.* The malaria TaqMan array card includes 87 assays for Plasmodium falciparum drug resistance, identification of species, and genotyping in a single reaction. *Antimicrob. Agents Chemother.* **61**, (2017).

16. Ansbro, M. R., Jacob, C. G., Amato, R., *et al.* Development of copy number assays for detection and surveillance of piperazine resistance associated plasmepsin 2/3 copy number variation in *Plasmodium falciparum*. *Malar. J.* **19**, 181 (2020).
17. Mahittikorn, A., Ramirez Masangkay, F., Uthaisar Kotepui, K., *et al.* Comparative performance of PCR using DNA extracted from dried blood spots and whole blood samples for malaria diagnosis: a meta-analysis. *Sci. Reports* | **11**, 4845 (2021).
18. Ishengoma, D. S., Saidi, Q., Sibley, C. H., *et al.* Deployment and utilization of next-generation sequencing of *Plasmodium falciparum* to guide anti-malarial drug policy decisions in sub-Saharan Africa: Opportunities and challenges. *Malar. J.* **18**, 1–10 (2019).
19. Tahar, R. & Basco, L. K. Molecular Epidemiology of Malaria in Cameroon. XXVI. Twelve-Year In Vitro and Molecular Surveillance of Pyrimethamine Resistance and Experimental Studies to Modulate Pyrimethamine Resistance. *Am. J. Trop. Med. Hyg.* **77**, 221–227 (2007).
20. Koehne, E., Adegnika, A. A., Held, J., *et al.* Pharmacotherapy for artemisinin-resistant malaria. *Expert Opin. Pharmacother.* **22**, 2483–2493 (2021).
21. Eldh, M., Hammar, U., Arnot, D., *et al.* Multiplicity of Asymptomatic *Plasmodium falciparum* Infections and Risk of Clinical Malaria: A Systematic Review and Pooled Analysis of Individual Participant Data. *J. Infect. Dis.* **221**, 775–785 (2020).
22. Rao, P. N., Uplekar, S., Kayal, S., *et al.* A method for amplicon deep sequencing of drug resistance genes in *Plasmodium falciparum* clinical isolates from India. *J. Clin. Microbiol.* **54**, 1500–1511 (2016).
23. Nag, S., Dalgaard, M. D., Kofoed, P. E., *et al.* High throughput resistance profiling of *Plasmodium falciparum* infections based on custom dual indexing and Illumina next generation sequencing-technology. *Sci. Reports* 2017 71 **7**, 1–13 (2017).
24. Talundzic, E., Ravishankar, S., Kelley, J., *et al.* Next-generation sequencing and bioinformatics protocol for malaria drug resistance marker surveillance. *Antimicrob. Agents Chemother.* **62**, (2018).
25. Guirou, E. A., Schindler, T., Hosch, S., *et al.* Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Nat. Sci. Reports* **10**, 1–14 (2020).
26. Zainabadi, K., Adams, M., Han, Z. Y., *et al.* A novel method for extracting nucleic acids from dried blood spots for ultrasensitive detection of low-density *Plasmodium falciparum* and *Plasmodium vivax* infections. *Malar. J.* **16**, 1–11 (2017).
27. Chugh, M., Scheurer, C., Sax, S., *et al.* Identification and Deconvolution of Cross-Resistance Signals from Antimalarial Compounds Using Multidrug-Resistant *Plasmodium falciparum* Strains. *Antimicrob. Agents Chemother.* **59**, 1110–1118 (2014).
28. Ariey, F., Witkowski, B., Amaratunga, C., *et al.* A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nat.* 2013 5057481 **505**, 50–55 (2013).
29. Runtuwene, L. R., Tuda, J. S. B., Mongan, A. E., *et al.* Nanopore sequencing of drug-resistance-associated genes in malaria parasites, *Plasmodium falciparum*. *Sci. Rep.* **8**, 1–13 (2018).
30. Ryan Wick. Filtlong.
31. Hathaway, N. J., Parobek, C. M., Juliano, J. J., *et al.* SeekDeep: Single-base resolution de novo clustering for amplicon deep sequencing. *Nucleic Acids Res.* **46**, 1–13 (2018).

32. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
33. Danecek, P., Bonfield, J. K., Liddle, J., *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).
34. Bonfield, J. K., Marshall, J., Danecek, P., *et al.* HTSlib: C library for reading/writing high-throughput sequencing data. *Gigascience* **10**, (2021).
35. Delahaye, C. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. *PLoS One* **16**, e0257521 (2021).
36. Tabata, Y., Matsuo, Y., Fujii, Y., *et al.* Rapid detection of single nucleotide polymorphisms using the MinION nanopore sequencer: a feasibility study for perioperative precision medicine. *JA Clin. Reports* **8**, (2022).
37. Ren, Z. L., Zhang, J. R., Zhang, X. M., *et al.* Forensic nanopore sequencing of STRs and SNPs using Verogen’s ForenSeq DNA Signature Prep Kit and MinION. *Int. J. Legal Med.* **135**, 1685–1693 (2021).
38. Edge, P. & Bansal, V. Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat. Commun.* **2019 101** **10**, 1–10 (2019).
39. Huang, N., Xu, M., Nie, F., *et al.* NanoSNP: a progressive and haplotype-aware SNP caller on low-coverage nanopore sequencing data. *Bioinformatics* **39**, (2023).
40. Zheng, Z., Li, S., Su, J., *et al.* Symphonizing pileup and full-alignment for deep learning-based long-read variant calling. *Nat. Comput. Sci.* **2022 212** **2**, 797–803 (2022).
41. Browne, P. D., Nielsen, T. K., Kot, W., *et al.* GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience* **9**, 1–14 (2020).
42. Oikonomopoulos, S., Wang, Y. C., Djambazian, H., *et al.* Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Sci. Rep.* **6**, (2016).
43. Krishnakumar, R., Sinha, A., Bird, S. W., *et al.* Systematic and stochastic influences on the performance of the MinION nanopore sequencer across a range of nucleotide bias. *Sci. Rep.* **8**, (2018).
44. Whitford, W., Hawkins, V., Moodley, K. S., *et al.* Proof of concept for multiplex amplicon sequencing for mutation identification using the MinION nanopore sequencer. *Sci. Rep.* **12**, (2022).
45. Triglia, T., Wang, P., Sims, P. F. G., *et al.* Allelic exchange at the endogenous genomic locus in *Plasmodium falciparum* proves the role of dihydropteroate synthase in sulfadoxine-resistant malaria. *EMBO J.* **17**, 3807–3815 (1998).
46. Illumina Inc. Maximum read length for Illumina sequencing platforms - Illumina Knowledge. https://knowledge.illumina.com/instrumentation/general/instrumentation-general-reference_material-list/000002826 (2023).
47. Stranneheim, H. & Lundeborg, J. Stepping stones in DNA sequencing. *Biotechnol. J.* **7**, 1063 (2012).

Supplementary methods

Control experiments using laboratory *P. falciparum* clones to validate the PAHRE pipeline

P. falciparum *in vitro* culture strains Dd2, HB3, K1 and NF54 were used to develop and test our bioinformatics analysis pipeline. For each strain, the cultured parasites were mixed with *P. falciparum* negative human blood and extracted with the QIAamp DNA Blood Mini Kit (Qiagen) according to the manufacturer's instructions. To generate the first dataset, the dihydrofolate reductase (*pfdhfr*), dihydropteroate synthase (*pfdhps*) and kelch 13 (*pfK13*) genes of strains HB3, NF54 and K1 were amplified and ONT sequenced on a R10.4 flow cell. After basecalling, reads were length filtered and aligned to the respective reference sequence to filter out contaminants. Then, to simulate mixed multiclonal samples, reads were randomly selected to generate 39 samples per gene. These 39 samples consisted of three different *in vitro* culture strain mixes (HB3-K1, HB3-NF54 and K1-NF54) at different mixing ratios, namely 1:99, 2.5:97.5, 5:95, 7.5:92.5, 10:90, 25:75, 50:50, 75:25, 90:10, 92.5:7.5, 95:5, 97.5:2.5, 99:1.

For the second control dataset, Dd2 and NF54 culture media with highly similar parasitaemia (ΔCq 0.24 as determined by qPCR on the *P. falciparum* ribonucleotide reductase R2_e2 gene¹) were mixed to generate mixtures ranging from 100% NF54 to 0% NF54 (steps: 100%, 99%, 90%, 50%, 25%, 5%, 0.5%, 0%). The culture medium mixes were added to parasite-free human whole blood and DNA was extracted individually. The *pfdhps* gene was amplified in three replicates for each mixture and sequenced on a R10.4 flow cell.

In a third control dataset to test the performance of detecting more than one clone in a single sample, DNA extracted from the *in vitro* control strains was mixed as follows: NF54-Dd2 (1:1), K1-HB3 (1:1) and a third sample with all four strains (1:1:1:1), followed by PCR amplification of the *pfdhfr*, *pfdhps* and *pfK13* genes and sequencing on a R10.4 flow cell.

SNP detection rate and performance of the PHARE pipeline compared to conventional methods using clinical samples

Blood samples were collected at the Paediatric Hospital and University Complex of Bangui (CHUPB) in the Central African Republic from children with fever between March 8th and 13th, 2021. Routinely collected data included malaria rapid diagnostic test (RDT) results, thick blood smear microscopy and full blood count. Dried blood spots were collected and then sent to the Swiss Tropical and Public Health Institute (Swiss TPH) for analysis of antimalarial resistance markers. DNA extraction was

performed using a previously established extraction protocol^{2,3}. To generate the clinical dataset, the *pfdhfr*, *pfdhps* and *pfK13* genes of twelve *P. falciparum* positive samples were amplified and sequenced on a R9.4.1 flow cell.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics and Scientific Committee from the University of Bangui (approval n°3/UB/FACSS/CSCVPER/PER) and by the Ministry of Health of the Central African Republic (approval n°0277/MSPP/CAB/DGSP/DMPM/ SMEE from 5 august 2002) as part of the communicable and endemic diseases surveillance diagnostic program. The patients were informed about the objectives of the study and nature of their participation. Then, written and signed informed consent was obtained from the parents on behalf of their children.

PCR amplification for Sanger sequencing

For the detection of antimalarial drug resistance markers *pfdhps*, *pfdhfr* and *pfK13*, previously established protocols were used^{4,5}. In brief, a primary PCR was performed with 5 µl DNA in a reaction volume of 25 µl, using 1 U HOT FIREPol DNA Polymerase (Solis Biodyne) and 0.25 µM primers. For *pfdhfr* (646 bp) and *pfdhps* (756 bp) primary PCRs, the following thermal cycling conditions were used: initial denaturation at 95 °C for 12 minutes, 30 cycles of 95 °C for 30 seconds, 52 °C for 90 seconds, 72 °C for 90 seconds, followed by final extension at 72 °C for 5 minutes. For *pfK13* primary PCR (2097 bp), the following cycling conditions were used: initial denaturation at 95 °C for 15 minutes, 30 cycles of 95 °C for 30 seconds, 58 °C for 2 minutes, 72 °C for 2 minutes, followed by final extension at 72 °C for 5 minutes. Nested PCRs were performed for each gene individually, using 1 µl DNA in a reaction volume of 50 µl, using 2 U HOT FIREPol DNA Polymerase (Solis Biodyne) and 0.25 µM primers. For *pfdhfr* nested PCR (638 bp), the following thermal cycling conditions were used: initial denaturation at 95 °C for 12 minutes, 30 cycles of 95 °C for 30 seconds, 52 °C for 90 seconds, 72 °C for 90 seconds, followed by final extension at 72 °C for 5 minutes. For *pfdhps* nested PCR (709 bp), the following thermal cycling conditions were used: initial denaturation at 95 °C for 12 minutes, 30 cycles of 95 °C for 30 seconds, 58 °C for 90 seconds, 72 °C for 90 seconds, followed by final extension at 72 °C for 5 minutes. For *pfK13* nested PCR (849 bp), the following cycling conditions were used: initial denaturation at 95 °C for 15 minutes, 40 cycles of 95 °C for 30 seconds, 60 °C for 1 minute, 72 °C for 1 minute, followed by final extension at 72 °C for 5 minutes. The final PCR product was checked for correct amplification on the QIAxcel Advanced Instrument and analysed with the corresponding QIAxcel ScreenGel 1.6.0 software (Qiagen). Samples with no band in the expected size were repeated. The successfully amplified samples were then sent to

Microsynth AG (Balgach, Switzerland) for Sanger sequencing using the same primers as the nested PCR.

PCR amplification, library preparation and ONT sequencing

The full-length drug resistance genes *pfdhfr*, *pfdhps* and *pfK13* were amplified separately using published primers ⁶ with a reaction concentration of 0.5 μ M. PCR amplification was performed with 3 μ l of template DNA in a reaction volume of 25 μ l, using 1x KAPA HiFi HotStart ReadyMix (Kapa Biosystems) according to the manufacturer's protocol. The conditions of the PCR were: 95 °C for 3 minutes, 35 cycles of 95 °C for 15 seconds, 50 °C for 30 seconds and 62 °C for 3 minutes, followed by final extension at 62 °C for 5 minutes. All PCR products were evaluated by agarose gel electrophoresis.

Amplified DNA was purified using 0.8x volume AMPure[®] XP beads (Beckman Coulter) and quantified with the Qubit dsDNA HS Assay Kit (Invitrogen). For the control datasets using laboratory strains, the Native Barcoding Kit 96 (ONT SQK-NBD112.96) and R10.4 flow cell (ONT FLO-MIN112) were used. For the clinical samples, Ligation Sequencing Kit (ONT SQK-LSK109), Native Barcoding Expansion 13-24 (EXP-NBD114) and R9.4.1 flow cell (ONT FLO-MIN106D) were used. Sequencing libraries were prepared according to manufacturer's instruction. Briefly, 200 fmol amplicon DNA were end-prepped, unique barcodes were ligated, and sequencing adapters were ligated. The libraries were loaded onto the flow cell and sequenced on the MinION Mk1C.

Data analysis and availability

Sanger sequencing data was analysed in SeqScape v4.0 (Applied Biosystems). The sequences and basecalling were manually checked. Mutations were marked and a mutation report was generated. The mutation report was then saved as a .csv file and imported into R. ONT sequencing data was extracted as raw data in FAST5 format. Basecalling was done using the "super high accuracy" basecalling in guppy (v6.4.2). To test the impact of the guppy basecalling model, the "high accuracy" and the "fast" model were also tested with the second control dataset. All sequencing data was uploaded to the NCBI Sequence Read Archive under BioProject ID PRJNA974955.

Programming of PHARE pipeline

The pipeline uses the tools Filtlong (v0.2.1) ⁷, SeekDeep extractor (v3.0.1) ⁸, minimap2 (v2.24) ⁹, samtools (v1.16.1) ¹⁰ and pysam (v0.20.0) ¹¹. The code is written in bash, R (v4.2.2) and python (v3.7.11).

Supplementary Table

Parameter	Description	Default value
SNP selection cut-off	Minimum % of sequencing reads with SNP at this site	10%
coverage minimum	Minimum % of reads covering a nucleotide position considered a SNP	80% for ONT sequencing chemistry R10.4 or 50% for ONT sequencing chemistry R9.4.1
minimum haplotype frequency	Minimum % of reads with certain haplotype	5%
Minqual	Minimum Phred score of reads at SNP sites	15
reference sequence	Gene reference in fasta format and additionally, genbank report in .jsonl format	
optional primer file	A file containing the primer sequences (according to the SeekDeep specification)	
Gene	Gene name	
Contig	Name of the contig in the fasta file	
Targetlen	Length of the PCR amplicon	
Minlen	Minimum length of reads	Targetlen - 50
Maxlen	Maximum length of reads	Targetlen + 350

Table S1. Adjustable parameters and input files of the PHARE pipeline.

Supplementary Figures

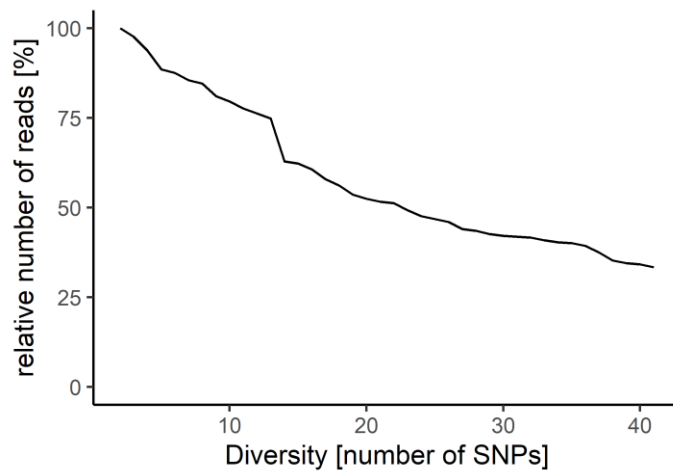


Figure S6. Relative number of reads compared with the number of SNPs used for variant calling. 100% is set at one SNP and corresponds to 5450 reads. To generate this figure, data of the second control experiment was used. The pipeline was run with default parameters 40 times, adding one SNP at a random site with every iteration, storing the mean number of reads over all samples.

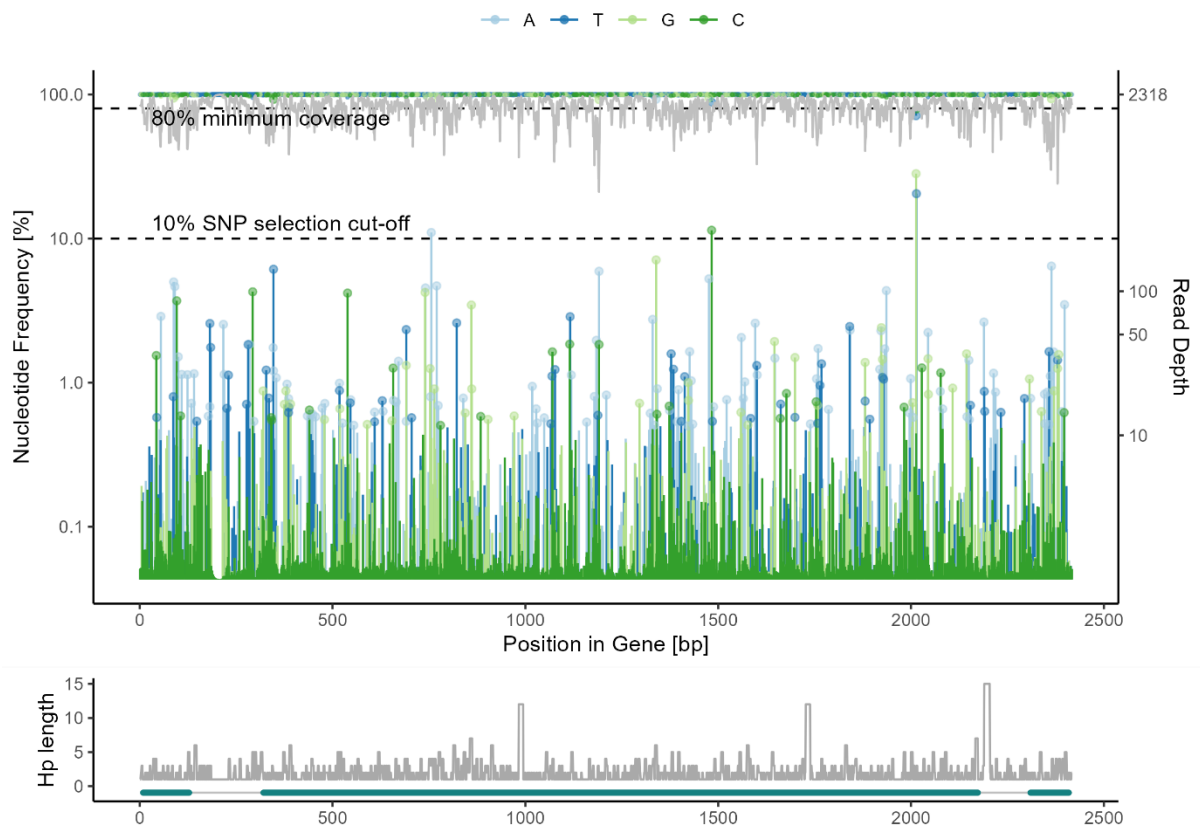


Figure S7. SNP calling using a dataset generated using “fast” basecalling of the 1:1 mix of in vitro culture strains NF54 and Dd2. The nucleotide position is on the x-axis. The read depth is shown in grey and the relative frequency of different bases in the respective colour. The lower dotted line is the SNP selection cut-off, which was set at 10%. The upper dotted line is the coverage minimum which was set at 80%.

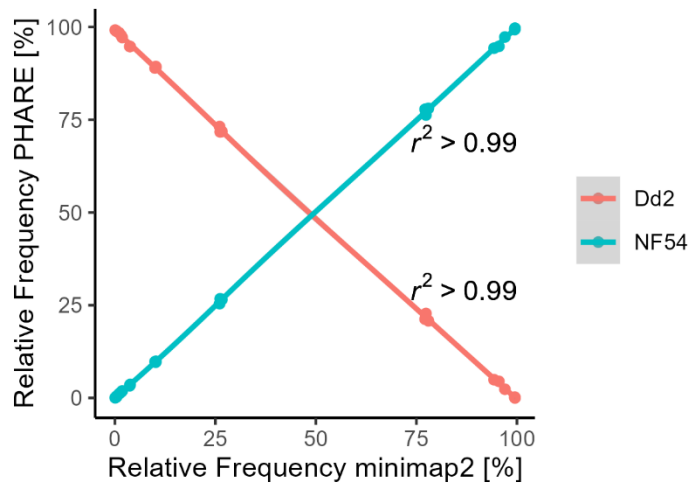


Figure S8. Relative frequencies of the two clones, NF54 and Dd2 when using the PHARE pipeline compared to mapping reads to the corresponding references using minimap2.

Supplementary References

1. Schindler, T., Deal, A. C., Fink, M., *et al.* A multiplex qPCR approach for detection of pfhrp2 and pfhrp3 gene deletions in multiple strain infections of *Plasmodium falciparum*. *Sci. Rep.* **9**, (2019).
2. Guirou, E. A., Schindler, T., Hosch, S., *et al.* Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Nat. Sci. Reports* **10**, 1–14 (2020).
3. Zainabadi, K., Adams, M., Han, Z. Y., *et al.* A novel method for extracting nucleic acids from dried blood spots for ultrasensitive detection of low-density *Plasmodium falciparum* and *Plasmodium vivax* infections. *Malar. J.* **16**, 1–11 (2017).
4. Chugh, M., Scheurer, C., Sax, S., *et al.* Identification and Deconvolution of Cross-Resistance Signals from Antimalarial Compounds Using Multidrug-Resistant *Plasmodium falciparum* Strains. *Antimicrob. Agents Chemother.* **59**, 1110–1118 (2014).
5. Ariey, F., Witkowski, B., Amaratunga, C., *et al.* A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nat. 2013 5057481* **505**, 50–55 (2013).
6. Runtuwene, L. R., Tuda, J. S. B., Mongan, A. E., *et al.* Nanopore sequencing of drug-resistance-associated genes in malaria parasites, *Plasmodium falciparum*. *Sci. Rep.* **8**, 1–13 (2018).
7. Ryan Wick. Filtlong.
8. Hathaway, N. J., Parobek, C. M., Juliano, J. J., *et al.* SeekDeep: Single-base resolution de novo clustering for amplicon deep sequencing. *Nucleic Acids Res.* **46**, 1–13 (2018).
9. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
10. Danecek, P., Bonfield, J. K., Liddle, J., *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).
11. Bonfield, J. K., Marshall, J., Danecek, P., *et al.* HTSlib: C library for reading/writing high-throughput sequencing data. *Gigascience* **10**, (2021).

4. One Health surveillance of foodborne bacteria

This chapter contains the following publication:

Abukhattab *et al.* **Whole Genome Sequencing for One Health Surveillance of Antimicrobial Resistance in Conflict Zones: A Case Study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine.** *Applied and Environmental Microbiology* 2023

4.1. Whole Genome Sequencing for One Health Surveillance of Antimicrobial Resistance in Conflict Zones: A Case Study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine

Published in Applied and Environmental Microbiology, 2023

Whole-genome sequencing for One Health surveillance of antimicrobial resistance in conflict zones: a case study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine

Said Abukhattab,^{1,2} Salome Hosch,^{1,2} Niveen M. E. Abu-Rmeileh,³ Shadi Hasan,⁴ Pascale Vonaesch,⁵ Lisa Crump,^{1,2} Jan Hattendorf,^{1,2} Claudia Daubenberger,^{1,2} Jakob Zinsstag,^{1,2} Tobias Schindler^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 10.

ABSTRACT Antimicrobial resistance (AMR) is a critical global concern driven by the overuse, misuse, and/or usage of inadequate antibiotics on humans, animals' agriculture, and as a result of contaminated environments. This study is the first One Health survey in the Middle East that incorporated whole-genome sequencing (WGS) to examine the spread of AMR in *Campylobacter* spp. and *Salmonella* spp. This cross-sectional study was conducted to examine the role of AMR at the human-animal-environmental interface and was performed in Ramallah/Al-Bireh and Jerusalem governorates of the central West Bank, Palestine. In 2021 and 2022, a total of 592 samples were collected and analyzed. From a total of 65 *Campylobacter jejuni* and 19 *Salmonella* spp. isolates, DNA was extracted for WGS using Oxford Nanopore Technologies MinION platform. We found that the dominant serotypes of *C. jejuni* and *Salmonella enterica* were present in chicken manure, chicken meat sold in markets, and feces of asymptomatic farm workers, with high genetic similarities between the isolates regardless of origin. Additionally, our results showed rapid strain turnover in *C. jejuni* from the same sites between 2021 and 2022. Most of the positive *Salmonella* spp. samples were multidrug-resistant (MDR) *S. enterica* serovar Muenchen carrying the plasmid of emerging *S. infantis* (pESI) megaplasmid, conferring resistance to multiple antibiotics. Our findings highlight the spread of MDR foodborne pathogens from animals to humans through the food chain, emphasizing the importance of a One Health approach that considers the interconnections between human, animal, and environmental health.

IMPORTANCE Prior to this study, there existed hardly an integrated human-animal-environmental study of Salmonellosis and Campylobacteriosis and related AMR in Middle Eastern countries. The few existing studies lack robust epidemiological study designs, adequate for a One Health approach, and did not use WGS to determine the circulating serotypes and their AMR profiles. Civil unrest and war in Middle Eastern countries drive AMR because of the breakdown of public health and food security services. This study samples simultaneously humans, animals, and the environment to comprehensively investigate foodborne pathogens in the broiler chicken production chain in Palestine using WGS. We show that identical serotypes of *C. jejuni* and *S. enterica* can be found in samples from chicken farms, chicken meat sold in markets, and asymptomatic broiler chicken production workers. The most striking feature is the rapid dynamic of change in the genetic profile of the detected species in the same sampling locations. The majority of positive *Salmonella* spp. samples are MDR *S. enterica* serovar Muenchen isolates carrying the pESI megaplasmid. The results demonstrate a close relationship between the *S. enterica* serovar Muenchen isolates found in our sample collection and those

Editor Christopher A. Elkins, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Address correspondence to Said Abukhattab, said.abukhattab@swisstph.ch.

Said Abukhattab and Salome Hosch contributed equally to this article. The order was established through consensus and in alphabetical order.

The authors declare no conflict of interest.

See the funding table on p. 11.

Received 21 April 2023

Accepted 13 July 2023

Published 1 September 2023

Copyright © 2023 Abukhattab et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

responsible for 40% of all clinical *Salmonella* spp. isolates in Israel as previously reported, with a sequence identity of over 99.9%. These findings suggest the transboundary spread of MDR *S. enterica* serovar Muenchen strains from animals to humans through the food chain. The study underscores the importance of combining integrated One Health studies with WGS for detecting environmental-animal-human transmission of foodborne pathogens that could not be detected otherwise. This study showcases the benefits of integrated environmental-animal-human sampling and WGS for monitoring AMR. Environmental samples, which may be more accessible in conflict-torn places where monitoring systems are limited and regulations are weak, can provide an effective AMR surveillance solution. WGS of bacterial isolates provides causal inference of the distribution and spread of bacterial serotypes and AMR in complex social-ecological systems. Consequently, our results point toward the expected benefits of operationalizing a One Health approach through closer cooperation of public and animal health and food safety authorities.

KEYWORDS Palestine, *Campylobacter*, *Salmonella*, AMR, WGS, One Health

The global spread of antimicrobial resistance (AMR), often referred to as the “silent pandemic” is one of the major challenges to public health in the 21st century. The O’Neill report estimates that by 2050, 10 million lives could be lost annually due to AMR (1). The problem of AMR is exacerbated in conflict-torn regions, such as the Middle East, where overuse of antibiotics, fragmented monitoring systems, inadequate infrastructure, and a lack of regulations and controls contribute to the rise of AMR. This includes both community and nosocomial transmission, which raises the incidence of AMR in these areas (2). Evidence is accumulating that prolonged and intense conflicts, as in the Palestine territories, lead to social and environmental conditions that foster the emergence of AMR (3). In Palestine, the lack of comprehensive national surveillance for AMR and weak published data hinder the ability to assess the extent of the problem and its risk factors. Furthermore, as a religious tourist destination, Palestine poses a potential risk for the emergence and international spread of AMR bacteria through the visitors.

Salmonellosis and Campylobacteriosis are leading causes of foodborne diseases (4). With an estimated 100 million individuals falling ill from foodborne infections each year, the Middle East has the third-highest prevalence globally (5). Due to the rapid spread of AMR, fluoroquinolone-resistant *Salmonellae* and *Campylobacter* were declared as high-priority pathogens for research and development of new antibiotics by the World Health Organization (6). Multidrug-resistant (MDR) foodborne pathogens are found widespread across the entire ecosystem and can spread to humans through food, environmental contamination, or direct contact with animals (7). The excessive use of antimicrobials in the food-producing animal industry fuels the emergence and spread of AMR (8). In 2017, approximately 73% of all antimicrobial usage worldwide was reported in animal agriculture (9). Based on recent projections, the global usage of antimicrobials in food-producing animals will further increase by 2030 (10). According to the agricultural census in 2021 by the Palestinian Central Bureau of Statistics (PCBS), there has been a significant growth in the industrial production of broiler chicken meat in Palestine. From 2010 to 2021, a remarkable increase of 128% was reported. This surge in production resulted in a total of 71 million broiler chickens being produced in 2021 (PCBS report in Arabic is available at <https://www.pcbs.gov.ps/Downloads/book2606.pdf>). The industrialized meat production in Palestine goes along with the excessive use of antibiotics (5, 10) and likely contributes to the spread of AMR.

Early detection and comprehensive understanding of drug-resistant pathogens, including their reservoir, spread, and genetic diversity, are, therefore, crucial to adopt intervention measures to combat AMR (11). This complexity highlights the importance of the One Health approach, which recognizes the interconnection between human, animal, and environmental health and focuses on demonstrating an incremental benefit of closer cooperation between human and animal health and related sectors (12).

Implementing this approach in the surveillance, prevention, and control of AMR in food animal production is vital in safeguarding human health and curbing the spread of AMR via the food supply chain. An integrated surveillance response system (iSRS) (13) for AMR should include continuous collection and testing of bacteria from various sources, such as animals, environment, food, and humans, in combination with surveillance tools like whole-genome sequencing (WGS). WGS data enable in-depth investigations into the transmission dynamics of AMR strains that circulate between animals, food, the environment, and humans (14). We used an iSRS approach complemented with WGS to investigate in the Ramallah/Al-Bireh and Jerusalem districts the two leading endemic foodborne pathogens, *Salmonella enterica* and *Campylobacter jejuni* in humans, chickens, and the environment.

RESULTS AND DISCUSSION

To our knowledge, this is the first integrated One Health survey supported by next-generation sequencing (NGS) that reports on the emergence and spread of *Campylobacter* spp. and *Salmonella* spp. AMR in the Middle East. AMR is a serious global public health concern, and surveillance has historically been concentrated in clinical settings in high-income countries. Outside clinical settings, resistant bacteria can circulate largely undetected in healthy humans, animals, and the environment, particularly in low- and middle-income countries (15). Very few studies exist that study the human-animal-environmental interface simultaneously (16). Our approach combines a One Health study design and WGS to identify the source and spread of foodborne pathogens and their AMR. Integrating WGS data from pathogens collected in the same place and time from humans, animals, and their environment using a One Health iSRS can lead to effective AMR control policies by identifying critical transmission routes and population dynamics (12). NGS technologies, such as Oxford Nanopore Technologies' MinION sequencing, have revolutionized clinical microbiology (17). Switching to NGS from traditional microbiological methods has a wide range of benefits, such as simultaneous identification of pathogens, assignment of serotypes, and detection of AMR markers. Therefore, integrating WGS-derived information into iSRS will help to design and guide control interventions and support monitoring of their effectiveness leading to reduced morbidity and mortality as well as saving health system resources (18).

A high positivity rate and distinct risk factors associated with *Salmonella* spp. and *Campylobacter* spp. characterize the broiler chicken production chain in Palestine

Three hundred and nineteen (319) specimens of *C. jejuni* were collected from chicken manure ($n = 126$), chicken meat ($n = 92$), and human stool ($n = 101$). Two hundred and seventy-three (273) specimens of *S. enterica* were collected from chicken manure ($n = 91$), chicken meat ($n = 81$), and human stool ($n = 101$). All samples which were culture positive for *Campylobacter* spp. were identified as *C. jejuni*, and all samples which were culture positive for *Salmonella* spp. were identified as *S. enterica*. The species identification of all positive samples was confirmed by PCR (Fig. 1A). In chicken manure samples, the positivity rate for *C. jejuni* was 30.1% (39/126) and for *S. enterica* 7.7% (7/91). Among chicken meat samples, the positivity rate for *C. jejuni* was 19.6% (18/92) and 9.9% (8/81) for *S. enterica*. Among fecal samples of broiler chicken production workers, we found a positivity rate of 19.8% (20/101) for *C. jejuni* and 1.0% (1/101) for *S. enterica*, respectively. Fig. 1B shows risk factors for infection with *C. jejuni* among workers' sociodemographic characteristics, with higher education decreasing the risk of infection (OR = 0.23, $P < 0.005$).

The high positivity rate of *Campylobacter* spp. and *Salmonella* spp. in the investigated samples, especially from chicken manure and meat, highlights the need for better hygiene control in chicken farms and the need for better food safety measures throughout the food production chain. A study on food safety in Palestinian territories found that poor hygiene practices, monitoring system fragmentation, a lack of regulations and

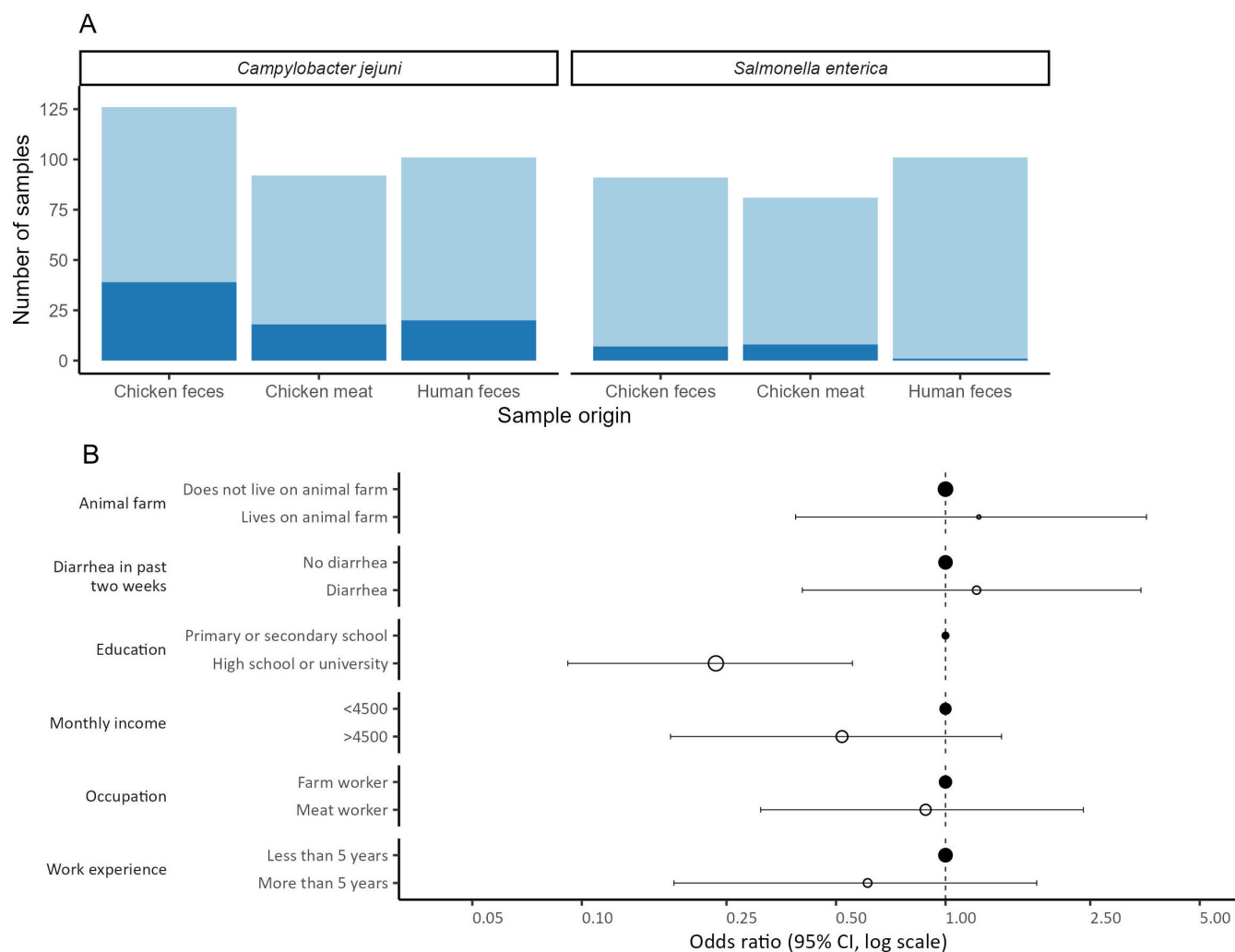


FIG 1 Epidemiology of *C. jejuni* and *S. enterica* in broiler chicken production chain in Palestine. (A) Positivity rate of *C. jejuni* and *S. enterica* among samples tested from different sources. (B) Univariable analysis of risk factors for *C. jejuni* infection among workers' sociodemographic characteristics and health status. filled circles are the reference group.

controls, and overuse of antimicrobials were obstacles to improving the food safety system (5). These results reflect the socioeconomic challenges in the region, including chronic conflict, population pressure, limited economic development, rapid urbanization, intensive agriculture, and difficulty enforcing policies and regulations (19). In conflict-torn places, collecting samples from the environment can have significant advantages, especially when access to human patients is limited. Environmental samples, such as chicken manure, can provide important information about the presence of foodborne pathogens and AMR, as these pathogens often persist in the environment long after the animals have left (20).

The survey result showed that meat production workers with higher education level have a lower risk of infection with *Campylobacter*. These results are consistent with a recent study that showed that education level was a determinant of zoonotic disease knowledge among workers in the production chain in Palestine and significantly influenced hygiene practices among them (5). Therefore, to address the challenges in food safety, zoonotic diseases, and AMR in Palestine, we recommend several strategies focusing on education, training, and awareness campaigns. Specific groups such as farm workers, slaughterhouse, and meat shop employees, families, and health professionals

should be the focus of targeted education and training initiatives to enhance their knowledge and skills.

A rapid strain and AMR profile turnover are observed in samples positive for *C. jejuni*

Regardless of the sample origin, all 47 *C. jejuni* isolates collected in 2021 were highly similar with an average nucleotide identity (ANI) of 99.95% (range: 99.90%–99.98%). The same was observed for 16 *C. jejuni* isolates collected in 2022, for which an ANI of 99.98% (range: 99.97%–99.99%) was determined. Comparing the isolates collected in 2021 and 2022 revealed an ANI of only 97.84% (range: 97.80%–97.93%) (Fig. 2A). A phylogenetic tree based on 1,000 single copy genes was created using 10 representative isolates from Palestine (Fig. 2B). For comparison, 29 closely related genomes from Bacterial and Viral Bioinformatics Resource Center (BV-BRC) and 22 genomes with the same sequence type (ST) as identified in PubMLST were added. All 47 isolates collected in 2021 were assigned to ST 11040 by Multi Locus Sequence Typing (MLST). Only one other isolate with the same ST was found in the PubMLST database, a human blood culture isolate collected in 2015 in Israel. These isolates did not match closely with any known genomes and showed a distinct AMR pattern. All samples collected in 2021 consisted of a single chromosome with a median length of 1,670,867 base pairs (range: 1,670,766–1,671,976 bp) and had identical drug resistance gene patterns, consisting of the aminoglycoside resistance genes *APH(2'')-IIIa*, *ANT(4)*, and *ANT(6)-Ia*, OXA β -lactamase *bla-OXA184*, macrolide resistance gene *ermB*, quinolone resistance mutation T86I in *gyrA*, and tetracycline resistance gene *tetL*. No plasmid was detected. The closest WGS we found had an ANI of only 98.39% (range: 98.27%–98.44%). The co-resistance conferred by the combination of AMR markers to both aminoglycosides and macrolides is worrisome, as these drugs are typically the first line of treatment for human cases (21). The 16 isolates collected in 2022 were assigned to ST 305, which has a total of 95 entries in PubMLST. The majority of sequences with available WGS data (53.1%, 17/32) were submitted by the United Kingdom, while only one isolate originated from the Middle East. These isolates consisted of a chromosome with a median length of 1,677,345 bp (range: 1,677,331–1,677,357 bp) and a plasmid with a median length of 51,587 bp (range: 51,584–51,589 bp). Isolates collected in 2022 had a resistance gene pattern consisting of OXA β -lactamase *bla-OXA61* and quinolone resistance mutation T86I in *gyrA* on the chromosome and the plasmid-encoded tetracycline resistance gene *tetO*. The *C. jejuni* isolates collected in 2022 did not have AMR markers conferring resistance to aminoglycosides or macrolides. Establishing an iSRS to address large annual fluctuations in bacterial strains and resistance characteristics requires enhanced collaboration and partnerships among relevant stakeholders (22). Implementation of standardized protocols for data collection, testing, and reporting is essential, supported by a robust data management system for centralized data collection and analysis (22). Regular monitoring activities also play a vital role in observing dynamic patterns, while collaborative research efforts can help identify factors that contribute to these changes (23). In addition, capacity building through introducing new laboratory techniques such as WGS, training programs, and effective sharing of information and communication among stakeholders is pivotal (18, 24). If successfully implemented, these approaches could result in an immediate response and intervention to effectively address emerging AMR threats. A recent study identified Palestine as a global hotspot for veterinary antimicrobial consumption (10), underscoring the problem of antibiotic overuse in intensive animal production. This issue is particularly prevalent in broiler production, where lower levels of biosecurity measures lead to heavy reliance on antibiotics. This excessive use could be a contributing factor to the rapid shift in the *C. jejuni* strain observed between 2021 and 2022.

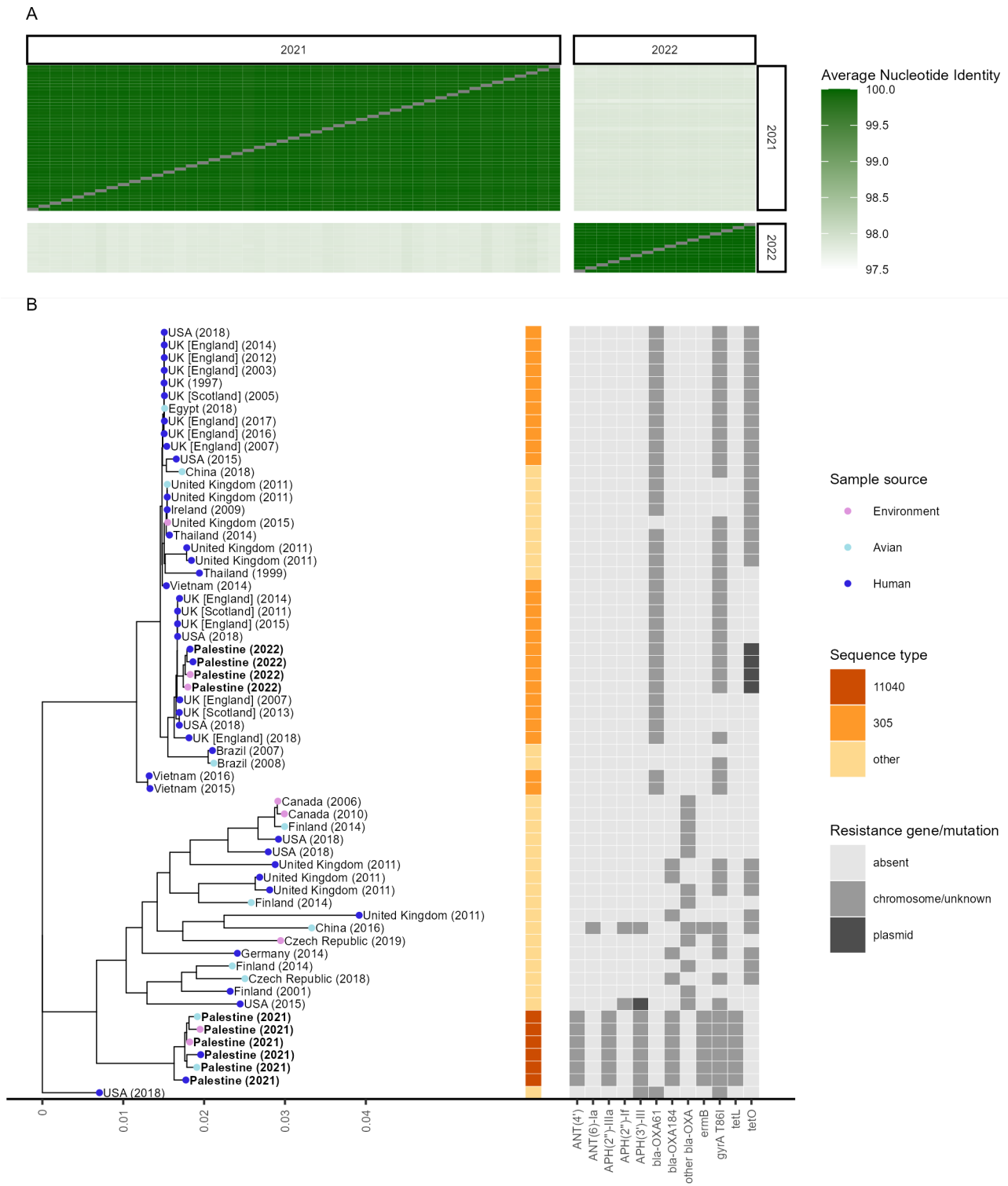


FIG 2 Phylogenetic analysis of *C. jejuni* isolates. (A) Average nucleotide identities between all 47 isolates collected in 2021 and all 16 isolates collected in 2022. (B) *C. jejuni* phylogenetic tree with 10 representative genomes from Palestine and 51 closely related genomes from different countries collected between 1997 and 2019. MLST-based sequence types are shown for the sequence types represented by Palestinian isolates. The presence of AMR markers is shown in light gray if located in the genome or an unidentified contig or dark gray if located on a plasmid. sample sources are depicted by colored, filled circles. Genomes generated in this study are in bold text. All human isolates collected in 2021 are from broiler chicken production chain workers, while all four human isolates collected in 2022 are from hospitalized gastroenteritis patients.

The MDR isolates *S. enterica* serovar Muenchen are found across the entire broiler chicken production chain

A phylogenetic analysis based on 1,000 single-copy genes for *S. enterica* revealed that 12 out of 14 *S. enterica* isolates collected in 2021 were assigned to *S. enterica* serovar Muenchen (Fig. 3A). These 12 isolates had a median chromosome length of 4,857,056 bp (range: 4,856,898–4,857,070 bp) and a median plasmid length of 285,076 bp (range: 285,070–285,085 bp). The isolates show a high similarity across all samples with an ANI of 100.00% (range: 99.99%–100.00%). The two other isolates were sampled from chicken meat in 2021 and assigned to *S. enterica* serovar Paratyphi B. The two *S. enterica* serovar Paratyphi B isolates had chromosome lengths of 4,697,034 bp and 4,697,043 bp and chromosome encoded aminoglycoside resistance genes *AAC(6)-ly* and *ANT(3)-Ia*, quinolone resistance mutation S83F in *gyrA*, trimethoprim resistance gene *dfrA1*, and the *mdsABC* efflux pump. Additionally, they have the plasmid-encoded β -lactamase *blaTEM-1*, phenicol resistance gene *floR*, sulfonamide resistance gene *sul2*, and tetracycline resistance gene *tetA*. They were closely related to *S. enterica* serovar Paratyphi B isolates collected from environmental and avian sources in Europe and Israel. Additionally, three confirmed *S. enterica* clinical isolates obtained from humans in 2022 were added as positive controls. These three isolates had a median chromosome size of 4,720,112 bp (range: 4,720,107–4,720,159 bp), a plasmid with a median length of 37,698 bp (range: 37,698–37,699 bp) as well as a plasmid with a median length of 59,372 bp (range: 59,372–59,373 bp). The *S. enterica* serovar Enteritidis isolates contain the chromosome-encoded aminoglycoside resistance gene *AAC(6)-ly* and the *mdsABC* efflux pump. They were closely related to *S. enterica* serovar Enteritidis strains isolated from humans globally. All 12 isolates assigned to *S. enterica* serovar Muenchen shared an identical pattern of molecular markers for AMR with resistance markers, including the aminoglycoside resistance gene *AAC(6)-Iaa*, the fluoroquinolone resistance mutation T57S of *parC*, and the *mdsABC* efflux pump encoded on their chromosome (Fig. 3B). Additionally, on the 285,077 bp (range: 285,070–285,085 bp) megaplasmid that was 99.99% (99.97–100.00) similar to “plasmid of emerging *S. Infantis* (pESI)” described from Israel (25), aminoglycoside resistance gene *ANT(3)-Ia*, sulfonamide resistance gene *sul1*, tetracycline resistance gene *tetA*, and trimethoprim resistance gene *dfrA14* are located (Fig. 3C).

Our study shows that the MDR *S. enterica* serovar Muenchen isolates are found in the entire chicken meat production chain. Among the *S. enterica* serovar Muenchen isolates, we found genotypical resistance markers to at least five classes of antibiotics, including aminoglycosides, fluoroquinolones, aminocoumarines, sulfonamides, and tetracyclines. Given the significant clinical importance of *S. enterica* infections, with a global annual incidence of over 27 million cases of enteric fever (26) and 78.7 million cases of gastroenteritis (4), the AMR monitoring as part of a One Health iSRS is of utmost importance.

All 12 isolates identified as *S. enterica* serovar Muenchen in our study showed a close relationship to an emerging clinical isolate reported in Israel (25). In addition, this study identified 19 additional *S. enterica* serovar Muenchen isolates from the United States, the United Kingdom, and South Africa containing the pESI plasmid with high genetic similarity (25). These *S. enterica* serovar Muenchen isolates were obtained from human clinical samples or avian sources. This noteworthy similarity among globally sourced isolates is likely to be attributed to the widespread and synchronized dissemination of breeding stocks contaminated with the bacteria due to the centralized sourcing practices and international trade. Similar patterns have been observed for *S. enterica* serotype Enteritidis (27), highlighting the role of global trade in facilitating the spread of such MDR bacterial strains.

The researchers from Israel noted an increasing prevalence of an MDR strain of *S. enterica* serovar Muenchen among their clinical *Salmonella* spp. isolates (25). These isolates displayed significant sequence similarity to the isolates described in our study. This suggests that there may be a regional spread of this particular MDR strain. This could be due to importing some of the raw materials used in animal feed production from

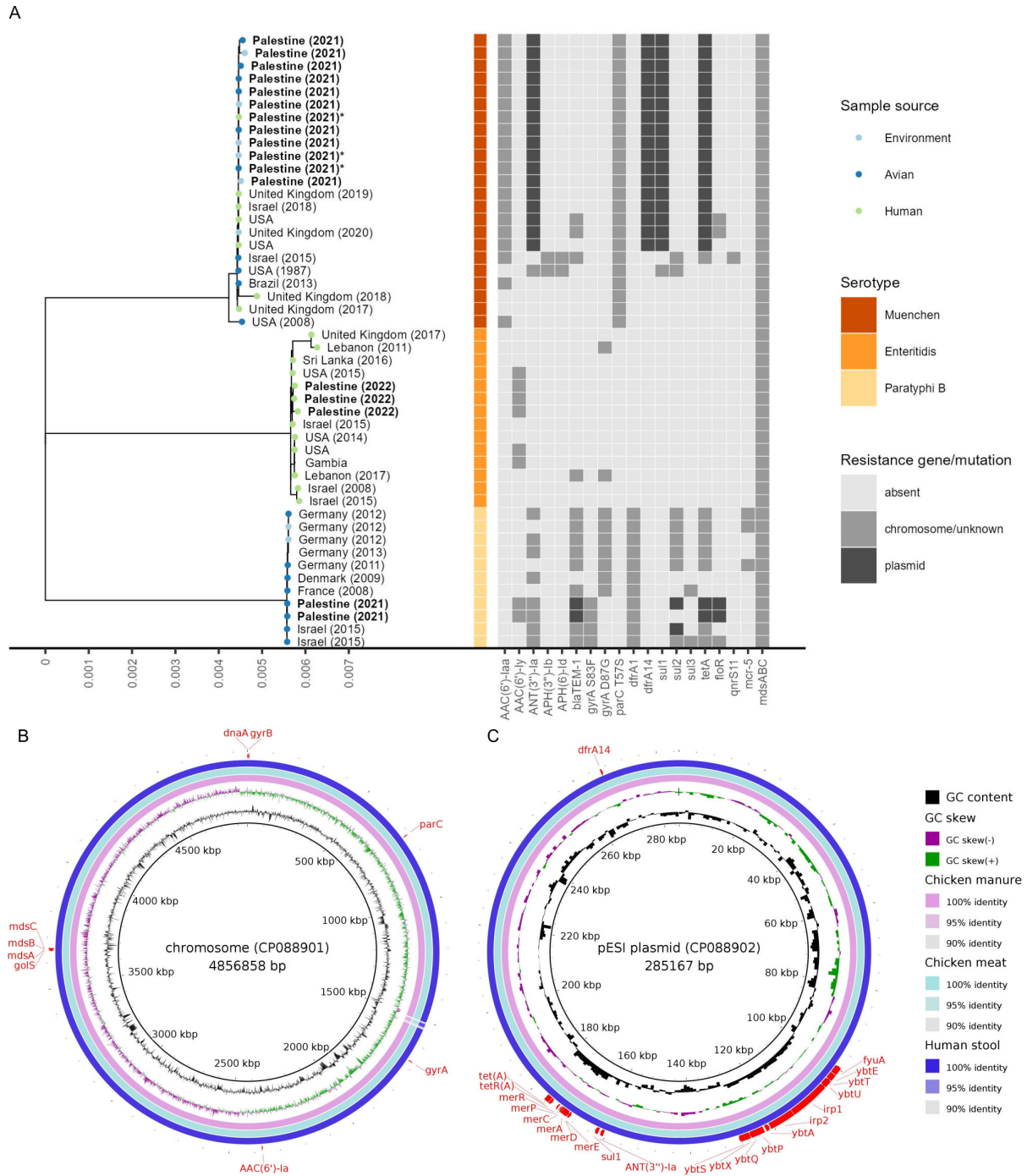


FIG 3 Phylogenetic analysis of *S. enterica* isolates. (A) Phylogenetic tree for *S. enterica* with 17 genomes from Palestine and 31 closely related genomes belonging to the three serotypes found in this study from different countries collected between 1987 and 2020. The presence of AMR markers is shown in light gray if located in the genome or an unidentified plasmid or dark gray if located on a plasmid. Sample sources are depicted by colored, filled circles. Genomes generated in this study are in bold text. Genomes used for Fig. 3B and C are marked with an asterisk. Comparison of Palestinian *S. enterica* serotype Muenchen chromosome (B) and pESI plasmid (C) from chicken manure and human feces as well as meat with a clinical isolate from Israel collected in 2018. The presence and location of antibiotic-resistance genes are shown in red. The human isolate collected in 2021 is from a broiler chicken production chain worker, while all three human isolates collected in 2022 are from hospitalized gastroenteritis patients.

Israel, intersecting distribution networks, and sharing the same ecosystem (19, 28). Moreover, the prolonged and intense conflicts in the Palestinian territories contribute to a weak control system, which leads to cross-border smuggling (5, 28). The discovery of a

dominant MDR strain emphasizes the need for continuous monitoring and open data sharing in real time to understand the epidemiology of AMR across borders and political systems.

MATERIALS AND METHODS

Study design, setting, and participants

This cross-sectional study was performed in Ramallah/Al-Bireh and Jerusalem governorates of the central West Bank, Palestine. Between June and October 2021, 557 samples were collected, and 284 samples were tested for the presence of *Campylobacter* and 273 samples for *Salmonella*. The samples came from abattoirs ($n = 50$, fresh chicken meat), large-scale broiler chicken farms with single batch at a time (all-in and all-out) system ($n = 91$), chicken manure, and asymptomatic chicken meat production workers ($n = 101$, fecal samples). In August 2022, an additional 35 chicken manure samples were collected from the same farms addressed in 2021 to investigate the persistence or replacement of *C. jejuni* strains over time. Furthermore, four *C. jejuni* isolates and three *S. enterica* isolates were isolated from gastroenteritis patients in 2022 after confirmation by Vitek 2 and PCR, preserved in the hospital laboratory to be used as positive control isolates at one of the targeted region's hospitals to compare the pathogens isolated in our study to those reported in the hospitals. To investigate the risk factors for the presence of *Salmonella* spp. and *Campylobacter* spp. along the broiler chicken production chain, a questionnaire in Arabic was used to collect information on workers' sociodemographic characteristics and health status related to infection with gastroenteritis infections. All sampling sites were selected using a random sampling approach and using authorities' records about farms, abattoirs, meat stores, and workers. All study participants signed a written consent form.

Microbiological laboratory procedures

The fecal samples from humans and chicken manure were grown directly on *Campylobacter* Selective Agar (Oxoid, UK) for *Campylobacter* selection and Xylose-Lysin-Desoxycholat Agar (Oxoid, UK), MacConkey (Oxoid, UK), and Sheep Blood Agar Base (Oxoid, UK) for *Salmonella* selection. Agar plates were incubated at 37°C and checked for the presence of colonies after 24–48 hours. The chicken meat samples were pre-processed according to the ISO 10272-1:2017 guidelines for *Campylobacter* and ISO 6579-1:2017 guidelines for *Salmonella*. From all samples, suspected colonies for each of the species were selected for microbiological confirmation by the Vitek 2 NH ID card for *Campylobacter* or Vitek 2 GN ID card for *Salmonella* using the Vitek 2 compact automated system (BioMerieux, France). All isolates which were confirmed as *Salmonella* or *Campylobacter* by the Vitek 2 microbial testing system were subjected to further confirmation by PCR. Molecular species identification for *Campylobacter* spp. was performed using the PCR protocol published by Nayak et al. (29), while for *Salmonella* spp., the procedure published by Paião et al. was used (30).

DNA extraction and WGS

From a total of 65 *C. jejuni* and 19 *Salmonella* spp. isolates, DNA was extracted using QIAamp DNA Mini Kit (Qiagen, Germany) and Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research, USA), respectively. Extracted DNA was shipped to the Swiss Tropical and Public Health Institute. DNA was quantified with the Qubit dsDNA HS Assay Kit (Invitrogen, Germany). Samples with DNA concentrations >33 ng/ μ L were selected for WGS using MinION platform (Oxford Nanopore Technologies, UK). The sequencing library was prepared according to the manufacturer's instructions using the Native Barcoding Kit 96 and loaded onto the R10.4 flow cell and sequenced on the MinION Mk1C using super-accurate basecalling.

Bioinformatics and statistical analysis

De novo assembly was conducted using Flye 2.9.1 (31) and Tricycler v0.5.3 (32) at the scientific computing core facility of the University of Basel. The BV-BRC was used for annotation and phylogenetic analysis of the assemblies. The assemblies were annotated using the RAST 2.0 toolkit (33). Only assemblies with less than 10 contigs were considered for further analysis. ST and serotype of assembled contigs were determined using MLST (34) and SeqSero (35), respectively. Phylogenetic trees were generated by aligning protein and nucleotide sequences using MUSCLE (36), MAFFT (37), and RAxML (38). The 20 most closely related genomes from BV-BRC were selected for each *C. jejuni* ST. Additionally, all available WGS for the respective ST were obtained from the public databases for molecular typing and microbial genome diversity (PubMLST). The 10 most closely related genomes from BV-BRC were selected for each *S. enterica* serotype. Additionally, four selected *S. enterica* serotype Muenchen genomes containing the pESI plasmid were obtained from the National Center for Biotechnology Information ([GCA_008248485](https://ncbi.nlm.nih.gov/nuccore/GCA_008248485), [GCA_009444355](https://ncbi.nlm.nih.gov/nuccore/GCA_009444355), [GCA_010825065](https://ncbi.nlm.nih.gov/nuccore/GCA_010825065), [GCA_019543295](https://ncbi.nlm.nih.gov/nuccore/GCA_019543295)). *C. jejuni* and *S. enterica* genomes were filtered to contain high-quality genomes with available data for collection year, country, and sample origin.

Closely related isolates from the Middle East collected after 2008 were selected for *S. enterica*. Genome-wide ANIs were determined with FastANI (39). Resistome analysis was performed using CARD (6.0.0) (40) and ResFinder (4.2.2) (41). The comparison of *S. enterica* serotype Muenchen genomes was visualized using BRIG (42). Risk factors for *C. jejuni* infections among study participants were calculated using median-unbiased estimates to calculate the univariate odds ratio using the R software environment version 4.2.2 and the package epitools.

ACKNOWLEDGMENTS

The authors would like to thank the study participants for their contribution as well as Mamoun Ibaideya for helping in the bacterial isolation, and Birzeit University laboratories staff for their commitment and support during sample collection.

This work was funded by the Swiss Agency for Development and Cooperation – Jerusalem Office (proposal no. 7F-04229.08.09). S.A. was supported by Federal Commission for Scholarships for Foreign Students (FCS) (ESKAS-Nr: 2020.0684). P.V. was supported by a Return Grant (P3P3PA_17877) from the Swiss National Science Foundation and the Forschungsfonds of the University of Basel.

S.A., N.M.E.A.-R., P.V., J.H., and J.Z. conceptualized the epidemiological study. S.A., S.H.o., N.M.E.A.-R., S.H., J.H., C.D., and T.S. curated the data. S.A., S.H.o., and T.S. did the formal analysis. S.H.o. and T.S. analyzed and visualized the genomic data. S.A., N.M.E.A.-R., and J.Z. acquired funding. S.A., N.M.E.A.-R., S.H.o., and T.S. did the investigation. S.A., S.H., S.H.o., and T.S. wrote the methodology. S.A., N.M.E.A.-R., J.Z., and T.S. did the project administration. S.A., N.M.E.A.-R., J.Z., C.D., and T.S. supervised the study. S.A., S.H.o., and T.S. wrote the original draft. S.H., P.V., L.C., J.H., C.D., and J.Z. critically revised the manuscript. All authors were responsible for the reviewing and editing of the manuscript. All authors had access to the data presented in this study and had final responsibility for the decision to submit for publication.

This study was supported by the Swiss Agency for Development and Cooperation—Jerusalem Office.

The authors declare no conflict of interest.

AUTHOR AFFILIATIONS

¹Swiss Tropical and Public Health Institute, Allschwil, Switzerland

²University of Basel, Basel, Switzerland

³Institute of Community and Public Health, Birzeit University, Birzeit, Palestine

⁴Master program in Clinical Laboratory Sciences, Birzeit University, Birzeit, Palestine

⁵Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland

AUTHOR ORCID*s*

Said Abukhattab  <http://orcid.org/0000-0003-1670-0143>

Salome Hosch  <http://orcid.org/0000-0001-9290-3589>

FUNDING

Funder	Grant(s)	Author(s)
Swiss Agency for Development and Cooperation- Jerusalem office	7F-04229.08.09	Niveen M. E. Abu-Rmeileh
Federal Commission for Scholarship for Foreign Students (FCS)- Switzerland	ESKAS-Nr: 2020.0684	Said Abukhattab
Swiss National Science Foundation and the Forschungsfonds of the University of Basel	2020 (12 - 1)	Pascale Vonaesch

DATA AVAILABILITY

The data that support the findings of this study are available from the authors upon reasonable request. The genome assemblies and raw sequence reads are available on Genbank under BioProjects [PRJNA942086](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA942086) (*S. enterica*) and [PRJNA942088](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA942088) (*C. jejuni*).

ETHICS APPROVAL

The study was approved by the Northwestern and Central Switzerland Ethics Committee (Ethikkommission Nordwest- und Zentralschweiz, EKNZ) (Reference No: AO_2021-00021) and the ethical review committee at the Institute of Community and Public Health (ICPH) at Birzeit University (Reference No: 2020 (12 – 1)).

REFERENCES

- Grande B, O'Neill J, Review on Antimicrobial Resistance(London). 2014. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. Review on Antimicrobial Resistance. <https://books.google.ch/books?id=b1E0kAEACAAJ>.
- Devi S. 2019. AMR in the middle east: a perfect storm. The Lancet 394:1311–1312. [https://doi.org/10.1016/S0140-6736\(19\)32306-2](https://doi.org/10.1016/S0140-6736(19)32306-2)
- Truppa C, Abo-Shehada MN. 2020. Antimicrobial resistance among GLASS pathogens in conflict and non-conflict affected settings in the middle east: a systematic review. BMC Infect Dis 20:936. <https://doi.org/10.1186/s12879-020-05503-8>
- Havelaar AH, Kirk MD, Torgerson PR, Gibb HJ, Hald T, Lake RJ, Praet N, Bellinger DC, Silva NR, Gargouri N, Speybroeck N, Cawthorne A, Mathers C, Stein C, Angulo FJ, Devleeschauwer B, World health organization foodborne disease burden epidemiology reference group. 2015. 2010. World health organization global estimates and regional comparisons of the burden of foodborne disease in 2010. PLoS Med 12. <https://doi.org/10.1371/journal.pmed.1001923>
- Abukhattab S, Kull M, Abu-Rmeileh NME, Cissé G, Crump L, Hattendorf J, Zinsstag J. 2022. Towards a one health food safety strategy for Palestine: a mixed-method study. Antibiotics 11:1359. <https://doi.org/10.3390/antibiotics11101359>
- WHO publishes list of bacteria for which new antibiotics are urgently needed. 2017 Available from: <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>
- Ngobese B, Zishiri OT, El Zowalaty ME. 2020. Molecular detection of virulence genes in campylobacter species isolated from livestock production systems in South Africa. J Integ Agr 19:1656–1670. [https://doi.org/10.1016/S2095-3119\(19\)62844-3](https://doi.org/10.1016/S2095-3119(19)62844-3)
- Chantziaras I, Boyen F, Callens B, Dewulf J. 2014. Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. J Antimicrob Chemother 69:827–834. <https://doi.org/10.1093/jac/dkt443>
- Van Boeckel TP, Glennon EE, Chen D, Gilbert M, Robinson TP, Grenfell BT, Levin SA, Bonhoeffer S, Laxminarayan R. 2017. Reducing antimicrobial use in food animals. Science 357:1350–1352. <https://doi.org/10.1126/science.aao1495>
- Mulchandani R, Wang Y, Gilbert M, Van Boeckel TP. 2023. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. PLOS Glob Public Health 3:e0001305. <https://doi.org/10.1371/journal.pgph.0001305>
- Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, Guerin PJ, Piddock LJV. 2016. Understanding the mechanisms and drivers of antimicrobial resistance. Lancet 387:176–187. [https://doi.org/10.1016/S0140-6736\(15\)00473-0](https://doi.org/10.1016/S0140-6736(15)00473-0)
- Zinsstag Jakob, Kaiser-Grolimund A, Heitz-Tokpa K, Sreedharan R, Lubroth J, Caya F, Stone M, Brown H, Bonfoh B, Dobell E, Morgan D, Homaira N, Kock R, Hattendorf J, Crump L, Mauti S, Del Rio Vilas V, Saikat S, Zumla A, Heymann D, Dar O, de la Rocque S. 2023. Advancing one human-animal-environment health for global health security: what does the evidence say Lancet 401:591–604. [https://doi.org/10.1016/S0140-6736\(22\)01595-1](https://doi.org/10.1016/S0140-6736(22)01595-1)
- Zinsstag J, Utzinger J, Probst-Hensch N, Shan L, Zhou X-N. 2020. Towards integrated surveillance-response systems for the prevention of future pandemics. Infect Dis Poverty 9:140. <https://doi.org/10.1186/s40249-020-00757-5>
- Wee BA, Muloi DM, van Bunnik BAD. 2020. Quantifying the transmission of antimicrobial resistance at the human and livestock interface with genomics. Clin Microbiol Infect 26:1612–1616. <https://doi.org/10.1016/j.cmi.2020.09.019>

15. Maciel-Guerra A, Baker M, Hu Y, Wang W, Zhang X, Rong J, Zhang Y, Zhang J, Kaler J, Renney D, Loose M, Emes RD, Liu L, Chen J, Peng Z, Li F, Dottorini T. 2023. Dissecting microbial communities and resistomes for interconnected humans, soil, and livestock. 1. ISME J 17:21–35. <https://doi.org/10.1038/s41396-022-01315-7>
16. Meier H, Spinner K, Crump L, Kuenzli E, Schuepbach G, Zinsstag J. 2023. State of knowledge on the acquisition, diversity, interspecies attribution and spread of antimicrobial resistance between humans animals and the environment: a systematic review. *Antibiotics* 12:73. <https://doi.org/10.3390/antibiotics12010073>
17. Leggett RM, Alcon-Giner C, Heavens D, Caim S, Brook TC, Kujawska M, Martin S, Peel N, Axford-Palmer H, Hoyles L, Clarke P, Hall LJ, Clark MD. 2020. Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nat Microbiol* 5:430–442. <https://doi.org/10.1038/s41564-019-0626-z>
18. Abukhattab S, Taweel H, Awad A, Crump L, Vonaesch P, Zinsstag J, Hattendorf J, Abu-Rmeileh NME. 2022. Systematic review and meta-analysis of integrated studies on *Salmonella* and campylobacter prevalence, serovar, and phenotyping and genetic of antimicrobial resistance in the middle east—a one health perspective. *Antibiotics* (Basel) 11:536. <https://doi.org/10.3390/antibiotics11050536>
19. EnvironmentUN 2020. State of Environment and Outlook Report for the occupied Palestinian territory 2020. UNEP - UN Environment Programme.
20. Parvin MS, Ali MY, Mandal AK, Talukder S, Islam MT. 2022. Sink survey to investigate multidrug resistance pattern of common foodborne bacteria from wholesale chicken markets in Dhaka city of Bangladesh. *Sci Rep* 12:10818. <https://doi.org/10.1038/s41598-022-14883-7>
21. Lopez-Chavarria V, Ugarte-Ruiz M, Barcena C, Olarra A, Garcia M, Saez JL, de Frutos C, Serrano T, Perez I, Moreno MA, Dominguez L, Alvarez J. 2021. Monitoring of antimicrobial resistance to aminoglycosides and macrolides in *Campylobacter coli* and *Campylobacter jejuni* from healthy livestock in Spain (2002–2018). *Front Microbiol* 12:689262. <https://doi.org/10.3389/fmicb.2021.689262>
22. Organization WH. 2017. Integrated surveillance of antimicrobial resistance in Foodborne bacteria: Application of a one health approach: guidance from the WHO advisory group on integrated Surveillance of antimicrobial resistance (AGISAR). World Health Organization
23. Hattendorf J, Bardosh KL, Zinsstag J. 2017. One health and its practical implications for surveillance of endemic zoonotic diseases in resource limited settings. *Acta Trop* 165:268–273. <https://doi.org/10.1016/j.actatropica.2016.10.009>
24. Seale AC, Gordon NC, Islam J, Peacock SJ, Scott JAG. 2017. AMR surveillance in low and middle-income settings - a roadmap for participation in the global antimicrobial surveillance system (GLASS). *Wellcome Open Res* 2:92. <https://doi.org/10.12688/wellcomeopenres.12527.1>
25. Cohen E, Kriger O, Amit S, Davidovich M, Rahav G, Gal-Mor O. 2022. The emergence of a multidrug resistant *Salmonella* muenchen in Israel is associated with horizontal acquisition of the epidemic pESI plasmid. *Clin Microbiol Infect* 28:1499. <https://doi.org/10.1016/j.cmi.2022.05.029>
26. Crump JA, Luby SP, Mintz ED. 2004. The global burden of typhoid fever. *Bull World Health Organ* 82:346–353.
27. Li S, He Y, Mann DA, Deng X. 2021. Global spread of *Salmonella enteritidis* via centralized sourcing and international trade of poultry breeding stocks. 1. *Nat Commun* 12:5109. <https://doi.org/10.1038/s41467-021-25319-7>
28. NataliaG. 2021. Bodies that count: administering multispecies in Palestine/Israel's borderlands. <https://doi.org/https://journals.sagepub.com/doi/full/10.1177/2514848620901445>
29. Nayak R, Stewart TM, Nawaz MS. 2005. PCR identification of *Campylobacter coli* and *Campylobacter Jejuni* by partial sequencing of virulence genes. *Mol Cell Probes* 19:187–193. <https://doi.org/10.1016/j.mcp.2004.11.005>
30. Paião FG, Arisitides LGA, Murate LS, Vilas-Bôas GT, Vilas-Boas LA, Shimokomaki M. 2013. Detection of salmonella Spp, *Salmonella enteritidis* and *Typhimurium* in naturally infected broiler chickens by a multiplex PCR-based assay. *Braz J Microbiol* 44:37–41. <https://doi.org/10.1590/S1517-83822013005000002>
31. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. 5. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
32. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 22:266. <https://doi.org/10.1186/s13059-021-02483-z>
33. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. Rasttk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>
34. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* 50:1355–1361. <https://doi.org/10.1128/JCM.06094-11>
35. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C, Fields PI, Deng X, Ledebore NA. 2015. *Salmonella* serotype determination utilizing high-throughput genome sequencing data. *J Clin Microbiol* 53:1685–1692. <https://doi.org/10.1128/JCM.00323-15>
36. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
37. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res*. 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>
38. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
39. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
40. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, Edalatmand A, Petkau A, Syed SA, Tsang KK, Baker SJC, Dave M, McCarthy MC, Mukiri KM, Nasir JA, Golbon B, Imtiaz H, Jiang X, Kaur K, Kwong M, Liang ZC, Niu KC, Shan P, Yang JYJ, Gray KL, Hoad GR, Jia B, Bhandu T, Carfrae LA, Farha MA, French S, Gordzevich R, Rachwalski K, Tu MM, Bordeleau E, Dooley D, Griffiths E, Zubyk HL, Brown ED, Maguire F, Beiko RG, Hsiao WWL, Brinkman FSL, Van Domselaar G, McArthur AG. 2023. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res* 51:D690–D699. <https://doi.org/10.1093/nar/gkac920>
41. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wiczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. Resfinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* 75:3491–3500. <https://doi.org/10.1093/jac/ckaa345>
42. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. <https://doi.org/10.1186/1471-2164-12-402>

5. Discussion

5.1. ONT MinION is a state of the art next-generation sequencing platform

In my work, I have sequenced and analysed genes and genomes of AT-rich organisms, like *P. falciparum* and *C. jejuni*, as well as genes and genomes of GC-rich genomes like nontuberculous mycobacteria using the ONT MinION. It is well known that GC-rich genomes can be difficult to sequence with multiple NGS platforms and library preparation methods (Browne et al., 2020). For example, Illumina MiSeq quality scores were significantly lower for sequences with GC contents above 65% (Browne et al., 2020). Low GC regions have lower coverage in Illumina and PacBio sequencing, however ONT sequencing remained unbiased (Browne et al., 2020). The choice of PCR conditions and enzymes introduces sequencing bias by preferably amplifying GC-rich or AT-rich targets and further bias can be introduced later in the library preparation for Illumina sequencing (Aird et al., 2011). Whitford *et al.* reported read depth variation in ONT sequencing of equimolar ratios of amplicons, with preferential sequencing of longer amplicons and amplicons with lower GC content (Whitford et al., 2022). Li *et al.* compared ONT to PacBio and Ion Torrent for viral detection and observed a lower mean GC content among ONT reads compared to the other technologies (Li et al., 2020). Read lengths did not have an effect on ONT basecalling quality, but the beginning and end of the sequence had lower basecalling quality (Krishnakumar et al., 2018).

It is commonly accepted that a combination of short-read and long-read sequencing technologies is best to obtain high quality reference genomes for vertebrates, since long reads enable resolution of complex repeats and remove errors, as observed in previous short-read assemblies (Rhie et al., 2021). Polishing of long-read assemblies with short reads obtained from the same sample usually improves per-base accuracy, however reads need to properly be split between heterozygous haplotypes in diploid genomes (Rhie et al., 2021). Structural variants (SV) such as insertions and deletions can be detected from human genome sequencing data for ONT and PacBio with similar performance, in the case of ONT sequencing, each genome was run on two PromethION flow cells generating 100 Megabases to achieve 33x coverage of the human reference genome (Fatima et al., 2020). ONT sequencing was used for a large-scale analysis of 3,622 individuals from Iceland for unbiased detection of SV and linking of SV with phenotypic traits (Beyter et al., 2021).

With the right setup, ONT data can be analysed during a sequencing run that can be stopped once sufficient data has been generated. The current main ONT basecaller *Guppy* based on Bonito's neural network architecture performed well in a recent basecaller benchmark analysis (Pagès-Gallego and de Ridder, 2023). Different basecalling models exist for *Guppy*, fast basecalling can be done in real time

on an Mk1C device, high accuracy basecalling is about eight times slower, but still possible on an Mk1C and the highest fidelity super-accuracy basecalling requires eight fold more computational power than high accuracy basecalling. Super-accuracy basecalling for flow cell R9.4.1 had similar accuracy as high accuracy basecalling with additional short read polishing for *Klebsiella pneumoniae* WGS and was able to determine resistance and virulence factors as well as assign lineages, but failed to determine SNP distances for outbreak cluster assignment (Foster-Nyarko et al., 2023).

Most analysis tools are based on command-line tools run on UNIX systems. Cloud-based analysis platforms such as BugSeq (Fan et al., 2021) and ONT EPI2ME are easy to use and fast, but cannot be adapted to specific needs. Minimap2 is widely used for mapping reads to a reference database (Li, 2018). The ARTIC Network released a bioinformatics protocol for automated SARS-CoV-2 consensus generation that combined different existing tools (Loman et al., 2020). *De novo* assembly tools such as Flye (Kolmogorov et al., 2019) and Canu (Koren et al., 2017) work best with read lengths of several kilobases, requiring high quality input DNA.

5.2. Molecular surveillance of emerging viruses

We generated over 200 high-quality SARS-CoV-2 genomes from upper respiratory tract samples collected between March 2020 and January 2022 in Equatorial Guinea using an amplicon sequencing approach. The first version of the SARS-CoV-2 amplicon sequencing protocol was provided by the ARTIC network in January 2020 and was updated and improved multiple times to adapt to circulating variants (Tyson et al., 2020). Reagent costs below £10 per sample were achieved using reduced volumes, cheaper reagents and a simplified workflow (Tyson et al., 2020). A comparison of SARS-CoV-2 sequencing approaches showed the superior performance of amplicon-based sequencing for low viral loads (Charre et al., 2020). Hybridisation capture virus enrichment provided an even depth of coverage across the genome including genome ends for higher viral loads (Charre et al., 2020). Metagenomics sequencing should be used to investigate loss of coverage of targeted approaches due to virus evolution (Charre et al., 2020). Similarly, direct sequencing of Zika virus from clinical material was difficult due to high host background and low viral loads in serum samples (Faria et al., 2016).

A mean 1.27% of reported SARS-CoV-2 infections were sequenced from African isolates, comparable to the global mean (1.31%), but infection detection rates were ten times lower in sub-Saharan Africa than the global average (Barber et al., 2022, Tegally et al., 2022). Different recommendations exist for the proportion of SARS-CoV-2 positive samples to be sequenced to detect and follow the spread of emerging lineages. Brito *et al.* set a benchmark for genomic surveillance, to sequence 0.5% of positive cases within 21 days after sample collection (Brito et al., 2022). Han *et al.* recommended testing 0.1% of the population per day and sequencing 5 to 10% of positive SARS-CoV-2 samples at every fourth

tertiary health facility within one week to detect emerging variants within 30 days (Han et al., 2023). In sub-Saharan Africa, the median sequencing turnaround time for positive SARS-CoV-2 samples varied from 51 days for locally sequenced samples to 113 days for isolates sequenced outside of Africa (Tegally et al., 2022).

An alternative to WGS is targeted sequencing of regions of interest. For SARS-CoV-2, sequencing a region of the Spike gene was used to detect key mutations of VOC (Singh et al., 2022). For Mpox surveillance, a similar protocol where regions containing identified SNPs were amplified by PCR and subsequently sequenced using ONT was used (Buenestado-Serrano et al., 2023). Results whether the sample belonged to the outbreak or not were generated within five hours and for €14 per sample, providing more information than RT-qPCR with similar turnaround time and costs (Buenestado-Serrano et al., 2023).

Additionally, mutation-specific RT-qPCR can be used to screen larger numbers of SARS-CoV-2 positive samples for known spike gene mutations at low cost and high speed to complement WGS. Our work has shown that primers and probes with locked nucleic acids (LNAs) enhance the discriminatory power of RT-qPCR at SNP level (Bechtold et al., 2021). Different VOC and VOI can be distinguished using distinct spike gene mutations and deletions (Wang et al., 2021a). LNA-based RT-qPCR assays can be adapted to emerging variants with novel SNPs rapidly (Bechtold et al., 2021). Mutation-specific RT-qPCR assays were successfully used to detect a novel VOC-like lineage with an unusual combination of SNPs in a SARS-CoV-2 isolate from Equatorial Guinea and this SARS-CoV-2 lineage (B.1.620) was later found to have originated in Central Africa and to be introduced to Europe by travellers with subsequent local transmission (Dudas et al., 2021). Since wild type and mutated sequence were detected simultaneously using two different fluorophores, co-infections could be identified in 2.1% of analysed samples (Hosch et al., 2022a). For the sample with the highest viral load determined by RT-qPCR, WGS confirmed the presence of almost all lineage-defining mutations of Beta VOC and Delta VOC (Hosch et al., 2022a). We transferred the mutation-specific RT-qPCR to the portable diaxxoPCR platform with ready-to-use cartridges containing lyophilized reagents. Little adaptation was required and the performance was comparable to a standard qPCR machine, however the PCR runtime was reduced by half and no cold chain was required, rendering it an attractive future approach for decentralized testing (Bechtold et al., 2021).

The unbiased approach of metagenomics NGS allows the timely identification of a known or novel pathogen in an outbreak situation, improving the public health emergency response (Yek et al., 2022). Norovirus was identified from a stool sample stemming from a diarrhoea outbreak within 10 hours using ONT, 20 hours using Ion Torrent and 66 hours using PacBio (Li et al., 2020). Although the

consensus sequence generated with ONT differed from the other sequencing platforms, it was sufficient to identify the viral pathogen (Li et al., 2020). Within ten days after confirmation of the Ebola virus outbreak in DRC, ONT sequencing identified the virus strain as Zaire Ebola virus (Mbala-Kingebeni et al., 2019). In the Ebola epidemic from 2013 to 2016 in West Africa, WGS was done for about 5% of clinical cases, however most sequencing was performed by international scientists and local public health laboratories received little capacity building (Holmes et al., 2016b). During the Ebola outbreak in DRC in 2018-2019, Illumina sequencing became available in two laboratories and 24% of positive samples were sequenced in country (Kinganda-Lusamaki et al., 2021). The local sequencing and bioinformatics capacity was later used to support the SARS-CoV-2 response (Kinganda-Lusamaki et al., 2021).

The co-circulation of multiple pathogens resulting in similar clinical symptoms in diseased people can result in misclassification and inaccurate reporting of diseases (Wilson et al., 2019). One example is the co-circulation of Dengue virus and Zika virus during a Chikungunya virus outbreak in Thailand, where all three arboviruses were found in serum of patients suspected to be infected with Chikungunya virus (Khongwichit et al., 2022). SARS-CoV-2 and Influenza present with flu-like symptoms and are difficult to distinguish based on clinical symptoms alone (Havasi et al., 2022). About 19% of febrile patients in an outpatient fever clinic in Japan were SARS-CoV-2 positive and among febrile patients presenting with a sore throat and a cough the SARS-CoV-2 positivity rate was as high as 45% (Inaba et al., 2023). The high rate of asymptomatic SARS-CoV-2 infections that are able to transmit SARS-CoV-2 has negatively impacted control measures. Cross-sectional studies reported a median of 45.6% asymptomatic infections at the time of testing, in longitudinal studies 72.3% did not develop any symptoms during a follow-up period, resulting in the conclusion that at least one third of total SARS-CoV-2 infections remain asymptomatic (Oran and Topol, 2021). Therefore, it is relevant to identify the pathogen in persons with clinical symptoms to provide adequate treatment and to monitor circulating pathogens in asymptomatic carriers to prevent transmission to at-risk populations.

5.3. Molecular monitoring of diagnostic escape in multiclonal infections

We extracted nucleic acids from over 2,500 used malaria RDTs collected during the 2018 malaria indicator survey (MIS) on Bioko Island. These samples were tested for *Plasmodium* spp. parasites by RT-qPCR, malaria species were identified by qPCR and the *pfk13* propeller region was amplified and sequenced in *P. falciparum* positive samples with sufficiently high parasite densities (Guirou et al., 2020). The extracted nucleic acids were used to assess the diagnostic performance of the RDTs used in the 2018 MIS. We found a high proportion of false-positive RDTs and were able to associate this observation with PfHRP2 persistence due to recent use of anti-malarial drugs (Hosch et al., 2022b). Recently developed highly sensitive HRP2/LDH RDTs have five-fold lower limits of antigen detection

compared to conventional RDTs, resulting in a higher sensitivity, however detection of false-positive RDTs are more common due to persisting PfHRP2, resulting in lower specificity (Niyukuri et al., 2022).

A side-by-side comparison of DNA extraction protocols for *P. falciparum* showed that five- to ten-fold more DNA can be recovered from fresh whole blood compared to dried blood spots (DBS) of the same blood volume, with the magnetic bead-based method outperforming the spin-column based method (Holzschuh and Koepfli, 2022). The low proportion of 4.5% false-negative RDTs in the 2018 MIS can be explained by the low amount of dried blood retained on a RDT and the low parasite densities of these asymptomatic infections, limiting the sensitivity of the RT-qPCR-based *Plasmodium* spp. detection (Hosch et al., 2022b). We have successfully adapted the RT-qPCR assay used to screen for *Plasmodium* spp. in DBS to the portable, rapid and mobile PCR cycler previously used for SARS-CoV-2 mutation detection (Bechtold et al., 2023). Tween-Chelex has been used for DNA extraction from DBS for whole genome sequencing of *P. falciparum* (Teyssier et al., 2021). Even though Tween-Chelex isolation does not remove all contaminants, less parasite DNA is lost during purification compared to commercial extraction kits, resulting in similar detection limits for *P. falciparum* at significantly lower costs (Holzschuh and Koepfli, 2022). To further facilitate the workflow for molecular *Plasmodium* spp. detection, we have developed a method based on blood from DBS. A punch of the DBS is boiled for three minutes in Chelex solution before applying the supernatant on a preloaded cartridge containing RT-qPCR master mix, primers and probes in a freeze-dried form and running a RT-qPCR on the rapid PCR cycler (Bechtold et al., 2023).

In recent years, *P. falciparum* with *pfhrp2/3* gene deletions have been reported from the Americas, Africa and Asia using different methods (Martíáñez-Vendrell et al., 2022). Surveillance of these gene deletions is highly relevant, as parasites with *pfhrp2/3* gene deletions are no longer detected by HRP2-based RDTs, threatening reliable malaria diagnosis (Verma et al., 2018). In malaria-endemic settings, infections with multiple different clones are common (Koepfli and Mueller, 2017). In multiclonal infections where only one strain carries a gene deletion, qualitative protocols for detection of *pfhrp2/3* deletions fail to detect the deletion that is masked by other strains (Schindler et al., 2019). The implementation of a validated protocol with high sensitivity and the ability to detect masked deletions in multiclonal infections based on molecular methods would improve the surveillance of *pfhrp2/3* gene deletions (Martíáñez-Vendrell et al., 2022). We ran a multiplex qPCR assay to detect *pfhrp2/3* gene deletions in samples collected during the 2018 MIS (Schindler et al., 2019). To avoid false reporting of *pfhrp2/3* gene deletions, we restricted the analysis to samples with sufficiently high parasite densities, successful amplification of the internal control gene and independent amplification by nested PCR of one *P. falciparum* merozoite surface protein gene (Hosch et al., 2022b). We found evidence of masked *pfhrp2/3* gene deletions and *pfhrp3* gene deletions but no *pfhrp2* gene deletions

in this setting with a high proportion of multiclonal infections, indicating the presence of *P. falciparum* strains carrying *pfhrp2* deletions on Bioko Island (Hosch et al., 2022b).

The nucleic acids extracted from these RDTs were additionally screened for the presence of DNA from the highly neglected filarial nematodes *M. perstans* and *L. loa* (Yoboue et al., 2022). Even though only a single drop blood is retained on a RDT, *M. perstans* and *L. loa* positivity rates in RDTs were similar to a cross-sectional study studying filarial nematodes on Bioko Island in 2014 using whole blood (Ta et al., 2018). However, we were not able to study filarial nematodes residing in the skin nor were we able to study *W. bancrofti* with their nocturnal periodicity due to the design of the MIS (Yoboue et al., 2022). In contrast to malaria, where an international standard for *P. falciparum* with defined parasitaemia exists, approximate quantification as well as determination of sensitivity and specificity would require paired samples with qPCR quantification cycle value and microscopy-based number of microfilariae and was therefore not feasible.

In summary, this series of manuscripts have shown that malaria RDTs which are routinely used by malaria control programs for MIS in sub-Saharan African countries can be used to store nucleic acids at room temperature for months to years. Extracted nucleic acids can be used for molecular screening of blood-borne pathogens as well as further characterization by sequencing. In Equatorial Guinea, yearly MIS are conducted and the collection site for each volunteer tested by RDT is georeferenced (Guerra et al., 2019). The molecular detection of *M. perstans* on these georeferenced samples allowed us to understand the age and geographical distribution of this under-researched parasite at very limited additional cost (manuscript in preparation).

5.4. Molecular monitoring of *P. falciparum* malaria drug resistance markers

SNPs conferring resistance to different antimalarials are distributed across multiple molecular resistance markers (Murmu et al., 2021). Sanger sequencing has limited sensitivity to detect minority clones in multiclonal *P. falciparum* infections that are encountered in medium to high malaria transmission settings (Sharma et al., 2016). A recent comparison of Illumina and Ion Torrent targeted amplicon deep sequencing of molecular resistance markers and showed the ability of both methods to detect minority alleles at 1% density with a read depth of 500x (Kunasol et al., 2022). In some genes, SNPs conferring drug resistance are distributed at a larger distance on the molecular resistance marker than the maximum read length of Illumina sequencing, making it necessary to amplify multiple fragments to cover all SNPs of interest (Nag et al., 2017). If SNPs are distributed across multiple fragments, only SNP frequencies can be obtained, but no haplotype can be assigned unequivocally (Nag et al., 2017). A different protocol that amplified the full length molecular resistance marker but then fragmented the PCR products for Ion Torrent sequencing found heterozygous variants at some

positions indicative of multiclonal infections but could not identify haplotypes from SNPs (Rao et al., 2016).

Our work likely fills an important gap of drug resistance marker monitoring tools at full length markers and in multiclonal infections by developing the PHARE pipeline. The PHARE pipeline for detection of drug resistance conferring SNPs in *P. falciparum* shows the ability of ONT sequencing to accurately identify SNPs using novel bioinformatics tools in combination with sufficient read depth. Different bioinformatics tools exist for SNP analysis from ONT sequencing data. An ONT sequencing approach for six SNPs related to perioperative outcomes mapped a probe containing the SNP site to the reads *in silico* and estimated allele frequencies from sequences with perfect match (Tabata et al., 2022). With this approach using 1000 reads per site and sample, allele frequencies for heterozygote and homozygote variants were close to 50% and 100% respectively. The entire workflow from DNA isolation, PCR, ONT sequencing to bioinformatics could be performed in less than four hours (Tabata et al., 2022). Others have used the ONT reads of thirty genetic loci associated with the autism spectrum disorder that were mapped to custom references and called variants were compared to Sanger sequencing (Whitford et al., 2022). After excluding variants in low complexity regions such as homopolymers, a coverage of 130x per amplicon was sufficient for accurate genotyping (Whitford et al., 2022). Forensic SNP and short tandem repeat (STR) profiling for forensic applications like human identity and kinship testing with ONT reads resulted in successful SNP profiling at all 94 loci with an overall accuracy of 99.96% (Ren et al., 2021). STR typing was possible for 32 out of 55 loci due to short insertion and deletion errors in ONT reads that hampered repeat recognition (Ren et al., 2021). The PHARE pipeline has the advantage over other pipelines that it is not limited to predefined SNPs but determines SNP sites using an iterative approach. In summary, the combination of targeted whole gene amplification of markers of interest with ONT sequencing resulting in high sequencing coverage will be able to overcome current problems of detecting minority malaria clones carrying drug resistance conferring combinations of SNPs in the field.

5.5. One Health pathogen surveillance using whole genome sequencing

While traditional laboratory workflows require multiple complex laboratory processes that might take days to weeks for bacterial species identification and susceptibility testing, WGS can be performed in a single step after pathogen isolation by culture (Didelot et al., 2012). The genome sequence of bacterial pathogens contains the information required for species identification and can guide treatment decisions and public health measures (Didelot et al., 2012).

Previous studies of genotypic resistance conducted in the Middle East for *C. jejuni* and *S. enterica* have relied on amplification and sequencing of few resistance genes of interest (Abukhattab et al., 2022).

Here, we have used a collection of *C. jejuni* and *S. enterica* isolated from chicken manure, chicken meat and human faeces in the white meat production chain to sequence by ONT. We found high genetic similarities between isolates of *C. jejuni* and *S. enterica* from all three sample types collected during the same time period. All *C. jejuni* collected in the same year showed the same sequence type and shared AMR genes and point mutations. However the AMR markers of *C. jejuni* isolated from chicken manure of the same farms but from consecutive years were completely different. Importantly, *C. jejuni* isolates from hospitalized gastroenteritis patients in Palestine were closely related to isolates from chicken faeces highlighting the power of WGS in the context of One Health approaches (Abukhattab et al., 2023).

In Israel, *C. jejuni* was isolated from human, poultry and cattle between 2003 and 2012 and the isolates clustered by WGS independent of sample source but no clear epidemiological link was established between isolates of a cluster (Rokney et al., 2018). Closely related *C. jejuni* isolates shared acquired AMR genes and point mutations and the accordance between AMR prediction from WGS data and phenotypic resistance was high (Rokney et al., 2020). However, WGS-based prediction was unable to find some resistant phenotypes and could not determine minimum inhibitory concentrations, requiring further optimization (Rokney et al., 2020). Similarly, *Campylobacter* spp. from chicken and human diarrhoeal disease cases in Botswana were highly related using core genome phylogenetic analysis and the isolates carried a high number of antimicrobial resistance genes (de Vries et al., 2018).

A study comparing bioinformatics pipelines for AMR prediction from WGS data of clinical carbapenem-resistant *Enterobacterales* between different laboratories with the same sequence data resulted in discordant results between laboratories (Doyle et al., 2020). Doyle *et al.* report that sequence quality, read depth and choice of reference database impacted resistance prediction the most, showing the need for sufficient coverage and strict quality control (Doyle et al., 2020). There is a need for an up-to-date curated database with stringent sequence identity cut-offs, as databases may differ and exact gene variants are required for reliable phenotypic AMR prediction of clinically relevant bacterial pathogens (Doyle et al., 2020). A catalogue of genomic mutations associated with AMR exists for the *M. tuberculosis* complex, where sensitivities were above 80% and specificities above 90% for first-line drugs (Walker et al., 2022). A machine learning model to predict minimum inhibitory concentration for 15 antibiotics in non-typhoidal *Salmonella* species had an accuracy of 95% (Nguyen et al., 2019).

Apart from AMR prediction, the generation of WGS data also allows for improved bacterial species classification. MLST based on several housekeeping genes using first-generation sequencing approaches like Sanger sequencing is time consuming and expensive, therefore tools have been developed to extract MLST alleles from raw or assembled WGS data (Larsen et al., 2012). While

phenotypic serotyping is labour-intensive and time-consuming, genetic markers can be extracted from WGS data of *S. enterica* and predict the serotype rapidly and accurately at no additional costs (Zhang et al., 2015). Further distinction of isolates can be done from WGS data using SNP-based or allele-based approaches. The core genome MLST approach showed high concordance with epidemiological data for clustering outbreak isolates, whole genome MLST and high quality SNP approaches were able to further differentiate *Campylobacter* spp. isolates (Joseph et al., 2023).

In summary, with decreasing sequencing costs and improved bioinformatics pipelines, the use of WGS in microbiology based on ONT MinION has the potential to become an important tool outside of academia for regular disease outbreak tracking and monitoring in combination with drug resistance surveillance. ONT based disease monitoring might become an important tool particularly in fragile socio-political environments where severe limitations on equipment and lack of human resources are evident (Truppa and Abo-Shehada, 2020).

5.6. Molecular diagnosis of fever causing viruses, bacteria and parasites

A diagnostic platform able to support a range of molecular tests would allow the rapid differentiation of fever causing pathogens (Mehta et al., 2019). A RT-qPCR for rabies virus with a limit of detection of 10 viral copies per microliter has been developed for a portable rapid qPCR device, however prior RNA extraction is required (Demetria et al., 2023). Inhibitor resistant RT-qPCR reagents have been shown to be able to directly amplify targets from clinical and environmental samples containing PCR inhibitors (Trombley Hall et al., 2013). SARS-CoV-2 detection directly from respiratory samples could be performed within 36 minutes on a portable PCR device with minimal hands-on-time and low costs (Wee et al., 2020). RT-qPCR for Dengue virus detection directly from plasma samples on a low-cost portable instrument had a sensitivity of 78.3% and a specificity of 93.3% compared to a commercial RT-qPCR kit with extracted RNA on a conventional laboratory qPCR instrument (Mehta et al., 2019). RT-qPCR based *P. falciparum* detection directly from blood with the same instrument had a sensitivity and specificity of 100%, with a limit of detection of 0.1 *Plasmodium* spp. per microliter (Taylor et al., 2017). Our protocol for simplified extraction from DBS combined with RT-qPCR to detect *Plasmodium* spp. on a rapid qPCR cyclers with preloaded cartridges had comparable performance to a laboratory-based RT-qPCR approach (Bechtold et al., 2023). The same device has been used for detection of *Schistosoma haematobium*, Mpox virus, bacteria causing sexually transmitted diseases (gonorrhoea and chlamydia), Influenza A and Influenza B (unpublished data).

6. Conclusion and future directions

We successfully developed and applied tools for molecular surveillance to a variety of infectious diseases. The work described in this thesis was carried out during the COVID-19 pandemic, tools were implemented as they were developed. Therefore, the work on SARS-CoV-2 had a direct impact on the public health response in Equatorial Guinea.

Molecular testing for pathogens or variants on a mobile RT-qPCR device combined with ONT sequencing increases the speed to respond to challenges. RT-qPCR approaches are suitable to rapidly screen large numbers of samples and can be used to select relevant subsets for ONT sequencing. Genetic changes impacting the performance of primers and probes used for RT-qPCR are detected in ONT sequencing data and can optimize primer design.

Many public health laboratories in sub-Saharan Africa have been supplied with qPCR devices and NGS platforms as a response to the COVID-19 pandemic. Now that COVID-19 is no longer a public health emergency of international concern, these platforms should be used for routine pathogen surveillance to avoid loss of laboratory and bioinformatics expertise. A decentralized qPCR platform and a portable sequencing platform like the ONT MinION require little additional equipment. They can therefore easily be implemented in basic laboratories or even independent of existing laboratories, filling essential gaps in socioeconomically fragile environments. Acquisition costs of ONT MinION are negligible and running costs comparable to other sequencing platforms. Protocols can easily be adapted to novel research questions, making it possible to quickly react to changing circumstances.

These characteristics make the ONT MinION a suitable tool for the molecular disease surveillance of any pathogen in any environment. Outbreaks of known and novel pathogens can be confirmed and their evolution can be monitored, helping public health authorities to an informed decision with a short turnaround time. Variants that may escape vaccine- or natural exposure-derived immunity can be characterised, as has been shown for SARS-CoV-2 VOC. Analysis of known markers of antibiotic or antimalarial resistance can inform on suitable drugs to be recommended in the local context.

Laboratory workflows for library preparation take one or two days, sequencing up to three days, but the longest time is spent on bioinformatics analyses of the generated data. Data transfer from the ONT device to the computing cluster could be optimized by integrating the sequencer into the network, as raw signal data has high data storage needs. The most accurate basecalling is currently only possible on a high performance computer. Data cleaning and filtering need to be rigorously applied. Most tools only exist for UNIX system and require significant computational resources due to the high amount of data. Automated bioinformatics pipelines for rapid analysis of general questions

exist, however in-depth analysis requires adaptation of pipelines and integration of different tools. After establishing a pipeline for a specific research question, little manual work is required.

We have implemented ONT sequencing in the Centre Suisse de Recherches Scientifiques en Côte d'Ivoire in spring 2022, where seven students and three laboratory technicians have been trained in library preparation and sequencing data analysis. As next steps, we plan to implement the ONT sequencing technologies in laboratories in Palestine and Equatorial Guinea. This technology transfer will help to generate local molecular testing and sequencing capacity preparing for future outbreaks of emerging or re-emerging pathogens.

7. References

ABUKHATTAB, S., HOSCH, S., ABU-RMEILEH, N. M. E., HASAN, S., VONAESCH, P., CRUMP, L., HATTENDORF, J., DAUBENBERGER, C., ZINSSTAG, J. & SCHINDLER, T. 2023. Whole-genome sequencing for One Health surveillance of antimicrobial resistance in conflict zones: a case study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine. *Appl Environ Microbiol*, e0065823.

ABUKHATTAB, S., TAWHEEL, H., AWAD, A., CRUMP, L., VONAESCH, P., ZINSSTAG, J., HATTENDORF, J. & ABU-RMEILEH, N. M. E. 2022. Systematic Review and Meta-Analysis of Integrated Studies on *Salmonella* and *Campylobacter* Prevalence, Serovar, and Phenotyping and Genetic of Antimicrobial Resistance in the Middle East—A One Health Perspective. *Antibiotics*, 11, 536.

AIRD, D., ROSS, M. G., CHEN, W. S., DANIELSSON, M., FENNELL, T., RUSS, C., JAFFE, D. B., NUSBAUM, C. & GNIRKE, A. 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol*, 12, R18.

ALHASSAN, A., LI, Z., POOLE, C. B. & CARLOW, C. K. S. 2015. Expanding the MDx toolbox for filarial diagnosis and surveillance. *Trends in Parasitology*, 31, 391-400.

ANAHTAR, M. N., YANG, J. H. & KANJILAL, S. 2021. Applications of Machine Learning to the Problem of Antimicrobial Resistance: an Emerging Model for Translational Research. *J Clin Microbiol*, 59, e0126020.

ASHLEY, E. A., DHORDA, M., FAIRHURST, R. M., AMARATUNGA, C., LIM, P., SUON, S., SRENG, S., ANDERSON, J. M., MAO, S., SAM, B., SOPHA, C., CHUOR, C. M., NGUON, C., SOVANNAROTH, S., PUKRITTAYAKAMEE, S., JITTAMALA, P., CHOTIVANICH, K., CHUTASMIT, K., SUCHATSOONTHORN, C., RUNCHAROEN, R., HIEN, T. T., THUY-NHIEN, N. T., THANH, N. V., PHU, N. H., HTUT, Y., HAN, K.-T., AYE, K. H., MOKUOLU, O. A., OLAOSEBIKAN, R. R., FOLARANMI, O. O., MAYXAY, M., KHANTHAVONG, M., HONGVANTHONG, B., NEWTON, P. N., ONYAMBOKO, M. A., FANELLO, C. I., TSHEFU, A. K., MISHRA, N., VALECHA, N., PHYO, A. P., NOSTEN, F., YI, P., TRIPURA, R., BORRMANN, S., BASHRAHEIL, M., PESHU, J., FAIZ, M. A., GHOSE, A., HOSSAIN, M. A., SAMAD, R., RAHMAN, M. R., HASAN, M. M., ISLAM, A., MIOTTO, O., AMATO, R., MACINNIS, B., STALKER, J., KWIATKOWSKI, D. P., BOZDECH, Z., JEEYAPANT, A., CHEAH, P. Y., SAKULTHAEW, T., CHALK, J., INTHARABUT, B., SILAMUT, K., LEE, S. J., VIHOKHERN, B., KUNASOL, C., IMWONG, M., TARNING, J., TAYLOR, W. J., YEUNG, S., WOODROW, C. J., FLEGG, J. A., DAS, D., SMITH, J., VENKATESAN, M., PLOWE, C. V., STEPNIIEWSKA, K., GUERIN, P. J., DONDORP, A. M., DAY, N. P., WHITE, N. J. & TRACKING RESISTANCE TO ARTEMISININ, C. 2014. Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *The New England journal of medicine*, 371, 411-423.

ASHLEY, E. A., PYAE PHYO, A. & WOODROW, C. J. 2018. Malaria. *The Lancet*, 391, 1608-1621.

ASIO, S. M., SIMONSEN, P. E. & ONAPA, A. W. 2009. Mansonella perstans filariasis in Uganda: patterns of microfilaraemia and clinical manifestations in two endemic communities. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 103, 266-273.

BAKAJKA, D. K., NIGO, M. M., LOTSIMA, J. P., MASIKINI, G. A., FISCHER, K., LLOYD, M. M., WEIL, G. J. & FISCHER, P. U. 2014. Filarial antigenemia and Loa loa night blood microfilaremia in an area without bancroftian filariasis in the Democratic Republic of Congo. *Am J Trop Med Hyg*, 91, 1142-1148.

BALIKAGALA, B., FUKUDA, N., IKEDA, M., KATURO, O. T., TACHIBANA, S.-I., YAMAUCHI, M., OPIO, W., EMOTO, S., ANYWAR, D. A., KIMURA, E., PALACPAC, N. M. Q., ODONGO-AGINYA, E. I., OGWANG, M., HORII, T. & MITA, T. 2021. Evidence of Artemisinin-Resistant Malaria in Africa. *New England Journal of Medicine*, 385, 1163-1171.

BARBER, R. M., SORENSEN, R. J. D., PIGOTT, D. M., BISIGNANO, C., CARTER, A., AMLAG, J. O., COLLINS, J. K., ABBAFATI, C., ADOLPH, C., ALLORANT, A., ARAVKIN, A. Y., BANG-JENSEN, B. L., CASTRO, E., CHAKRABARTI, S., COGEN, R. M., COMBS, E., COMFORT, H., COOPERRIDER, K., DAI, X., DAUD, F., DEEN, A., EARL, L., ERICKSON, M., EWALD, S. B., FERRARI, A. J., FLAXMAN, A. D., FROSTAD, J. J., FULLMAN, N., GILES, J. R., GUO, G., HE, J., HELAK, M., HULLAND, E. N., HUNTLEY, B. M., LAZZARATWOOD, A., LEGRAND, K. E., LIM, S. S., LINDSTROM, A., LINEBARGER, E., LOZANO, R., MAGISTRO, B., MALTA, D. C., MÅNSSON, J., MANTILLA HERRERA, A. M., MOKDAD, A. H., MONASTA, L., NAGHAVI, M., NOMURA, S., ODELL, C. M., OLANA, L. T., OSTROFF, S. M., PASOVIC, M., PEASE, S. A., REINER JR, R. C., REINKE, G., RIBEIRO, A. L. P., SANTOMAURO, D. F., SHOLOKHOV, A., SPURLOCK, E. E., SYAILENDRAWATI, R., TOPOR-MADRY, R., VO, A. T., VOS, T., WALCOTT, R., WALKER, A., WIENS, K. E., WIYSONGE, C. S., WORKU, N. A., ZHENG, P., HAY, S. I., GAKIDOU, E. & MURRAY, C. J. L. 2022. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. *The Lancet*, 399, 2351-2380.

BECHTOLD, P., WAGNER, P., HOSCH, S., GREGORINI, M., STARK, W. J., GODY, J. C., KODIALENGUETAMA, E. R., PAGONENDJI, M. S., DONFACK, O. T., PHIRI, W. P., GARCÍA, G. A., NSANZANBANA, C., DAUBENBERGER, C. A., SCHINDLER, T. & VICKOS, U. 2023. Development and evaluation of Plasmopod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance. *PLOS Glob Public Health*, 3, e0001516.

BECHTOLD, P., WAGNER, P., HOSCH, S., SIEGRIST, D., RUIZ-SERRANO, A., GREGORINI, M., MPINA, M., ONDÓ, F. A., OBAMA, J., AYEKABA, M. O., ENGLER, O., STARK, W. J., DAUBENBERGER, C. A. & SCHINDLER, T. 2021. Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable peakPCR Platform. *Anal Chem*, 93, 16350-16359.

BEIER, J. C., DAVIS, J. R., VAUGHAN, J. A., NODEN, B. H. & BEIER, M. S. 1991. Quantitation of Plasmodium falciparum Sporozoites Transmitted in Vitro by Experimentally Infected Anopheles gambiae and Anopheles stephensi. *The American Journal of Tropical Medicine and Hygiene*, 44, 564-570.

BEYTER, D., INGIMUNDARDOTTIR, H., ODDSSON, A., EGGERTSSON, H. P., BJORNSSON, E., JONSSON, H., ATLASON, B. A., KRISTMUNDSDOTTIR, S., MEHRINGER, S., HARDARSON, M. T., GUDJONSSON, S. A., MAGNUSDOTTIR, D. N., JONASDOTTIR, A., JONASDOTTIR, A., KRISTJANSSON, R. P., SVERRISSON, S. T., HOLLEY, G., PALSSON, G., STEFANSSON, O. A., EYJOLFSSON, G., OLAFSSON, I., SIGURDARDOTTIR, O., TORFASON, B., MASSON, G., HELGASON, A., THORSTEINSDOTTIR, U., HOLM, H., GUDBJARTSSON, D. F., SULEM, P., MAGNUSSON, O. T., HALLDORSSON, B. V. & STEFANSSON, K. 2021. Long-read sequencing of 3,622 Icelanders provides insight into the role of structural variants in human diseases and other traits. *Nature Genetics*, 53, 779-786.

BHATT, S., WEISS, D. J., CAMERON, E., BISANZIO, D., MAPPIN, B., DALRYMPLE, U., BATTLE, K., MOYES, C. L., HENRY, A., ECKHOFF, P. A., WENGER, E. A., BRIËT, O., PENNY, M. A., SMITH, T. A., BENNETT, A., YUKICH, J., EISELE, T. P., GRIFFIN, J. T., FERGUS, C. A., LYNCH, M., LINDGREN, F., COHEN, J. M., MURRAY, C. L. J., SMITH, D. L., HAY, S. I., CIBULSKIS, R. E. & GETTING, P. W. 2015. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature*, 526, 207-211.

BLASER, M. J. & ENGBERG, J. 2008. Clinical Aspects of Campylobacter jejuni and Campylobacter coli Infections. *Campylobacter*.

BLOLAND, P. B., LACKRITZ, E. M., KAZEMBE, P. N., WERE, J. B. O., STEKETEE, R. & CAMPBELL, C. C. 1993. Beyond Chloroquine: Implications of Drug Resistance for Evaluating Malaria Therapy Efficacy and Treatment Policy in Africa. *The Journal of Infectious Diseases*, 167, 932-937.

BÖGER, B., FACHI, M. M., VILHENA, R. O., COBRE, A. F., TONIN, F. S. & PONTAROLO, R. 2021. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. *American journal of infection control*, 49, 21-29.

BOGITSH, B. J., CARTER, C. E. & OELTMANN, T. N. 2019a. Chapter 17 - Blood and Tissue Nematodes. In: BOGITSH, B. J., CARTER, C. E. & OELTMANN, T. N. (eds.) *Human Parasitology (Fifth Edition)*. Academic Press.

BOGITSH, B. J., CARTER, C. E. & OELTMANN, T. N. 2019b. Chapter 18 - Arthropods as Vectors. In: BOGITSH, B. J., CARTER, C. E. & OELTMANN, T. N. (eds.) *Human Parasitology (Fifth Edition)*. Academic Press.

BOUSSINESQ, M. 2006. Loiasis. *Ann Trop Med Parasitol*, 100, 715-31.

BOUSSINESQ, M., GARDON, J., GARDON-WENDEL, N. & CHIPPAUX, J.-P. 2003. Clinical picture, epidemiology and outcome of Loa-associated serious adverse events related to mass ivermectin treatment of onchocerciasis in Cameroon. *Filaria Journal*, 2, S4.

BRITO, A. F., SEMENOVA, E., DUDAS, G., HASSLER, G. W., KALINICH, C. C., KRAEMER, M. U. G., HO, J., TEGALLY, H., GITHINJI, G., AGOTI, C. N., MATKIN, L. E., WHITTAKER, C., HOWDEN, B. P., SINTCHENKO, V., ZUCKERMAN, N. S., MOR, O., BLANKENSHIP, H. M., DE OLIVEIRA, T., LIN, R. T. P., SIQUEIRA, M. M., RESENDE, P. C., VASCONCELOS, A. T. R., SPILKI, F. R., AGUIAR, R. S., ALEXIEV, I., IVANOV, I. N., PHILIPOVA, I., CARRINGTON, C. V. F., SAHADEO, N. S. D., BRANDA, B., GURRY, C., MAURER-STROH, S., NAIDOO, D., VON EIJE, K. J., PERKINS, M. D., VAN KERKHOVE, M., HILL, S. C., SABINO, E. C., PYBUS, O. G., DYE, C., BHATT, S., FLAXMAN, S., SUCHARD, M. A., GRUBAUGH, N. D., BAELE, G. & FARIA, N. R. 2022. Global disparities in SARS-CoV-2 genomic surveillance. *Nat Commun*, 13, 7003.

BROWNE, P. D., NIELSEN, T. K., KOT, W., AGGERHOLM, A., GILBERT, M. T. P., PUETZ, L., RASMUSSEN, M., ZERVAS, A. & HANSEN, L. H. 2020. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience*, 9.

BUENESTADO-SERRANO, S., HERRANZ, M., PALOMINO-CABRERA, R., RODRÍGUEZ-GRANDE, C., PEÑAS-UTRILLA, D., MOLERO-SALINAS, A., VEINTIMILLA, C., CATALÁN, P., ALONSO, R., MUÑOZ, P., PÉREZ-LAGO, L. & GARCÍA DE VIEDMA, D. 2023. Rapid Identification of Relevant Microbial Strains by Identifying Multiple Marker Single Nucleotide Polymorphisms via Amplicon Sequencing: Epidemic Monkeypox Virus as a Proof of Concept. *Microbiol Spectr*, 11, e0419622.

CAMERONI, E., BOWEN, J. E., ROSEN, L. E., SALIBA, C., ZEPEDA, S. K., CULAP, K., PINTO, D., VANBLARGAN, L. A., DE MARCO, A., DI IULIO, J., ZATTA, F., KAISER, H., NOACK, J., FARHAT, N., CZUDNOCHOWSKI, N., HAVENAR-DAUGHTON, C., SPROUSE, K. R., DILLEN, J. R., POWELL, A. E., CHEN, A., MAHER, C., YIN, L., SUN, D., SORIAGA, L., BASSI, J., SILACCI-FREGNI, C., GUSTAFSSON, C., FRANKO, N. M., LOGUE, J., IQBAL, N. T., MAZZITELLI, I., GEFFNER, J., GRIFANTINI, R., CHU, H., GORI, A., RIVA, A., GIANNINI, O., CESCHI, A., FERRARI, P., CIPPÀ, P. E., FRANZETTI-PELLANDA, A., GARZONI, C., HALFMANN, P. J., KAWAOKA, Y., HEBNER, C., PURCELL, L. A., PICCOLI, L., PIZZUTO, M. S., WALLS, A. C., DIAMOND, M. S., TELENTI, A., VIRGIN, H. W., LANZAVECCHIA, A., SNELL, G., VEESLER, D. & CORTI, D. 2022. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature*, 602, 664-670.

CHAKRABORTY, C., BHATTACHARYA, M., SHARMA, A. R. & DHAMA, K. 2022. Recombinant SARS-CoV-2 variants XD, XE, and XF: The emergence of recombinant variants requires an urgent call for research - Correspondence. *Int J Surg*, 102, 106670.

CHARRE, C., GINEVRA, C., SABATIER, M., REGUE, H., DESTRAS, G., BRUN, S., BURFIN, G., SCHOLTES, C., MORFIN, F., VALETTE, M., LINA, B., BAL, A. & JOSSET, L. 2020. Evaluation of NGS-based approaches for SARS-CoV-2 whole genome characterisation. *Virus Evol*, 6, veaa075.

CHERF, G. M., LIEBERMAN, K. R., RASHID, H., LAM, C. E., KARPLUS, K. & AKESON, M. 2012. Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nat Biotechnol*, 30, 344-8.

CHESNAIS, C. B., TAKOUGANG, I., PAGUÉLÉ, M., PION, S. D. & BOUSSINESQ, M. 2017. Excess mortality associated with loiasis: a retrospective population-based cohort study. *The Lancet Infectious Diseases*, 17, 108-116.

CHIARA, M., D'ERCHIA, A. M., GISSI, C., MANZARI, C., PARISI, A., RESTA, N., ZAMBELLI, F., PICARDI, E., PAVESI, G., HORNER, D. S. & PESOLE, G. 2021. Next generation sequencing of SARS-CoV-2 genomes: challenges, applications and opportunities. *Brief Bioinform*, 22, 616-630.

CHIN, W., CONTACOS, P. G., COATNEY, G. R. & KIMBALL, H. R. 1965. A Naturally Acquired Quotidian-Type Malaria in Man Transferable to Monkeys. *Science*, 149, 865-865.

CHIU, C. Y. 2013. Viral pathogen discovery. *Curr Opin Microbiol*, 16, 468-78.

CLARKE, J., WU, H.-C., JAYASINGHE, L., PATEL, A., REID, S. & BAYLEY, H. 2009. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*, 4, 265-270.

COLLIGNON, P. J. & MCEWEN, S. A. 2019. One Health-Its Importance in Helping to Better Control Antimicrobial Resistance. *Trop Med Infect Dis*, 4.

COREY, L., BEYRER, C., COHEN, M. S., MICHAEL, N. L., BEDFORD, T. & ROLLAND, M. 2021. SARS-CoV-2 Variants in Patients with Immunosuppression. *New England Journal of Medicine*, 385, 562-566.

CROMWELL, E. A., SCHMIDT, C. A., KWONG, K. T., PIGOTT, D. M., MUPFASONI, D., BISWAS, G., SHIRUDE, S., HILL, E., DONKERS, K. M., ABDOLI, A., ABRIGO, M. R. M., ADEKANMBI, V., ADETOKUNBOH SR, O. O., ADINARAYANAN, S., AHMADPOUR, E., AHMED, M. B., AKALU, T. Y., ALANEZI, F. M., ALANZI, T. M., ALINIA, C., ALIPOUR, V., AMIT SR, A. M. L., ANBER, N. H., ANCUCEANU, R., ANDUALEM, Z., ANJOMSHOA, M., ANSARI, F., ANTONIO, C. A. T., ANVARI, D., APPIAH, S. C. Y., ARABLOO, J., ARNOLD, B. F., AUSLOOS, M., AYANORE SR, M. A., BADIRZADEH, A., BAIG JR, A. A., BANACH SR, M., BARAKI SR, A. G., BÄRNIGHAUSEN, T. W., BAYATI, M., BHATTACHARYYA SR, K., BHUTTA, Z. A., BIJANI, A., BISANZIO, D., BOCKARIE, M. J., BOHLOULI, S., BOHLULI, M., BUTT, Z. A., CANO, J., CARVALHO, F., CHATTU, V. K.,

CHAVSHIN, A. R., CORMIER, N. M., DAMIANI, G., DANDONA, L., DANDONA, R., DARWESH, A. M., DARYANI, A., DASH, A. P., DERIBE, K., DESHPANDE, A., DESSU, B. K., DHIMAL, M., DIANATINASAB, M., DIAZ, D., DO, H. T., EARL, L., EL TANTAWI, M., FARAJ, A., FATTAHI, N., FERNANDES, E., FISCHER, F., FOIGT, N. A., FOROUTAN, M., GUO, Y., HAILU, G. B., HASABALLAH, A. I., HASSANKHANI, H., HERTELIU, C., HIDRU, H. D. D., HOLE, M. K., HON, J., HOSSAIN, N., HOSSEINZADEH, M., HOUSEH, M., HUMAYUN, A., ILESANMI, O. S., ILIC, I. M., ILIC, M. D., IQBAL, U., IRVANI, S. S. N., ISLAM, M. M., JHA, R. P., JI, J. S., JOHNSON, K. B., JOZWIAK, J. J., KABIR, A., KALANKESH, L. R., KALHOR, R., KARAMI MATIN, B., et al. 2020. The global distribution of lymphatic filariasis, 2000–18: a geospatial analysis. *The Lancet Global Health*, 8, e1186-e1194.

CROSS, J. H. 1996. Filarial Nematodes. In: BARON, S. (ed.) *Medical Microbiology*. Galveston (TX): University of Texas Medical Branch at Galveston

Copyright © 1996, The University of Texas Medical Branch at Galveston.

CRUMP, J. A. 2019. Progress in Typhoid Fever Epidemiology. *Clinical Infectious Diseases*, 68, S4-S9.

CRUMP, J. A., SJÖLUND-KARLSSON, M., GORDON, M. A. & PARRY, C. M. 2015. Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial Resistance, and Antimicrobial Management of Invasive Salmonella Infections. *Clinical Microbiology Reviews*, 28, 901-937.

D'ACREMONT, V., KILOWOKO, M., KYUNGU, E., PHILIPINA, S., SANGU, W., KAHAMA-MARO, J., LENGELER, C., CHERPILLOD, P., KAISER, L. & GENTON, B. 2014. Beyond malaria--causes of fever in outpatient Tanzanian children. *N Engl J Med*, 370, 809-17.

DAI, L., SAHIN, O., GROVER, M. & ZHANG, Q. 2020. New and alternative strategies for the prevention, control, and treatment of antibiotic-resistant *Campylobacter*. *Transl Res*, 223, 76-88.

DALRYMPLE, U., ARAMBEPOLA, R., GETHING, P. W. & CAMERON, E. 2018. How long do rapid diagnostic tests remain positive after anti-malarial treatment? *Malaria Journal*, 17, 228.

DE VRIES, S. P. W., VURAYAI, M., HOLMES, M., GUPTA, S., BATEMAN, M., GOLDFARB, D., MASKELL, D. J., MATSHEKA, M. I. & GRANT, A. J. 2018. Phylogenetic analyses and antimicrobial resistance profiles of *Campylobacter* spp. from diarrhoeal patients and chickens in Botswana. *PLoS One*, 13, e0194481.

DEBRUYNE, L., GEVERS, D. & VANDAMME, P. 2008. Taxonomy of the Family *Campylobacteraceae*. *Campylobacter*.

DEMETRIA, C., KIMITSUKI, K., YAHIRO, T., SAITO, N., HASHIMOTO, T., KHAN, S., CHU, M. Y. J., MANALO, D., MANANGGIT, M., QUIAMBAO, B. & NISHIZONO, A. 2023. Evaluation of a real-time mobile PCR device (PCR 1100) for the detection of the rabies gene in field samples. *Trop Med Health*, 51, 17.

DIDELOT, X., BOWDEN, R., WILSON, D. J., PETO, T. E. A. & CROOK, D. W. 2012. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet*, 13, 601-612.

DINNES, J., DEEKS, J. J., BERHANE, S., TAYLOR, M., ADRIANO, A., DAVENPORT, C., DITTRICH, S., EMPERADOR, D., TAKWOINGI, Y., CUNNINGHAM, J., BEESE, S., DOMEN, J., DRETZKE, J., FERRANTE DI RUFFANO, L., HARRIS, I. M., PRICE, M. J., TAYLOR-PHILLIPS, S., HOOFT, L., LEEFLANG, M. M., MCINNES, M. D., SPIJKER, R. & VAN DEN BRUEL, A. 2021. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev*, 3, Cd013705.

DOCHERTY, A. B., HARRISON, E. M., GREEN, C. A., HARDWICK, H. E., PIUS, R., NORMAN, L., HOLDEN, K. A., READ, J. M., DONDELINGER, F., CARSON, G., MERSON, L., LEE, J., PLOTKIN, D., SIGFRID, L., HALPIN, S., JACKSON, C., GAMBLE, C., HORBY, P. W., NGUYEN-VAN-TAM, J. S., HO, A., RUSSELL, C. D., DUNNING, J., OPENSHAW, P. J., BAILLIE, J. K. & SEMPLE, M. G. 2020. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *Bmj*, 369, m1985.

DONDORP, A., NOSTEN, F., STEPNIEWSKA, K., DAY, N. & WHITE, N. 2005a. Artesunate versus quinine for treatment of severe falciparum malaria: a randomised trial. *Lancet*, 366, 717-25.

DONDORP, A. M., DESAKORN, V., PONGTAVORNPINYO, W., SAHASSANANDA, D., SILAMUT, K., CHOTIVANICH, K., NEWTON, P. N., PITISUTTITHUM, P., SMITHYMAN, A. M., WHITE, N. J. & DAY, N. P. J. 2005b. Estimation of the Total Parasite Biomass in Acute Falciparum Malaria from Plasma PfHRP2. *PLOS Medicine*, 2, e204.

DONDORP, A. M., FANELLO, C. I., HENDRIKSEN, I. C., GOMES, E., SENI, A., CHHAGANLAL, K. D., BOJANG, K., OLAOSEBIKAN, R., ANUNOBI, N., MAITLAND, K., KIVAYA, E., AGBENYEGA, T., NGUAH, S. B., EVANS, J., GESASE, S., KAHABUKA, C., MTOVE, G., NADJM, B., DEEN, J., MWANGA-AMUMPAIRE, J., NANSUMBA, M., KAREMA, C., UMULISA, N., UWIMANA, A., MOKUOLU, O. A., ADEDOYIN, O. T., JOHNSON, W. B., TSHEFU, A. K., ONYAMBOKO, M. A., SAKULTHAEW, T., NGUM, W. P., SILAMUT, K., STEPNIEWSKA, K., WOODROW, C. J., BETHELL, D., WILLS, B., ONEKO, M., PETO, T. E., VON SEIDLEIN, L., DAY, N. P. & WHITE, N. J. 2010. Artesunate versus quinine in the treatment of severe falciparum malaria in African children (AQUAMAT): an open-label, randomised trial. *Lancet*, 376, 1647-57.

DONDORP, A. M., LEE, S. J., FAIZ, M. A., MISHRA, S., PRICE, R., TJITRA, E., THAN, M., HTUT, Y., MOHANTY, S., YUNUS, E. B., RAHMAN, R., NOSTEN, F., ANSTEY, N. M., DAY, N. P. J. & WHITE, N. J. 2008. The Relationship between Age and the Manifestations of and Mortality Associated with Severe Malaria. *Clinical Infectious Diseases*, 47, 151-157.

DONDORP, A. M., NOSTEN, F., YI, P., DAS, D., PHYO, A. P., TARNING, J., LWIN, K. M., ARIEY, F., HANPITHAKPONG, W., LEE, S. J., RINGWALD, P., SILAMUT, K., IMWONG, M., CHOTIVANICH, K., LIM, P., HERDMAN, T., AN, S. S., YEUNG, S., SINGHASIVANON, P., DAY, N. P. J., LINDEGARDH, N., SOCHEAT, D. & WHITE, N. J. 2009. Artemisinin resistance in *Plasmodium falciparum* malaria. *The New England journal of medicine*, 361, 455-467.

DOYLE, R. M., O'SULLIVAN, D. M., ALLER, S. D., BRUCHMANN, S., CLARK, T., COELLO PELEGRIN, A., CORMICAN, M., DIEZ BENAVENTE, E., ELLINGTON, M. J., MCGRATH, E., MOTRO, Y., PHUONG THUY NGUYEN, T., PHELAN, J., SHAW, L. P., STABLER, R. A., VAN BELKUM, A., VAN DORP, L., WOODFORD, N., MORAN-GILAD, J., HUGGETT, J. F. & HARRIS, K. A. 2020. Discordant bioinformatic predictions of antimicrobial resistance from whole-genome sequencing data of bacterial isolates: an inter-laboratory study. *Microb Genom*, 6.

DROSTEN, C., GÜNTHER, S., PREISER, W., VAN DER WERF, S., BRODT, H.-R., BECKER, S., RABENAU, H., PANNING, M., KOLESNIKOVA, L., FOUCHIER, R. A. M., BERGER, A., BURGUIÈRE, A.-M., CINATL, J., EICKMANN, M., ESCRIOU, N., GRYWNA, K., KRAMME, S., MANUGUERRA, J.-C., MÜLLER, S., RICKERTS, V., STÜRMER, M., VIETH, S., KLENK, H.-D., OSTERHAUS, A. D. M. E., SCHMITZ, H. & DOERR, H. W. 2003. Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome. *New England Journal of Medicine*, 348, 1967-1976.

DUDAS, G., HONG, S. L., POTTER, B. I., CALVIGNAC-SPENCER, S., NIATOU-SINGA, F. S., TOMBOLOMAKO, T. B., FUH-NEBA, T., VICKOS, U., ULRICH, M., LEENDERTZ, F. H., KHAN, K., HUBER, C., WATTS, A., OLENDRAITÉ, I., SNIJDER, J., WIJNANT, K. N., BONVIN, A. M. J. J., MARTRES, P., BEHILLIL, S., AYOUBA, A., MAIDADI, M. F., DJOMSI, D. M., GODWE, C., BUTEL, C., ŠIMAITIS, A., GABRIELAITĖ, M., KATĖNAITĖ, M., NORVILAS, R., RAUGAITĖ, L., KOYAWEDA, G. W., KANDOU, J. K., JONIKAS, R., NASVYTIENĖ, I., ŽEMECKIENĖ, Ž., GEČYS, D., TAMUŠAUSKAITĖ, K., NORKIENĖ, M., VASILIŪNAITĖ, E., ŽIOGIENĖ, D., TIMINSKAS, A., ŠUKYS, M., ŠARAUSKAS, M., ALZBUTAS, G., AZIZA, A. A., LUSAMAKI, E. K., CIGOLO, J.-C. M., MAWETE, F. M., LOFIKO, E. L., KINGEBENI, P. M., TAMFUM, J.-J. M., BELIZAIRE, M. R. D., ESSOMBA, R. G., ASSOUMOU, M. C. O., MBORINGONG, A. B., DIENG, A. B., JUOZAPAITĖ, D., HOSCH, S., OBAMA, J., AYEKABA, M. O. O., NAUMOVAS, D., PAUTIENIUS, A., RAFAĪ, C. D., VITKAUSKIENĖ, A., UGENSKIENĖ, R., GEDVILAITĖ, A., ČEREŠKEVIČIUS, D., LESAUSKAITĖ, V., ŽEMAITIS, L., GRIŠKEVIČIUS, L. & BAELE, G. 2021. Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions. *Nature Communications*, 12, 5769.

DUPONT, A., ZUE-N'DONG, J. & PINDER, M. 1988. Common occurrence of amicrofilaraemic *Loa loa* filariasis within the endemic region. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 82, 730-730.

ECKER, A., LEHANE, A. M., CLAIN, J. & FIDOCK, D. A. 2012. PfCRT and its role in antimalarial drug resistance. *Trends in parasitology*, 28, 504-514.

EID, J., FEHR, A., GRAY, J., LUONG, K., LYLE, J., OTTO, G., PELUSO, P., RANK, D., BAYBAYAN, P., BETTMAN, B., BIBILLO, A., BJORNSON, K., CHAUDHURI, B., CHRISTIANS, F., CICERO, R., CLARK, S., DALAL, R., DEWINTER, A., DIXON, J., FOQUET, M., GAERTNER, A., HARDENBOL, P., HEINER, C., HESTER, K., HOLDEN, D., KEARNS, G., KONG, X., KUSE, R., LACROIX, Y., LIN, S., LUNDQUIST, P., MA, C., MARKS, P., MAXHAM, M., MURPHY, D., PARK, I., PHAM, T., PHILLIPS, M., ROY, J., SEBRA, R., SHEN, G., SORENSON, J., TOMANEY, A., TRAVERS, K., TRULSON, M., VIECELI, J., WEGENER, J., WU, D., YANG, A., ZACCARIN, D., ZHAO, P., ZHONG, F., KORLACH, J. & TURNER, S. 2009. Real-Time DNA Sequencing from Single Polymerase Molecules. *Science*, 323, 133-138.

ENGELS, D. & ZHOU, X.-N. 2020. Neglected tropical diseases: an effective global response to local poverty-related disease priorities. *Infectious Diseases of Poverty*, 9, 10.

FAN, J., HUANG, S. & CHORLTON, S. D. 2021. BugSeq: a highly accurate cloud platform for long-read metagenomic analyses. *BMC Bioinformatics*, 22, 160.

FARIA, N. R., MELLAN, T. A., WHITTAKER, C., CLARO, I. M., CANDIDO, D. D. S., MISHRA, S., CRISPIM, M. A. E., SALES, F. C. S., HAWRYLUK, I., MCCRONE, J. T., HULSWIT, R. J. G., FRANCO, L. A. M., RAMUNDO, M. S., DE JESUS, J. G., ANDRADE, P. S., COLETTI, T. M., FERREIRA, G. M., SILVA, C. A. M., MANULI, E. R., PEREIRA, R. H. M., PEIXOTO, P. S., KRAEMER, M. U. G., GABURO, N., CAMILO, C. D. C., HOELTGEBAUM, H., SOUZA, W. M., ROCHA, E. C., DE SOUZA, L. M., DE PINHO, M. C., ARAUJO, L. J. T., MALTA, F. S. V., DE LIMA, A. B., SILVA, J. D. P., ZAULI, D. A. G., FERREIRA, A. C. D. S., SCHNEKENBERG, R. P., LAYDON, D. J., WALKER, P. G. T., SCHLÜTER, H. M., DOS SANTOS, A. L. P., VIDAL, M. S., DEL CARO, V. S., FILHO, R. M. F., DOS SANTOS, H. M., AGUIAR, R. S., PROENÇA-MODENA, J. L., NELSON, B., HAY, J. A., MONOD, M., MISCOURIDOU, X., COUPLAND, H., SONABEND, R., VOLLMER, M., GANDY, A., PRETE, C. A., NASCIMENTO, V. H., SUCHARD, M. A., BOWDEN, T. A., POND, S. L. K., WU, C.-H., RATMANN, O., FERGUSON, N. M., DYE, C., LOMAN, N. J., LEMEY, P., RAMBAUT, A., FRAIJI, N. A., CARVALHO, M. D. P. S. S., PYBUS, O. G., FLAXMAN, S., BHATT, S. & SABINO, E. C. 2021. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*, 372, 815-821.

FARIA, N. R., SABINO, E. C., NUNES, M. R., ALCANTARA, L. C., LOMAN, N. J. & PYBUS, O. G. 2016. Mobile real-time surveillance of Zika virus in Brazil. *Genome Med*, 8, 97.

FATIMA, N., PETRI, A., GYLLENSTEN, U., FEUK, L. & AMEUR, A. 2020. Evaluation of Single-Molecule Sequencing Technologies for Structural Variant Detection in Two Swedish Human Genomes. *Genes (Basel)*, 11.

FEHR, A. R. & PERLMAN, S. 2015. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol*, 1282, 1-23.

FEIKIN, D. R., HIGDON, M. M., ABU-RADDAD, L. J., ANDREWS, N., ARAOS, R., GOLDBERG, Y., GROOME, M. J., HUPPERT, A., O'BRIEN, K. L., SMITH, P. G., WILDER-SMITH, A., ZEGER, S., DELORIA KNOLL, M. & PATEL, M. K. 2022. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *The Lancet*, 399, 924-944.

FEIKIN, D. R., OLACK, B., BIGOGO, G. M., AUDI, A., COSMAS, L., AURA, B., BURKE, H., NJENGA, M. K., WILLIAMSON, J. & BREIMAN, R. F. 2011. The burden of common infectious disease syndromes at the clinic and household level from population-based surveillance in rural and urban Kenya. *PLoS One*, 6, e16085.

FERGUSON, S., MCLAY, T., ANDREW, R. L., BRUHL, J. J., SCHWESSINGER, B., BOREVITZ, J. & JONES, A. 2022. Species-specific basecallers improve actual accuracy of nanopore sequencing in plants. *Plant Methods*, 18, 137.

FINK, G., D'ACREMONT, V., LESLIE, H. H. & COHEN, J. 2020. Antibiotic exposure among children younger than 5 years in low-income and middle-income countries: a cross-sectional study of nationally representative facility-based and household-based surveys. *The Lancet Infectious Diseases*, 20, 179-187.

FLORES-GARCIA, Y., NASIR, G., HOPP, C. S., MUNOZ, C., BALABAN, A. E., ZAVALA, F. & SINNIS, P. 2018. Antibody-Mediated Protection against *Plasmodium* Sporozoites Begins at the Dermal Inoculation Site. *mBio*, 9, e02194-18.

FOSTER-NYARKO, E., COTTINGHAM, H., WICK, R. R., JUDD, L. M., LAM, M. M. C., WYRES, K. L., STANTON, T. D., TSANG, K. K., DAVID, S., AANENSEN, D. M., BRISSE, S. & HOLT, K. E. 2023. Nanopore-only assemblies for genomic surveillance of the global priority drug-resistant pathogen, *Klebsiella pneumoniae*. *Microbial Genomics*, 9.

FOURNIER, P.-E., DRANCOURT, M. & RAOULT, D. 2007. Bacterial genome sequencing and its use in infectious diseases. *The Lancet Infectious Diseases*, 7, 711-723.

FOX, T., GEPPERT, J., DINNES, J., SCANDRETT, K., BIGIO, J., SULIS, G., HETTIARACHCHI, D., MATHANGASINGHE, Y., WEERATUNGA, P., WICKRAMASINGHE, D., BERGMAN, H., BUCKLEY, B. S., PROBYN, K., SGUASSERO, Y., DAVENPORT, C., CUNNINGHAM, J., DITTRICH, S., EMPERADOR, D., HOOFT, L., LEEFLANG, M. M., MCINNES, M. D., SPIJKER, R., STRUYF, T., VAN DEN BRUEL, A., VERBAKEL, J. Y., TAKWOINGI, Y., TAYLOR-PHILLIPS, S., DEEKS, J. J. & COCHRANE, C.-D. T. A. G. 2022. Antibody tests

for identification of current and past infection with SARS-CoV-2. *The Cochrane database of systematic reviews*, 11, CD013652-CD013652.

FRIED, M. & DUFFY, P. E. 2017. Malaria during Pregnancy. *Cold Spring Harb Perspect Med*, 7.

GARDON, J., GARDON-WENDEL, N., DEMANGA, N., KAMGNO, J., CHIPPAUX, J.-P. & BOUSSINESQ, M. 1997. Serious reactions after mass treatment of onchocerciasis with ivermectin in an area endemic for Loa loa infection. *The Lancet*, 350, 18-22.

GENDROT, M., FAWAZ, R., DORMOI, J., MADAMET, M. & PRADINES, B. 2019. Genetic diversity and deletion of Plasmodium falciparum histidine-rich protein 2 and 3: a threat to diagnosis of P. falciparum malaria. *Clinical Microbiology and Infection*, 25, 580-585.

GILMAN, R. H., TERMINEL, M., LEVINE, M. M., HERNANDEZ-MENDOZA, P. & HORNICK, R. B. 1975. Relative efficacy of blood, urine, rectal swab, bone-marrow, and rose-spot cultures for recovery of Salmonella typhi in typhoid fever. *Lancet*, 1, 1211-3.

GOODWIN, S., MCPHERSON, J. D. & MCCOMBIE, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17, 333-351.

GRAÑA, C., GHOSN, L., EVRENOGLOU, T., JARDE, A., MINOZZI, S., BERGMAN, H., BUCKLEY, B. S., PROBYN, K., VILLANUEVA, G., HENSCHKE, N., BONNET, H., ASSI, R., MENON, S., MARTI, M., DEVANE, D., MALLON, P., LELIEVRE, J. D., ASKIE, L. M., KREDO, T., FERRAND, G., DAVIDSON, M., RIVEROS, C., TOVEY, D., MEERPOHL, J. J., GRASSELLI, G., RADA, G., HRÓBJARTSSON, A., RAVAUD, P., CHAIMANI, A. & BOUTRON, I. 2022. Efficacy and safety of COVID-19 vaccines. *Cochrane Database Syst Rev*, 12, Cd015477.

GREIG, J. D. & RAVEL, A. 2009. Analysis of foodborne outbreak data reported internationally for source attribution. *International Journal of Food Microbiology*, 130, 77-87.

GRIMONT, P. A. & WEILL, F.-X. 2007. Antigenic formulae of the Salmonella serovars. *WHO collaborating centre for reference and research on Salmonella*, 9, 1-166.

GU, W., DENG, X., LEE, M., SUCU, Y. D., AREVALO, S., STRYKE, D., FEDERMAN, S., GOPEZ, A., REYES, K., ZORN, K., SAMPLE, H., YU, G., ISHPUNIANI, G., BRIGGS, B., CHOW, E. D., BERGER, A., WILSON, M. R., WANG, C., HSU, E., MILLER, S., DERISI, J. L. & CHIU, C. Y. 2021. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat Med*, 27, 115-124.

GU, W., MILLER, S. & CHIU, C. Y. 2019. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu Rev Pathol*, 14, 319-338.

- GUERRA, C. A., CITRON, D. T., GARCÍA, G. A. & SMITH, D. L. 2019. Characterising malaria connectivity using malaria indicator survey data. *Malaria Journal*, 18, 440.
- GUIROU, E. A., SCHINDLER, T., HOSCH, S., DONFACK, O. T., YOBOUE, C. A., KRÄHENBÜHL, S., DEAL, A., COSI, G., GONDWE, L., MWANGOKA, G., MASUKI, H., SALIM, N., MPINA, M., SAID, J., ABDULLA, S., HOFFMAN, S. L., NLAVO, B. M., MAAS, C., FALLA, C. C., PHIRI, W. P., GARCIA, G. A., TANNER, M. & DAUBENBERGER, C. 2020. Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests. *Scientific Reports*, 10, 12305.
- GUPTA, S., SNOW, R. W., DONNELLY, C. A., MARSH, K. & NEWBOLD, C. 1999. Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nature Medicine*, 5, 340-343.
- GUTTERY, D. S., ZEESHAN, M., FERGUSON, D. J. P., HOLDER, A. A. & TEWARI, R. 2022. Division and Transmission: Malaria Parasite Development in the Mosquito. *Annual Review of Microbiology*, 76, 113-134.
- HADFIELD, J., MEGILL, C., BELL, S. M., HUDDLESTON, J., POTTER, B., CALLENDER, C., SAGULENKO, P., BEDFORD, T. & NEHER, R. A. 2018. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*, 34, 4121-4123.
- HAEUSLER, G. M. & CURTIS, N. 2013. Non-typhoidal Salmonella in Children: Microbiology, Epidemiology and Treatment. In: CURTIS, N., FINN, A. & POLLARD, A. J. (eds.) *Hot Topics in Infection and Immunity in Children IX*. New York, NY: Springer New York.
- HAN, A. X., TOPOROWSKI, A., SACKS, J. A., PERKINS, M. D., BRIAND, S., VAN KERKHOVE, M., HANNAY, E., CARMONA, S., RODRIGUEZ, B., PARKER, E., NICHOLS, B. E. & RUSSELL, C. A. 2023. SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs. *Nature Genetics*, 55, 26-33.
- HANBOONKUNUPAKARN, B., TARNING, J., PUKRITTAYAKAMEE, S. & CHOTIVANICH, K. 2022. Artemisinin resistance and malaria elimination: Where are we now? *Frontiers in pharmacology*, 13, 876282-876282.
- HARRIS, J. B. & BROOKS, W. A. 2020. 74 - Typhoid and Paratyphoid (Enteric) Fever. In: RYAN, E. T., HILL, D. R., SOLOMON, T., ARONSON, N. E. & ENDY, T. P. (eds.) *Hunter's Tropical Medicine and Emerging Infectious Diseases (Tenth Edition)*. London: Elsevier.
- HARRISON, A. G., LIN, T. & WANG, P. 2020. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends in Immunology*, 41, 1100-1115.

HASAN, M. R., RAWAT, A., TANG, P., JITHESH, P. V., THOMAS, E., TAN, R. & TILLEY, P. 2016. Depletion of Human DNA in Spiked Clinical Specimens for Improvement of Sensitivity of Pathogen Detection by Next-Generation Sequencing. *J Clin Microbiol*, 54, 919-27.

HAVASI, A., VISAN, S., CAINAP, C., CAINAP, S. S., MIHAILA, A. A. & POP, L. A. 2022. Influenza A, Influenza B, and SARS-CoV-2 Similarities and Differences - A Focus on Diagnosis. *Front Microbiol*, 13, 908525.

HAVELAAR, A. H., KIRK, M. D., TORGERSON, P. R., GIBB, H. J., HALD, T., LAKE, R. J., PRAET, N., BELLINGER, D. C., DE SILVA, N. R., GARGOURI, N., SPEYBROECK, N., CAWTHORNE, A., MATHERS, C., STEIN, C., ANGULO, F. J. & DEVLEESSCHAUWER, B. 2015. World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010. *PLoS Med*, 12, e1001923.

HE, S., WURTZEL, O., SINGH, K., FROULA, J. L., YILMAZ, S., TRINGE, S. G., WANG, Z., CHEN, F., LINDQUIST, E. A., SOREK, R. & HUGENHOLTZ, P. 2010. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Methods*, 7, 807-12.

HILT, E. E. & FERRIERI, P. 2022. Next Generation and Other Sequencing Technologies in Diagnostic Microbiology and Infectious Diseases. *Genes (Basel)*, 13.

HOFMANN, N., MWINGIRA, F., SHEKALAGHE, S., ROBINSON, L. J., MUELLER, I. & FELGER, I. 2015. Ultra-Sensitive Detection of Plasmodium falciparum by Amplification of Multi-Copy Subtelomeric Targets. *PLOS Medicine*, 12, e1001788.

HOLMES, A. H., MOORE, L. S. P., SUNDSFJORD, A., STEINBAKK, M., REGMI, S., KARKEY, A., GUERIN, P. J. & PIDDOCK, L. J. V. 2016a. Understanding the mechanisms and drivers of antimicrobial resistance. *The Lancet*, 387, 176-187.

HOLMES, E. C., DUDAS, G., RAMBAUT, A. & ANDERSEN, K. G. 2016b. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. *Nature*, 538, 193-200.

HOLZSCHUH, A. & KOEPFLI, C. 2022. Tenfold difference in DNA recovery rate: systematic comparison of whole blood vs. dried blood spot sample collection for malaria molecular surveillance. *Malaria Journal*, 21, 88.

HOSCH, S., MPINA, M., NYAKURUNGU, E., BORICO, N. S., OBAMA, T. M. A., OVONA, M. C., WAGNER, P., RUBIN, S. E., VICKOS, U., MILANG, D. V. N., AYEKABA, M. O. O., PHIRI, W. P., DAUBENBERGER, C. A. & SCHINDLER, T. 2022a. Genomic Surveillance Enables the Identification of Co-infections With Multiple SARS-CoV-2 Lineages in Equatorial Guinea. *Frontiers in Public Health*, 9.

HOSCH, S., YOBOUE, C. A., DONFACK, O. T., GUIROU, E. A., DANGY, J.-P., MPINA, M., NYAKURUNGU, E., BLÖCHLIGER, K., GUERRA, C. A., PHIRI, W. P., AYEKABA, M. O. O., GARCÍA, G. A., TANNER, M.,

DAUBENBERGER, C. & SCHINDLER, T. 2022b. Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment. *Malaria Journal*, 21, 23.

INABA, S., NAKAO, Y., IKEDA, S., MIZUMOTO, Y., UTSUNOMIYA, T., HONJO, M., TAKADA, Y., NOGAMI, N., ISHII, E. & YAMAGUCHI, O. 2023. Simple Symptom-Based Prediction of COVID-19: A Single-Center Study of Outpatient Fever Clinic in Japan. *Cureus*, 15, e36614.

JAIN, M., KOREN, S., MIGA, K. H., QUICK, J., RAND, A. C., SASANI, T. A., TYSON, J. R., BEGGS, A. D., DILTHEY, A. T., FIDDES, I. T., MALLA, S., MARRIOTT, H., NIETO, T., O'GRADY, J., OLSEN, H. E., PEDERSEN, B. S., RHIE, A., RICHARDSON, H., QUINLAN, A. R., SNUTCH, T. P., TEE, L., PATEN, B., PHILLIPPY, A. M., SIMPSON, J. T., LOMAN, N. J. & LOOSE, M. 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology*, 36, 338-345.

JANDA, J. M. & ABBOTT, S. L. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol*, 45, 2761-4.

JOSEPH, L. A., GRISWOLD, T., VIDYAPRAKASH, E., IM, S. B., WILLIAMS, G. M., POUSEELE, H. A., HISE, K. B. & CARLETON, H. A. 2023. Evaluation of core genome and whole genome multilocus sequence typing schemes for *Campylobacter jejuni* and *Campylobacter coli* outbreak detection in the USA. *Microbial Genomics*, 9.

JOSLING, G. A. & LLINÁS, M. 2015. Sexual development in Plasmodium parasites: knowing when it's time to commit. *Nature Reviews Microbiology*, 13, 573-587.

KAAKOUSH, N. O., CASTAÑO-RODRÍGUEZ, N., MITCHELL, H. M. & MAN, S. M. 2015. Global Epidemiology of *Campylobacter* Infection. *Clin Microbiol Rev*, 28, 687-720.

KABORÉ, B., POST, A., LOMPO, P., BOGNINI, J. D., DIALLO, S., KAM, B. T. D., RAHAMAT-LANGENDOEN, J., WERTHEIM, H. F. L., VAN OPZEELAND, F., LANGEREIS, J. D., DE JONGE, M. I., TINTO, H., JACOBS, J., VAN DER VEN, A. J. & DE MAST, Q. 2021. Aetiology of acute febrile illness in children in a high malaria transmission area in West Africa. *Clin Microbiol Infect*, 27, 590-596.

KAMAU, E., TOLBERT, L. S., KORTEPETER, L., PRATT, M., NYAKOE, N., MURINGO, L., OGUTU, B., WAITUMBI, J. N. & OCKENHOUSE, C. F. 2011. Development of a highly sensitive genus-specific quantitative reverse transcriptase real-time PCR assay for detection and quantitation of plasmodium by amplifying RNA and DNA of the 18S rRNA genes. *J Clin Microbiol*, 49, 2946-53.

KAMGNO, J., PION, S. D., CHESNAIS, C. B., BAKALAR, M. H., D'AMBROSIO, M. V., MACKENZIE, C. D., NANA-DJEUNGA, H. C., GOUNOUE-KAMKUMO, R., NJITCHOUANG, G.-R., NWANE, P., TCHATCHUENG-

MBOUGA, J. B., WANJI, S., STOLK, W. A., FLETCHER, D. A., KLION, A. D., NUTMAN, T. B. & BOUSSINESQ, M. 2017. A Test-and-Not-Treat Strategy for Onchocerciasis in Loa loa–Endemic Areas. *New England Journal of Medicine*, 377, 2044-2052.

KAYIBA, N. K., YOBI, D. M., TCHAKOUNANG, V. R. K., MVUMBI, D. M., KABUTUTU, P. Z., DEVLEESSCHAUWER, B., MUKOMENA, E. S., DEMOL, P., HAYETTE, M.-P., MVUMBI, G. L., ROSAS-AGUIRRE, A., LUSAMBA, P. D. & SPEYBROECK, N. 2021. Evaluation of the usefulness of intermittent preventive treatment of malaria in pregnancy with sulfadoxine-pyrimethamine in a context with increased resistance of Plasmodium falciparum in Kingasani Hospital, Kinshasa in the Democratic Republic of Congo. *Infection, Genetics and Evolution*, 94, 105009.

KELLER, M. W., RAMBO-MARTIN, B. L., WILSON, M. M., RIDENOUR, C. A., SHEPARD, S. S., STARK, T. J., NEUHAUS, E. B., DUGAN, V. G., WENTWORTH, D. E. & BARNES, J. R. 2018. Direct RNA Sequencing of the Coding Complete Influenza A Virus Genome. *Sci Rep*, 8, 14408.

KHONGWICHIT, S., CHUCHAONA, W., VONGPUNSAWAD, S. & POOVORAWAN, Y. 2022. Molecular surveillance of arboviruses circulation and co-infection during a large chikungunya virus outbreak in Thailand, October 2018 to February 2020. *Sci Rep*, 12, 22323.

KINGANDA-LUSAMAKI, E., BLACK, A., MUKADI, D. B., HADFIELD, J., MBALA-KINGEBENI, P., PRATT, C. B., AZIZA, A., DIAGNE, M. M., WHITE, B., BISENTO, N., NSUNDA, B., AKONGA, M., FAYE, M., FAYE, O., EDIDI-ATANI, F., MATONDO-KUAMFUMU, M., MAMBU-MBIKA, F., BULABULA, J., DI PAOLA, N., PAUTHNER, M. G., ANDERSEN, K. G., PALACIOS, G., DELAPORTE, E., SALL, A. A., PEETERS, M., WILEY, M. R., AHUKA-MUNDEKE, S., BEDFORD, T. & TAMFUM, J. M. 2021. Integration of genomic sequencing into the response to the Ebola virus outbreak in Nord Kivu, Democratic Republic of the Congo. *Nat Med*, 27, 710-716.

KINGRY, L., SHELDON, S., OATMAN, S., PRITT, B., ANACKER, M., BJORK, J., NEITZEL, D., STRAIN, A., BERRY, J., SLOAN, L., RESPICIO-KINGRY, L., DIETRICH, E., BLOCH, K., MONCAYO, A., SRINIVASAMOORTHY, G., HU, B., HINCKLEY, A., MEAD, P., KUGELER, K. & PETERSEN, J. 2020. Targeted Metagenomics for Clinical Detection and Discovery of Bacterial Tick-Borne Pathogens. *J Clin Microbiol*, 58.

KNODLER, L. A. & ELFENBEIN, J. R. 2019. Salmonella enterica. *Trends in Microbiology*, 27, 964-965.

KNOPP, S., STEINMANN, P., HATZ, C., KEISER, J. & UTZINGER, J. 2012. Nematode Infections:: Filariases. *Infectious Disease Clinics of North America*, 26, 359-381.

KNYAZEYEV, S., CHHUGANI, K., SARWAL, V., AYYALA, R., SINGH, H., KARTHIKEYAN, S., DESHPANDE, D., BAYKAL, P. I., COMAROVA, Z., LU, A., POROZOV, Y., VASYLYEVA, T. I., WERTHEIM, J. O., TIERNEY, B. T., CHIU, C. Y., SUN, R., WU, A., ABEDALTHAGAFI, M. S., PAK, V. M., NAGARAJ, S. H., SMITH, A. L., SKUMS, P., PASANIUC, B., KOMISSAROV, A., MASON, C. E., BORTZ, E., LEMEY, P., KONDRASHOV, F., BEERENWINKEL, N., LAM, T. T.-Y., WU, N. C., ZELIKOVSKY, A., KNIGHT, R., CRANDALL, K. A. & MANGUL, S. 2022. Unlocking capacities of genomics for the COVID-19 response and future pandemics. *Nature methods*, 19, 374-380.

KOEPFLI, C. & MUELLER, I. 2017. Malaria Epidemiology at the Clone Level. *Trends Parasitol*, 33, 974-985.

KOLMOGOROV, M., YUAN, J., LIN, Y. & PEVZNER, P. A. 2019. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, 37, 540-546.

KORBER, B., FISCHER, W. M., GNANAKARAN, S., YOON, H., THEILER, J., ABFALTERER, W., HENGARTNER, N., GIORGI, E. E., BHATTACHARYA, T., FOLEY, B., HASTIE, K. M., PARKER, M. D., PARTRIDGE, D. G., EVANS, C. M., FREEMAN, T. M., DE SILVA, T. I., MCDANAL, C., PEREZ, L. G., TANG, H., MOON-WALKER, A., WHELAN, S. P., LABRANCHE, C. C., SAPHIRE, E. O. & MONTEFIORI, D. C. 2020. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*, 182, 812-827.e19.

KOREN, S., WALENZ, B. P., BERLIN, K., MILLER, J. R., BERGMAN, N. H. & PHILLIPPY, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*, 27, 722-736.

KÖSER, C. U., ELLINGTON, M. J., CARTWRIGHT, E. J., GILLESPIE, S. H., BROWN, N. M., FARRINGTON, M., HOLDEN, M. T., DOUGAN, G., BENTLEY, S. D., PARKHILL, J. & PEACOCK, S. J. 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog*, 8, e1002824.

KRISHNAKUMAR, R., SINHA, A., BIRD, S. W., JAYAMOCHAN, H., EDWARDS, H. S., SCHOENIGER, J. S., PATEL, K. D., BRANDA, S. S. & BARTSCH, M. S. 2018. Systematic and stochastic influences on the performance of the MinION nanopore sequencer across a range of nucleotide bias. *Sci Rep*, 8, 3159.

KUNASOL, C., DONDORP, A. M., BATTY, E. M., NAKHONSRI, V., SINJANAKHOM, P., DAY, N. P. J. & IMWONG, M. 2022. Comparative analysis of targeted next-generation sequencing for Plasmodium falciparum drug resistance markers. *Sci Rep*, 12, 5563.

LAM, T. T.-Y., JIA, N., ZHANG, Y.-W., SHUM, M. H.-H., JIANG, J.-F., ZHU, H.-C., TONG, Y.-G., SHI, Y.-X., NI, X.-B., LIAO, Y.-S., LI, W.-J., JIANG, B.-G., WEI, W., YUAN, T.-T., ZHENG, K., CUI, X.-M., LI, J., PEI, G.-Q., QIANG, X., CHEUNG, W. Y.-M., LI, L.-F., SUN, F.-F., QIN, S., HUANG, J.-C., LEUNG, G. M., HOLMES, E. C., HU, Y.-L., GUAN, Y. & CAO, W.-C. 2020. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, 583, 282-285.

LARSEN, M. V., COSENTINO, S., RASMUSSEN, S., FRIIS, C., HASMAN, H., MARVIG, R. L., JELSBK, L., SICHERITZ-PONTÉN, T., USSERY, D. W., AARESTRUP, F. M. & LUND, O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol*, 50, 1355-61.

LATINNE, A., HU, B., OLIVAL, K. J., ZHU, G., ZHANG, L., LI, H., CHMURA, A. A., FIELD, H. E., ZAMBRANA-TORRELIO, C., EPSTEIN, J. H., LI, B., ZHANG, W., WANG, L. F., SHI, Z. L. & DASZAK, P. 2020. Origin and cross-species transmission of bat coronaviruses in China. *Nat Commun*, 11, 4235.

LAU, H., KHOSRAWIPOUR, V., KOCBACH, P., MIKOLAJCZYK, A., SCHUBERT, J., BANIA, J. & KHOSRAWIPOUR, T. 2020. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *J Travel Med*, 27.

LAWN, J. E., BLENCOWE, H., WAISWA, P., AMOUZOU, A., MATHERS, C., HOGAN, D., FLENADY, V., FRØEN, J. F., QURESHI, Z. U., CALDERWOOD, C., SHIEKH, S., JASSIR, F. B., YOU, D., MCCLURE, E. M., MATHAI, M. & COUSENS, S. 2016. Stillbirths: rates, risk factors, and acceleration towards 2030. *Lancet*, 387, 587-603.

LEE, I., RAZAGHI, R., GILPATRICK, T., MOLNAR, M., GERSHMAN, A., SADOWSKI, N., SEDLAZECK, F. J., HANSEN, K. D., SIMPSON, J. T. & TIMP, W. 2020. Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing. *Nat Methods*, 17, 1191-1199.

LEVENE, M. J., KORLACH, J., TURNER, S. W., FOQUET, M., CRAIGHEAD, H. G. & WEBB, W. W. 2003. Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations. *Science*, 299, 682-686.

LI, B., SUN, Z., LI, X., LI, X., WANG, H., CHEN, W., CHEN, P., QIAO, M. & MAO, Y. 2017. Performance of pfHRP2 versus pLDH antigen rapid diagnostic tests for the detection of *Plasmodium falciparum*: a systematic review and meta-analysis. *Archives of Medical Science*, 13, 541-549.

LI, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094-3100.

LI, Y., HE, X.-Z., LI, M.-H., LI, B., YANG, M.-J., XIE, Y., ZHANG, Y. & MA, X.-J. 2020. Comparison of third-generation sequencing approaches to identify viral pathogens under public health emergency conditions. *Virus Genes*, 56, 288-297.

LOMAN, N. J., RAMBAUT, A., QUICK, J., ROWE, W., SIMPSON, J. A., BEAULAUER, J., FLOREK, K. & WILSON, S. J. 2020. *ARTIC nanopore protocol for nCoV2019 novel coronavirus* [Online]. Github: artic-network. Available: <https://github.com/artic-network/artic-ncov2019.git> [Accessed 04.05.2023].

MACLEAN, O. A., LYTRAS, S., WEAVER, S., SINGER, J. B., BONI, M. F., LEMEY, P., KOSAKOVSKY POND, S. L. & ROBERTSON, D. L. 2021. Natural selection in the evolution of SARS-CoV-2 in bats created a generalist virus and highly capable human pathogen. *PLoS Biology*, 19, e3001115.

MAN, S. M. 2011. The clinical importance of emerging *Campylobacter* species. *Nature Reviews Gastroenterology & Hepatology*, 8, 669-685.

MARGULIES, M., EGHOLM, M., ALTMAN, W. E., ATTIYA, S., BADER, J. S., BEMBEN, L. A., BERKA, J., BRAVERMAN, M. S., CHEN, Y.-J., CHEN, Z., DEWELL, S. B., DU, L., FIERRO, J. M., GOMES, X. V., GODWIN, B. C., HE, W., HELGESEN, S., HO, C. H., IRZYK, G. P., JANDO, S. C., ALLENQUER, M. L. I., JARVIE, T. P., JIRAGE, K. B., KIM, J.-B., KNIGHT, J. R., LANZA, J. R., LEAMON, J. H., LEFKOWITZ, S. M., LEI, M., LI, J., LOHMAN, K. L., LU, H., MAKHIJANI, V. B., MCDADE, K. E., MCKENNA, M. P., MYERS, E. W., NICKERSON, E., NOBILE, J. R., PLANT, R., PUC, B. P., RONAN, M. T., ROTH, G. T., SARKIS, G. J., SIMONS, J. F., SIMPSON, J. W., SRINIVASAN, M., TARTARO, K. R., TOMASZ, A., VOGT, K. A., VOLKMER, G. A., WANG, S. H., WANG, Y., WEINER, M. P., YU, P., BEGLEY, R. F. & ROTHBERG, J. M. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437, 376-380.

MARSH, K. & KINYANJUI, S. 2006. Immune effector mechanisms in malaria. *Parasite Immunology*, 28, 51-60.

MARTIÁÑEZ-VENDRELL, X., SKJEFTE, M., SIKKA, R. & GUPTA, H. 2022. Factors Affecting the Performance of HRP2-Based Malaria Rapid Diagnostic Tests. *Tropical Medicine and Infectious Disease*, 7, 265.

MARTIN, D. P., LYTRAS, S., LUCACI, A. G., MAIER, W., GRÜNING, B., SHANK, S. D., WEAVER, S., MACLEAN, O. A., ORTON, R. J., LEMEY, P., BONI, M. F., TEGALLY, H., HARKINS, G. W., SCHEEPERS, C., BHIMAN, J. N., EVERATT, J., AMOAKO, D. G., SAN, J. E., GIANDHARI, J., SIGAL, A., NGS, S. A., WILLIAMSON, C., HSIAO, N.-Y., VON GOTTEBERG, A., DE KLERK, A., SHAFER, R. W., ROBERTSON, D. L., WILKINSON, R. J., SEWELL, B. T., LESSELLS, R., NEKRUTENKO, A., GREANEY, A. J., STARR, T. N., BLOOM, J. D., MURRELL, B., WILKINSON, E., GUPTA, R. K., DE OLIVEIRA, T. & KOSAKOVSKY POND, S. L. 2022. Selection Analysis Identifies Clusters of Unusual Mutational Changes in Omicron Lineage BA.1 That Likely Impact Spike Function. *Molecular biology and evolution*, 39, msac061.

MARTIN, D. P., WEAVER, S., TEGALLY, H., SAN, J. E., SHANK, S. D., WILKINSON, E., LUCACI, A. G., GIANDHARI, J., NAIDOO, S., PILLAY, Y., SINGH, L., LESSELLS, R. J., GUPTA, R. K., WERTHEIM, J. O.,

NEKTURENKO, A., MURRELL, B., HARKINS, G. W., LEMEY, P., MACLEAN, O. A., ROBERTSON, D. L., DE OLIVEIRA, T. & KOSAKOVSKY POND, S. L. 2021. The emergence and ongoing convergent evolution of the SARS-CoV-2 N501Y lineages. *Cell*, 184, 5189-5200.e7.

MATHIEU, E., RITCHIE, H., ORTIZ-OSPINA, E., ROSER, M., HASELL, J., APPEL, C., GIATTINO, C. & RODÉS-GUIRAO, L. 2021. A global database of COVID-19 vaccinations. *Nature Human Behaviour*, 5, 947-953.

MAURIN, M., FENOLLAR, F., MEDIANNIKOV, O., DAVOUST, B., DEVAUX, C. & RAOULT, D. 2021. Current Status of Putative Animal Sources of SARS-CoV-2 Infection in Humans: Wildlife, Domestic Animals and Pets. *Microorganisms*, 9, 868.

MBALA-KINGEBENI, P., VILLABONA-ARENAS, C. J., VIDAL, N., LIKOFATA, J., NSIO-MBETA, J., MAKIALA-MANDANDA, S., MUKADI, D., MUKADI, P., KUMAKAMBA, C., DJOKOLO, B., AYOUBA, A., DELAPORTE, E., PEETERS, M., MUYEMBE TAMFUM, J. J. & AHUKA-MUNDEKE, S. 2019. Rapid Confirmation of the Zaire Ebola Virus in the Outbreak of the Equateur Province in the Democratic Republic of Congo: Implications for Public Health Interventions. *Clin Infect Dis*, 68, 330-333.

MEHTA, N., PERRAIS, B., MARTIN, K., KUMAR, A., HOBMAN, T. C., CABALFIN-CHUA, M. N., DONALDO, M. E., SIOSE PAINAGA, M. S., GAITE, J. Y., TRAN, V., KAIN, K. C., HAWKES, M. T. & YANOW, S. K. 2019. A Direct from Blood/Plasma Reverse Transcription-Polymerase Chain Reaction for Dengue Virus Detection in Point-of-Care Settings. *Am J Trop Med Hyg*, 100, 1534-1540.

METZGER, W. G. & MORDMÜLLER, B. 2014. Loa loa—does it deserve to be neglected? *The Lancet Infectious Diseases*, 14, 353-357.

METZKER, M. L. 2010. Sequencing technologies — the next generation. *Nature Reviews Genetics*, 11, 31-46.

MITCHELL, S. L. & SIMNER, P. J. 2019. Next-Generation Sequencing in Clinical Microbiology: Are We There Yet? *Clinics in Laboratory Medicine*, 39, 405-418.

MLCOCHOVA, P., KEMP, S. A., DHAR, M. S., PAPA, G., MENG, B., FERREIRA, I. A. T. M., DATIR, R., COLLIER, D. A., ALBECKA, A., SINGH, S., PANDEY, R., BROWN, J., ZHOU, J., GOONAWARDANE, N., MISHRA, S., WHITTAKER, C., MELLAN, T., MARWAL, R., DATTA, M., SENGUPTA, S., PONNUSAMY, K., RADHAKRISHNAN, V. S., ABDULLAHI, A., CHARLES, O., CHATTOPADHYAY, P., DEVI, P., CAPUTO, D., PEACOCK, T., WATTAL, C., GOEL, N., SATWIK, A., VAISHYA, R., AGARWAL, M., CHAUHAN, H., DIKID, T., GOGIA, H., LALL, H., VERMA, K., DHAR, M. S., SINGH, M. K., SONI, N., MEENA, N., MADAN, P., SINGH, P., SHARMA, R., SHARMA, R., KABRA, S., KUMAR, S., KUMARI, S., SHARMA, U., CHAUDHARY, U., SIVASUBBU, S., SCARIA, V., OBEROI, J. K., RAVEENDRAN, R., DATTA, S., DAS, S., MAITRA, A.,

CHINNASWAMY, S., BISWAS, N. K., PARIDA, A., RAGHAV, S. K., PRASAD, P., SARIN, A., MAYOR, S., RAMAKRISHNAN, U., PALAKODETI, D., SESHASAYEE, A. S. N., THANGARAJ, K., BASHYAM, M. D., DALAL, A., BHAT, M., SHOUCHE, Y., PILLAI, A., ABRAHAM, P., POTDAR, V. A., CHERIAN, S. S., DESAI, A. S., PATTABIRAMAN, C., MANJUNATHA, M. V., MANI, R. S., UDUPI, G. A., NANDICOORI, V., TALLAPAKA, K. B., SOWPATI, D. T., KAWABATA, R., MORIZAKO, N., SADAMASU, K., ASAKURA, H., NAGASHIMA, M., YOSHIMURA, K., ITO, J., KIMURA, I., URIU, K., KOSUGI, Y., SUGANAMI, M., OIDE, A., YOKOYAMA, M., CHIBA, M., SAITO, A., et al. 2021. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature*, 599, 114-119.

MOODY, A. 2002. Rapid diagnostic tests for malaria parasites. *Clin Microbiol Rev*, 15, 66-78.

MOODY, A. H. & CHIODINI, P. L. 2000. Methods for the detection of blood parasites. *Clinical & Laboratory Haematology*, 22, 189-201.

MORSE, S. S., MAZET, J. A., WOOLHOUSE, M., PARRISH, C. R., CARROLL, D., KARESH, W. B., ZAMBRANA-TORRELIO, C., LIPKIN, W. I. & DASZAK, P. 2012. Prediction and prevention of the next pandemic zoonosis. *Lancet*, 380, 1956-65.

MOUATCHO, J. C. & GOLDRING, J. P. D. 2013. Malaria rapid diagnostic tests: challenges and prospects. *Journal of Medical Microbiology*, 62, 1491-1505.

MOUREMBOU, G., FENOLLAR, F., LEKANA-DOUKI, J. B., NDJOYI MBIGUINO, A., MAGHENDJI NZONDO, S., MATSIEGUI, P. B., ZOLEKO MANEGO, R., EHOUNOUD, C. H., BITTAR, F., RAOULT, D. & MEDIANNIKOV, O. 2015. Mansonella, including a Potential New Species, as Common Parasites in Children in Gabon. *PLoS Negl Trop Dis*, 9, e0004155.

MURMU, L. K., SAHU, A. A. & BARIK, T. K. 2021. Diagnosing the drug resistance signature in Plasmodium falciparum: a review from contemporary methods to novel approaches. *J Parasit Dis*, 45, 869-876.

NAG, S., DALGAARD, M. D., KOFOED, P.-E., URSING, J., CRESPO, M., ANDERSEN, L. O. B., AARESTRUP, F. M., LUND, O. & ALIFRANGIS, M. 2017. High throughput resistance profiling of Plasmodium falciparum infections based on custom dual indexing and Illumina next generation sequencing-technology. *Scientific Reports*, 7, 2398.

NGUYEN, M., LONG, S. W., MCDERMOTT, P. F., OLSEN, R. J., OLSON, R., STEVENS, R. L., TYSON, G. H., ZHAO, S. & DAVIS, J. J. 2019. Using Machine Learning To Predict Antimicrobial MICs and Associated Genomic Features for Nontyphoidal *Salmonella*. *Journal of Clinical Microbiology*, 57, e01260-18.

NIYUKURI, D., SINZINKAYO, D., TROTH, E. V., ODUMA, C. O., BARENGAYABO, M., NDEREYIMANA, M., HOLZSCHUH, A., VERA-ARIAS, C. A., GEBRE, Y., BADU, K., NYANDWI, J., BAZA, D., JUMA, E. & KOEPFLI, C. 2022. Performance of highly sensitive and conventional rapid diagnostic tests for clinical and subclinical Plasmodium falciparum infections, and hrp2/3 deletion status in Burundi. *PLOS Global Public Health*, 2, e0000828.

OBERMEYER, F., JANKOWIAK, M., BARKAS, N., SCHAFFNER, S. F., PYLE, J. D., YURKOVETSKIY, L., BOSSO, M., PARK, D. J., BABADI, M., MACINNIS, B. L., LUBAN, J., SABETI, P. C. & LEMIEUX, J. E. 2022. Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with fitness. *Science*, 376, 1327-1332.

OHLSEN, E. C., HAWKSWORTH, A. W., WONG, K., GUAGLIARDO, S. A. J., FULLER, J. A., SLOAN, M. L. & O'LAUGHLIN, K. 2022. Determining Gaps in Publicly Shared SARS-CoV-2 Genomic Surveillance Data by Analysis of Global Submissions. *Emerg Infect Dis*, 28, S85-s92.

OKELL, L. C., GHANI, A. C., LYONS, E. & DRAKELEY, C. J. 2009. Submicroscopic Infection in Plasmodium falciparum-Endemic Populations: A Systematic Review and Meta-Analysis. *The Journal of Infectious Diseases*, 200, 1509-1517.

ORAN, D. P. & TOPOL, E. J. 2021. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic : A Systematic Review. *Ann Intern Med*, 174, 655-662.

PADGETT, J. J. & JACOBSEN, K. H. 2008. Loiasis: African eye worm. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 102, 983-989.

PAGÈS-GALLEGO, M. & DE RIDDER, J. 2023. Comprehensive benchmark and architectural analysis of deep learning models for nanopore sequencing basecalling. *Genome Biol*, 24, 71.

PALUMBO, E. 2008. Filariasis: diagnosis, treatment and prevention. *Acta bio-medica : Atenei Parmensis*, 79, 106-109.

PARKHILL, J., WREN, B. W., MUNGALL, K., KETLEY, J. M., CHURCHER, C., BASHAM, D., CHILLINGWORTH, T., DAVIES, R. M., FELTWELL, T., HOLROYD, S., JAGELS, K., KARLYSHEV, A. V., MOULE, S., PALLEN, M. J., PENN, C. W., QUAIL, M. A., RAJANDREAM, M. A., RUTHERFORD, K. M., VAN VLIET, A. H., WHITEHEAD, S. & BARRELL, B. G. 2000. The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature*, 403, 665-8.

PARRY, C. M., WIJEDORU, L., ARJYAL, A. & BAKER, S. 2011. The utility of diagnostic tests for enteric fever in endemic locations. *Expert Review of Anti-infective Therapy*, 9, 711-725.

- PAYNE, D. 1987. Spread of chloroquine resistance in *Plasmodium falciparum*. *Parasitology Today*, 3, 241-246.
- PEGUES, D. A. & MILLER, S. I. 2015. 225 - *Salmonella* Species. In: BENNETT, J. E., DOLIN, R. & BLASER, M. J. (eds.) *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition)*. Philadelphia: W.B. Saunders.
- PEKAR, J., WOROBEY, M., MOSHIRI, N., SCHEFFLER, K. & WERTHEIM, J. O. 2021. Timing the SARS-CoV-2 index case in Hubei province. *Science*, 372, 412-417.
- PETERS, W. Chemotherapy and drug resistance in malaria. 2nd edition. Vols 1 and 2 continued. 1987.
- PRUDÊNCIO, M., RODRIGUEZ, A. & MOTA, M. M. 2006. The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nature Reviews Microbiology*, 4, 849-856.
- PRYCE, J., MEDLEY, N. & CHOI, L. 2022. Indoor residual spraying for preventing malaria in communities using insecticide-treated nets. *Cochrane Database Syst Rev*, 1, Cd012688.
- PRYCE, J., RICHARDSON, M. & LENGELER, C. 2018. Insecticide-treated nets for preventing malaria. *Cochrane Database Syst Rev*, 11, Cd000363.
- QAMAR, F. N., HUSSAIN, W. & QURESHI, S. 2022. Salmonellosis Including Enteric Fever. *Pediatric Clinics of North America*, 69, 65-77.
- QINGHAOSU, A. C. R. G. 1979. Antimalaria studies on Qinghaosu. *Chin Med J (Engl)*, 92, 811-6.
- QUAN, T. P., BAWA, Z., FOSTER, D., WALKER, T., ELIAS, C. D. O., RATHOD, P., IQBAL, Z., BRADLEY, P., MOWBRAY, J., WALKER, A. S., CROOK, D. W., WYLLIE, D. H., PETO, T. E. A. & SMITH, E. G. 2018. Evaluation of Whole-Genome Sequencing for Mycobacterial Species Identification and Drug Susceptibility Testing in a Clinical Setting: a Large-Scale Prospective Assessment of Performance against Line Probe Assays and Phenotyping. *Journal of Clinical Microbiology*, 56, e01480-17.
- RAMBAUT, A., HOLMES, E. C., O'TOOLE, Á., HILL, V., MCCRONE, J. T., RUIS, C., DU PLESSIS, L. & PYBUS, O. G. 2020a. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*, 5, 1403-1407.
- RAMBAUT, A., LOMAN, N., PYBUS, O., BARCLAY, W., BARRETT, J., CARABELLI, A., CONNOR, T., PEACOCK, T., ROBERTSON, D. L., VOLZ, E. & (COG-UK), C.-G. C. U. 2020b. *Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations* [Online]. Virological. Available: <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations>

[emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563](#) [Accessed 23.03.2023].

RANG, F. J., KLOOSTERMAN, W. P. & DE RIDDER, J. 2018. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biology*, 19, 90.

RANSON, H., N'GUESSAN, R., LINES, J., MOIROUX, N., NKUNI, Z. & CORBEL, V. 2011. Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control? *Trends in Parasitology*, 27, 91-98.

RAO, P. N., UPLEKAR, S., KAYAL, S., MALLICK, P. K., BANDYOPADHYAY, N., KALE, S., SINGH, O. P., MOHANTY, A., MOHANTY, S., WASSMER, S. C. & CARLTON, J. M. 2016. A Method for Amplicon Deep Sequencing of Drug Resistance Genes in Plasmodium falciparum Clinical Isolates from India. *Journal of Clinical Microbiology*, 54, 1500-1511.

REN, Z.-L., ZHANG, J.-R., ZHANG, X.-M., LIU, X., LIN, Y.-F., BAI, H., WANG, M.-C., CHENG, F., LIU, J.-D., LI, P., KONG, L., BO, X.-C., WANG, S.-Q., NI, M. & YAN, J.-W. 2021. Forensic nanopore sequencing of STRs and SNPs using Verogen's ForenSeq DNA Signature Prep Kit and MiniON. *International Journal of Legal Medicine*, 135, 1685-1693.

RHIE, A., MCCARTHY, S. A., FEDRIGO, O., DAMAS, J., FORMENTI, G., KOREN, S., ULIANO-SILVA, M., CHOW, W., FUNGTAMMASAN, A., KIM, J., LEE, C., KO, B. J., CHAISSON, M., GEDMAN, G. L., CANTIN, L. J., THIBAUD-NISSEN, F., HAGGERTY, L., BISTA, I., SMITH, M., HAASE, B., MOUNTCASTLE, J., WINKLER, S., PAEZ, S., HOWARD, J., VERNES, S. C., LAMA, T. M., GRUTZNER, F., WARREN, W. C., BALAKRISHNAN, C. N., BURT, D., GEORGE, J. M., BIEGLER, M. T., IORNS, D., DIGBY, A., EASON, D., ROBERTSON, B., EDWARDS, T., WILKINSON, M., TURNER, G., MEYER, A., KAUTT, A. F., FRANCHINI, P., DETRICH, H. W., 3RD, SVARDAL, H., WAGNER, M., NAYLOR, G. J. P., PIPPEL, M., MALINSKY, M., MOONEY, M., SIMBIRSKY, M., HANNIGAN, B. T., PESOUT, T., HOUCK, M., MISURACA, A., KINGAN, S. B., HALL, R., KRONENBERG, Z., SOVIĆ, I., DUNN, C., NING, Z., HASTIE, A., LEE, J., SELVARAJ, S., GREEN, R. E., PUTNAM, N. H., GUT, I., GHURYE, J., GARRISON, E., SIMS, Y., COLLINS, J., PELAN, S., TORRANCE, J., TRACEY, A., WOOD, J., DAGNEW, R. E., GUAN, D., LONDON, S. E., CLAYTON, D. F., MELLO, C. V., FRIEDRICH, S. R., LOVELL, P. V., OSIPOVA, E., AL-AJLI, F. O., SECOMANDI, S., KIM, H., THEOFANOPOULOU, C., HILLER, M., ZHOU, Y., HARRIS, R. S., MAKOVA, K. D., MEDVEDEV, P., HOFFMAN, J., MASTERSON, P., CLARK, K., MARTIN, F., HOWE, K., FLICEK, P., WALENZ, B. P., KWAK, W., CLAWSON, H., et al. 2021. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592, 737-746.

- RHOADS, A. & AU, K. F. 2015. PacBio Sequencing and Its Applications. *Genomics, Proteomics & Bioinformatics*, 13, 278-289.
- ROKNEY, A., VALINSKY, L., MORAN-GILAD, J., VRANCKX, K., AGMON, V. & WEINBERGER, M. 2018. Genomic Epidemiology of *Campylobacter jejuni* Transmission in Israel. *Front Microbiol*, 9, 2432.
- ROKNEY, A., VALINSKY, L., VRANCKX, K., FELDMAN, N., AGMON, V., MORAN-GILAD, J. & WEINBERGER, M. 2020. WGS-Based Prediction and Analysis of Antimicrobial Resistance in *Campylobacter jejuni* Isolates From Israel. *Front Cell Infect Microbiol*, 10, 365.
- ROTHBERG, J. M., HINZ, W., REARICK, T. M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J. H., JOHNSON, K., MILGREW, M. J., EDWARDS, M., HOON, J., SIMONS, J. F., MARRAN, D., MYERS, J. W., DAVIDSON, J. F., BRANTING, A., NOBILE, J. R., PUC, B. P., LIGHT, D., CLARK, T. A., HUBER, M., BRANCIFORTE, J. T., STONER, I. B., CAWLEY, S. E., LYONS, M., FU, Y., HOMER, N., SEDOVA, M., MIAO, X., REED, B., SABINA, J., FEIERSTEIN, E., SCHORN, M., ALANJARY, M., DIMALANTA, E., DRESSMAN, D., KASINSKAS, R., SOKOLSKY, T., FIDANZA, J. A., NAMSARAEV, E., MCKERNAN, K. J., WILLIAMS, A., ROTH, G. T. & BUSTILLO, J. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475, 348-352.
- SAH, P., FITZPATRICK, M. C., ZIMMER, C. F., ABDOLLAHI, E., JUDEN-KELLY, L., MOGHADAS, S. M., SINGER, B. H. & GALVANI, A. P. 2021. Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. *Proc Natl Acad Sci U S A*, 118.
- SALZBERGER, B., BUDER, F., LAMPL, B., EHRENSTEIN, B., HITZENBICHLER, F., HOLZMANN, T., SCHMIDT, B. & HANSES, F. 2021. Epidemiology of SARS-CoV-2. *Infection*, 49, 233-239.
- SCHADT, E. E., TURNER, S. & KASARSKIS, A. 2010. A window into third-generation sequencing. *Human Molecular Genetics*, 19, R227-R240.
- SCHINDLER, T., DEAL, A. C., FINK, M., GUIROU, E., MOSER, K. A., MWAKASUNGULA, S. M., MIHAYO, M. G., JONGO, S. A., CHAKI, P. P., ABDULLA, S., VALVERDE, P. C. M., TORRES, K., BIJERI, J. R., SILVA, J. C., HOFFMAN, S. L., GAMBOA, D., TANNER, M. & DAUBENBERGER, C. 2019. A multiplex qPCR approach for detection of pfhrp2 and pfhrp3 gene deletions in multiple strain infections of *Plasmodium falciparum*. *Sci Rep*, 9, 13107.
- SCHÜRCH, A. C. & SIEZEN, R. J. 2010. Genomic tracing of epidemics and disease outbreaks. *Microb Biotechnol*, 3, 628-33.
- SEREIKA, M., KIRKEGAARD, R. H., KARST, S. M., MICHAELSEN, T. Y., SØRENSEN, E. A., WOLLENBERG, R. D. & ALBERTSEN, M. 2022. Oxford Nanopore R10.4 long-read sequencing enables the generation of

near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nature Methods*, 19, 823-826.

SHARMA, D., LATHER, M., DYKES, C. L., DANG, A. S., ADAK, T. & SINGH, O. P. 2016. Disagreement in genotyping results of drug resistance alleles of the *Plasmodium falciparum* dihydrofolate reductase (Pfdhfr) gene by allele-specific PCR (ASPCR) assays and Sanger sequencing. *Parasitology Research*, 115, 323-328.

SHU, Y. & MCCAULEY, J. 2017. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*, 22, 30494.

SIBLEY, C. H., HYDE, J. E., SIMS, P. F. G., PLOWE, C. V., KUBLIN, J. G., MBERU, E. K., COWMAN, A. F., WINSTANLEY, P. A., WATKINS, W. M. & NZILA, A. M. 2001. Pyrimethamine–sulfadoxine resistance in *Plasmodium falciparum*: what next? *Trends in Parasitology*, 17, 582-588.

SIMONSEN, P., FISCHER, P., HOERAUF, A. & WEIL, G. 2014. *Manson's Tropical Diseases*, 23rd Edition: The Filariases.

SIMONSEN, P. E., ONAPA, A. W. & ASIO, S. M. 2011. *Mansonella perstans* filariasis in Africa. *Acta Tropica*, 120, S109-S120.

SIMPSON, J. A., AARONS, L., COLLINS, W. E., JEFFERY, G. M. & WHITE, N. J. 2002. Population dynamics of untreated *Plasmodium falciparum* malaria within the adult human host during the expansion phase of the infection. *Parasitology*, 124, 247-263.

SINGH, B., SUNG, L. K., MATUSOP, A., RADHAKRISHNAN, A., SHAMSUL, S. S. G., COX-SINGH, J., THOMAS, A. & CONWAY, D. J. 2004. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *The Lancet*, 363, 1017-1024.

SINGH, L., SAN, J. E., TEGALLY, H., BRZOSKA, P. M., ANYANEJI, U. J., WILKINSON, E., CLARK, L., GIANDHARI, J., PILLAY, S., LESSELLS, R. J., MARTIN, D. P., FURTADO, M., KIRAN, A. M. & DE OLIVEIRA, T. 2022. Targeted Sanger sequencing to recover key mutations in SARS-CoV-2 variant genome assemblies produced by next-generation sequencing. *Microbial Genomics*, 8.

STANAWAY, J. D., PARISI, A., SARKAR, K., BLACKER, B. F., REINER, R. C., HAY, S. I., NIXON, M. R., DOLECEK, C., JAMES, S. L., MOKDAD, A. H., ABEBE, G., AHMADIAN, E., ALAHDAB, F., ALEMNEW, B. T. T., ALIPOUR, V., ALLAH BAKESHEI, F., ANIMUT, M. D., ANSARI, F., ARABLOO, J., ASFAW, E. T., BAGHERZADEH, M., BASSAT, Q., BELAYNEH, Y. M. M., CARVALHO, F., DARYANI, A., DEMEKE, F. M., DEMIS, A. B. B., DUBEY, M., DUKEN, E. E., DUNACHIE, S. J., EFTEKHARI, A., FERNANDES, E., FOULADI FARD, R., GEDEFW, G. A., GETA, B., GIBNEY, K. B., HASANZADEH, A., HOANG, C. L., KASAEIAN, A.,

KHATER, A., KIDANEMARIAM, Z. T., LAKEW, A. M., MALEKZADEH, R., MELESE, A., MENGISTU, D. T., MESTROVIC, T., MIAZGOWSKI, B., MOHAMMAD, K. A., MOHAMMADIAN, M., MOHAMMADIAN-HAFSHEJANI, A., NGUYEN, C. T., NGUYEN, L. H., NGUYEN, S. H., NIRAYO, Y. L., OLAGUNJU, A. T., OLAGUNJU, T. O., POURJAFAR, H., QORBANI, M., RABIEE, M., RABIEE, N., RAFAY, A., REZAPOUR, A., SAMY, A. M., SEPANLOU, S. G., SHAIKH, M. A., SHARIF, M., SHIGEMATSU, M., TESSEMA, B., TRAN, B. X., ULLAH, I., YIMER, E. M., ZAIDI, Z., MURRAY, C. J. L. & CRUMP, J. A. 2019a. The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Diseases*, 19, 1312-1324.

STANAWAY, J. D., REINER, R. C., BLACKER, B. F., GOLDBERG, E. M., KHALIL, I. A., TROEGER, C. E., ANDREWS, J. R., BHUTTA, Z. A., CRUMP, J. A., IM, J., MARKS, F., MINTZ, E., PARK, S. E., ZAIDI, A. K. M., ABEBE, Z., ABEJIE, A. N., ADEDEJI, I. A., ALI, B. A., AMARE, A. T., ATALAY, H. T., AVOKPAHO, E. F. G. A., BACHA, U., BARAC, A., BEDI, N., BERHANE, A., BROWNE, A. J., CHIRINOS, J. L., CHITHEER, A., DOLECEK, C., EL SAYED ZAKI, M., ESHRATI, B., FOREMAN, K. J., GEMECHU, A., GUPTA, R., HAILU, G. B., HENOK, A., HIBSTU, D. T., HOANG, C. L., ILESANMI, O. S., IYER, V. J., KAHSAY, A., KASAEIAN, A., KASSA, T. D., KHAN, E. A., KHANG, Y.-H., MAGDY ABD EL RAZEK, H., MELKU, M., MENGISTU, D. T., MOHAMMAD, K. A., MOHAMMED, S., MOKDAD, A. H., NACHEGA, J. B., NAHEED, A., NGUYEN, C. T., NGUYEN, H. L. T., NGUYEN, L. H., NGUYEN, N. B., NGUYEN, T. H., NIRAYO, Y. L., PANGESTU, T., PATTON, G. C., QORBANI, M., RAI, R. K., RANA, S. M., RANABHAT, C. L., ROBA, K. T., ROBERTS, N. L. S., RUBINO, S., SAFIRI, S., SARTORIUS, B., SAWHNEY, M., SHIFERAW, M. S., SMITH, D. L., SYKES, B. L., TRAN, B. X., TRAN, T. T., UKWAJA, K. N., VU, G. T., VU, L. G., WELDEGEBREAL, F., YENIT, M. K., MURRAY, C. J. L. & HAY, S. I. 2019b. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Diseases*, 19, 369-381.

STEEL, C., GOLDEN, A., STEVENS, E., YOKOBE, L., DOMINGO, G. J., DE LOS SANTOS, T. & NUTMAN, T. B. 2015. Rapid Point-of-Contact Tool for Mapping and Integrated Surveillance of *Wuchereria bancrofti* and *Onchocerca volvulus* Infection. *Clin Vaccine Immunol*, 22, 896-901.

SU, M., SATOLA, S. W. & READ, T. D. 2019. Genome-Based Prediction of Bacterial Antibiotic Resistance. *J Clin Microbiol*, 57.

SU, S., WONG, G., SHI, W., LIU, J., LAI, A. C. K., ZHOU, J., LIU, W., BI, Y. & GAO, G. F. 2016. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends in Microbiology*, 24, 490-502.

SUTHERLAND, C. J., TANOMSING, N., NOLDER, D., OGUIKE, M., JENNISON, C., PUKRITTAYAKAMEE, S., DOLECEK, C., HIEN, T. T., DO ROSÁRIO, V. E., AREZ, A. P., PINTO, J., MICHON, P., ESCALANTE, A. A., NOSTEN, F., BURKE, M., LEE, R., BLAZE, M., OTTO, T. D., BARNWELL, J. W., PAIN, A., WILLIAMS, J.,

WHITE, N. J., DAY, N. P., SNOUNOU, G., LOCKHART, P. J., CHIODINI, P. L., IMWONG, M. & POLLEY, S. D. 2010. Two nonrecombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. *J Infect Dis*, 201, 1544-50.

SZOBOSZLAY, M., SCHRAMM, L., PINZAUTI, D., SCERRI, J., SANDIONIGI, A. & BIAZZO, M. 2023. Nanopore Is Preferable over Illumina for 16S Amplicon Sequencing of the Gut Microbiota When Species-Level Taxonomic Classification, Accurate Estimation of Richness, or Focus on Rare Taxa Is Required. *Microorganisms*, 11.

TA-TANG, T. H., CRAINEY, J. L., POST, R. J., LUZ, S. L. & RUBIO, J. M. 2018. Mansonellosis: current perspectives. *Res Rep Trop Med*, 9, 9-24.

TA, T.-H., MOYA, L., NGUEMA, J., APARICIO, P., MIGUEL-OTEO, M., CENZUAL, G., CANOREA, I., LANZA, M., BENITO, A., CRAINEY, J. L. & RUBIO, J. M. 2018. Geographical distribution and species identification of human filariasis and onchocerciasis in Bioko Island, Equatorial Guinea. *Acta Tropica*, 180, 12-17.

TABATA, Y., MATSUO, Y., FUJII, Y., OHTA, A. & HIROTA, K. 2022. Rapid detection of single nucleotide polymorphisms using the MinION nanopore sequencer: a feasibility study for perioperative precision medicine. *JA Clinical Reports*, 8, 17.

TAKOUGANG, I., MEREMIKWU, M., WANDJI, S., YENSHU, E. V., ARIPKO, B., LAMLENN, S. B., EKA, B. L., ENYONG, P., MELI, J. & KALE, O. 2002. Rapid assessment method for prevalence and intensity of *Loa loa* infection. *Bulletin of the World Health Organization*, 80, 852-858.

TALIC, S., SHAH, S., WILD, H., GASEVIC, D., MAHARAJ, A., ADEMI, Z., LI, X., XU, W., MESA-EGUIAGARAY, I., ROSTRON, J., THEODORATOU, E., ZHANG, X., MOTEE, A., LIEW, D. & ILIC, D. 2021. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. *Bmj*, 375, e068302.

TAYLOR, B. J., LANKE, K., BANMAN, S. L., MORLAIS, I., MORIN, M. J., BOUSEMA, T., RIJPMA, S. R. & YANOW, S. K. 2017. A Direct from Blood Reverse Transcriptase Polymerase Chain Reaction Assay for Monitoring *Falciparum* Malaria Parasite Transmission in Elimination Settings. *Am J Trop Med Hyg*, 97, 533-543.

TAYLOR, E. V., HERMAN, K. M., AILES, E. C., FITZGERALD, C., YODER, J. S., MAHON, B. E. & TAUXE, R. V. 2013. Common source outbreaks of *Campylobacter* infection in the USA, 1997–2008. *Epidemiology & Infection*, 141, 987-996.

TAYLOR, M. J., BANDI, C. & HOERAUF, A. 2005. Wolbachia. Bacterial Endosymbionts of Filarial Nematodes. In: BAKER, J. R., MULLER, R. & ROLLINSON, D. (eds.) *Advances in Parasitology*. Academic Press.

TAYLOR, M. J., HOERAUF, A. & BOCKARIE, M. 2010. Lymphatic filariasis and onchocerciasis. *The Lancet*, 376, 1175-1185.

TEGALLY, H., SAN, J. E., COTTEN, M., MOIR, M., TEGOMOH, B., MBOOWA, G., MARTIN, D. P., BAXTER, C., LAMBISIA, A. W., DIALLO, A., AMOAKO, D. G., DIAGNE, M. M., SISAY, A., ZEKRI, A.-R. N., GUEYE, A. S., SANGARE, A. K., OUEDRAOGO, A.-S., SOW, A., MUSA, A. O., SESAY, A. K., ABIAS, A. G., ELZAGHEID, A. I., LAGARE, A., KEMI, A.-S., ABAR, A. E., JOHNSON, A. A., FOWOTADE, A., OLUWAPELUMI, A. O., AMURI, A. A., JURU, A., KANDEIL, A., MOSTAFA, A., REBAI, A., SAYED, A., KAZEEM, A., BALDE, A., CHRISTOFFELS, A., TROTTER, A. J., CAMPBELL, A., KEITA, A. K., KONE, A., BOUZID, A., SOUISSI, A., AGWEYU, A., NAGUIB, A., GUTIERREZ, A. V., NKESHIMANA, A., PAGE, A. J., YADOLETON, A., VINZE, A., HAPPI, A. N., CHOUIKHA, A., IRANZADEH, A., MAHARAJ, A., BATCHI-BOUYOU, A. L., ISMAIL, A., SYLVERKEN, A. A., GOBA, A., FEMI, A., SIJUWOLA, A. E., MARYCELIN, B., SALAKO, B. L., ODERINDE, B. S., BOLAJOKO, B., DIARRA, B., HERRING, B. L., TSOFA, B., LEKANA-DOUKI, B., MVULA, B., NJANPOP-LAFOURCADE, B.-M., MARONDERA, B. T., KHAIREH, B. A., KOURIBA, B., ADU, B., POOL, B., MCINNIS, B., BROOK, C., WILLIAMSON, C., NDUWIMANA, C., ANSCOMBE, C., PRATT, C. B., SCHEEPERS, C., AKOUA-KOFFI, C. G., AGOTI, C. N., MAPANGUY, C. M., LOUCOUBAR, C., ONWUAMAH, C. K., IHEKWEAZU, C., MALAKA, C. N., PEYREFITTE, C., GRACE, C., OMORUYI, C. E., RAFAÏ, C. D., MORANG'A, C. M., ERAMEH, C., LULE, D. B., BRIDGES, D. J., MUKADI-BAMULEKA, D., PARK, D., RASMUSSEN, D. A., et al. 2022. The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance. *Science*, 378, eabq5358.

TEGALLY, H., WILKINSON, E., GIOVANETTI, M., IRANZADEH, A., FONSECA, V., GIANDHARI, J., DOOLABH, D., PILLAY, S., SAN, E. J., MSOMI, N., MLISANA, K., VON GOTTBERG, A., WALAZA, S., ALLAM, M., ISMAIL, A., MOHALE, T., GLASS, A. J., ENGELBRECHT, S., VAN ZYL, G., PREISER, W., PETRUCCIONE, F., SIGAL, A., HARDIE, D., MARAIS, G., HSIAO, N.-Y., KORSMAN, S., DAVIES, M.-A., TYERS, L., MUDAU, I., YORK, D., MASLO, C., GOEDHALS, D., ABRAHAMS, S., LAGUDA-AKINGBA, O., ALISOLTANI-DEHKORDI, A., GODZIK, A., WIBMER, C. K., SEWELL, B. T., LOURENÇO, J., ALCANTARA, L. C. J., KOSAKOVSKY POND, S. L., WEAVER, S., MARTIN, D., LESSELLS, R. J., BHIMAN, J. N., WILLIAMSON, C. & DE OLIVEIRA, T. 2021. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*, 592, 438-443.

TEMMAM, S., VONGPHAYLOTH, K., BAQUERO, E., MUNIER, S., BONOMI, M., REGNAULT, B., DOUANGBOUBPHA, B., KARAMI, Y., CHRÉTIEN, D., SANAMXAY, D., XAYAPHET, V., PAPHAPHANH, P., LACOSTE, V., SOMLOR, S., LAKEOMANY, K., PHOMMAVANH, N., PÉROT, P., DEHAN, O., AMARA, F.,

- DONATI, F., BIGOT, T., NILGES, M., REY, F. A., VAN DER WERF, S., BREY, P. T. & ELOIT, M. 2022. Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*, 604, 330-336.
- TEYSSIER, N. B., CHEN, A., DUARTE, E. M., SIT, R., GREENHOUSE, B. & TESSEMA, S. K. 2021. Optimization of whole-genome sequencing of *Plasmodium falciparum* from low-density dried blood spot samples. *Malar J*, 20, 116.
- TRAPE, J.-F., PISON, G., PREZIOSI, M.-P., ENEL, C., DU LOÛ, A. D., DELAUNAY, V., SAMB, B., LAGARDE, E., MOLEZ, J.-F. & SIMONDON, F. 1998. Impact of chloroquine resistance on malaria mortality. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*, 321, 689-697.
- TROMBLEY HALL, A., MCKAY ZOVANYI, A., CHRISTENSEN, D. R., KOEHLER, J. W. & DEVINS MINOGUE, T. 2013. Evaluation of inhibitor-resistant real-time PCR methods for diagnostics in clinical and environmental samples. *PLoS One*, 8, e73845.
- TRUPPA, C. & ABO-SHEHADA, M. N. 2020. Antimicrobial resistance among GLASS pathogens in conflict and non-conflict affected settings in the Middle East: a systematic review. *BMC Infectious Diseases*, 20, 936.
- TYSON, J. R., JAMES, P., STODDART, D., SPARKS, N., WICKENHAGEN, A., HALL, G., CHOI, J. H., LAPOINTE, H., KAMELIAN, K., SMITH, A. D., PRYSTAJECKY, N., GOODFELLOW, I., WILSON, S. J., HARRIGAN, R., SNUTCH, T. P., LOMAN, N. J. & QUICK, J. 2020. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *bioRxiv*.
- TZOU, P. L., ARIYARATNE, P., VARGHESE, V., LEE, C., RAKHMANALIEV, E., VILLY, C., YEE, M., TAN, K., MICHEL, G., PINSKY, B. A. & SHAFER, R. W. 2018. Comparison of an *In Vitro* Diagnostic Next-Generation Sequencing Assay with Sanger Sequencing for HIV-1 Genotypic Resistance Testing. *Journal of Clinical Microbiology*, 56, e00105-18.
- VAN DE MAAT, J., DE SANTIS, O., LUWANDA, L., TAN, R. & KEITEL, K. 2021. Primary Care Case Management of Febrile Children: Insights From the ePOCT Routine Care Cohort in Dar es Salaam, Tanzania. *Frontiers in Pediatrics*, 9.
- VARO, R., BALANZA, N., MAYOR, A. & BASSAT, Q. 2021. Diagnosis of clinical malaria in endemic settings. *Expert Review of Anti-infective Therapy*, 19, 79-92.
- VERMA, A. K., BHARTI, P. K. & DAS, A. 2018. HRP-2 deletion: a hole in the ship of malaria elimination. *Lancet Infect Dis*, 18, 826-827.
- VIANA, R., MOYO, S., AMOAKO, D. G., TEGALLY, H., SCHEEPERS, C., ALTHAUS, C. L., ANYANEJI, U. J., BESTER, P. A., BONI, M. F., CHAND, M., CHOGA, W. T., COLQUHOUN, R., DAVIDS, M., DEFORCHE, K.,

DOOLABH, D., DU PLESSIS, L., ENGELBRECHT, S., EVERATT, J., GIANDHARI, J., GIOVANETTI, M., HARDIE, D., HILL, V., HSIAO, N. Y., IRANZADEH, A., ISMAIL, A., JOSEPH, C., JOSEPH, R., KOOPILE, L., KOSAKOVSKY POND, S. L., KRAEMER, M. U. G., KUATE-LERE, L., LAGUDA-AKINGBA, O., LESETEDI-MAFOKO, O., LESSELLS, R. J., LOCKMAN, S., LUCACI, A. G., MAHARAJ, A., MAHLANGU, B., MAPONGA, T., MAHLAKWANE, K., MAKATINI, Z., MARAIS, G., MARUAPULA, D., MASUPU, K., MATSHABA, M., MAYAPHI, S., MBHELE, N., MBULAWA, M. B., MENDES, A., MLISANA, K., MNGUNI, A., MOHALE, T., MOIR, M., MORUISI, K., MOSEPELE, M., MOTSATSI, G., MOTSWALEDI, M. S., MPHUYAKGOSI, T., MSOMI, N., MWANGI, P. N., NAIDOO, Y., NTULI, N., NYAGA, M., OLUBAYO, L., PILLAY, S., RADIBE, B., RAMPHAL, Y., RAMPHAL, U., SAN, J. E., SCOTT, L., SHAPIRO, R., SINGH, L., SMITH-LAWRENCE, P., STEVENS, W., STRYDOM, A., SUBRAMONEY, K., TEBEILA, N., TSHIABUILA, D., TSUI, J., VAN WYK, S., WEAVER, S., WIBMER, C. K., WILKINSON, E., WOLTER, N., ZAREBSKI, A. E., ZUZE, B., GOEDHALS, D., PREISER, W., TREURNICHT, F., VENTER, M., WILLIAMSON, C., PYBUS, O. G., BHIMAN, J., GLASS, A., MARTIN, D. P., RAMBAUT, A., GASEITSIWE, S., VON GOTTBURG, A. & DE OLIVEIRA, T. 2022. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*, 603, 679-686.

WAGNER, K., SPRINGER, B., PIRES, V. P. & KELLER, P. M. 2018. Molecular detection of fungal pathogens in clinical specimens by 18S rDNA high-throughput screening in comparison to ITS PCR and culture. *Sci Rep*, 8, 6964.

WALKER, T. M., MIOTTO, P., KÖSER, C. U., FOWLER, P. W., KNAGGS, J., IQBAL, Z., HUNT, M., CHINDELEVITCH, L., FARHAT, M., CIRILLO, D. M., COMAS, I., POSEY, J., OMAR, S. V., PETO, T. E., SURESH, A., UPLEKAR, S., LAURENT, S., COLMAN, R. E., NATHANSON, C. M., ZIGNOL, M., WALKER, A. S., CROOK, D. W., ISMAIL, N. & RODWELL, T. C. 2022. The 2021 WHO catalogue of Mycobacterium tuberculosis complex mutations associated with drug resistance: A genotypic analysis. *Lancet Microbe*, 3, e265-e273.

WANG, H., JEAN, S., ELTRINGHAM, R., MADISON, J., SNYDER, P., TU, H., JONES, D. M. & LEBER, A. L. 2021a. Mutation-Specific SARS-CoV-2 PCR Screen: Rapid and Accurate Detection of Variants of Concern and the Identification of a Newly Emerging Variant with Spike L452R Mutation. *J Clin Microbiol*, 59, e0092621.

WANG, Y., ZHAO, Y., BOLLAS, A., WANG, Y. & AU, K. F. 2021b. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*, 39, 1348-1365.

WEE, S. K., SIVALINGAM, S. P. & YAP, E. P. H. 2020. Rapid Direct Nucleic Acid Amplification Test without RNA Extraction for SARS-CoV-2 Using a Portable PCR Thermocycler. *Genes (Basel)*, 11.

WEIL, G. J., CURTIS, K. C., FAKOLI, L., FISCHER, K., GANKPALA, L., LAMMIE, P. J., MAJEWSKI, A. C., PELLETREAU, S., WON, K. Y., BOLAY, F. K. & FISCHER, P. U. 2013. Laboratory and field evaluation of a new rapid test for detecting *Wuchereria bancrofti* antigen in human blood. *Am J Trop Med Hyg*, 89, 11-15.

WEIL, G. J., LAMMIE, P. J. & WEISS, N. 1997. The ICT Filariasis Test: A rapid-format antigen test for diagnosis of bancroftian filariasis. *Parasitology Today*, 13, 401-404.

WENGER, A. M., PELUSO, P., ROWELL, W. J., CHANG, P.-C., HALL, R. J., CONCEPCION, G. T., EBLER, J., FUNGTAMMASAN, A., KOLESNIKOV, A., OLSON, N. D., TÖPFER, A., ALONGE, M., MAHMOUD, M., QIAN, Y., CHIN, C.-S., PHILLIPPY, A. M., SCHATZ, M. C., MYERS, G., DEPRISTO, M. A., RUAN, J., MARSCHALL, T., SEDLAZECK, F. J., ZOOK, J. M., LI, H., KOREN, S., CARROLL, A., RANK, D. R. & HUNKAPILLER, M. W. 2019. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, 37, 1155-1162.

WHITE, N. J. 2004. Antimalarial drug resistance. *The Journal of clinical investigation*, 113, 1084-1092.

WHITE, N. J. 2011. Determinants of relapse periodicity in *Plasmodium vivax* malaria. *Malar J*, 10, 297.

WHITE, N. J. 2022. Severe malaria. *Malar J*, 21, 284.

WHITE, N. J. & OLLIARO, P. L. 1996. Strategies for the prevention of antimalarial drug resistance: rationale for combination chemotherapy for malaria. *Parasitol Today*, 12, 399-401.

WHITE, N. J., PUKRITTAYAKAMEE, S., HIEN, T. T., FAIZ, M. A., MOKUOLU, O. A. & DONDORP, A. M. 2014. Malaria. *The Lancet*, 383, 723-735.

WHITFORD, W., HAWKINS, V., MOODLEY, K. S., GRANT, M. J., LEHNERT, K., SNELL, R. G. & JACOBSEN, J. C. 2022. Proof of concept for multiplex amplicon sequencing for mutation identification using the MinION nanopore sequencer. *Sci Rep*, 12, 8572.

WHITTAKER, C., WALKER, M., PION, S. D. S., CHESNAIS, C. B., BOUSSINESQ, M. & BASÁÑEZ, M.-G. 2018. The Population Biology and Transmission Dynamics of *Loa loa*. *Trends in Parasitology*, 34, 335-350.

WILSON, M. R., SAMPLE, H. A., ZORN, K. C., AREVALO, S., YU, G., NEUHAUS, J., FEDERMAN, S., STRYKE, D., BRIGGS, B., LANGELIER, C., BERGER, A., DOUGLAS, V., JOSEPHSON, S. A., CHOW, F. C., FULTON, B. D., DERISI, J. L., GELFAND, J. M., NACCACHE, S. N., BENDER, J., DIEN BARD, J., MURKEY, J., CARLSON, M., VESPA, P. M., VIJAYAN, T., ALLYN, P. R., CAMPEAU, S., HUMPHRIES, R. M., KLAUSNER, J. D., GANZON, C. D., MEMAR, F., OCAMPO, N. A., ZIMMERMANN, L. L., COHEN, S. H., POLAGE, C. R., DEBIASI, R. L., HALLER, B., DALLAS, R., MARON, G., HAYDEN, R., MESSACAR, K., DOMINGUEZ, S. R.,

MILLER, S. & CHIU, C. Y. 2019. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl J Med*, 380, 2327-2340.

WINAND, R., BOGAERTS, B., HOFFMAN, S., LEFEVRE, L., DELVOYE, M., VAN BRAEKEL, J., FU, Q., ROOSENS, N. H., DE KEERSMAECKER, S. C. & VANNESTE, K. 2020. Targeting the 16S rRNA Gene for Bacterial Identification in Complex Mixed Samples: Comparative Evaluation of Second (Illumina) and Third (Oxford Nanopore Technologies) Generation Sequencing Technologies. *International Journal of Molecular Sciences*, 21, 298.

WOLFE, N. D., DUNAVAN, C. P. & DIAMOND, J. 2007. Origins of major human infectious diseases. *Nature*, 447, 279-83.

WOOLHOUSE, M. E. J. & WARD, M. J. 2013. Sources of Antimicrobial Resistance. *Science*, 341, 1460-1461.

WORLD HEALTH, O. 2010. Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases. Geneva: World Health Organization.

WORLD HEALTH ORGANIZATION 2022a. Elimination of human onchocerciasis: progress report, 2021–Élimination de l'onchocercose humaine: rapport de situation, 2021. *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire*, 97, 591-598.

WORLD HEALTH ORGANIZATION 2022b. Global programme to eliminate lymphatic filariasis: progress report, 2021–Programme mondial pour l'élimination de la filariose lymphatique: rapport de situation, 2021. *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire*, 97, 513-524.

WORLD HEALTH ORGANIZATION 2022c. World malaria report 2022. Geneva.

WU, F., ZHAO, S., YU, B., CHEN, Y.-M., WANG, W., SONG, Z.-G., HU, Y., TAO, Z.-W., TIAN, J.-H., PEI, Y.-Y., YUAN, M.-L., ZHANG, Y.-L., DAI, F.-H., LIU, Y., WANG, Q.-M., ZHENG, J.-J., XU, L., HOLMES, E. C. & ZHANG, Y.-Z. 2020. A new coronavirus associated with human respiratory disease in China. *Nature*, 579, 265-269.

WU, Y., KANG, L., GUO, Z., LIU, J., LIU, M. & LIANG, W. 2022. Incubation Period of COVID-19 Caused by Unique SARS-CoV-2 Strains: A Systematic Review and Meta-analysis. *JAMA Netw Open*, 5, e2228008.

XIA, S., WANG, L., ZHU, Y., LU, L. & JIANG, S. 2022. Origin, virological features, immune evasion and intervention of SARS-CoV-2 Omicron sublineages. *Signal Transduct Target Ther*, 7, 241.

YAMASOBA, D., KIMURA, I., NASSER, H., MORIOKA, Y., NAO, N., ITO, J., URIU, K., TSUDA, M., ZAHRADNIK, J., SHIRAKAWA, K., SUZUKI, R., KISHIMOTO, M., KOSUGI, Y., KOBIYAMA, K., HARA, T.,

TOYODA, M., TANAKA, Y. L., BUTLERTANAKA, E. P., SHIMIZU, R., ITO, H., WANG, L., ODA, Y., ORBA, Y., SASAKI, M., NAGATA, K., YOSHIMATSU, K., ASAKURA, H., NAGASHIMA, M., SADAMASU, K., YOSHIMURA, K., KURAMOCHI, J., SEKI, M., FUJIKI, R., KANEDA, A., SHIMADA, T., NAKADA, T. A., SAKAO, S., SUZUKI, T., UENO, T., TAKAORI-KONDO, A., ISHII, K. J., SCHREIBER, G., SAWA, H., SAITO, A., IRIE, T., TANAKA, S., MATSUNO, K., FUKUHARA, T., IKEDA, T. & SATO, K. 2022. Virological characteristics of the SARS-CoV-2 Omicron BA.2 spike. *Cell*, 185, 2103-2115.e19.

YEK, C., PACHECO, A. R., VANAERSCHOT, M., BOHL, J. A., FAHSBENDER, E., ARANDA-DÍAZ, A., LAY, S., CHEA, S., OUM, M. H., LON, C., TATO, C. M. & MANNING, J. E. 2022. Metagenomic Pathogen Sequencing in Resource-Scarce Settings: Lessons Learned and the Road Ahead. *Front Epidemiol*, 2.

YOBOUE, C. A., HOSCH, S., DONFACK, O. T., GUIROU, E. A., NLAVO, B. M., AYEKABA, M. O. O., GUERRA, C., PHIRI, W. P., GARCIA, G. A., SCHINDLER, T. & DAUBENBERGER, C. A. 2022. Characterising co-infections with Plasmodium spp., Mansonella perstans or Loa loa in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests. *PLOS Neglected Tropical Diseases*, 16, e0009798.

ZAKI, A. M., VAN BOHEEMEN, S., BESTEBROER, T. M., OSTERHAUS, A. D. M. E. & FOUCHIER, R. A. M. 2012. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine*, 367, 1814-1820.

ZHANG, J. J., DONG, X., LIU, G. H. & GAO, Y. D. 2023. Risk and Protective Factors for COVID-19 Morbidity, Severity, and Mortality. *Clin Rev Allergy Immunol*, 64, 90-107.

ZHANG, S., YIN, Y., JONES, M. B., ZHANG, Z., DEATHERAGE KAISER, B. L., DINSMORE, B. A., FITZGERALD, C., FIELDS, P. I. & DENG, X. 2015. Salmonella serotype determination utilizing high-throughput genome sequencing data. *J Clin Microbiol*, 53, 1685-92.

ZHOU, H., JI, J., CHEN, X., BI, Y., LI, J., WANG, Q., HU, T., SONG, H., ZHAO, R., CHEN, Y., CUI, M., ZHANG, Y., HUGHES, A. C., HOLMES, E. C. & SHI, W. 2021. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell*, 184, 4380-4391.e14.

ZHOU, P., YANG, X.-L., WANG, X.-G., HU, B., ZHANG, L., ZHANG, W., SI, H.-R., ZHU, Y., LI, B., HUANG, C.-L., CHEN, H.-D., CHEN, J., LUO, Y., GUO, H., JIANG, R.-D., LIU, M.-Q., CHEN, Y., SHEN, X.-R., WANG, X., ZHENG, X.-S., ZHAO, K., CHEN, Q.-J., DENG, F., LIU, L.-L., YAN, B., ZHAN, F.-X., WANG, Y.-Y., XIAO, G.-F. & SHI, Z.-L. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579, 270-273.

ZHU, N., ZHANG, D., WANG, W., LI, X., YANG, B., SONG, J., ZHAO, X., HUANG, B., SHI, W., LU, R., NIU, P., ZHAN, F., MA, X., WANG, D., XU, W., WU, G., GAO, G. F. & TAN, W. 2020. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, 382, 727-733.

ZINSSTAG, J., KAISER-GROLIMUND, A., HEITZ-TOKPA, K., SREEDHARAN, R., LUBROTH, J., CAYA, F., STONE, M., BROWN, H., BONFOH, B., DOBELL, E., MORGAN, D., HOMAIRA, N., KOCK, R., HATTENDORF, J., CRUMP, L., MAUTI, S., DEL RIO VILAS, V., SAIKAT, S., ZUMLA, A., HEYMANN, D., DAR, O. & DE LA ROCQUE, S. 2023. Advancing One human-animal-environment Health for global health security: what does the evidence say? *The Lancet*, 401, 591-604.

ZOURÉ, H. G. M., WANJI, S., NOMA, M., AMAZIGO, U. V., DIGGLE, P. J., TEKLE, A. H. & REMME, J. H. F. 2011. The Geographic Distribution of Loa loa in Africa: Results of Large-Scale Implementation of the Rapid Assessment Procedure for Loiasis (RAPLOA). *PLOS Neglected Tropical Diseases*, 5, e1210.

8. Appendix

Curriculum vitae Salome Hosch

CURRICULUM VITAE – SALOME HOSCH

Date of birth: 28.02.1996

Nationality: Swiss

Address: Dornacherstrasse 174, 4053 Basel, Switzerland

Phone: +41 77 459 89 14

E-mail: salome.hosch@gmx.ch

Education

March 2020 – June 2023	PhD in Microbiology	<i>University of Basel</i>
Sep. 2018 – Feb. 2020	Master of Science in Infection Biology	<i>University of Basel</i>
Sep. 2015 – June 2018	Bachelor of Science in Biology	<i>University of Basel</i>

Work experience

March 2020 – June 2023	PhD thesis	<i>Swiss Tropical and Public Health Institute</i>
	PhD Thesis: Investigating drug resistance markers, evolution and co-infections of pathogens of public health concern using portable molecular detection and sequencing tools	
Feb. 2019 – Feb. 2020	Master thesis	<i>Swiss Tropical and Public Health Institute</i>
	Master thesis: Molecular characterization of <i>Plasmodium</i> spp. prevalence and investigation of <i>k13</i> mutants and <i>hrp2</i> / <i>hrp3</i> gene deletion carrying <i>P. falciparum</i> on Bioko Island	
July 2018 – Oct. 2018	Assistant in block course plant biology	<i>University of Basel</i>
	Preparation of experiments, supervision of students	
April 2015 – June 2017	Part-time job as student assistant	<i>b'SYS GmbH</i>
	Preparation of buffers, general laboratory duties	

Relevant skills

Languages	German	Mother tongue
	English	Fluent
	French	Intermediate
Bioanalytics	Strong technical background in various laboratory techniques.	

Publications

1. Krähenbühl S, Studer F, Guirou E, Deal A, Mächler P, **Hosch S**, Mpina M, Mswata S, Daubenberger C, Schindler T. **ELIMU-MDx: a web-based, open-source platform for storage, management and analysis of diagnostic qPCR data**. *Biotechniques*. 2020 Jan;68(1):22-27. doi: 10.2144/btn-2019-0064. Epub 2019 Oct 7.
2. Guirou EA, Schindler T, **Hosch S**, Donfack OT, Yoboue CA, Krähenbühl S, Deal A, Cosi G, Gondwe L, Mwangoka G, Masuki H, Salim N, Mpina M, Said J, Abdulla S, Hoffman SL, Nlavo BM, Maas C, Falla CC, Phiri WP, Garcia GA, Tanner M, Daubenberger C. **Molecular malaria surveillance using a novel protocol for extraction and analysis of nucleic acids retained on used rapid diagnostic tests**. *Sci Rep*. 2020 Jul 23;10(1):12305. doi: 10.1038/s41598-020-69268-5.
3. Jongo SA, Urbano V, Church LWP, Olotu A, Manock SR, Schindler T, Mtoro A, Kc N, Hamad A, Nyakarungu E, Mpina M, Deal A, Bijeri JR, Ondo Mangué ME, Ntutumu Pasiolo BE, Nguema GN, Owono SN, Rivas MR, Chemba M, Kassim KR, James ER, Stabler TC, Abebe Y, Saverino E, Sax J, **Hosch S**, Tumbo AM, Gondwe L, Segura JL, Falla CC, Phiri WP, Hergott DEB, García GA, Schwabe C, Maas CD, Murshedkar T, Billingsley PF, Tanner M, Ayekaba MO, Sim BKL, Daubenberger C, Richie TL, Abdulla S, Hoffman SL. **Immunogenicity and Protective Efficacy of Radiation-Attenuated and Chemo-Attenuated PfSPZ Vaccines in Equatoguinean Adults**. *Am J Trop Med Hyg*. 2021 Jan;104(1):283-293. doi: 10.4269/ajtmh.20-0435.
4. Dudas G, Hong SL, Potter BI, Calvignac-Spencer S, Niatou-Singa FS, Tombolomako TB, Fuh-Neba T, Vickos U, Ulrich M, Leendertz FH, Khan K, Huber C, Watts A, Olendraitė I, Snijder J, Wijnant KN, Bonvin AMJJ, Martres P, Behillil S, Ayoubia A, Maidadi MF, Djomsi DM, Godwe C, Butel C, Šimaitis A, Gabrielaitė M, Katėnaitė M, Norvilas R, Raugaitė L, Koyaweda GW, Kandou JK, Jonikas R, Nasvytienė I, Žemeckienė Ž, Gečys D, Tamušauskaitė K, Norkienė M, Vasiliūnaitė E, Žiogienė D, Timinskas A, Šukys M, Šarauskas M, Alzbutas G, Aziza AA, Lusamaki EK, Cigolo JM, Mawete FM, Lofiko EL, Kingebeni PM, Tamfum JM, Belizaire MRD, Essomba RG, Assoumou MCO, Mboringong AB, Dieng AB, Juozapaitė D, **Hosch S**, Obama J, Ayekaba MO, Naumovas D, Pautienius A, Rafaï CD, Vitkauskienė A, Ugenskienė R, Gedvilaitė A, Čereškevičius D, Lesauskaitė V, Žemaitis L, Griškevičius L, Baele G. **Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions**. *Nat Commun*. 2021 Oct 1;12(1):5769. doi: 10.1038/s41467-021-26055-8.
5. Bechtold P, Wagner P, **Hosch S**, Siegrist D, Ruiz-Serrano A, Gregorini M, Mpina M, Ondó FA, Obama J, Ayekaba MO, Engler O, Stark WJ, Daubenberger CA, Schindler T. **Rapid Identification of SARS-CoV-2 Variants of Concern Using a Portable peakPCR Platform**. *Anal Chem*. 2021 Dec 14;93(49):16350-16359. doi: 10.1021/acs.analchem.1c02368. Epub 2021 Dec 1.
6. **Hosch S**, Mpina M, Nyakarungu E, Borico NS, Obama TMA, Ovona MC, Wagner P, Rubin SE, Vickos U, Milang DVN, Ayekaba MO, Phiri WP, Daubenberger CA, Schindler T. **Genomic Surveillance Enables the Identification of Co-infections With Multiple SARS-CoV-2 Lineages**

in Equatorial Guinea. Front Public Health. 2022 Jan 4;9:818401. doi: 10.3389/fpubh.2021.818401.

7. **Hosch S**, Yoboue CA, Donfack OT, Guirou EA, Dangy JP, Mpina M, Nyakurungu E, Blöchliger K, Guerra CA, Phiri WP, Ayekaba MO, García GA, Tanner M, Daubenberger C, Schindler T. **Analysis of nucleic acids extracted from rapid diagnostic tests reveals a significant proportion of false positive test results associated with recent malaria treatment.** Malar J. 2022 Jan 24;21(1):23. doi: 10.1186/s12936-022-04043-7.
8. Yoboue CA, **Hosch S**, Donfack OT, Guirou EA, Nlavo BM, Ayekaba MO, Guerra C, Phiri WP, Garcia GA, Schindler T, Daubenberger CA. **Characterising co-infections with Plasmodium spp., Mansonella perstans or Loa loa in asymptomatic children, adults and elderly people living on Bioko Island using nucleic acids extracted from malaria rapid diagnostic tests.** PLoS Negl Trop Dis. 2022 Jan 31;16(1):e0009798. doi: 10.1371/journal.pntd.0009798.
9. Tegally H, San JE, Cotten M, Moir M, Tegomoh B, Mboowa G, Martin DP, Baxter C, Lambisia AW, Diallo A, Amoako DG, Diagne MM, Sisay A, Zekri AN, Gueye AS, Sangare AK, Ouedraogo AS, Sow A, Musa AO, Sesay AK, Abias AG, Elzagheid AI, Lagare A, Kemi AS, Abar AE, Johnson AA, Fowotade A, Oluwapelumi AO, Amuri AA, Juru A, Kandeil A, Mostafa A, Rebai A, Sayed A, Kazeem A, Balde A, Christoffels A, Trotter AJ, Campbell A, Keita AK, Kone A, Bouzid A, Souissi A, Agweyu A, Naguib A, Gutierrez AV, Nkeshimana A, Page AJ, Yadouleton A, Vinze A, Happi AN, Chouikha A, Iranzadeh A, Maharaj A, Batchi-Bouyou AL, Ismail A, Sylverken AA, Goba A, Femi A, Sijuwola AE, Marycelin B, Salako BL, Oderinde BS, Bolajoko B, Diarra B, Herring BL, Tsofa B, Lekana-Douki B, Mvula B, Njanpop-Lafourcade BM, Marondera BT, Khaireh BA, Kouriba B, Adu B, Pool B, McInnis B, Brook C, Williamson C, Nduwimana C, Anscombe C, Pratt CB, Scheepers C, Akoua-Koffi CG, Agoti CN, Mapanguy CM, Loucoubar C, Onwuamah CK, Ihekweazu C, Malaka CN, Peyrefitte C, Grace C, Omoruyi CE, Rafai CD, Morang'a CM, Erameh C, Lule DB, Bridges DJ, Mukadi-Bamuleka D, Park D, Rasmussen DA, Baker D, Nokes DJ, Ssemwanga D, Tshiabuila D, Amuzu DSY, Goedhals D, Grant DS, Omuoyo DO, Maruapula D, Wanjohi DW, Foster-Nyarko E, Lusamaki EK, Simulundu E, Ong'era EM, Ngabana EN, Abworo EO, Otieno E, Shumba E, Barasa E, Ahmed EB, Ahmed EA, Lokilo E, Mukantwari E, Philomena E, Belarbi E, Simon-Lorriere E, Anoh EA, Manuel E, Leendertz F, Taweh FM, Wasfi F, Abdelmoula F, Takawira FT, Derrar F, Ajogbasile FV, Treurnicht F, Onikepe F, Ntoumi F, Muyembe FM, Ragomzingba FEZ, Dratibi FA, Iyanu FA, Mbunsu GK, Thilliez G, Kay GL, Akpede GO, van Zyl GU, Awandare GA, Kpeli GS, Schubert G, Maphalala GP, Ranaivoson HC, Omunakwe HE, Onywera H, Abe H, Karray H, Nansumba H, Triki H, Kadjo HAA, Elgahzaly H, Gumbo H, Mathieu H, Kavunga-Membo H, Smeti I, Olawoye IB, Adetifa IMO, Odia I, Ben Boubaker IB, Muhammad IA, Ssewanyana I, Wurie I, Konstantinus IS, Halatoko JWA, Ayei J, Sonoo J, Makangara JC, Tamfum JM, Heraud JM, Shaffer JG, Giandhari J, Musyoki J, Nkurunziza J, Uwanibe JN, Bhiman JN, Yasuda J, Morais J, Kiconco J, Sandi JD, Huddleston J, Odoom JK, Morobe JM, Gyapong JO, Kayiwa JT, Okolie JC, Xavier JS, Gyamfi J, Wamala JF, Bonney JHK, Nyandwi J, Everatt J, Nakaseegu J, Ngoi JM, Namulondo J, Oguzie JU, Andeko JC, Lutwama JJ, Mogga JJH, O'Grady J, Siddie KJ, Victoir K, Adeyemi KT, Tumedi KA, Carvalho KS, Mohammed KS, Dellagi K, Musonda KG, Duedu KO, Fki-Berrajah L, Singh L, Kepler LM, Biscornet L, de Oliveira Martins L, Chabuka L, Olubayo L, Ojok LD, Deng LL, Ochola-Oyier LI, Tyers L, Mine M, Ramuth M, Mastouri M, ElHefnawi M, Mbanne M, Matsheka MI, Kebabonye M, Diop M, Momoh M, Lima Mendonça MDL, Venter M, Paye MF, Faye M, Nyaga MM, Mareka M, Damaris MM, Mburu MW, Mpina MG, Owusu M, Wiley MR, Tatfeng MY, Ayekaba MO, Abouelhoda M, Beloufa MA, Seadawy MG, Khalifa MK, Matobo MM, Kane M,

Salou M, Mbulawa MB, Mwenda M, Allam M, Phan MVT, Abid N, Rujeni N, Abuzaid N, Ismael N, Elguindy N, Top NM, Dia N, Mabunda N, Hsiao NY, Silochi NB, Francisco NM, Saasa N, Bbosa N, Murunga N, Gumede N, Wolter N, Sitharam N, Ndodo N, Ajayi NA, Tordo N, Mbhele N, Razanajatovo NH, Iguosadolo N, Mba N, Kingsley OC, Sylvanus O, Femi O, Adewumi OM, Testimony O, Ogunsanya OA, Fakayode O, Ogah OE, Oludayo OE, Faye O, Smith-Lawrence P, Ondo P, Combe P, Nabisubi P, Semanda P, Oluniyi PE, Arnaldo P, Quashie PK, Okokhere PO, Bejon P, Dussart P, Bester PA, Mbala PK, Kaleebu P, Abechi P, El-Shesheny R, Joseph R, Aziz RK, Essomba RG, Ayivor-Djanie R, Njouom R, Phillips RO, Gorman R, Kingsley RA, Neto Rodrigues RMDESA, Audu RA, Carr RAA, Gargouri S, Masmoudi S, Bootsma S, Sankhe S, Mohamed SI, Femi S, Mhalla S, **Hosch S**, Kassim SK, Metha S, Trabelsi S, Agwa SH, Mwangi SW, Doumbia S, Makiala-Mandanda S, Aryeetey S, Ahmed SS, Ahmed SM, Elhamoumi S, Moyo S, Lutucuta S, Gaseitsiwe S, Jalloh S, Andriamandimby SF, Oguntope S, Grayo S, Lekana-Douki S, Prosolek S, Ouangraoua S, van Wyk S, Schaffner SF, Kanyerezi S, Ahuka-Mundeye S, Rudder S, Pillay S, Nabadda S, Behillil S, Budiaki SL, van der Werf S, Mashe T, Mohale T, Le-Viet T, Velavan TP, Schindler T, Maponga TG, Bedford T, Anyaneji UJ, Chinedu U, Ramphal U, George UE, Enouf V, Nene V, Gorova V, Roshdy WH, Karim WA, Ampofo WK, Preiser W, Choga WT, Ahmed YA, Ramphal Y, Bediako Y, Naidoo Y, Butera Y, de Laurent ZR; Africa Pathogen Genomics Initiative (Africa PGI)†; Ouma AEO, von Gottberg A, Githinji G, Moeti M, Tomori O, Sabeti PC, Sall AA, Oyola SO, Tebeje YK, Tessema SK, de Oliveira T, Happi C, Lessells R, Nkengasong J, Wilkinson E. The evolving SARS-CoV-2 epidemic in Africa: **Insights from rapidly expanding genomic surveillance**. *Science*. 2022 Oct 7;378(6615):eabq5358. doi: 10.1126/science.abq5358. Epub 2022 Oct 7.

10. Jongo SA, Urbano Nsue Ndong Nchama V, Church LWP, Olotu A, Manock SR, Schindler T, Mtoro A, Kc N, Devinsky O, Zan E, Hamad A, Nyakarungu E, Mpina M, Deal A, Bijeri JR, Ondo Mangu ME, Ntutumu Pasialo BE, Nguema GN, Rivas MR, Chemba M, Ramadhani KK, James ER, Stabler TC, Abebe Y, Riyahi P, Saverino ES, Sax J, **Hosch S**, Tumbo A, Gondwe L, Segura JL, Falla CC, Phiri WP, Hergott DEB, García GA, Maas C, Murshedkar T, Billingsley PF, Tanner M, Ayekaba MO, Sim BKL, Daubenberger C, Richie TL, Abdulla S, Hoffman SL. **Safety and Immunogenicity of Radiation-Attenuated PfSPZ Vaccine in Equatoguinean Infants, Children, and Adults**. *Am J Trop Med Hyg*. 2023 May 9:tpmd220773. doi: 10.4269/ajtmh.22-0773.
11. Abukhattab S, **Hosch S**, Abu-Rmeileh NME, Hasan S, Vonaesch P, Crump L, Hattendorf J, Daubenberger C, Zinsstag J, Schindler T. **Whole-genome sequencing for One Health surveillance of antimicrobial resistance in conflict zones: a case study of *Salmonella* spp. and *Campylobacter* spp. in the West Bank, Palestine**. *Appl Environ Microbiol*. 2023 Sep 1:e0065823. doi: 10.1128/aem.00658-23. Epub ahead of print. PMID: 37655921.
12. Bechtold P, Wagner P, **Hosch S**, Gregorini M, Stark WJ, Gody JC, Kodia-Lenguetama ER, Pagonendji MS, Donfack OT, Phiri WP, García GA, Nsanzanbana C, Daubenberger CA, Schindler T, Vickos U. **Development and evaluation of PlasmoPod: A cartridge-based nucleic acid amplification test for rapid malaria diagnosis and surveillance**. *PLOS Glob Public Health*. 2023 Sep 27;3(9):e0001516. doi: 10.1371/journal.pgph.0001516. PMID: 37756280; PMCID: PMC10529553.