



OPEN

Genomic investigation of a dengue virus outbreak in Thiès, Senegal, in 2018

Amy Gaye^{1,7}, Tolla Ndiaye^{1,7}, Mouhamad Sy^{1,7}, Awa B. Deme¹, Alphonse B. Thiaw¹, Aita Sene¹, Cheikh Ndiaye¹, Younouss Diedhiou¹, Amadou M. Mbaye¹, Ibrahima Ndiaye¹, Christopher Tomkins-Tinch^{2,3}, Jules F. Gomis¹, Aida S. Badiane¹, Bronwyn MacInnis², Daniel J. Park², Mouhamadou Ndiaye¹, Ngayo Sy⁴, Pardis C. Sabeti^{2,3,5,6}, Katherine J. Siddle^{2,3,8} & Daouda Ndiaye^{1,8}✉

Dengue virus is a major and rapidly growing public health concern in tropic and subtropic regions across the globe. In late 2018, Senegal experienced its largest dengue virus outbreak to date, covering several regions. However, little is known about the genetic diversity of dengue virus (DENV) in Senegal. Here we report complete viral genomes from 17 previously undetected DENV cases from the city of Thiès. In total we identified 19 cases of DENV in a cohort of 198 individuals with fever collected in October and November 2018. We detected 3 co-circulating serotypes; DENV 3 was the most frequent accounting for 11/17 sequences (65%), 4 (23%) were DENV2 and 2 (12%) were DENV1. Sequences were most similar to recent sequences from West Africa, suggesting ongoing local circulation of viral populations; however, detailed inference is limited by the scarcity of available genomic data. We did not find clear associations with reported clinical signs or symptoms, highlighting the importance of testing for diagnosing febrile diseases. Overall, these findings expand the known range of DENV in Senegal, and underscore the need for better genomic characterization of DENV in West Africa.

Dengue virus is a major and growing problem across the world, with an estimated 390 million infections per year¹. Recent years have seen large outbreaks in endemic areas across Asia, Africa and the Americas as well as the appearance of the disease in new areas. Climatic and demographic changes, including urbanization and increasing temperatures, may be contributing to this emergence, creating more favourable conditions for the *Aedes* mosquitoes, the disease vector^{2,3}. The majority of dengue virus infections are either mild or asymptomatic; these can be difficult to diagnose as symptoms—commonly including fever, myalgia, arthralgia, and headaches—are usually non-specific, meaning diagnostic tests are needed to confirm infection. A small fraction of cases result in a severe disease, especially following secondary infection with a different serotype of the virus via antibody-dependent enhancement⁴.

Dengue is endemic to 34 countries across Africa, and outbreaks of this disease have been reported in all regions⁵. Dengue is caused by a virus of the Flaviviridae family and there are 4 distinct serotypes (DENV1, DENV2, DENV3 and DENV4). All four serotypes have been reported in Africa, with DENV2 being associated with the greatest number of epidemics on the continent⁵. In Senegal, dengue virus was first isolated in 1970 in the rural Bandia area⁶. *Aedes* (*Ae.*) *aegypti*, the main mosquito vector of DENV, has been isolated in the three different ecological areas in Senegal⁷, as have other potential DENV vectors⁸. All serotypes of DENV have since been detected in Senegal, with DENV3 causing an epidemic in 2009 Dakar⁹ and DENV2, often confined to the forest cycle, spilling over into urban areas in Senegal in 2014–2015⁸. Since then, recurrent urban outbreaks associated with multiple serotypes have been reported in both 2017 (138 cases) and 2018 (314 cases)^{8,10}, raising

¹Laboratoire de Parasitologie-Mycologie, Hôpital Aristide le Dantec, Université Cheikh Anta Diop de Dakar, BP 16477, Dakar, Senegal. ²Broad Institute of MIT and Harvard, Cambridge, MA, USA. ³Department of Organismal and Evolutionary Biology, Harvard University, Cambridge, MA, USA. ⁴Service de Lutte AntiParasitaire, Thiès, Senegal. ⁵Department of Immunology and Infectious Disease, Harvard School of Public Health, Boston, MA, USA. ⁶Howard Hughes Medical Institute (HHMI), Chevy Chase, MD, USA. ⁷These authors contributed equally: Amy Gaye, Tolla Ndiaye and Mouhamad Sy. ⁸These authors jointly supervised this work: Katherine J. Siddle and Daouda Ndiaye. ✉email: daouda.ndiaye@ucad.edu.sn

the possibility that environmental changes and high mobility of people has led to a more widespread distribution of dengue virus across the country.

In late 2018, Senegal experienced the largest dengue outbreak to date, with 314 confirmed cases detected in several regions of the country. Over 60% of cases were in the region of Diourbel, concentrated in the city of Touba, where in late October the Grand Magal—a 2-day religious pilgrimage that gathers over 3 million people in this holy city—took place^{11,12}. Given the risk of disease transmission at this large gathering, additional measures to enhance disease surveillance were put in place, including a mobile laboratory from the Institut Pasteur equipped with molecular tests for rapid diagnosis¹¹. However, several factors indicate that many cases were undetected: the incubation period of the disease is typically 5–7 days, dengue fever was detected elsewhere in Senegal even before the Magal, and the similarity in symptoms and seasonality of dengue fever and malaria suggest that many more cases, especially of mild disease, remained undetected. Dengue is, therefore, likely to be an underdiagnosed cause of non-malarial fever.

Genomics can provide important information regarding the origin, evolution and spread of a viral outbreak, such as dengue^{13–15}. Here we used metagenomic sequencing to characterize the genetic diversity of 17 dengue virus cases from the city of Thiès; part of a cluster of 19 previously undetected cases identified in the course of surveillance for the causes of acute febrile disease in the region, and evaluated clinical metadata to identify risk factors associated with dengue virus infection. Despite the growing burden of dengue fever in Senegal and West Africa more broadly, little is known about the genetic diversity of the virus and few viral genomes are available. This information could help explain disease transmission dynamics and enable a better understanding of dengue virus circulation and transmission in Senegal.

Results

Study participation and location. Thiès is the third largest city in Senegal and the administrative capital of the region of Thiès. It is located 72 km east of Dakar, Senegal's capital and largest city, and 123 km west of Touba, Senegal's second largest city. Thiès has a very low malaria endemicity with an annual entomological inoculation rate of 1–5¹⁶. Individuals enrolled in this study presented to the SLAP (Service de Lutte AntiParasitaire) clinic (Thiès) with either an acute fever (temperature > 38 °C) or a recent history of fever in the last 4 days. All samples were collected between October 2nd and November 12th, a period just following the rainy season in this area of Senegal and when malaria transmission is at its annual peak¹⁷. In total, 198 individuals were enrolled into the study, of whom 36 tested positive for malaria by Rapid Diagnostic Test. Individuals were aged between 2 and 72 years of age with 121 males and 77 females.

Molecular detection of dengue virus among febrile cases. Following RNA extraction, all 198 samples were tested for the presence of flavivirus RNA using an in house RT-qPCR assay with previously published primers (Fig. 1A, see Methods). 38 samples that were suggestive of viral presence were further tested using the RealStar Dengue RT-PCR Kit. This approach identified 17 positive samples (Fig. 1B) collected between October 2nd and November 12th (Fig. 1C). Coinfections of malaria and arboviruses have previously been reported in other regions of Senegal¹⁸. Here, no samples that tested positive for malaria (N = 36) also tested positive for dengue virus infection, suggesting no co-infections were present in our dataset.

Metagenomic sequencing of dengue virus genomes. Using metagenomic sequencing we assembled complete viral genomes from 15 of the 17 positive samples (1 sample was excluded due to a low concentration of input material and another did not yield sufficient reads to assemble a high quality genome). The average coverage ranged from 28× to 5370× and the genome assembly length varied between 10,525 bp and 10,680 bp (Table S1). Four samples were from serotype DENV2 and 11 samples were from serotype DENV3 (Fig. 2A). No other pathogenic virus was detected in any of the samples sequenced (Fig. S1).

We also performed metagenomic sequencing on 17 samples that tested negative by the pan-flavivirus RT-qPCR assay (Fig. 1A). In 2 of these samples we detected sufficient reads to assemble complete genomes for DENV1, indicating that these were missed by the RT-qPCR assay. Indeed, sequence comparison of these 2 newly assembled genomes identified 4 mutations in the forward primer, likely impacting the performance of this assay (Fig. S2). We confirmed this result using the RealStar Dengue RT-PCR Kit. We also re-tested all available samples using a third RT-qPCR assay for Dengue virus that was able to detect these two positive samples and did not identify any other samples that were positive for DENV1. In total, dengue virus was detected in 19/198 (9.6%) of all febrile patients and 19/162 (11.7%) of all cases of non-malarial fevers.

Phylogenetics of dengue virus in Senegal. We next investigated the phylogenetic relationships of these sequences to existing dengue virus genomes. DENV3 was the most frequent serotype in our data. All 11 genomes were highly similar and fell within an African cluster of samples in genotype III that also includes 5 sequences from Gabon and 1 from Senegal collected between 2009 and 2017 (Fig. 2B). However, the estimated TMRCA of these sequences with the most closely related sequences—from the Gabonese outbreak in 2016–17—was ~ 13 years ago (95% HPD 11.3–14.2 years; Fig. 3A), suggesting that much of the recent evolutionary history of DENV3 is absent from our tree. To increase resolution to infer the recent history DENV3 in West Africa we repeated our phylogenetic analyses using a 935nt fragment of the envelope gene. This allowed the inclusion of 15 additional sequences from Africa including Cote d'Ivoire, and Cap Verde. These sequences were all part of the African genotype III cluster, but shared more similarity with the Gabonese and older Senegal sequence than the present sequences (Fig. S3).

For DENV2, the four genomes all belonged to the Cosmopolitan genotype (II). All the new DENV2 genomes are highly similar and are closest to other available sequences from Africa, particularly those from Burkina Faso

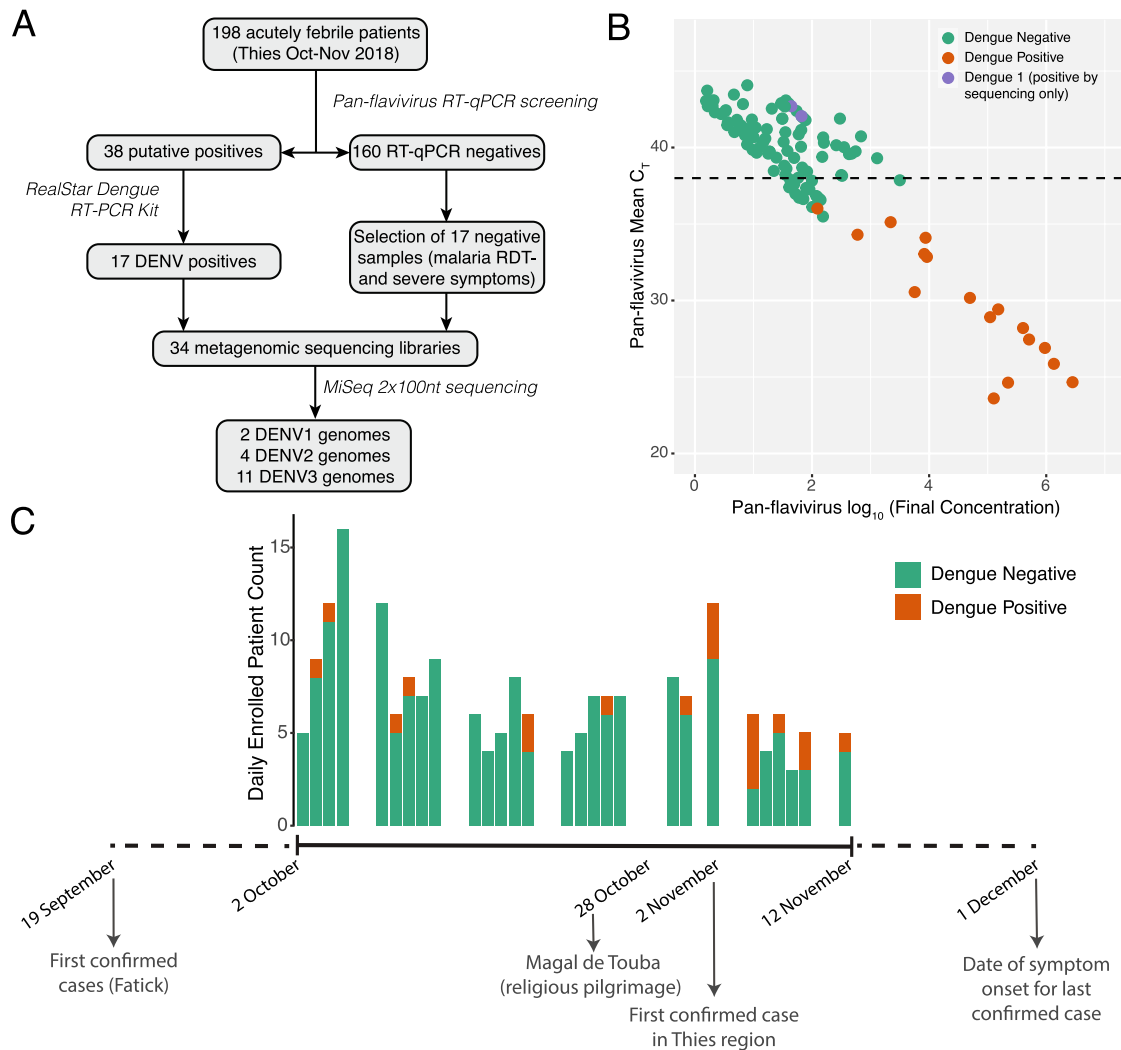


Figure 1. Detection of dengue virus in individuals with acute fever. **(A)** Flowchart of the study showing the major assays performed and numbers of samples at each step. **(B)** Plot showing the C_T and viral concentration of all 198 samples tested. Dengue-positive samples, as determined by the RealStar assay, are colored in orange, and DENV1 samples identified by sequencing only are colored in purple. The C_T cutoff used to identify putative positives for further testing is shown by a black dashed line. **(C)** Timeline of the 2018 dengue virus outbreak in Senegal showing the number of patients enrolled per day and the number of individuals that tested positive for dengue. Below the plot are marked several important events during the outbreak.

in 2016 (Fig. 2C). We estimated the time to the most recent common ancestor (TMRCA) of the Senegalese and Burkina Faso sequences to be 4 years), indicating a recent shared ancestor (Fig. 3B). It is notable that Burkina Faso is the only other West African country from which complete genomes of the Cosmopolitan genotype are available. A shorter sequence from Cote d'Ivoire in 2017 (LC310791) also clustered with these sequences in a maximum likelihood tree of only the envelope gene (not shown). Notably, these samples are not part of the sylvatic clade, unlike all previous DENV2 sequences that are available from Senegal (Fig. 2), consistent with the urban nature of this outbreak. However, further investigation of the nonsynonymous changes that separated the 4 new genomes from existing sequences identified a valine to isoleucine change at codon 58 of the NS2A protein, just upstream of the first membrane-spanning region of the protein¹⁹, which is shared exclusively with all West African sequences of the sylvatic clade.

The two DENV1 genomes detected by metagenomic sequencing were from genotype III and most similar to the only other African complete genomes in this genotype (1 from Benin and 5 from Burkina Faso; Fig. 2D) with a TMRCA of ~5 years ago (95% HPD 3.9–6.1 years; Fig. 3C). Notably, these sequences cluster distinctly from the only other DENV1 sequences from Senegal—two short fragments of the polymerase gene from the Dakar suburb of Medina Gounass collected in 2015²⁰, which cluster with Asian sequences of this genotype (Fig. S4). However, the short length of this sequence prevents more detailed phylogenetic relationships from being resolved.

Association with clinical, demographic and epidemiological characteristics. Participants with presumptive dengue fever were of all ages (8–49; mean 27.79 years) with an equal number of males and females

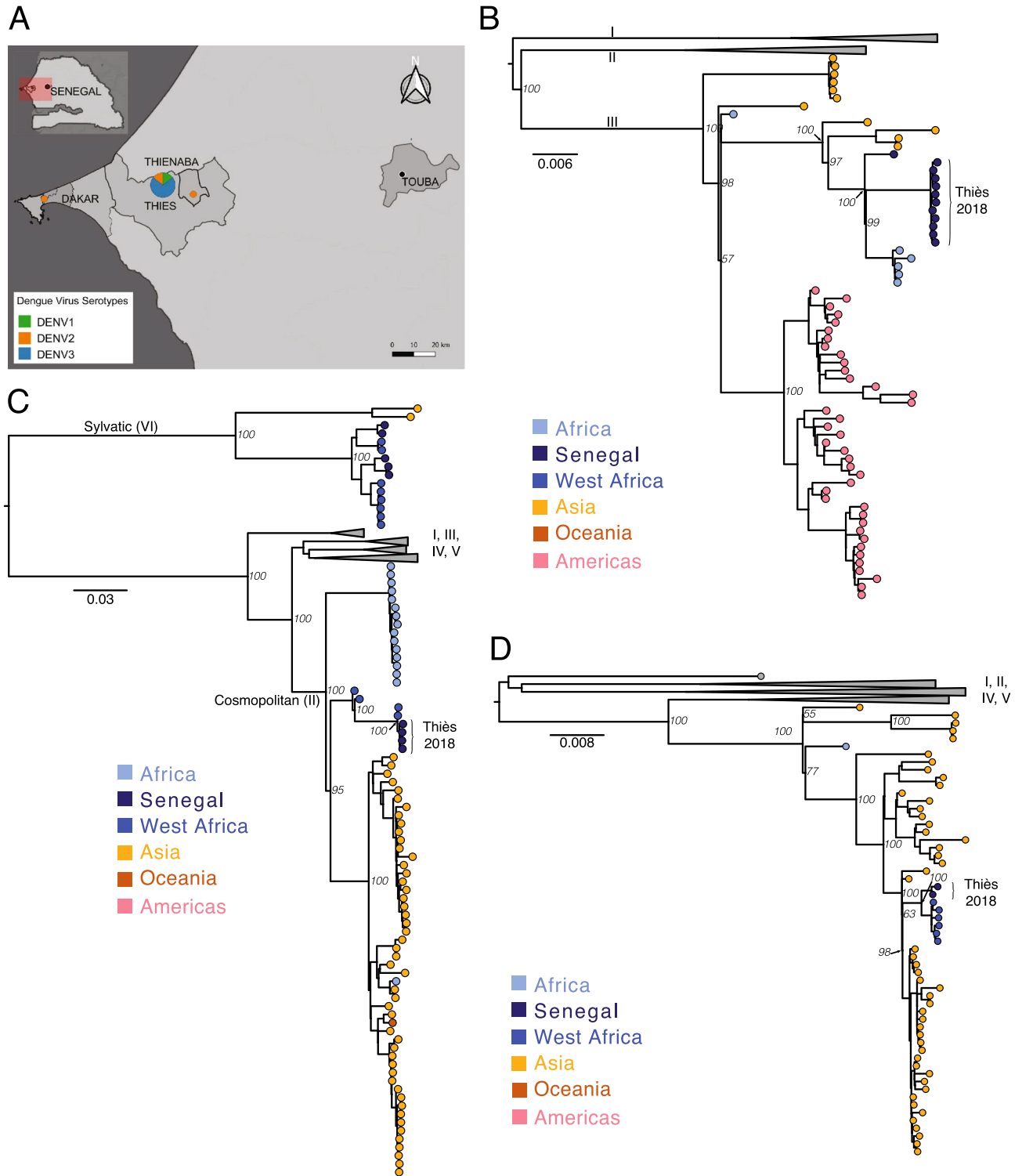


Figure 2. Phylogenetic analysis of dengue virus genomes from Senegal. (A) Map showing the locations of DENV cases identified in this study by virus serotype the size of the dot is proportional to the number of samples. Major cities, including Thiès—the sampling site—and Touba—the location of the Magal—are also shown. This map was created using QGIS v3.4.7 (<http://www.qgis.osgeo.org>). (B) DENV3 maximum likelihood (ML) phylogeny: genomes from Thiès are in genotype III and are most similar to genomes from Africa. (C) DENV2 ML phylogeny: genomes from Thiès are in the Cosmopolitan genotype and are most similar to other West African genomes. (D) DENV1 ML phylogeny: genomes from Thiès are in genotype III and are most similar to genomes from West Africa. Bootstrap values for key nodes are shown.

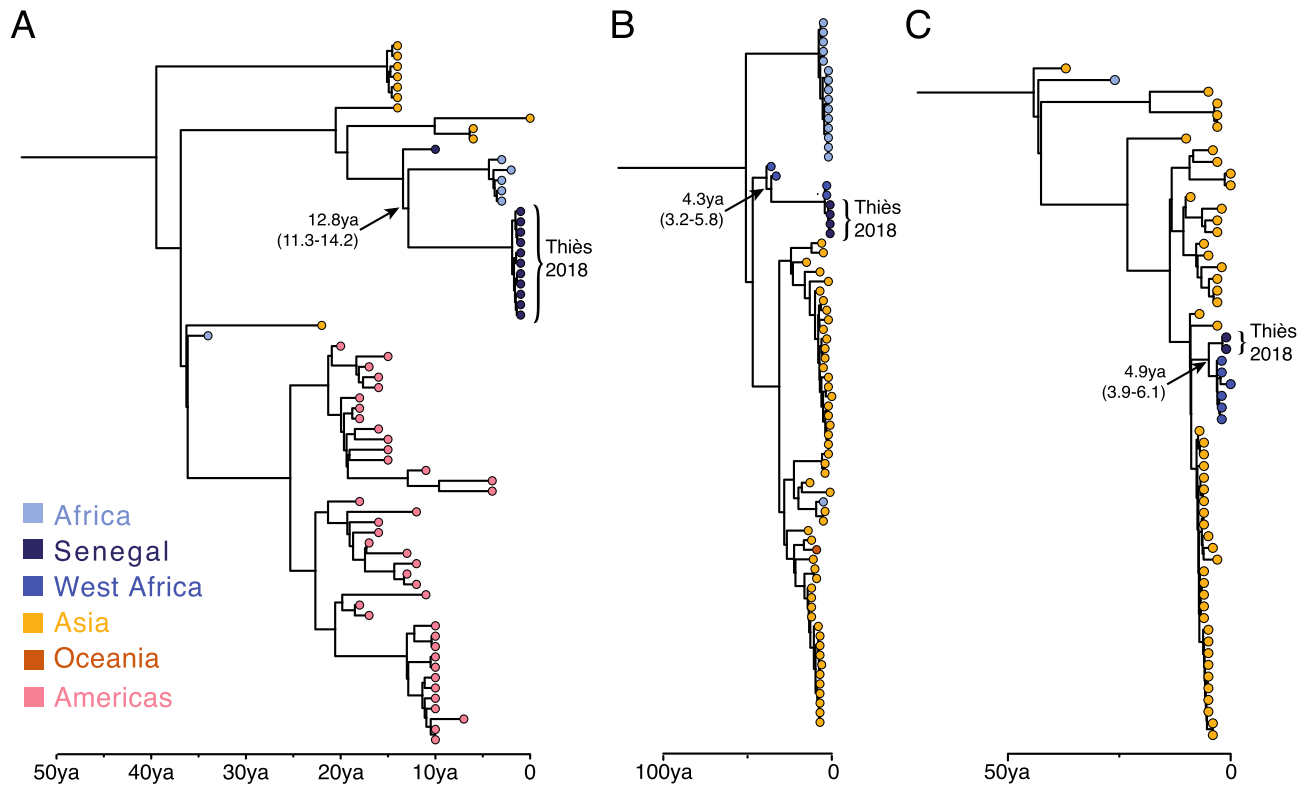


Figure 3. Maximum clade credibility trees for complete genomes of (A) DENV3 genotype III, (B) DENV2 Cosmopolitan genotype and (C) DENV1 genotype III. BEAST analysis was performed on the complete CDS from 136, 188 and 118 genomes for DENV1, DENV2 and DENV3, respectively. Estimated date and 95% posterior densities are given for nodes of interest described in the text.

(Table 1). No clinical signs or symptoms were found to be differentially reported between presumptive dengue cases and dengue-negative febrile individuals. Headaches were the most common sign reported by all participants (197/198 participants including all presumptive dengue cases). Among participants with presumptive dengue fever, 89% (17/19) reported muscle aches and 79% (15/19) reported back aches. However, neither was significantly more common among these individuals, even after excluding individuals with a known malaria diagnosis (N=36) (Table 1).

Recent travel was significantly more frequent among individuals with presumptive dengue fever (52% vs 26%, $p=0.041$, Table 1). Nine of the ten individuals with presumptive dengue who reported recent travel had traveled to Touba (the other individual had traveled to Saint Louis). However, only 6 of these were after the dates of the Magal (Fig. 1C). All 6 individuals reported onset of fever within 10 days of the end of the Magal (the earliest on October 30th, the latest on November 6th). This timing could be consistent with the dates of the Magal, however, we do not know if the individuals attended the Magal and cannot rule out that they were infected elsewhere.

Discussion

In this study we identify and describe the genomic diversity of a cluster of previously undetected dengue virus cases from the city of Thiès, Senegal. We report 17 complete DENV genomes (of serotypes 1, 2 and 3). This data enables the investigation of risk factors associated with dengue virus infection and is useful for the surveillance of repetitive outbreaks recorded in Senegal, where, as in other regions, urban outbreaks of multiple serotypes are increasingly common. These findings expand the known range of local transmission of dengue virus in late 2018 in Senegal and point to a much wider distribution of three serotypes of the virus than previously reported.

During the 2018 outbreak DENV1 and DENV2 were reported respectively from the southwest (Fatick) and north of Senegal (Saint Louis and Richard Toll); however, these findings suggest that it was more widespread. Of the 2 individuals with DENV1 infection, 1 reported travel to Touba, and the other reported no recent travel. Of the 4 individuals with DENV2 infection, 1 reported travel to Saint Louis prior to developing symptoms, 1 reported travel to Touba and the other 2 reported no travel, suggesting that DENV1 and DENV2 were circulating in the Thiès region. All four DENV2 isolates were genetically very similar and fell into the cosmopolitan (II) genotype, distinct from the 5 other genomes available from Senegal that are part of the sylvatic clade. The cosmopolitan genotype is widespread across Africa and Asia; however the present sequences are most similar to recent sequences from Burkina Faso in 2016. This suggests that the lineage of DENV2 associated with this outbreak was related to viruses that have been circulating in the West African region in recent years, however, the scarcity of sequence data prevents higher resolution on this.

Binary variable	Prevalence NEG (%)	Prevalence POS (%)	Prevalence overall (%)	P value	Odds ratio	OR 95% CI	Risk ratio	RR 95% CI	Missing (%)
Recent travel	37/138 (26)	10/19(52)	47/157(29)	0.04166822	3.03	(1.14, 8.05)	1.55	(0.95, 2.51)	5/162(3)
Back ache	76/143 (53)	15/19(78)	91/162 (56)	0.05963917	331	(1.05, 10.45)	2.23	(0.92, 5.41)	0/162(0)
Dizziness	59/143 (41)	12/19 (63)	71/162 (43)	0.11841943	2.44	(0.91, 6.57)	1.59	(0.87, 2.92)	0/162(0)
Tiredness/weakness	85/143 (59)	15/19 (78)	100/162(61)	0.16380211	2.56	(0.81, 8.10)	1.93	(0.79, 4.71)	0/162(0)
Muscle aches	103/143 (72)	17/19 (89)	120/162 (74)	0.17646169	3.3	(0.73, 14.94)	2.66	(0.70, 10.12)	0/162 (0)
Cough/sore throat	30/143 (20)	1/19 (5)	31/162(19)	0.18491174	0.21	(0.03, 1.63)	0.83	(0.73, 0.96)	0/162(0)
Headache	142/143 (99)	19/19 (100)	161/162 (99)	0.23281451	100	(0.00, 100.00)	0.27	(0.01, 7.86)	0/162(0)
Bleeding	1/143 (0)	0/19(0)	1/162 (0)	0.23281451	0	(0.00, 100.00)	1.02	(0.95, 1.10)	0/162(0)
Chest pain	59/143 (41)	11/19 (57)	70/162(43)	0.25895297	1.96	(0.74, 5.16)	1.4	(0.81, 2.41)	0/162(0)
Inflammation	4/143 (2)	2/19(10)	6/162(3)	0.30320234	4.09	(0.70, 24.01)	1.09	(0.93, 1.27)	0/162 (0)
Vomiting	42/143 (29)	8/19 (42)	50/162 (30)	0.38720446	1.75	(0.66, 4.66)	1.22	(0.82, 1.82)	0/162(0)
Difficulty breathing	25/143 (17)	5/19 (26)	30/162(18)	0.53725623	1.69	(0.56, 5.11)	1.12	(0.85, 1.48)	0/162 (0)
Rash	9/143 (6)	0/19(0)	9/162 (5)	0.55369857	0	(0.00, 100.00)	0.96	(0.88, 1.05)	0/162(0)
Sex	61/143 (42)	9/19 (47)	70/162 (43)	0.88628197	1.21	(0.46, 3.16)	1.09	(0.70, 1.71)	0/162 (0)
Continuous variable	Mean (95% CI)	Mean POS (95% CI)	Pearson's R	P value	Odds ratio	OR 95% CI	Risk ratio	RR 95% CI	Missing (%)
Age	21.71 (0.00, 51.18)	27.79 (5.81, 49.77)	0.13	0.09147119	1.638	(0.92, 2.93)	0.59	(0.44, 0.79)	0/162 (0)
Temperature	38.33 (36.60, 40.06)	38.62 (37.06, 40.18)	0.11	0.17980714	1.563	(0.81, 3.01)	0.66	(0.50, 0.88)	0/162 (0)
Red blood cell count	12.52 (8.76, 16.28)	13.06 (10.41, 15.71)	0.09	0.23461531	1.41	(0.81, 2.46)	0.88	(0.58, 1.34)	2/162 (1)
Heart rate	92.63 (51.49, 133.77)	97.53 (53.69, 141.36)	0.08	0.34641639	1.295	(0.76, 2.21)	1.06	(0.64, 1.76)	8/162 (4)
Respiratory rate	32.42 (19.73, 45.10)	31.47(19.51, 43.44)	-0.05	0.55091079	0.89	(0.61, 1.30)	1.13	(0.75, 1.68)	11/162 (6)
Blood sugar	1.66 (0.00, 15.00)	1.04 (0.59, 1.50)	-0.03	0.69625529	0.741	(0.06, 8.82)	0.96	(0.61, 1.51)	2/162 (1)

Table 1. Association values for clinical and demographic variables. Top: binary variables and bottom: continuous variables.

The majority of cases in this study (65%) were from serotype 3. This serotype has become increasingly common in West and Central Africa in the last decade and has been responsible for several outbreaks in Senegal⁹, Côte d'Ivoire⁵, Gabon²¹ and elsewhere. Our phylogenetic analysis suggests that this is a closely related but distinct lineage to that responsible for the 2009 outbreak in Senegal, as the samples cluster more closely to those from a 2016–17 outbreak in Gabon²¹ than the single sequence from Senegal in 2009, sequenced from a traveler returning to Germany²². However, we lack genomic data from this region; only 7 complete DENV3 genomes from Africa were available in GenBank prior to this study, in addition to 24 short sequences of 220–1923nt. This scarcity of data makes phylogenetic inference, especially around the recent evolutionary history of DENV3 in West Africa, challenging.

In Senegal, the DENV3 serotype was most frequent in Touba (Diourbel) during the 2018 outbreak and the timing of the outbreak overlapped with the Grand Magal religious pilgrimage. 45% of DENV3 positive patients reported travel to Touba (N = 5), of which 3 developed a fever in a time frame consistent with attendance at the Magal. It should be noted that we do not know when participants travelled to Touba or whether they attended the Magal. The other 6 individuals reported no recent travel. Even if we postulate that all individuals with recent travel were infected outside of Thiès, almost half of cases are consistent with locally-acquired infection.

Despite using a pan-flavivirus RT-qPCR for screening of all samples prior to sequencing, the two DENV1 infections were missed, owing to 4 mismatches in the forward primer binding region that compromised the sensitivity of the test. This finding demonstrates the utility of genomic surveillance, especially with metagenomic sequencing²³ for finding genetically distinct lineages that may be missed by PCR-based approaches. Where knowledge of the prevalence of different pathogens is scarce, this unbiased approach can serve to identify the species of greatest concern, providing recommendations on the priorities for implementing more rapid, low-cost, targeted diagnostic testing, and characterize the genetic diversity of circulating lineages to ensure that these tests are locally suited and sensitive. However, to enable this genomics-informed surveillance and diagnostics more widespread and ongoing sequencing is needed to provide high-resolution information on the extent of genetic variation. Developing local capacity for this work in endemic countries is central to this effort and to ensuring pathogen genomics can be translated into actionable insights.

We found very little association between the clinical and demographic variables investigated and presumptive dengue virus infection. Individuals with presumptive dengue were equally distributed with regards to sex and age. Associations were stronger after removing patients with a confirmed malaria diagnosis, but still did not reach statistical significance. These findings are consistent with the mild symptoms of these individuals and the broader difficulty of clinically differentiating dengue fever from other febrile diseases without diagnostic testing. Although dengue fever does not have a specific therapy and the majority of infections are mild, diagnosis is still important to rule out other causes that would have different courses of treatment. The availability of rapid, sensitive and widely available diagnostic tests for common pathogens, other than malaria, is still severely limited in Senegal and many other regions and presents a major challenge for public health.

In conclusion, we report the analysis of 17 complete genomes from 19 previously undiagnosed dengue cases from the 2018 outbreak in Senegal. This work expands our understanding of the source and spread of this outbreak but also of dengue virus genetic diversity in West Africa more broadly. Our findings highlight the need for enhanced diagnosis and genomic surveillance to better understand the past and track the future course of this disease.

Methods

Sample collection. Patient samples were obtained through a study approved by the Institutional Review Boards at the Senegalese Ministry of Health and Harvard University. Patient samples were obtained from individuals presenting to the Service de Lutte AntiParasitaire in Thiès, Senegal. Study staff obtained written, informed consent from all participants enrolled in the study (or a patient/legal guardian for participants age less than 18 years) following an explanation of the study objectives. All methods were performed in accordance with the relevant guidelines and regulations. A 5 mL venous blood draw was collected from each enrolled patient in a BD Vacutainer EDTA tube. Demographic information as well as clinical signs and symptoms were also recorded. Blood samples were kept on ice and transported daily to the Laboratoire Parasitologie et Mycologie at Aristide Le Dantec Hospital, where plasma was isolated on the same day and frozen prior to RNA extraction.

RNA extraction and PCR. Total RNA was extracted from patient plasma using the QiAmp viral RNA mini kit (Qiagen) according to the manufacturer's instructions. Detection of Flavivirus RNA was performed by RT-qPCR using the Applied Biosystems Power SYBR RNA-to-CT 1-Step Kit on a Roche LightCycler 96 for 45 cycles. Primer sequences have been shown to capture all known flavivirus species²⁴. Qualitative detection of Dengue virus RNA in samples with suspected flavivirus infections was performed using the RealStar Dengue RT-PCR Kit 3.0 from Altona-Diagnostics on an Applied Biosystems 7500 Fast Real-Time PCR System according to manufacturer's instructions.

Sequencing and data analysis. Sequencing libraries were prepared from all dengue virus positive samples (N = 17). We additionally sequenced 17 samples with no known diagnosis, selected as follows; negative for both the pan-flavivirus RT-qPCR and RealStar Dengue RT-PCR, and negative for malaria by Rapid Diagnostic Test.

Briefly, DNA was removed by Turbo DNase treatment (Thermo Fisher Scientific), cDNA was synthesized with random primers and sequencing libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina) as previously described²⁵. All samples were sequenced on an Illumina MiSeq instrument at the Laboratoire Parasitologie et Mycologie in Dakar, Senegal, with paired end 100nt sequencing.

Sequencing data was analysed using viral-ngs v1.25.0²⁶, implemented on the DNAnexus cloud platform. First, samples were demultiplexed and an initial taxonomic classification of metagenomic sequencing reads was performed using KrakenUniq²⁷. Reads were compared against a database that encompassed the known diversity of all viruses that infect humans. Next, sequences were filtered by removing reads aligning to the human genome, synthetic spike-in sequences and other known contaminants. The remaining reads were filtered against a database of dengue virus sequences, and de novo assembly was performed with Trinity²⁸. Contigs were then scaffolded against one of four RefSeq genomes, representing the four serotypes (NC_001477.1, NC_001474.2, NC_001475.2, NC_002640.1). Short read sequence data and assembled genomes have been deposited in NCBI's SRA and Genbank databases under BioProject PRJNA662334 and accession numbers MW288024–MW288040.

Code availability. Computational workflows used to analyze viral sequencing data are have been made publicly available through dockstore: <https://dockstore.org/organizations/BroadInstitute/collections/pgs>.

Phylogenetic tree construction. Genotype Detective (https://www.genomedetective.com/app/typin_tool/dengue/) was used to determine the genotype and serotype of the newly assembled dengue virus genomes, as well as the genotypes of all whole dengue virus genomes (> 10,000 nt) available on GenBank on February 10th 2020 (DENV2 and DENV3) or June 24th 2020 (DENV1). We further filtered genome records to remove those where collection date or country of origin were not recorded. Due to the large number of non-African dengue genomes, a subset of these sequences was selected for phylogenetic analysis. Specifically, for each serotype, we randomly selected 50 published genomes from the same genotype as the current samples as well as up to 25 genomes from each other genotype for context. All complete sequences from Africa were also included (DENV1; N = 10, DENV2; N = 33, DENV3; N = 7) yielding a final dataset of 136 genomes for DENV1, 188 genomes for DENV2 and 118 for DENV3.

Sequences were aligned using MAFFT v7.388 implemented in Geneious and trimmed with Trimal. Maximum likelihood phylogenies were estimated separately for each serotype using IQTREE v1.5.5²⁹, using a (GTR) + Gamma substitution model and ultrafast bootstrapping (N = 1000)³⁰. This model was selected as it is the most general and makes few assumptions about substitution rates³¹, and was previously used in phylogenetic analyses of DENV outbreaks¹⁵. Time-scaled Bayesian phylogenetic analysis was performed using a Markov chain Monte Carlo (MCMC) algorithm implemented in the BEAST v1.10.4³² on the CIPRES Science Gateway³³. Briefly, BEAST was run for 50 million MCMC steps using a GTR + Gamma substitution model with 1,2,3 codon partitioning, an uncorrelated relaxed clock with log-normal distribution and a Skyline tree prior (See Supplemental files 1–3 for complete XML files used to perform BEAST analysis for DENV1, DENV2, and DENV3, respectively). For each serotype we used the full coding sequence and sampled tip dates. Trees were sampled every 5,000 generations. The convergence was assessed using Tracer v1.7.1 and TreeAnnotator v1.10.4 was used to generate maximum clade credibility trees with a burn-in of 10%. All trees were viewed in FigTree v1.4.4³⁴.

Demographic and clinical data analysis. A univariate correlation analysis was performed on 14 categorical and 6 continuous variables to determine the significance of association with a presumptive diagnosis of dengue fever (determined by the presence of dengue virus RNA in the plasma detectable by RT-qPCR or sequencing). Analyses were performed as previously described³⁵. Briefly, the biserial correlation coefficient and the two-tailed Fisher Exact test were used to calculate the significance of continuous and categorical variables, respectively at a p-value cut-off of 0.05. Categorical variables tested were; sex, recent travel and the presence of 12 clinical signs or symptoms reported by at least 1 participant; muscle aches, dizziness, vomiting, weakness, headache, backache, chest pain, difficulty breathing, rash, sore throat, inflammation and bleeding. Continuous variables tested were; age, temperature, heart rate, respiratory rate, hemoglobin level and blood sugar level.

Data availability

Short read sequence data and assembled genomes have been deposited in NCBI's SRA and Genbank databases under BioProject PRJNA662334 and accession numbers MW288024–MW288040.

Received: 19 November 2020; Accepted: 16 March 2021

Published online: 14 May 2021

References

- Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
- Murray, N. E. A., Quam, M. B. & Wilder-Smith, A. Epidemiology of dengue: Past, present and future prospects. *Clin. Epidemiol.* **5**, 299–309 (2013).
- Gubler, D. J. Dengue, urbanization and globalization: The unholy trinity of the 21(st) century. *Trop. Med. Health* **39**, 3–11 (2011).
- Katzelnick, L. C. *et al.* Antibody-dependent enhancement of severe dengue disease in humans. *Science* **358**, 929–932 (2017).
- Amarasinghe, A., Kuritsk, J. N., Letson, G. W. & Margolis, H. S. Dengue virus infection in Africa. *Emerg. Infect. Dis.* **17**, 1349–1354 (2011).
- Robin, Y., Cornet, M., Heme, G. & Le Gonidec, G. Isolement du virus de la dengue au Sénégal. *Annales de l'Institut Pasteur/Virologie* **131**, 149–154 (1980).
- Huber, K. *et al.* Aedes aegypti in Senegal: genetic diversity and genetic structure of domestic and sylvatic populations. *Am. J. Trop. Med. Hyg.* **79**, 218–229 (2008).
- Gaye, A. *et al.* Potential for sylvatic and urban Aedes mosquitoes from Senegal to transmit the new emerging dengue serotypes 1, 3 and 4 in West Africa. *PLoS Negl. Trop. Dis.* **13**, e0007043 (2019).
- Faye, O. *et al.* Urban epidemic of dengue virus serotype 3 infection, Senegal, 2009. *Emerg. Infect. Dis.* **20**, 456–459 (2014).
- WHO. Outbreaks and Emergencies Bulletin, Week 52: 22–28 December 2018. WHO | Regional Office for Africa <https://www.afro.who.int/publications/outbreaks-and-emergencies-bulletin-week-52-22-28-december-2018-0> (2018).
- Diagne, C. T., Barry, M. A., Ba, Y., Faye, O. & Sall, A. A. Dengue epidemic in Touba, Senegal: Implications for the Grand Magal Pilgrimage for travellers. *J. Travel Med.* **26**, (2019).
- Sokhna, C., Goumballa, N., Diaw & Gautret, P. The Grand Magal of Touba in the time of a dengue outbreak in Senegal. *Travel Med. Infect. Dis.* **28**, 107–108 (2019).
- Salje, H. *et al.* Dengue diversity across spatial and temporal scales: Local structure and the effect of host population size. *Science* **355**, 1302–1306 (2017).
- Faria, N. R. *et al.* Genomic and epidemiological characterisation of a dengue virus outbreak among blood donors in Brazil. *Sci. Rep.* **7**, 15216 (2017).
- Hill, S. C. *et al.* Early genomic detection of cosmopolitan genotype of dengue virus serotype 2, Angola, 2018. *Emerg. Infect. Dis. J.* **25**, 784 (2019).
- Bei, A. K. *et al.* Dramatic changes in Malaria population genetic complexity in Dielmo and Ndiop, Senegal, revealed using genomic surveillance. *J. Infect. Dis.* **217**, 622–627 (2018).
- Thwing, J. *et al.* Declines in Malaria burden and all-cause child mortality following increases in control interventions in Senegal, 2005–2010. *Am. J. Trop. Med. Hyg.* **97**, 89–98 (2017).
- Sow, A. *et al.* Concurrent malaria and arbovirus infections in Kedougou, southeastern Senegal. *Malar. J.* **15**, 47 (2016).
- Xie, X., Gayen, S., Kang, C., Yuan, Z. & Shi, P.-Y. Membrane topology and function of dengue virus NS2A protein. *J. Virol.* **87**, 4609–4622 (2013).
- Dieng, I. *et al.* Mobile Laboratory reveals the circulation of dengue virus serotype I of Asian origin in Medina Gounass (Guediawaye), Senegal. *Diagnostics (Basel)* **10**, (2020).
- Abe, H. *et al.* Re-emergence of dengue virus serotype 3 infections in Gabon in 2016–2017, and evidence for the risk of repeated dengue virus infections. *Int. J. Infect. Dis.* **91**, 129–136 (2020).
- Shihada, S. *et al.* Genetic diversity and new lineages of dengue virus serotypes 3 and 4 in returning travelers, Germany, 2006–2015. *Emerg. Infect. Dis.* **23**, 272–275 (2017).
- Grubaugh, N. D. *et al.* Tracking virus outbreaks in the twenty-first century. *Nat. Microbiol.* **4**, 10–19 (2019).
- Vina-Rodriguez, A. *et al.* A novel pan-flavivirus detection and identification assay based on RT-qPCR and microarray. *Biomed. Res. Int.* **2017**, 4248756 (2017).
- Matranga, C. B. *et al.* Unbiased deep sequencing of RNA viruses from clinical samples. *J. Vis. Exp.* <https://doi.org/10.3791/54117> (2016).
- Park, D. *et al.* *broadinstitute/viral-ngs: v1.25.0.* (2019). <https://doi.org/10.5281/zenodo.3509008>.
- Breitwieser, F. P., Baker, D. N. & Salzberg, S. L. KrakenUniq: Confident and fast metagenomics classification using unique k-mer counts. *Genome Biol.* **19**, 198 (2018).
- Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
- Wohl, S., Schaffner, S. F. & Sabeti, P. C. Genomic analysis of viral outbreaks. *Annu. Rev. Virol.* **3**, 173–195 (2016).
- Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
- Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES science gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (2010).
- Rambaut, A. FigTree v1.4.4. <http://tree.bio.ed.ac.uk/software/figtree/> (2018).

35. Okokhere, P. *et al.* Clinical and laboratory predictors of Lassa fever outcome in a dedicated treatment facility in Nigeria: a retrospective, observational cohort study. *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(18\)30121-X](https://doi.org/10.1016/S1473-3099(18)30121-X) (2018).

Acknowledgements

We thank the participants of this study and the staff of the Service de Lutte AntiParasitaire for their contributions to this work. We thank Andres Colubri for sharing code for the bivariate analyses, and Chris Edwards and Monica Schreiber for review of the manuscript. This work is supported by grants from the National Institute of Allergy and Infectious Diseases; NIH-H3Africa (U54HG007480) and the West African Emerging Infectious Diseases Research Center (U01 AI151812). K.J.S. is supported by the ASTMH Centennial Award. P.C.S. is an investigator supported by the Howard Hughes Medical Institute (HHMI).

Author contributions

A.B.D., A.S.B., B.L.M., D.J.P., M.N., N.S., P.C.S., K.J.S., D.N. conceived and designed the study; A.G., T.N., M.S., A.T., A.S., C.N., Y.D., A.M., I.N. acquired the data; A.G., T.N., M.S., A.B.D., J.G., K.J.S. analyzed the data; D.P., C.T-T. created software used in this work; A.G., T.N., M.S., A.B.D., K.J.S. drafted and revised the manuscript with input from all authors. All authors approved the manuscript before submission.

Competing interests

P.C.S. is a co-founder and shareholder of Sherlock Biosciences, and a Board member and shareholder of Danaher Corporation. The other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-89070-1>.

Correspondence and requests for materials should be addressed to D.N.

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

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



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

© The Author(s) 2021