Genomic characterization of Dengue virus circulation in Ethiopia.

Adugna Abera\textsuperscript{1,2†}, Houriiyah Tegally\textsuperscript{3,4}*\textsuperscript{†}, Geremew Tasew\textsuperscript{1}, Eduan Wilkinson\textsuperscript{3}, Abraham Ali\textsuperscript{1}, Feyisa Regasa\textsuperscript{1}, Molalegne Bitew\textsuperscript{5}, Lucious Chabuka\textsuperscript{3}, Gaspary Mwanyika\textsuperscript{3}, Derek Tshiabuila\textsuperscript{3}, Jennifer Giandhari\textsuperscript{4}, Sureshnee Pillay\textsuperscript{4}, Jenicca Poogavanan\textsuperscript{3}, Monika Moir\textsuperscript{3}, CLIMADE Consortium\textsuperscript{#}, Moritz U.G. Kraemer\textsuperscript{7,8}, Kamran Khan\textsuperscript{9,10}, Carmen Huber\textsuperscript{9}, Getachew Tollera\textsuperscript{1}, Tobias F. Rinke de Wit\textsuperscript{6}, Cheryl Baxter\textsuperscript{3}, Richard Lessells\textsuperscript{4}, Dawit Wolday\textsuperscript{1,11}, Dereje Beyene\textsuperscript{2$}, Tulio de Oliveira\textsuperscript{3,4$}

Authors affiliations:

1 Ethiopian Public Health Institute, Addis Ababa, Ethiopia
2 Department of Microbial, Cellular, and Molecular Biology, College of Natural and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia
3 Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, South Africa
4 KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa
5 Bio and Emerging Technology Institute, Addis Ababa, Ethiopia
6 Amsterdam Institute for Global Health and Development (AIGHD), Department of Global Health, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
7 Department of Biology, University of Oxford, Oxford, UK
8 Pandemic Sciences Institute, University of Oxford, Oxford, UK
9 BlueDot, Toronto, ON, Canada
10 Department of Medicine, Division of Infectious Diseases, University of Toronto, Toronto, ON, Canada
11 Department of Biochemistry and Biomedical Sciences, Michael G. DeGroote Institute for Infectious Diseases Research (IIDR), and McMaster Immunology Research Center (MIRC), Health Sciences, McMaster University, Hamilton, Ontario, Canada

†These authors contributed equally to this work.
$ Co-senior authors
# CLIMADE consortium author list and affiliations in supplementary material

Corresponding Author Contact Information

*Authors to whom correspondence should be addressed.
Adugna Abera, Department of Microbial, Cellular, and Molecular Biology, College of Natural and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia and Ethiopian Public Health Institute, Arbagnoch Street, Addis Ababa, Ethiopia; email: adugnabe@yahoo.com.

Houriiyah Tegally, Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, South Africa; email: houriiyah@sun.ac.za.
ABSTRACT

In Ethiopia, dengue virus (DENV) infections have been reported in several regions, however, little is known about the circulating genetic diversity. Here, we conducted clinical surveillance for DENV during the 2023 nationwide outbreak and sequenced DENV whole genomes for the first time in Ethiopia. We enrolled patients at three sentinel hospital sites. Using RT-PCR, we screened serum samples for three arboviruses followed by serotyping and sequencing for DENV-positive samples (10.4% of samples). We detected two DENV serotypes (DENV1 and DENV3). Phylogenetic analysis identified one transmission cluster of DENV1 (genotype III major lineage A), and two clusters of DENV3 (genotype III major lineage B). The first showed close evolutionary relationship to the 2023 Italian outbreak and the second cluster to Indian isolates. Co-circulation of DENV1 and DENV3 in some regions of Ethiopia highlights the potential for severe dengue. Intensified surveillance and coordinated public health response are needed to address the threat of severe dengue outbreaks.

Key words: Serotypes, Genotypes, Phylogenetic analysis, Seasonality, Human mobility, Dengue, Ethiopia.
Introduction

Dengue virus (DENV) is primarily transmitted by *Aedes* mosquitoes and causes significant epidemics in tropical and subtropical regions. Since the beginning of 2023, over 5 million cases and 5,000 dengue-related deaths have been reported in over 80 countries (1). In 2023, the WHO African region reported 171,991 suspected cases, 70,223 of which were confirmed cases with 753 fatalities. Outbreaks have been reported in 15 African countries: Burkina Faso has accounted for 85% of cases, followed by Ethiopia (8.2%), Mali (2.5%), and Côte d'Ivoire (2.2%) (2).

Ethiopia has been facing several outbreaks of dengue fever since 2013. The first recorded outbreak was reported in Dire Dawa City, which affected 9,441 individuals (3). A second outbreak occurred in Somale Godey Town and the Afar Region in 2014 and 2015, respectively (4). Since 2015, there has been an increase in the number of severe febrile illness cases in Godey, Dire Dawa town, and the Afar area, with no apparent cause. The first outbreak started in April 2023 in the Afar region, followed by Dire Dawa and the neighboring regions of Amhara and Oromia. There have been 27,577 reported cases and 21 deaths in 12 districts in the eastern part of Ethiopia (1,5).

DENV is a single-stranded positive-sense RNA virus, and its genome encodes three structural and seven nonstructural proteins. There are four serotypes of DENV, namely DENV1, DENV2, DENV3, and DENV4 (8, 9). When an individual is infected with one serotype, they develop immunity to that specific serotype for a long time. However, they only develop partial immunity to the other serotypes. In case of subsequent infections with a different serotype, the risk of severe dengue increases. Each serotype of the virus is further divided into several genotypes, which are categorized based on the envelope gene (E gene). However, different genotypes of the same serotype may vary in their ability to infect host cells and cause severe forms of the disease. Based on the genetic characterization, five genotypes have been defined for DENV1, six for DENV2, five for DENV3, and four for DENV4 (10-13). Recently, a nomenclature system has been proposed to further subdivide genotypes into major and minor lineages to aid global monitoring efforts (14).

Studies have shown that during dengue outbreaks, the emergence of new serotypes or changes in circulating genotypes of the virus in a particular region can lead to more severe outbreaks (15,16). Given that there are currently no approved medical countermeasures for treating severe dengue, and vaccines are currently not available to most countries, it is crucial to continuously monitor the
circulating dengue genetic diversity in endemic areas. This information will be instrumental in
developing and evaluating vaccines and treatments (17), and in responding effectively to dengue
outbreaks.

Although there have been multiple outbreaks and geographical expansion of DENV in Ethiopia,
there is currently no documented characterization of the circulating diversity of DENV, and no
publicly available whole genome sequences. In this study, we address this gap by identifying
circulating DENV serotypes and genotypes among individuals with febrile illness in three hospital-
based sentinel sites in Ethiopia and conduct phylogenetic analysis to determine spatial and
temporal patterns of dengue transmission in Ethiopia.

Materials and methods

Patient characteristics and study settings

We conducted a cross-sectional hospital-based study in Afar, Gambela, and Dire Dawa city
administration (Supplementary Figure 1). Between December 2022 and November 2023, we
enrolled outpatients and inpatients presenting with fever (>37.5°C) at three major public hospitals
in Dubti (Afar region), Dilchora (Dire Dawa), and Gambela. These hospitals serve as sentinel sites
for acute febrile illnesses in the country and were selected due to their previous history of arbovirus
outbreaks and serological evidence for arboviruses (5, 18).

Recruitment of participants

We obtained informed consent from all participants aged 18 years and older. For individuals aged
1 to 17 years, we obtained informed consent from a parent or guardian and, where appropriate,
assent from the child.

Eligibility/exclusion criteria

The inclusion criteria for the study were: a fever of more than 37.5 °C, following the 2009 WHO
Dengue classification (19), joint pain, rash, headache, myalgia, retro-ocular discomfort, stomach
pain, and hemorrhagic presentation.

Patients with severe, and established chronic clinical illnesses such as people living with HIV,
malignancies, and known metabolic disorders were excluded from the study.
Sample collection and processing

Each participant provided five milliliters of whole blood, which was collected into serum separator tubes. At the Afar site, blood sample collection took place between April 25, 2023, and May 15, 2023. The serum was isolated and collected in tubes containing 1 ml DNA/RNA shield and stored at -20 °C during field collections. It was then transported in a cold chain to the laboratory of the Ethiopian Public Health Institute (EPHI) and stored at -20 °C until analysis for PCR and sequencing. Blood samples from Dire Dawa and Gambela sites were collected between June 2023 and October 2023, stored at -20 °C, and then shipped in a cold chain to the laboratory of the EPHI, where they were stored at -80 °C until analysis.

RNA extraction and PCR amplification

We performed RNA extraction within five days of receiving the serum samples at the EPHI laboratory from each site. The RNA extraction was carried out using a Bioer NPA-32P instrument (Bior Technology, Zhejiang, China), which is a high-throughput automated workstation. We used the MagaBio Plus RNA Purification (Bior Technology, Zhejiang, China) for the extraction process, following the manufacturer's instructions. Briefly, we prepared 96-well plates and added 300 µl of thawed and briefly vortexed serum samples to columns #1 and #7. The plate was then placed into the instrument, which agitated the plates every 10 seconds for 9 minutes until the extraction was completed, resulting in 70 µl RNA extracts.

The remaining serum samples from the Afar region were stored at -20°C for later whole genome sequencing. Meanwhile, samples from the Dire Dawa and Gambela sites were stored at -80°C. In February 2024, some of the qualifying serum samples were sent to the CERI and KRISP laboratories in South Africa for whole genome sequencing.

On the same day that we extracted the RNA, we used the CDC Trioplex real-time RT-PCR assay to screen for arboviral infections such as DENV, CHIKV, and ZIKV. This assay includes a set of published oligonucleotide primers and dual-labeled hydrolysis Taqman probes. To run the test, we combined 10 µl of sample RNA with 12.5 µl of PCR master mix reaction buffer (SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase, ThermoFischer Scientific), 1 µM virus-specific primers, 0.3 µM dengue-specific probe, 0.15 µM chikungunya-specific probe,
0.15 µM Zika virus-specific probe, and nuclease-free water in a 96-well optical PCR plate to make a final reaction volume of 25 µl. We then carried out reverse transcription (RT) for 30 minutes at 50°C, followed by RT inactivation for 2 minutes at 95°C, fluorescence detection for 15 seconds at 95°C, and annealing for 1 minute at 60°C. Additionally, the assay includes an internal control reaction targeting human endogenous ribonuclease P (RP) to ensure the extracted test specimen contains amplifiable RNA. The fluorescent signal intensity was captured using a real-time PCR instrument called QuantStudio5 Real-Time PCR system (ThermoFisher Scientific).

Dengue serotyping

Following PCR screening, samples positive for DENV were serotyped from the same RNA extract using CDC DENV 1-4 rRT-PCR Multiplex assay on the QuantStudio5 Real-Time PCR system. The assay employed DENV 1-4 specific primers and probes provided by the CDC, USA. To perform the assay, we mixed 5 µl of extracted RNA templates with 12.5 µl of the master mix (SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase, ThermoFischer Scientific). The assay was run in a multiplex format, with each serotype-specific probe labeled with different probes as per the manufacturer’s protocol.

Dengue sequencing

After serotyping, we selected specimens with PCR screening Ct values below 26 for sequencing. Sequencing was carried out with technical support from CERI and KRISP. The primer scheme for DENV sequencing and other arboviruses is available on the CLIMADE GitHub (https://github.com/CERI-KRISP/CLIMADE/tree/master/Protocols/Arboviruses). The COVIDSeq protocol was adapted to perform the library preparation, followed by sequencing on the Illumina Miseq platform. This protocol is available on the CLIMADE GitHub website, and it is described in detail by Elaine Vieira Santos, Debora Glenda Lima de La Roque in protocols.io (https://protocols.io/view/pathogen-whole-genome-sequencing-multiplexed-ampli-cgwbtzan).

The libraries were prepared using the Illumina COVIDSeq protocol (Illumina Inc, USA). Briefly, RNA was reverse transcribed to cDNA using random hexamers. The Dengue genome was amplified using the two pools of primers specific for the Dengue serotype 1-4. PCR amplicons were tagmented using the EBLTS (Enrichment BLT), which is a process that fragments and tags the PCR amplicons with adapter sequences. Adaptor ligated amplicons were further amplified.
using the unique 10 base pair Index 1 (i7) adapters and Index 2 (i5) adapters (IDT for Illumina-PCR Indexes Set 1) for each sample. The pooled amplicon library was quantified using a Qubit 2.0 fluorometer (Invitrogen, Inc.) and diluted to a 4 nM. The 4nM library was denatured and diluted to a final loading concentration of 12pM. Dual indexed paired-end sequencing was performed on the Illumina Miseq using a v3 600 cycle flow cell.

Bioinformatics analysis

Following base calling and demultiplexing of the sequence run, fastq files were processed using the Genome Detective analysis pipeline version 2.13.3 (https://www.genomedetective.com/). Upon completion, consensus fasta and Binary Alignment Map (BAM) files were retrieved from Genome Detective for each individual sample. Sequence genotyping and lineage classification was performed using the Genome Detective Dengue Typing Tool (https://www.genomedetective.com/app/typingtool/dengue/), based on a newly developed nomenclature for DENV classification (20). The resulting genomic sequences have been made available on the GISAID Epi-Arbo database (EPI_ISL_19229161 - EPI_ISL_19229193).

Sequence Alignment and Phylogenetic Analysis

DENV reference datasets were retrieved from two primary databases: the National Centre for Biotechnology Information (NCBI) GenBank and the GISAID Epi-Arbo web interface (https://gisaid.org/). Any duplicate entries were removed. Sequence genotyping and lineage classification was performed using the Genome Detective Dengue Typing Tool (https://www.genomedetective.com/app/typingtool/dengue/), based on a newly developed nomenclature for DENV classification (20). The retrieved sequences corresponding to the lineages detected from newly sequenced Ethiopian genomes were subjected to initial quality control to remove any unverified sequences and incomplete records (i.e. geographical location and sampling dates). The final dataset used for the phylogenetic analyses in this study consisted of 33 sequences generated in this study from Ethiopia, as well as 2348 publicly available sequences of DENV3 genotype III major lineage B (corresponding to one of the transmission lineages detected in Ethiopia) and 990 sequences of DENV1 genotype III major lineage A (corresponding to the other transmission lineage detected in Ethiopia). Our DENV1 and DENV3 datasets were aligned against the appropriate DENV serotype reference genome (DENV1: NC_001477.1; DENV3: NC_001475.2) using Nextalign v1.3.0 alignment tool (21).
Maximum likelihood trees for each of the two genotypes were generated using IQ-TREE v2.2.2.2 with 1,000 bootstraps (22). The nucleotide substitution model TIM2+F+I+G4 was selected for DENV1 and GTR+F+I+G4 for DENV3, respectively, according to Bayesian Information Criterion using Model Finder (23). The molecular clock signal was evaluated using TempEst v1.5.3 (24). Potential outliers that violated the molecular clock assumption were removed prior to inferring a time-scaled tree (TreeTime v 0.10.0) using a molecular clock rate of $5.015 \times 10^{-4}$ and $1.225 \times 10^{-4}$ for DENV1 and DENV3 respectively nucleotide substitutions per site per year, as determined by the clock function in TreeTime.

**Time-calibrated Bayesian phylogenies**

Time-calibrated Bayesian phylogenetic trees were constructed to estimate the time of emergence (time to the most common recent ancestor, TMRCA) of the transmission lineages circulating during the 2023 outbreak, and the likely introduction routes. The latest release of BEAST (v.1.10.4) was used along with the BEAGLE library v3.2.0 to improve computational speed (25, 26). A relaxed clock model was applied for all analyses along with the HKY+G+I nucleotide substitution model. A constant population coalescent model assumption was used. Subsequently, all analyses were performed in two independent runs of 100 million iterations each. Convergence of MCMC chains was checked using Tracer v.1.7.1 (27) and 10% of initial chains was discarded as burn-in. Post-burn-in samples were pooled to summarize parameter estimates using LogCombiner and TreeAnnotator, including posterior probability of each parameter and maximum clade credibility (MCC) trees.

**Visualisations**

The phylogenetic trees were visualized using Figtree v1.4.4 ([http://tree.bio.ed.ac.uk/software/figtree/](http://tree.bio.ed.ac.uk/software/figtree/)) and other figures using R and ggplot (28).

**Air Travel Data**

We evaluated travel data generated from the International Air Transport Association (IATA) to quantify passenger volumes originating from international airports into. IATA data accounts for ~90% of passenger travel itineraries on commercial flights, excluding transportation via unscheduled charter flights (the remainder is modeled using market intelligence).
Ethical considerations

The study obtained ethical approval from Addis Ababa University (Ref. no. CNCSDO/175/15/2023), and Institutional Review Boards (IRBs) of the EPHI (EPHI-IRB-536-2023).

Results

Participants were recruited during a recorded DENV outbreak in Ethiopia (Supplementary Figure 1). Cases were recorded in multiple regions of the country by the Ministry of Health (MoH) in Ethiopia from early 2023, with Afar and Dire Dawa being the two most affected regions (Figure 1A, C). There was a first peak in cases in the Afar region between April and August 2023, low level transmission in Dire Dawa over that period, and a second prominent peak in Dire Dawa from October 2023 to January 2024. Our clinical surveillance was able to detect dengue acute infections from both these regions, with detection in Dire Dawa intensifying prior to the peak in recorded cases (Figure 1A).

Figure 1. Spatiotemporal distribution of dengue epidemiology and ecology in Ethiopia.
A) Epidemiological curve of dengue cases in Ethiopia: Afar region (outbreak from April 2023 to August 2023), Dire Dawa City Administration (outbreak from June 2023 to April 2024), and national count from 2023 to 2024. B) Serotype distribution of sampled cases. C) Map of DENV cases in Ethiopia from 2023 to 2024, with the Afar region and Dire Dawa City administration outlined in red. D) Demographic distribution of sampled cases.
Sample collection and dengue epidemiology in Ethiopia

Out of 891 febrile patients screened for DENV, CHIKV, and ZIKV infections, only dengue virus was detected in this study. The overall proportion with a positive PCR for DENV was 10.4% (93 out of 891 patients). The test positivity was slightly higher in men than women (Table 1). DENV was isolated from two of the sites; the proportion with a positive DENV PCR was 13.7% (41/300) in Afar and 17.9% (52/291) in Dire Dawa. None of the 300 samples from Gambela were positive for DENV. Virus genotyping shows evidence of co-circulation of DENV1 and DENV3 throughout the outbreak period (Figure 1B). Of the 93 isolates, 88 (94.6%) were identified as belonging to the DENV3 serotype, while the remaining 5 (5.4%) were identified as DENV1. At the Afar site, only DENV3 was detected (n=36), whereas at the Dire Dawa site, both DENV1 (n=5) and DENV3 (n=47) were detected (Table 1).

Table 1: Socio-demographics characteristics, burden of DENV infection, and serotype circulation in Ethiopia.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Dengue test positive, n (%)</th>
<th>Total</th>
<th>Percentage positivity</th>
<th>Dengue serotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>400</td>
<td>449</td>
<td>10.91</td>
</tr>
<tr>
<td>Female</td>
<td>44</td>
<td>398</td>
<td>442</td>
<td>9.95</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afar</td>
<td>41</td>
<td>259</td>
<td>300</td>
<td>13.67</td>
</tr>
<tr>
<td>Dire Dawa</td>
<td>52</td>
<td>239</td>
<td>291</td>
<td>17.87</td>
</tr>
<tr>
<td>Gambela</td>
<td>0</td>
<td>300</td>
<td>300</td>
<td>0</td>
</tr>
<tr>
<td>Age Groups (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-4</td>
<td>0</td>
<td>27</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>5-14</td>
<td>2</td>
<td>40</td>
<td>42</td>
<td>4.8</td>
</tr>
<tr>
<td>15-29</td>
<td>41</td>
<td>393</td>
<td>434</td>
<td>9.4</td>
</tr>
<tr>
<td>30-44</td>
<td>28</td>
<td>228</td>
<td>256</td>
<td>10.9</td>
</tr>
<tr>
<td>&gt;= 45</td>
<td>21</td>
<td>111</td>
<td>132</td>
<td>15.9</td>
</tr>
</tbody>
</table>

Table 1: Socio-demographics characteristics, burden of DENV infection, and serotype circulation in Ethiopia.
Virus genome sequencing

A total of 55 viral isolates were selected based on their PCR screening Ct values from Afar region and Dire Dawa City administration. The Ct values of the isolates ranged from 15.5 to 25.6. A total of 20 DENV3 were selected from Afar and 35 (5 DENV1 and 30 DENV3) were selected from Dire Dawa sites. Out of these, 54 isolates were successfully sequenced, and only one sample failed to be sequenced due to low sample volume. Among the successfully sequenced isolates, 33 gave near whole genome sequences (>90% genome coverage). Of these, all of the 5 DENV1 identified as DENV1 genotype III major lineage A (DENV1_IIIA), while the remaining 28 were identified as DENV3 genotype III major lineage B (DENV3_IIIB), with 1 isolate from Afar and 27 from Dire Dawa (Figure 2).

Samples collected from Dire Dawa demonstrated an expected association between a high viral load (lower Ct values) and high sequence coverage whereas most samples from Afar resulted in low sequence coverage genomes despite low Ct values. The Ct values were not repeated after storing the serum samples; we only used the first screening PCR Ct values. The sample collection and storage freezers used were different for both sites. The samples from Afar site were collected in DNA/RNA shield during the outbreak in May 2023 and stored for a long time at -20°C, with intermittent temperature fluctuations of the freezer. In contrast, samples collected from Dire Dawa and Gambela sites were relatively recent and stored at -80°C until they were sequenced.
**Figure 2. Sequencing process and coverage results.** A) Sample selection, genotyping and sequencing workflow. B) CT values versus sequence coverage for all sequenced specimens.

**Evolutionary history of sequences belonging to Ethiopia’s 2023 outbreak**

A phylogenetic tree was reconstructed with 990 global DENV1_{IIIA} aligned sequences, including five from Ethiopia sequenced in this study (Figure 3A). For this lineage, all sequences from Ethiopia belonged to a single transmission cluster, which clustered monophyletically, with other African sequences (Democratic Republic of Congo and Djibouti), suggesting cryptic transmission in the region since 2019 (Figure 3B). TMRCA of the Ethiopian clade specifically to mid 2021 (Figure 3B). Basal to this African clade are sequences mostly of Asian origin suggested an introduction into Africa from Asia, which we can date back to the year 2018 (Figure 3B). For the DENV3_{IIIB} lineage, we constructed a phylogenetic tree with 2,348 global sequences, including the ones from Ethiopia (Figure 3C). Sequences from Ethiopia belong to two distinct clades on the tree (Figure 3D, E). While the first clade consisted of sequences generally of Asian and African origin, the Ethiopian sequences, with a TMRCA of early 2021 (95% HPD: 2020 – 2022) cluster
monophyletically with a sequence from Italy’s 2023 outbreak (29). We infer that the common ancestor between the Ethiopian isolates and the Italy one existed around mid-2019 (95% HPD: mid-2017 - 2021) with long branches separating the two locations. The Ethiopia-Italy clade is supported with >70% posterior support at the relevant internal nodes on the MCC tree. This suggests some level of cryptic transmission in unsampled areas from where both Ethiopia and Italy could have received viral introductions. However, given historical ties between the two countries and the remaining high connectivity, Italy being the European country with the second most air travel passengers into Ethiopia (Supplementary Figure 2), an introduction from there is plausible. The second relevant clade of this lineage shows a clear introduction from Asia to Ethiopia late 2021 (95% HPD: 2021 - 2023), with potential persistent transmission since then (Figure 3E). While the closest relatives to this Ethiopian clade are sequences from India, it remains possible, again given long branch lengths on the tree, that intermediate unsampled regions are involved, including country such as Saudi Arabia (30), which is the top countries with largest passenger volume to Ethiopia (Supplementary Figure 2).
Figure 3. Phylogenetic analysis of Ethiopian genomes. A) Timed maximum likelihood phylogeny of the DENV1 genotype III major lineage A clade, containing Ethiopian genomes sequenced in this study. The tree is coloured by sequence locations. B) Zoom-in of a sub-clade of the tree to show close evolutionary relationships related to the Ethiopian DENV1 genomes. C) Timed maximum likelihood phylogeny of the DENV3 genotype III major lineage B clade, containing Ethiopian genomes sequenced in this study. The tree is coloured by sequence locations. D) Maximum clade credibility (MCC) tree of sub-tree containing cluster 1 of Ethiopian DENV3
In this study, we investigated the genomic epidemiology of DENV in two areas of Ethiopia. This study presents the distribution of serotypes and first-ever report of the genotypes and lineages circulating in Ethiopia from the first ever whole genomes sequenced in the country.

Our analysis showed that two serotypes, DENV1 and DENV3 were isolated from Dire Dawa, while only DENV3 was isolated in the Afar area. These preliminary data suggest a need for more comprehensive epidemiological or genomic surveillance and strengthened systems for severe dengue surveillance due to the potential risk of severe dengue from multiple serotypes or genotypes. In Africa, all four serotypes of DENV have been detected in both humans and Aedes mosquitoes (31). DENV2 is the most prevalent serotype in the East African region. The virus caused outbreaks in Ethiopia in 2013, Kenya in 2013, Tanzania in 2014, and Mozambique in 2014-2015, and has continued to remain prevalent in these areas (32-34). DENV1 outbreaks have been detected at different times in Angola, Kenya, Senegal, and Somalia (35). The current DENV1 serotype may have been historically co-circulating with DENV2 or may have been a recent introduction into Dire Dawa.

The current outbreak of dengue fever first began in the Mile district of the Afar Region in Ethiopia in April 2023, in which the outbreak was caused by DENV3 and has since spread to at least five other regions in the Eastern part of the country (36). Between April 2, 2023, and February 26, 2024, dengue cases were reported to the Public Health Emergency Operations Center in Ethiopia from these affected regions.

Phylogenetic analysis revealed that the DENV1 serotype belongs to genotype III major lineage A. The analysis also showed that the circulation of this lineage during the 2023 outbreak corresponds to a single transmission cluster. It seems that this cluster was introduced to Ethiopia from Asia via other African countries. Dengue viruses belonging to the same genotype were sequenced during the 2023 dengue outbreak in Chad (2). However, the Ethiopian genomes generated in this study are not directly linked to the outbreak in Chad, which originated from a large outbreak in Tanzania in 2019, via Nigeria (37). The phylogenetic reconstruction from available DENV1 genomic data
in Africa, which we have supplemented with our sequences, show that several distinct lineages are currently circulating on the continent. All known Central African DENV1 strains belong to the genotype V African lineage, with the oldest strains of this lineage being isolated in Nigeria in 1968. However, it remains that the lack of publicly available data limits our understanding of the dynamics of DENV1 within the continent.

On the other hand, the DENV3 viruses sequenced during Ethiopia’s 2023 outbreak belonged to the genotype III major lineage B, and phylogenetic reconstructions revealed two transmission clusters circulating in Ethiopia. Cluster-1 is closely related to a DENV3 virus sequenced in 2023 sequence from Italy (as shown in Figure 3A) (29). This highlights the possibility that the virus may have been introduced into Ethiopia from Italy. Since we have no previous sequence information, we cannot exclude the possibility of transmission from Ethiopia to Italy.

Available data indicates that DENV3 genotype III exists in neighboring countries such as Sudan, Kenya, Djibouti, and Somalia. However, the current DENV3 genotype III in Ethiopia is not directly related to the African isolates. Cluster-2 appears to have been introduced from India perhaps via secondary locations (as shown in Figure 3E). Phylogenetic analysis of DENV3 in India has shown that outbreaks were caused by genotype III and the outbreak during 2017-2018 was similar to the strain that was circulating in 2016 (40).

While performing the whole genome sequencing, only 33 samples resulted in approximately 90% and above genome coverage. Thirty-two (DENV1 =5 and DENV3 =27) of these samples were from Dire Dawa site but only one sample gave approximately 90% genome coverage from Afar site. Serum samples from the Afar site was collected during the outbreak in tubes containing DNA/RNA shield solution and sent to the Ethiopian Public Health Laboratory for detection and serotyping. The remaining samples were then stored at -20°C for around 11 months before sequencing. Sequencing was carried out following this period of storage. The limited genome coverage observed could be attributed to the extended storage under unsuitable conditions. In contrast, serum samples obtained from Dire Dawa were gathered without DNA/RNA shield and were relatively recent, stored at -80°C for approximately 4 to 8 months, resulting in greater sequence coverage.
This study presents the first report on the genomic epidemiology of the 2023 dengue outbreak in Ethiopia. However, it is important to note some limitations: 1) Poor sequencing outcomes were encountered, possibly due to inadequate sample storage at Afar. Improved sample storage conditions could have led to better whole genome sequencing results. 2) The limited availability of DENV sequences continentally and globally restricts our ability to infer global transmission dynamics and direct introduction routes. In this study, we investigated the cocirculation of DENV1 and DENV3 in Ethiopia during the 2023 outbreak. We were able to identify circulation genotypes and lineages and infer possible introduction routes and timings. Our findings highlight the utility for comprehensive disease surveillance, including pathogen genome sequencing, of dengue transmission in Ethiopia and beyond to understand transmission dynamics of emerging outbreaks. Regular active surveillance through highly connected sentinel sites as well as at port of entry can improve response time to outbreaks and reduce the likelihood of dengue establishment.

Supplementary Materials

Supplementary Figure 1: Map showing the location of hospital-based sentinel sites where samples were collected for the study. The Afar region and Dire Dawa City administration are
located in the eastern part of Ethiopia, while the Gambela region is situated in the western part of Ethiopia.

**Supplementary Figure 2:** Monthly incoming travel into Ethiopia from countries with >2000 incoming passengers. The data shown is for February 2020 as a representative month.

**Supplementary Table 1:** Acknowledgement of Authors for data obtained from the GISAID EpiArbo database.

**Author Contributions**

Ad.A. and D.B. conceived and designed the experiments; AdA., H.T. and E.W. performed the data analysis and experiments; Ab.A., F.R., M.B., H.T., K.K., C.H., M.U.G.K, T.d.O. contributed materials/analysis tools; J.G., and S.P. participated in the wet laboratory work; Ge.Ta., Ge.To., D.B., R.de.W.T., T.d.O., D.W., supervised the work; and Ad.A., H.T. E.W., and D.B. wrote the original manuscript and the paper. All authors have read and agreed to the published version of the manuscript.

**Funding**
Support was provided by the Ethiopian Public Health Institute and Addis Ababa University. Sequencing and modelling activities at KRISP and CERI are supported in part by grants from the Rockefeller Foundation (HTH 017), the Abbott Pandemic Defense Coalition (APDC), the National Institute of Health USA (U01 AI151698) for the United World Antivirus Research Network (UWARN), the SAMRC South African mRNA Vaccine Consortium (SAMVAC), Global Health EDCTP3 Joint Undertaking and its members as well as Bill & Melinda Gates Foundation (101103171), European Union’s Horizon Europe Research and Innovation Programme (101046041), the Health Emergency Preparedness and Response Umbrella Program (HEPR Program), managed by the World Bank Group (TF0B8412), the GIZ commissioned by the Government of the Federal Republic of Germany, the UK’s Medical Research Foundation (MRF-RG-ICCH-2022-100069), and the Wellcome Trust for the Global health project (228186/Z/23/Z, also M.U.G.K.). The EU project EpiGen (https://epigenethiopia.org/) is gratefully acknowledged for complementary support. 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. M.U.G.K. acknowledges funding from The Rockefeller Foundation, Google.org, the Oxford Martin School Pandemic Genomics programme, European Union’s Horizon Europe programme projects MOOD (#874850) and E4Warning (#101086640), the John Fell Fund, a Branco Weiss Fellowship and Wellcome Trust grants 225288/Z/22/Z, United Kingdom Research and Innovation (#APP8583) and the Medical Research Foundation (MRF-RG-ICCH-2022-100069). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission or the other funders.

Informed Consent Statement

All participants gave written informed consent for participation and publication.

Data Availability Statement

All sequences used for this manuscript are publicly available on the GISAID EpiArbo database, subject to its terms and conditions (https://www.gisaid.org/), GISAID identifier: EPI_ISL_19229161 - EPI_ISL_19229193.

Acknowledgments
Concerning our reference data against which we analysed the genomes generated in this study, where a portion of those were obtained from the GISAID EpiArbo database, we gratefully acknowledge Authors from the Originating laboratories responsible for obtaining the specimens, as well as the Submitting laboratories where the genome data were generated and shared via GISAID (See Supplementary Table 1). Additionally, we extend our gratitude to CDC Atlanta for supplying Dengue detection and serotyping kits.

Conflicts of Interest

K.K. is the founder of BlueDot, a social enterprise that develops digital technologies for public health. C.H. is employed at BlueDot. All other authors declare no competing interests or personal relationships that could influence this research paper.

Abbreviations

- BAM: Binary Alignment Map
- CDC: Center for Diseases Control and prevention
- CERI: Center for Emerging and Re-emerging Infectious diseases
- CHIKV: Chikungunya virus
- CLIMADE: Climate Amplified Diseases and Epidemics
- COVID-19: Corona Virus Diseases-19
- DENV: Dengue virus
- EPHI: Ethiopian Public Health Institute
- GISAID: Global initiative on sharing all influenza data
- HPD: Highest Posterior Density
- IATA: International Air Transport Association
- IRB: Institutional Review Board
- KRISP: KwaZulu-Natal Research Innovation and Sequencing Platform
- MCC: Maximum Clade Credibility
NCBI  National Centre for Biotechnology Information
RT-PCR  Reverse transcription polymerase chain reaction
TMRCA  The Most Common Recent Ancestor
ZIKV  Zika virus


