JAN 27, 2023

OPEN ACCESS

DOI:
dx.doi.org/10.17504/protocols.io.kqdg39xxeg25/v1

Protocol Citation: Chantal Vogels, Chrispin Chaguza, Mallery I Breban, Afeez Sodeinde, Emma Taylor-Salmon, Abigail J. Porzucek, Nathan D Grubaugh 2023. DengueSeq: A pan-serotype whole genome amplicon sequencing protocol for dengue virus. **protocols.io** https://dx.doi.org/10.17504/protocols.io.kqdg39xxeg25/v1

Protocol status: In development
We are still developing and optimizing this protocol

Created: Nov 14, 2022

Last Modified: Jan 27, 2023

PROTOCOL integer ID: 72720

DengueSeq: A pan-serotype whole genome amplicon sequencing protocol for dengue virus

Chantal Vogels[1], Chrispin Chaguza[1], Mallery I Breban[1], Afeez Sodeinde[1], Emma Taylor-Salmon[1], Abigail J. Porzucek[1], Nathan D Grubaugh[1]

[1]Department of Epidemiology of Microbial Diseases, Yale School of Public Health

Nathan D Grubaugh
Department of Epidemiology of Microbial Diseases, Yale Schoo...

ABSTRACT

Background

Amplicon-based sequencing (PrimalSeq) was developed in response to the Zika virus epidemic due to difficulties generating complete genomes using metagenomic approaches [1,2]. Later this approach was adapted as the primary sequencing method for SARS-CoV-2 (i.e. the “ARTIC” protocol) [3]. During the COVID-19 pandemic, investments in genomic infrastructure have resulted in a significant increase in sequencing capacity leading to a record number of over 14 million publicly available SARS-CoV-2 genomes. This expanded sequencing capacity can be utilized to improve genomic surveillance of other viruses, by swapping out components such as primer schemes. The COVID-19 pandemic further exemplifies how genomic surveillance can help to track the emergence and spread of variants, facilitate vaccine development, and inform diagnostics and therapeutics. Increased genomic surveillance of other viruses of public health concern, such as the mosquito-borne dengue virus, is needed to reduce the future burden of disease. Particularly, genomic surveillance can help to monitor the roll out of novel control strategies such as vaccines and release of mosquitoes carrying the virus-inhibiting Wolbachia bacteria. However, the majority of currently available sequences are partial, making them unsuitable to monitor and refine novel control tools. Here, we developed and validated primer schemes for all four dengue virus serotypes to improve whole-genome sequencing using an amplicon-based sequencing approach. The serotype-specific primers can be used individually or combined as a universal pan-serotype dengue virus amplicon-based sequencing approach.

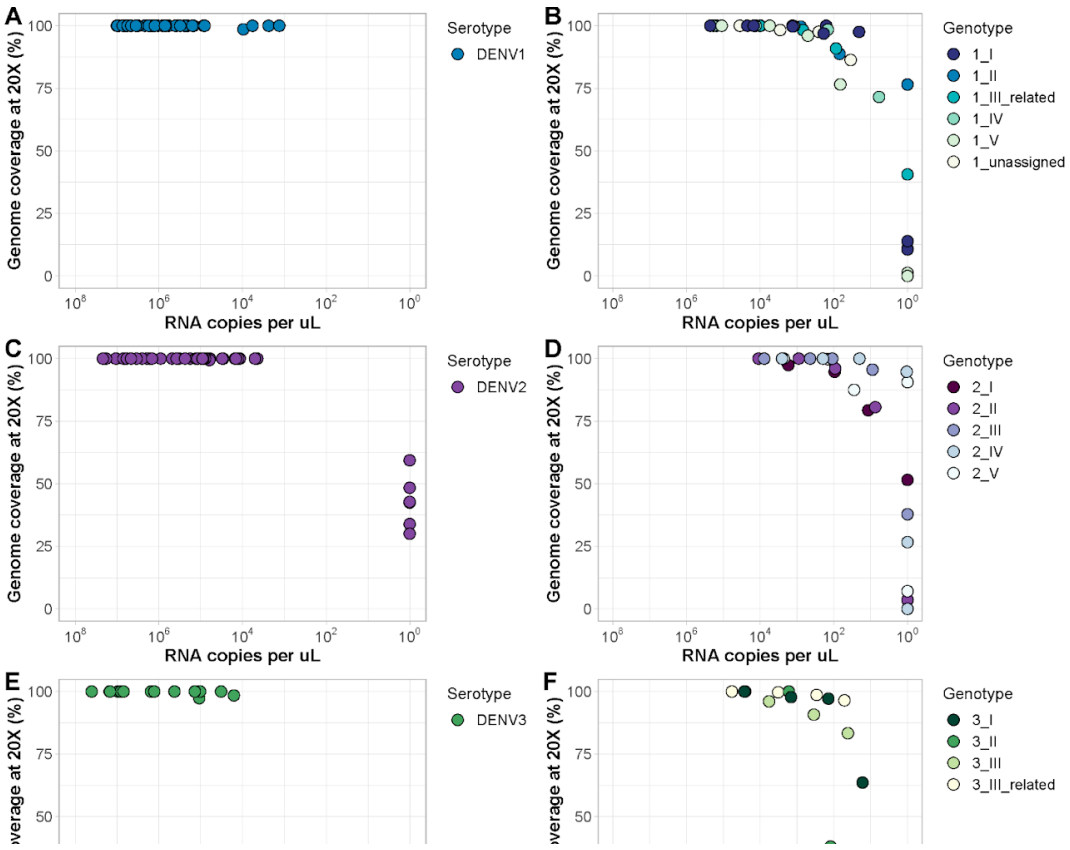
Overview of design

We used PrimalScheme (https://primalscheme.com/) to generate four serotype-specific dengue virus primer schemes (serotypes 1-4). For each primer scheme, we selected 10 publicly available genomes that represent the genetic diversity within each serotype. As reference genomes, we used the four serotype-specific

reference genomes available on GenBank (DENV1: NC_001477.1, DENV2: NC_001474.2, DENV3: NC_001475.2, DENV4: NC_002640.1). Each primer scheme consists of 35-37 primer pairs with an amplicon length of ~400 bp.

Initial validation

We validated the four primer schemes with dengue virus stocks that we obtained from the Yale Arbovirus Research Unit (YARU) and World Reference Center for Emerging Viruses and Arboviruses (WRCEVA, UTMB) collections. These virus stocks spanned the defined genotypes within each of the four serotypes. Serotypes were confirmed and PCR cycle threshold (Ct) values were determined with the CDC real-time RT-PCR assay [4]. We sequenced dengue viruses using the Illumina COVIDSeq test (RUO version) and generated consensus genomes at a depth of coverage of 20X using iVar (version 1.3.1). Initially, we sequenced undiluted virus stocks with the serotype-specific assay and found high coverage across all genotypes within each serotype, except for sylvatic dengue virus serotype 2 (genotype VI; **Figure 1**). Next, we selected 1-2 viruses for each genotype and diluted these viruses until no RNA was detected. By sequencing these diluted viruses we found that we were able to generate near-complete genomes for samples containing at least ~100 RNA copies per µL (PCR Ct value of approximately 30) for each serotype. Some selected viruses had mismatches with the real-time RT-PCR assay, resulting in reduced sensitivity of the assay. As a result, **there may be varying levels of genome coverage for samples that were not detected by the real-time RT-PCR assay**.



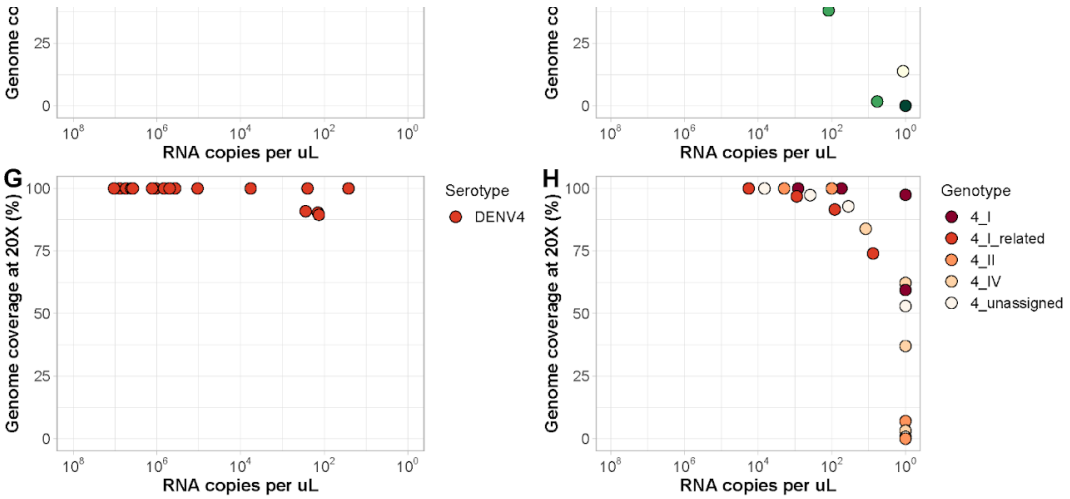
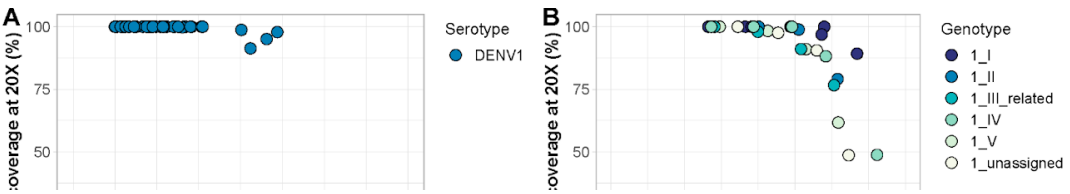


Figure 1: <u>Serotype-specific primers</u> - Percent genome coverage of undiluted and diluted dengue virus stocks sequenced with the serotype-specific amplicon-based sequencing approach. Dengue virus serotypes were determined using the CDC real-time RT-PCR assay. Viruses spanning the various genotypes within each serotype were sequenced with the serotype-specific primer schemes using the Illumina COVIDSeq Test (RUO version). Consensus genomes were generated at a depth of coverage of 20X using iVar (version 1.3.1). Genome coverage at 20X for undiluted dengue virus serotype 1 (A), 2 (C), 3 (E), and 4 (G) virus stocks. Each dot represents a different dengue virus stock. Genome coverage at 20X for selected dengue virus serotype 1 (B), 2 (D), 3 (F), and 4 (H) virus stocks diluted until no longer detected by the CDC real-time RT-PCR assay. Dots represent dengue virus stocks diluted to different concentrations. **Some samples have high coverage, while not detected by the real-time RT-PCR assay, which is due to mismatches with the primer or probe sequences (lower sensitivity).**

To facilitate high-throughput sequencing, we developed a single pan-serotype dengue virus approach that can be used for all four serotypes. By mixing the four primer schemes and increasing the working concentration of the primers to 20 µM, we were able to achieve similar genome coverage as compared to the serotype-specific approach (**Figure 2**). This pan-serotype protocol for dengue virus is particularly useful when the serotype is unknown, when there are co-infections with multiple serotypes, or when multiple serotypes are included in a single sequencing run.



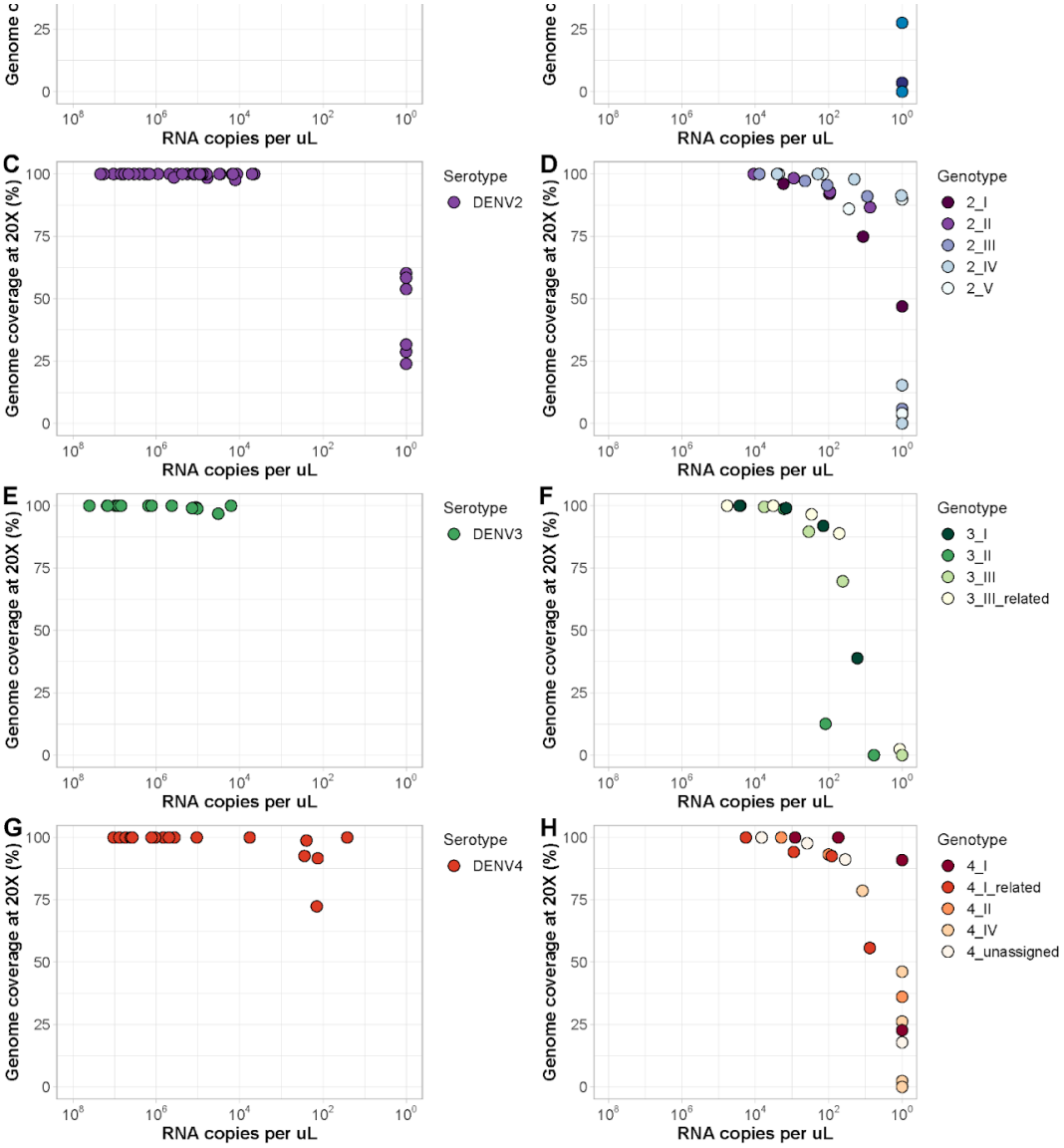


Figure 2: Pan-serotype primers - Percent genome coverage of undiluted and diluted dengue virus stocks sequenced with the pan-serotype dengue virus amplicon-based sequencing approach. The same dengue viruses as shown in Figure 1 were sequenced, by using the pan-serotype approach. The four individual primer schemes were mixed for use as a single universal pan-serotype dengue virus scheme. Consensus genomes were generated at a depth of coverage of 20X using iVar (version 1.3.1). Genome coverage at 20X for undiluted dengue virus serotype 1 (A), 2 (C), 3 (E), and 4 (G) virus stocks. Each dot represents a different dengue virus stock. Genome coverage at 20X for selected dengue virus serotype 1 (B), 2 (D), 3 (F), and 4 (H) virus stocks diluted until no longer detected by the CDC real-time RT-PCR assay. Dots represent dengue virus stocks diluted to different concentrations. **Some samples have high coverage, while not detected by the real-time RT-PCR assay, which is due to mismatches with the primer or probe sequences (lower sensitivity).**

Conclusion

In this protocol we present the initial validation of a whole genome amplicon-based sequencing approach for all four dengue virus serotypes. We show that the four primer schemes can be used as serotype-specific or combined pan-serotype dengue virus whole genome sequencing protocol. The newly developed primer schemes can be used with currently established amplicon-based sequencing workflows, like COVIDseq, with minimal change to the protocol. Further testing involves sequencing of clinical samples. Our ultimate goal is to help increase capacity for whole genome dengue virus sequencing.

Acknowledgements

We would like to acknowledge Philip Armstrong, Doug Brackney, and team (Connecticut Agricultural Experiment Station) for generously providing dengue virus samples from the Yale Arbovirus Research Unit collection, Kenneth Plante and team (The University of Texas Medical Branch) for generously providing dengue virus samples from the World Reference Center for Emerging Viruses and Arboviruses collection, and Mary Petrone and Isabel Ott for preparation of virus stocks.

References

1. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 2017;12: 1261–1276.
2. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol. 2019;20: 8.
3. Artic Network. [cited 25 Jul 2022]. Available: https://artic.network/ncov-2019
4. Santiago GA, Vergne E, Quiles Y, Cosme J, Vazquez J, Medina JF, et al. Analytical and clinical performance of the CDC real time RT-PCR assay for detection and typing of dengue virus. PLoS Negl Trop Dis. 2013;7: e2311.

GUIDELINES

It is recommended that steps performed up to amplicon generation be performed at a different workstation and with different equipment than steps post-amplicon generation.

MATERIALS

Primers

Primer schemes for the four dengue virus serotypes (1-4) can be used individually as serotype-specific or combined as a single pan-serotype approach. Primers may be ordered from any oligonucleotide company. For primer preparation instructions for

the serotype-specific and pan-serotype approaches, see Step 1 of the protocol. See files below for primer sequences.

Primer Schemes:

DENV1_Primer-Scheme.xlsx
DENV2_Primer-Scheme.xlsx
DENV3_Primer-Scheme.xlsx
DENV4_Primer-Scheme.xlsx

Equipment

- 96-well format Thermocycler (two needed for 49+ sample runs)
- Qubit
- Bioanalyzer
- Magnetic rack fit for a 96-well PCR plate
- Magnetic rack fit for 1.5 / 2-mL microcentrifuge tubes
- Cold block fit for a 96-well PCR plate
- Pipettes (assorted sizes)

Consumables

- 8-strip PCR tubes; may be substituted with PCR plates with heat-sealing film
- 1.5 / 2mL / 5 mL microcentrifuge tubes
- Filtered pipette tips (assorted sizes)
- Reservoirs
- Waste containers

Reagents

- Nuclease-free water
- Molecular-grade ethanol
- Illumina COVIDSeq Test Kit:

	A	B	C
	Illumina COVIDSeq Test Box 1 – 3072 Samples, Part # 20044408		
	Reagent	Description	Storage
	ITB	Illumina Tune Beads	Room Temperature
	ST2 HT	Stop Tagment Buffer 2 HT	Room temperature, post-amp environment
	Illumina COVIDSeqTest Box 2 – 3072 Samples, Part # 20044409		
	EBLTS HT	Enrichment BLT HT	2°C to 8°C post-amp environment
	TWB HT	Tagmentation Wash Buffer HT	2°C to 8°C post-amp environment

	A	B	C
	RSB HT	Resuspension Buffer HT	2°C to 8°C, post-amp environment
	Illumina COVIDSeq Test Box 3 – 3072 Samples, Part # 20044410		
	EPH3 HT	Elution Prime Fragment 3HC Mix HT	-25°C to -15°C, pre-amp environment
	FSM HT	Frist Strand Mix HT	-25°C to -15°C, pre-amp environment
	RVT HT	Reverse Transcriptase HT	-25°C to -15°C, pre-amp environment
	IPM HT		
	TB1 HT	Tagmentation Buffer 1 HT	-25°C to -15°C, post-amp environment
	EPM HT	Enhanced PCR Mix HT	-25°C to -15°C, pre-amp environment
	Index Adapater Part Numbers : 20043132, 20043133, 20043134, 20043135		
	Index Adapters	IDT for Illumina- PCR Indexes Set 1-4	Room Temperature

SAFETY WARNINGS

Before starting work with dengue virus samples, please contact your local EHS (environment, health and safety) or biosafety office for proper guidance on how to work with these samples in your laboratory.

BEFORE START INSTRUCTIONS

If substituting 8-strip PCR tubes for PCR plates, plates must be heat-sealed before each thermocycling step.

Dengue Serotype Identification / Plate Setup

1 There are two primer scheme options available for use: Single-Serotype and Pan-Serotype.
Step 1 includes a Step case.
Single-Serotype
Pan-Serotype

Prepare Working Primer Solutions

step case

Single-Serotype

Identify the serotype of all samples being sequenced.

We recommend using the CDC DENV-1-4 Real-Time RT-PCR Multiplex Assay for serotype identification if an in-house assay is not available.

Organize samples by serotype. Step 4 requires a serotype-specific master mix to be added to the samples; order the samples to best accommodate this.

Each serotype-specific master mix will require its own negative template controls, one cDNA control (added in Step 3) and one PCR control (added in step 4). Be sure to account for these extra controls in the plate design.

2

Briefly centrifuge all primers

2.1 Within each serotype, organize odd and even-numbered primers into two separate groups:
Odd-numbered primers = Pool 1
Even-numbered primers = Pool 2

> Note
>
> Recommended to organize numerically, i.e. 1_left, 1_right, 3-left, 3 right...

2.2 If primers were ordered lyophilized, resuspend to
[M] 100 micromolar (µM) in nuclease-free water

2.3 Add 90 µL nuclease-free water to two sets of 8-strip PCR tubes for each serotype:

	Serotype	Tubes for Pool 1	Tubes for Pool 2
	DENV1	36	34
	DENV2	38	36
	DENV3	36	34
	DENV4	36	34

2.4 Label the tops of each tube with the corresponding primer that will be added

- For readability, recommended this format: D1 1L, D1 1R, D1 3L, D2 3R...

2.5 Add 10 µL 100 µM primer to each corresponding tube to reach [M] 10 micromolar (µM) /primer

- Recommend to either mix tubes with a new pipette tip or gently vortex

2.6 Working pool-by-pool, transfer 10 µL 10µM primer into 1.5mL microcentrifuge tubes

- 8 tubes total, with 2 tubes (Pool 1 and Pool 2) per serotype
- Recommended to label each tube with serotype, concentration, pool #, and date of dilution
- This will be the working primer solution that will be added during the amplicon generation step

2.7 Store all primers at -20ºC until use

cDNA Synthesis

3 Prepare the following reagents:

	Reagent	Storage	Instructions
	EPH3 HT	-20°C	Thaw at room temperature, invert to mix
	FSM HT	-20°C	Thaw at room temperature, invert to mix
	RVT HT	-20°C	Invert to mix, keep on ice
	Nuclease-free water	Room Temp.	Keep at room temperature

3.1 Add 8.5 µL EPH3 HT to new PCR tubes according to the number of samples

- Include one additional reaction for a cDNA NTC for each serotype-specific master mix (NTC), which will be included through the entirety of the protocol

3.2 Add 8.5 µL RNA to each tube

Add 8.5 µL nuclease-free water to the cDNA NTC

- Mix by pipetting up and down 10 times

- Briefly centrifuge tubes

3.3 Load tubes into thermocycler and run the following program to generate first-strand cDNA:

	Temperature	Time
	65°C	3 minutes
	4°C	Hold

17µL reaction, lid temp = 105°C, preheat lid = on

3.4 In a new 1.5mL tube, prepare the following master mix according to the number of samples:

	Reagent	µL per Sample
	FSM HT	7.2
	RVT HT	0.8
	Total	8

> **Note**
>
> All master mix volumes in this protocol are for one reaction and do not account for lost volume due to pipetting. Multiply volumes by reaction number accordingly.

3.5 Add 8 µL Master Mix to each tube
Mix by pipetting up and down 10 times
Briefly centrifuge tubes

3.6 Load tubes into thermocycler and run the following program to generate second-strand cDNA:

	Temperature	Time
	25°C	5 minutes
	50°C	10 minutes
	80°C	5 minutes
	4°C	∞

25µL reaction, lid temp = 105°C, preheat lid = on

3.7 Remove tubes from thermocycler and briefly centrifuge

Note

This is a safe stopping point. cDNA can be stored long-term at -20°C.

Amplicon Generation

4 Prepare the following reagents:

	Reagent	Storage	Instructions
	Primer Pool 1	-20°C	Thaw at room temperature, vortex to mix
	Primer Pool 2	-20°C	Thaw at room temperature, vortex to mix
	IPM HT	-20°C	Thaw at room temperature, invert to mix
	Nuclease-free water	Room temp.	Keep at room temperature

4.1 In two separate tubes, prepare the following master mixes according to the number of samples:

- **If processing more than one serotype in a run, prepare separate master mixes for each serotype, using the proper primer pools.**
- Include one additional reaction for a PCR NTC for each serotype-specific master mix, which will be included through the entirety of the protocol.

Pool 1

	Reagent	µL per Sample
	IPM HT	12.5
	Primer Pool 1	3.6
	Nuclease-free Water	3.9
	Total	20

Pool 2

	Reagent	µL per Sample

	Reagent	µL per Sample
	IPM HT	12.5
	Primer Pool 2	3.6
	Nuclease-free Water	3.9
	Total	20

4.2 Distribute the master mixes:

Add 20 µL Pool 1 Master Mix to a new set of PCR tubes, according to the number of samples

Add 20 µL Pool 2 Master Mix to a new set of PCR tubes, according to the number of samples

- Include one additional tube at the end of each set for the PCR NTC
- **Be sure to match the serotype of each master mix to the previously-identified serotype of each sample**

4.3 Add 5 µL cDNA to each tube in both Pool 1 and Pool 2.

Add 5 µL nuclease-free water to the PCR NTC in both Pool 1 and Pool 2.

- Mix by pipetting up and down 10 times
- Briefly centrifuge tubes

4.4 Load tubes into thermocycler and run the following program:

	Step(s)	Temperature	Time	Cycles
	Initial denaturation	98°C	3 minutes	1
	Denaturation	98°C	15 seconds	35
	Anneal and extension	63°C	5 minutes	
	Hold	4°C	∞	∞

25µL reaction, lid temp = 105°C, preheat lid = on

4.5 When program is complete, samples can remain in the thermocycler at 4°C or be stored at -20°C

- Briefly centrifuge tubes before use

Note

This is a safe stopping point. Amplicons can be stored long-term at -20°C.

Amplicon Tagmentation and Cleanup

5 Prepare the following reagents:

	Reagent	Storage	Instructions
	EBLTS HT	4°C	Vortex to mix
	TB1 HT	-20°C	Thaw at room temperature, vortex to mix
	ST2 HT	Room temp.	Vortex before use
	TWB HT	4°C	Vortex before use
	Nuclease-free water	Room temp.	Keep at room temperature

5.1 Quantify samples on a Qubit to confirm proper amplification in both amplicon pools

5.2 Prepare the following master mix:

	Reagent	µL per sample
	TB1 HT	10
	EBLTS HT	3.3
	Nuclease-free water	16.7
	Total	30

5.3 Add 30 µL Master Mix to a new set of PCR tubes, according to the number of samples

5.4 For each sample, combine amplicon pools:

Add 10 µL Pool 1 to the Master Mix tubes

Add 10 µL Pool 2 to the Master Mix tubes

- Mix by pipetting up and down 10 times
- Briefly centrifuge tubes

Note

Remaining amplicon pools can be stored long-term at -20°C

5.5 Load tubes into thermocycler and run the following program:

	Temperature	Time
	55°C	3 minutes
	10°C	∞

50µL reaction, lid temp = 105°C, preheat lid = on

5.6 When the program is complete, **immediately** remove the tubes from the thermocycler

- Briefly centrifuge tubes

5.7 Add 10 µL ST2 to each tube

- Mix by pipetting up and down
- Briefly centrifuge tubes

5.8 Incubate at Room temperature for 5 minutes

5.9 Place tubes on a magnetic stand and wait until the liquid is clear before continuing

5.10 Remove and discard all supernatant

- Do not allow pipette tip to come into contact with the inner walls of the tube, as it can scrape away the magnetic beads

5.11 Remove tubes from the magnetic stand and add 100 µL TWB HT to each tube (first wash)

- Mix by pipetting up and down 10 times

- Briefly centrifuge tubes
- Be careful not to introduce bubbles

5.12 Repeat steps 5.9 - 5.11 (second wash)

- Do not remove supernatant after the second wash at this point

5.13 Place the tubes on the magnetic stand

Amplify Tagmented Amplicons

6 Prepare the following reagents:

	Reagent	Storage	Instructions
	EPM HT	-20°C	Thaw at room temperature, invert to mix
	Index adapters	-20°C	Thaw at room temperature, vortex to mix
	Nuclease-free water	Room temp.	Keep at room temperature

6.1 Prepare the following Master Mix:

	Reagent	µL per sample
	EPM	20
	Nuclease-free water	20
	Total	40

6.2 When the tubes with tagmented amplicons (on the magnetic stand) are clear, remove and discard all supernatant

- Remove any remaining supernatant with a smaller-sized pipette.

6.3 Remove tubes from the magnetic stand and add 40 µL Master Mix to each tube

6.4 Briefly centrifuge the plate containing the dual-barcoded index adapters
Clean the surface of the sealing foil with RNase Away
Remove residual RNase away by wiping with 70% EtOH

6.5 Add 10 µL dual-coded index adapters to each well

- Mix by pipetting up and down 10 times
- Briefly centrifuge tubes

6.6 Load tubes into thermocycler and run the following program:

	Temperature	Time	Cycles
	72°C	3 minutes	1
	98°C	3 minutes	1
	98°C	20 seconds	7
	60°C	30 seconds	
	72°C	1 minute	
	72°C	3 minutes	1
	10°C	∞	∞

When program is complete, samples can remain in the thermocycler at 10°C or be stored at -20°C.

Note

This is a safe stopping point. Individual libraries can be stored long-term at -20°C.

Pooling and Cleanup

7 Prepare the following reagents:

	Reagent	Storage	Instructions
	ITB HT	Room temp.	Vortex to mix
	RSB HT	4°C	Keep at room temperature, vortex to mix
	80% EtOH	Room temp.	Prepare 2.5 - 3mL using nuclease free water and molecular-grade 100% EtOH immediately before use

7.1 Place tubes on a magnetic stand and wait until the liquid is clear

7.2 Pool libraries together in a new 1.5mL microcentrifuge tube by equal volume:

	Number of samples	µL per sample
	1 - 24	15
	25 - 48	10
	49 - 96	5

Note

For easier pooling of a large number of samples, it is recommended to multichannel the libraries into a new 8-strip of 0.2mL PCR tubes, then transfer the total volume to the microcentrifuge tube.

7.3 Calculate the volume of ITB needed to reach **0.9x ITB : total pooled volume**

Note

Example:
34 samples; 10µL of each sample pooled
*34 * 10 = 340µL total pooled volume*
*340µL * 0.9 = 306µL ITB*

7.4 Add the calculated volume of ITB to the pooled sample tube

- Vortex for 10 seconds to mix
- Briefly centrifuge

7.5 Incubate at Room temperature for 5 minutes

7.6 Place tube on a single-tube magnetic stand and wait until the liquid is clear

7.7 Remove and discard all supernatant

7.8 Add 1000 µL 80% EtOH to the tube and incubate at Room temperature for 30 seconds

7.9 Repeat steps 7.7 - 7.8 and remove and discard supernatant after the second wash

- Remove any residual EtOH with a smaller-sized pipette

7.10 Add RSB HT to the tube according to the number of samples pooled:

- For 1 - 24 samples add 30 µL RSB HT
- For 25 - 48 samples add 40 µL RSB HT
- For 49 - 96 samples add 50 µL RSB HT

7.11 Vortex tube to mix
Briefly centrifuge tube

7.12 Incubate at Room temperature 2 minutes

7.13 Place on the magnetic stand and wait until the liquid is clear

7.14 Transfer **[total volume RSB] - 5µL** to a new 1.5mL microcentrifuge tube

7.15 Quantify library on a Qubit and obtain fragment distribution using a Bioanalyzer/Tape Station

Sequencing

8 Protocol validated on the Illumina NovaSeq (2x150).

> Note
>
> For sequencing we recommend generating at least 1 million reads per sample for optimal sequencing coverage. Sequencing may be performed on Illumina and Oxford Nanopore Technologies sequencing platforms following standard protocols.

Bioinformatics/Analysis

9 Sequencing results may be analyzed utilizing a standard amplicon sequencing bioinformatics pipeline, including those employed for SARS-CoV-2 sequencing (e.g. iVar).

Reference Sequences:

NC_001477.1_DENV1.fasta
NC_001474.2_DENV2.fasta
NC_001475.2_DENV3.fasta
NC_002640.1_DENV4.fasta

BED files:

DENV1.primer.bed
DENV2.primer.bed
DENV3.primer.bed
DENV4.primer.bed