Pathogen Genomics Centers of Excellence Network

Phase II (End-to-End) of PGCoE Benchmarking Study

Standard Operating Procedure (SOP)

I. Purpose & Scope

The purpose of this benchmarking study is to collect a standardized set of metrics and data files that enable comparison of sequencing library preparation and bioinformatic processing for viral respiratory pathogens of public health concern. This is crucial for increasing the availability of approaches for generating pathogen whole genome sequence data while maintaining the consistency and reliability required by public health. This study and any outcomes are applicable only for surveillance purposes (not clinical).

The scope of this document is to provide guidance for how to process inactivated virus RNA extracts for the 2025 Pathogen Genomics Centers of Excellence Phase II (End-to-End) Benchmarking Study and is applicable to all participating laboratories. The Springer Laboratory (NE PGCoE) will ship for overnight delivery on dry ice the following RNA preparations:

Pure virus RNA extract preparations:

- SARS-CoV-2 (composite of clinical specimens)
- RSVA (ATCC VR-1540)
- RSVB (ATCC VR-3381)
- Influenza A H1N1 (ATCC VR-1469)
- Influenza A H3N2 (ATCC VR-777)
- Influenza B (ATCC VR-1931)

Mixed virus RNA extract preparations in ratios of 4:1, 1:1, and 1:4:

- H13: H1N1 (ATCC VR-1469) + H3N2 (ATCC VR-777)
- RV2: RSVA2 (ATCC VR-1540) + RSVB1 (ATCC VR-3381)
- RVH1: RSVA2 (ATCC VR-1540) + H1N1 (ATCC VR-1469)

Mixed H13 (H1N1 (ATCC VR-1469) + H3N2 (ATCC VR-777)) RNA extract preparations in varied ratios (processing these samples is optional):

- 1:500
- 1:100
- 1:25

Mixed RVH1 (RSVA2 (ATCC VR-1540) + H1N1 (ATCC VR-1469)) RNA extract preparations in varied

ratios (processing these samples is optional):

- 2000:1
- 500:1
- 125:1

Completion of the entire End-to-End Benchmarking Study entails the following:

- Generate sequencing data (FASTQ files) using the RNA extracts provided by the Springer Laboratory in accordance with current laboratory procedures, as applicable to the panel provided. Laboratories are not required to perform sequencing for specific RNA preparations of viruses that are not currently encompassed in individual laboratory procedures.
- Generate genomic assembly data (FASTA files) using your standard bioinformatics workflow from the FASTQ data generated from each RNA extract or mixture.
- Submit sequencing read (FASTQ) and assembly (FASTA) files to the benchmarking submission portal: https://benchmarking.nepgcoe.org.
- Complete the Phase II Lab Tab of the Data Capture Table designed to collect information on RNA extract evaluation, storage, and testing procedures performed for the End-to-End Benchmarking Study.
- Complete the Phase II BFx Tab of the Data Capture Table. Metrics will be computed as defined in the Network Proposal (available here).
- By participating in the End-to-End Benchmarking Study, labs provide consent to use the exercise data in subsequent analyses, reports, and manuscript publications generated by PGCoE members. Participants will be acknowledged for their contribution in any publication that may arise from the information collected in the 2025 End-to-End Benchmarking Study.

II. Required Materials

A. Materials Provided:
   1. 100 µL of each RNA preparation (at least 15,000 copies/µL)
   2. Shipping manifest
   3. Instructions for the storage of materials provided and study participation (this document)

B. Materials Required and Not Provided:
   1. -80°C freezer
   2. Micropipettes
   3. Sterile pipette tips
   4. Library preparation reagents and consumables required for laboratory procedures
   5. Thermal cycler
   6. Sequencing instrumentation

C. Materials Needed as Applicable to Laboratory Procedures:
   1. Reagents, consumables, and instrumentation for assessing nucleic acid purity and quantity
   2. Reagents, consumables, and instrumentation for evaluation of viral load in RNA extracts
   3. Reagents, consumables, and instrumentation for fragment analysis during library preparation

III. Safety Considerations

The RNA extracts provided to study participants have been purified and extracted using standard methods by the Springer Laboratory by the QIAamp Viral RNA Kit. The chemical lysis step used for purification of the viral RNA renders the RNA non-infectious and it may be handled according to individual procedures for manipulation of nucleic acid extracts to ensure preservation of the quality of RNA. Typically, this would be at the lowest biosafety level (biosafety level-1 or biosafety level-2), at the discretion of the participant laboratory.

RNA extracts will be shipped on dry ice (Carbon Dioxide, solid). Solid carbon dioxide is not considered a hazardous substance; however, it is classified as a Dangerous Good for transport purposes under the Department of Transportation. Precautions to take when unpacking a shipment with dry ice include contact precautions to avoid direct contact with skin, which can cause severe frostbite if skin is unprotected. Loose-fitting thermally insulated gloves (leather or cloth) are required to directly handle dry ice. Nitrile gloves are not sufficient. The container should only be opened in a well-ventilated space to avoid potentially substantial volumes of CO2 that occur when dry ice changes from solid to gas, which occurs at any temperature above -109°F.

IV. Study Procedure

Section 1: RNA Receipt, Storage, and Optional Quality and Quantity Evaluation
Section 2: Library Preparation and Sequencing
Section 3: Sequencing Data Quality Assessment and Bioinformatics Analysis
Section 4: Data Transfer and Survey Completion

Section I: RNA Receipt, Storage, and Optional Quantity and Quality Evaluation

1. **RNA Receipt:** The shipment should be opened on the same day as it is received, preferably as soon as possible after receipt, to avoid the thawing of the pre-frozen RNA mixtures. Upon unpacking the shipper, document vial identifiers, preparation dates, and date of receipt at the laboratory for your records. Inspect tubes for any defects, including damage to the microcentrifuge tubes or other potential issues that might compromise the integrity of the RNA provided. Vials will be labeled with shortened identifiers, as described in the Sample ID Table.

   Defects or other observed issues with the RNA preparations should be reported to michael_springer@hms.harvard.edu and a new aliquot requested.

2. **RNA Storage:** Immediately after inspection, all RNA preparations should be stored at -80°C and handled minimally thereafter to avoid nucleic acid degradation prior to sequencing testing.

3. **Optional Quantity and Quality Evaluation:** Optionally and in accordance with laboratory processes upstream of sequencing testing, participant laboratories may choose to pre-assess RNA preparations prior to sequencing. Typical assessments may include detection and qualification assessments by real-time reverse transcriptase PCR, RNA quantification, and/or RNA purity by fluorometric or spectrophotometric measurements. If pre-sequencing assessment of the RNA preparations is performed, record the approach, values obtained, and

instrumentation used in the Phase II Lab Tab of the Data Capture Table alongside laboratory thresholds for acceptability. If RNA preparations do not meet minimum criteria for sample testing or sequencing according to your laboratory standards, please provide this notation in the Phase II Lab Tab, but proceed with library and sequencing regardless.

Any manipulation of RNA preparations that require freeze-thaw cycles prior to the sequencing test should be kept to a minimum to avoid deleterious RNA degradation. The number of freeze-thaw cycles each preparation is subjected to during study preparation should be recorded in the Phase II Lab Tab of the Data Capture Table.

Section II: Library Preparation and Sequencing

***Any manipulations to the RNA preparations to adjust input volume or corresponding viral copy input should be performed according to routine laboratory procedures and recorded in the Phase II Lab Tab of the Data Capture Table.*

*** The sequencing run for the RNA preparations should be performed and loaded consistently with any routine run and follow established loading requirements including the routine total sample load for the methodology, sequencing chemistry, and platform utilized.*

***The laboratory should perform routine quality checkpoints in the library preparation procedure that are used to assess library quality and/or quantity of the prepared libraries and record the method used, values obtained, instrumentation, and thresholds for acceptability in the Phase II Lab Tab of the Data Capture Table.*

*** Control/reference materials and contamination controls routinely used by the laboratory should be included in the run for assessment alongside the RNA extracts.*

1. **Library Preparation:** Perform library preparation according to your laboratories' normal workflow and SOPs, as applicable to the viral RNA extracts provided. Participant laboratories are encouraged to process the RNA preparations along with other routine viral extracts to meet established sample and genome loading requirements described in laboratory protocols. If the laboratory has insufficient samples for the run, one option is to prepare and sequence the RNA preparations in replicate and/or to prepare a 10-fold dilution series of the RNA preparations in DEPC-treated nuclease-free water and to prepare libraries from the diluted RNA to achieve an estimated limit of detection for the laboratory method. Please append "_DilX" where X is the dilution factor (e.g., "Dil10" for a 1:10 dilution) to the standardized sample names listed in the Data Capture Table if you sequence from diluted samples.
2. **Sequencing:** Sequence the libraries according to your laboratory's normal workflow and procedures. Refer to the Sample ID Table for the sequence file sample IDs to use.
3. **Run Metrics:** Assess post-sequencing run-level metrics and document in accordance with laboratory procedures. Provide metrics and thresholds applied to determine run-level acceptability (if applicable) in the Phase II Lab Tab of the Data Capture Table.

Section III: Sequencing Data Quality Assessment and Bioinformatics Analysis

1. **Quality Assessment:** Follow your laboratory workflows to assess the quality of each set of FASTQ files, including the generation of genome assemblies for the Benchmarking Study samples. Please complete Phase II BFx Tab of the Data Capture Table after genome assembly, as described in the Computational Implementation of the PGCoE Benchmarking Study SOP.

Section IV: Data Transfer and Survey Completion

Data Transfer: After checking the quality of FASTQ and FASTA records, transfer the data to the benchmarking submission portal: https://benchmarking.nepgcoe.org according to the Computational Implementation of the PGCoE Benchmarking Study SOP. Note any datasets that did not meet the minimum laboratory data acceptability criteria. Data to be uploaded should include:

- Completed Data Capture Table ("Phase II Lab" and "Phase II BFx" tabs only). You will only need to provide detailed protocol information for the viral RNA preparations that were completed by your laboratory (e.g., if you do not have a standard sequencing workflow for RSV, you do not need to provide a protocol for RSV sequencing or fill out these rows of the table.)
- FASTQ and FASTA data for all sequenced RNA extracts (please see Data Capture Table for requested file naming scheme).
- [Optional] All workflow outputs, including primary results, intermediate files, and log files describing the compute environment, software versions, and workflow parameters.
- If the protocol used for library preparation is not publicly linked and the protocol can be shared within the PGCoE, please upload an electronic copy of the document.