


# Laboratory validation of a clinical metagenomic next-generation sequencing assay for respiratory virus detection and discovery

Received: 28 May 2024

Accepted: 7 August 2024

Published online: 12 November 2024

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Tools for rapid identification of novel and/or emerging viruses are urgently needed for clinical diagnosis of unexplained infections and pandemic preparedness. Here we developed and clinically validated a largely automated metagenomic next-generation sequencing (mNGS) assay for agnostic detection of respiratory viral pathogens from upper respiratory swab and bronchoalveolar lavage samples in <24 h. The mNGS assay achieved mean limits of detection of 543 copies/mL, viral load quantification with 100% linearity, and 93.6% sensitivity, 93.8% specificity, and 93.7% accuracy compared to gold-standard clinical multiplex RT-PCR testing. Performance increased to 97.9% overall predictive agreement after discrepancy testing and clinical adjudication, which was superior to that of RT-PCR (95.0% agreement). To enable discovery of novel, sequence-divergent human viruses with pandemic potential, *de novo* assembly and translated nucleotide algorithms were incorporated into the automated SURPI+ computational pipeline used by the mNGS assay for pathogen detection. Using *in silico* analysis, we showed that after removal of all human viral sequences from the reference database, 70 (100%) of 70 representative human viral pathogens could still be identified based on homology to related animal or plant viruses. Our assay, which was granted breakthrough device designation from the US Food and Drug Administration (FDA) in August of 2023, demonstrates the feasibility of routine mNGS testing in clinical and public health laboratories, thus facilitating a robust and rapid response to the next viral pandemic.

Respiratory infections are among the most common infections globally and are associated with significant morbidity and mortality<sup>1–3</sup>. Despite their importance, half of adult patients hospitalized in the United States with community-acquired pneumonia, which is most commonly caused by respiratory viruses, have no causative pathogen identified<sup>2–5</sup>. Respiratory infections caused by viruses can be especially challenging to diagnose because of the diversity of potential agents<sup>6–8</sup>. In particular, emerging pandemic viruses represent an unpredictable

threat which traditional diagnostic tools such as nucleic acid amplification tests have not been designed to detect<sup>9</sup>. The importance of unbiased assays for rapid identification of viral pathogens, especially those with sequence-divergent genomes, became evident during the discovery of SARS-CoV-2<sup>10,11</sup>.

Metagenomic next-generation sequencing (mNGS) has emerged as an attractive diagnostic method for identifying causative agents in unexplained infections as it provides a comprehensive and agnostic

✉ e-mail: [charles.chiu@ucsf.edu](mailto:charles.chiu@ucsf.edu)

approach by which all potential pathogens can be identified in a single assay without the need for specific primers and probes<sup>12,13</sup>. mNGS has been used for broadly diagnosing infections, whether viral, bacterial, fungal, or parasitic, from multiple specimen types<sup>14–16</sup>, and its clinical utility has been demonstrated for neurological and bloodstream infections<sup>16–18</sup>. However, despite the favorable performance of mNGS testing as shown by multiple studies, general adoption of mNGS technologies in clinical microbiology laboratories has been hindered by high costs, complex protocols, lack of automation, insufficient standardization of bioinformatic pipelines, prolonged turnaround times (24–72 h), lack of regulatory guidelines for clinical validation, and overall lower sensitivity for detection of common pathogens relative to targeted approaches such as polymerase chain reaction (PCR) assays<sup>19</sup>.

Here we describe the development, optimization, and clinical validation of a streamlined and largely automated mNGS laboratory-developed test (LDT) with a sample-to-result turnaround time of less than 24 h for identification of common as well as unexpected and/or novel viral respiratory pathogens. The computational SURPI+ pipeline used by the mNGS assay was modified to provide enhanced analysis capabilities, including viral load quantification, incorporation of curated reference genome databases such as FDA dAtabase for Reference Grade micrObial Sequences (FDA-ARGOS), and sensitive identification of novel, sequence-divergent viruses by de novo assembly and translated nucleotide alignment. We comprehensively evaluated assay performance metrics, including limits of detection, linearity, precision, inclusivity and exclusivity, contamination, interference, matrix effect, stability, accuracy, and capacity to detect novel viruses.

## Results

### Development and optimization of an mNGS assay for detection of viral respiratory pathogens

We developed an mNGS assay for the detection of viral pathogens from respiratory secretions, including upper respiratory swab and bronchoalveolar lavage (BAL) fluid samples (Fig. 1). We leveraged our 7-year experience running clinical mNGS assays for pathogen detection from cerebrospinal fluid<sup>20</sup> by optimizing the sample preparation and bioinformatics analysis protocols to maximize sensitivity and decrease assay sample-to-result turnaround time. We tested different combinations of centrifugation, heat, and addition of a DNA/RNA stabilization medium prior to total nucleic acid extraction and found that centrifugation alone produced the highest yield of detected viral reads. To decrease turnaround times, we used a 15-min protocol for human rRNA depletion and reduced incubation times for the reverse transcription and second-strand cDNA synthesis steps to 15 and 9 min, respectively. The final assay used 450 µL of sample input volume and consisted of the following steps: (1) centrifugation (~15 min), total nucleic acid extraction and DNase treatment for isolation of total RNA (~1 h), (2) cDNA synthesis with ribosomal RNA (rRNA) depletion (~1 h), (3) barcoded adapter ligation, library PCR amplification and purification on an automated instrument (~6.5 h), (4) library pooling (~5 min), (5) Illumina (San Diego, CA) sequencing (5 or 13 h, depending on whether a MiniSeq or NextSeq sequencer is used), and (6) bioinformatics analysis for viral detection and quantification using the SURPI+ pipeline (~1 h). Overall sample-to-answer assay turnaround time was 14–24 h. We used MS2 phage and External RNA Controls Consortium (ERCC) RNA Spike-In Mix (Invitrogen, Waltham, MA) added into each sample as internal qualitative and quantitative controls, respectively. The MS2 phage and ERCC sequencing results were also used to evaluate and interpret the background level in the sample, generally originating from the human host (Supplementary Tables 1 and 2). A commercial reference panel (Accuplex Panel, SeraCare, Milford, MA) consisting of quantified SARS-CoV-2, influenza A, influenza B, and respiratory syncytial virus (RSV) was spiked into pooled virus-negative

nasopharyngeal swab matrix as an external positive control (PC) for the assay (see Methods for details), with pooled virus-negative nasopharyngeal swabs from healthy uninfected donors as the negative matrix serving as an external negative control (NC).

The SURPI+ computational pipeline, run as a container on either a server or cloud, was used for the identification of viral respiratory pathogens from mNGS data<sup>21,22</sup>. Three enhancements were made (Fig. 2A). First, we added the capability for viral load quantification using the PC and a standard curve generated for each sample from the ERCC. A standard curve is generated for each sample using the normalized ERCC results and absolute quantification by comparison of the ERCC data with the external PC. Second, “tagging” of Genbank accession numbers in the SURPI+ database was incorporated to allow inclusion of curated viral reference genomes, such as those deposited in the FDA-ARGOS database<sup>23</sup>, for virus identification by alignment and results reporting. Third, a custom algorithm consisting of de novo assembly of metagenomic reads and translated nucleotide or amino acid alignment of the reads to a viral protein database was developed to enable detection of novel, sequence-divergent viruses<sup>23</sup>.

Following clinical chart review, we investigated the correlation between viral load concentration, quantified in copies per milliliter (cp/mL) (Fig. 2B), and infection severity, which was categorized on a scale ranging from asymptomatic to mild, moderate, and severe (Supplementary Table 3). We observed significant differences in median viral loads between patients with asymptomatic or mild and moderately severe to severe infections ( $P < 0.001$  by the Mann-Whitney U test) (Fig. 2B, left). Further stratification of patients into asymptomatic, mild, moderate, and severe infections highlighted an increasing trend in viral load concentrations (Fig. 2B, right), with significant differences in median viral loads overall across all severity levels ( $P < 0.001$  by Kruskal-Wallis H test). Pairwise differences in median viral loads between asymptomatic or mild and moderately severe infections were also significant ( $P < 0.01$  by Mann-Whitney U test).

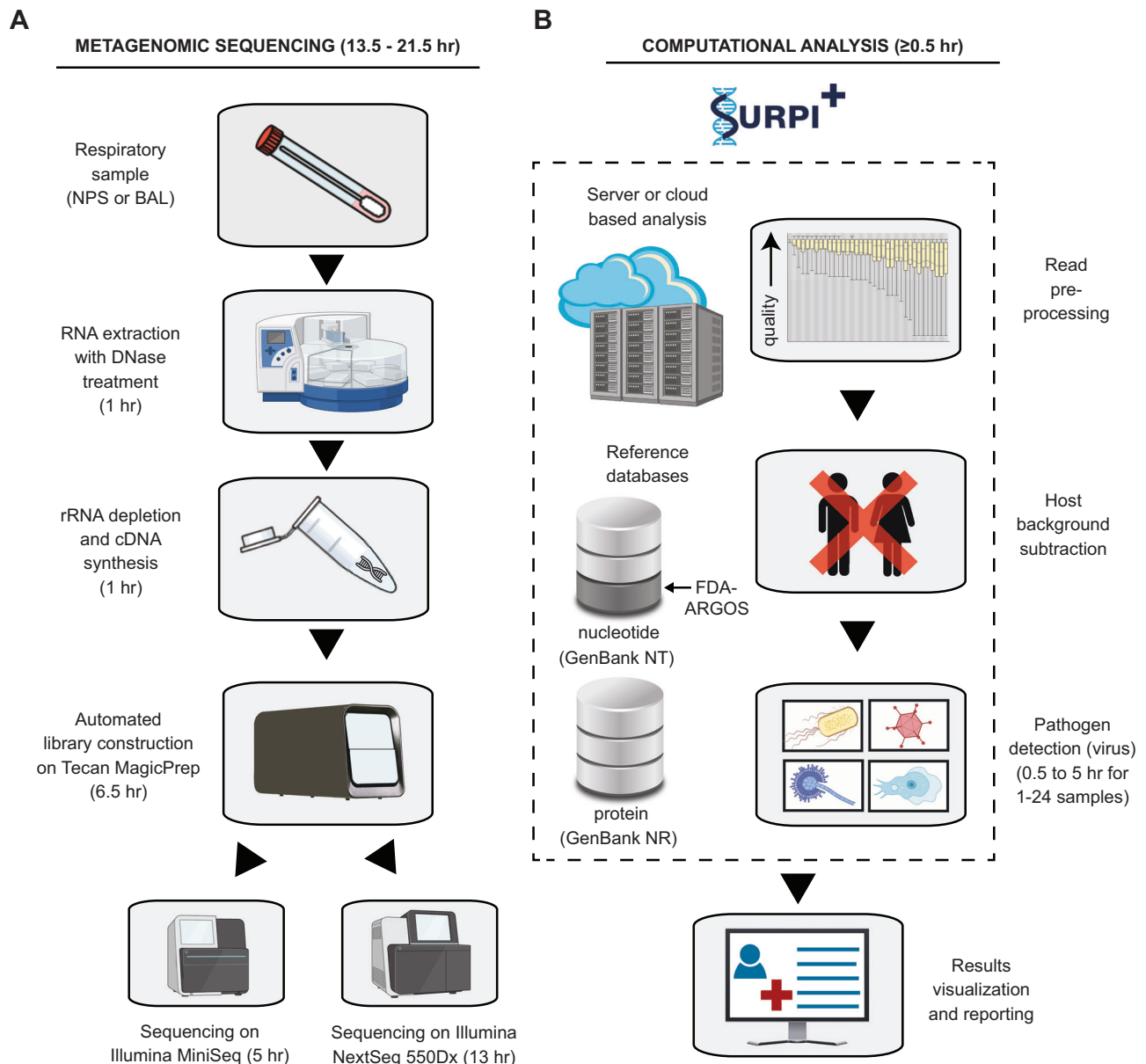
Quality control metrics were based on those previously established for a validated cerebrospinal fluid mNGS assay<sup>21</sup> and included a minimum of 5 million preprocessed reads per sample, >75% of data with quality score >30 ( $Q > 30$ ), and successful detection of the internal spiked MS2 phage control and all four respiratory viruses in the PC. A threshold criterion of  $\geq 3$  non-overlapping viral reads or contigs aligning to the target viral genome was considered a positive detection. Overall, 93% (156 of 167) of both positive ( $n = 111$ ) and negative ( $n = 56$ ) nasopharyngeal swab samples met QC metrics; those that did not meet QC metrics were excluded from the analysis.

### Analytical sensitivity

We adopted Clinical and Laboratory Standards Institute guidelines for NGS-based infectious diseases testing (MM24)<sup>24</sup> and validation of multiplex nucleic acid assays (MM17)<sup>25</sup> to conduct a comprehensive evaluation of assay performance metrics (Table 1). To determine limits of detection (LoD), negative nasopharyngeal swab matrix was spiked with the Accuplex Verification Panel and diluted at concentrations ranging from 5000 to 100 copies/mL, with 10 to 40 replicates at each concentration. By 95% probit analysis, the LoD was determined for each of the four representative organisms in the panel (SARS-CoV-2, Influenza A, Influenza B, and RSV). We found LoDs ranging from 439 to 706 copies/mL for the four respiratory viruses in the positive control (Fig. 3). The achieved average LoD of 550 copies/mL was comparable within one log to reported LoDs from specific reverse transcription-polymerase chain reaction (RT-PCR) assays for detection of viral respiratory pathogens<sup>26</sup>.

### Linearity

To evaluate the assay's capability to accurately quantitate viral load for detected viruses, a linearity panel was generated using five log



**Fig. 1 | Schematic of the mNGS assay workflow.** **A** RNA from respiratory samples is extracted and treated with DNase. Internal control is added to assess human background during sequencing. Human rRNA is depleted during cDNA synthesis. Libraries are generated on the automated Tecan MagicPrep NGS instrument. Libraries are normalized, pooled, and loaded onto the sequencer. **B** Sequences are processed using SURPI+ software for alignment and classification. Reads are pre-processed by trimming of adapters and removal of low-quality/low-complexity

sequences, followed by computational subtraction of human reads. Reads are mapped to the closest matched genome to identify non-overlapping regions using NCBI GenBank and FDA-ARGOS database. To aid in analysis, automated result summaries, heat maps of both raw and normalized read counts, and coverage and pairwise identity plots are generated for clinical interpretation. Total turnaround time is between 14 and 22 h depending on type of sequencer used.

dilutions of a quantified high-titer SARS-CoV-2 positive nasal swab sample and compared to a commercially available AccuSpan™ HCV RNA Linearity Panel. For both panels, the calculated linearity was 100% after running duplicates or triplicate replicates across a minimum of four 10-fold dilutions (Fig. 4). The absolute  $\log_{10}$  deviation of calculated from expected viral loads was  $<0.52 \log_{10}$ , which was favorable in comparison to the interquartile ranges for virus-specific qPCR assays between different laboratories<sup>27</sup>.

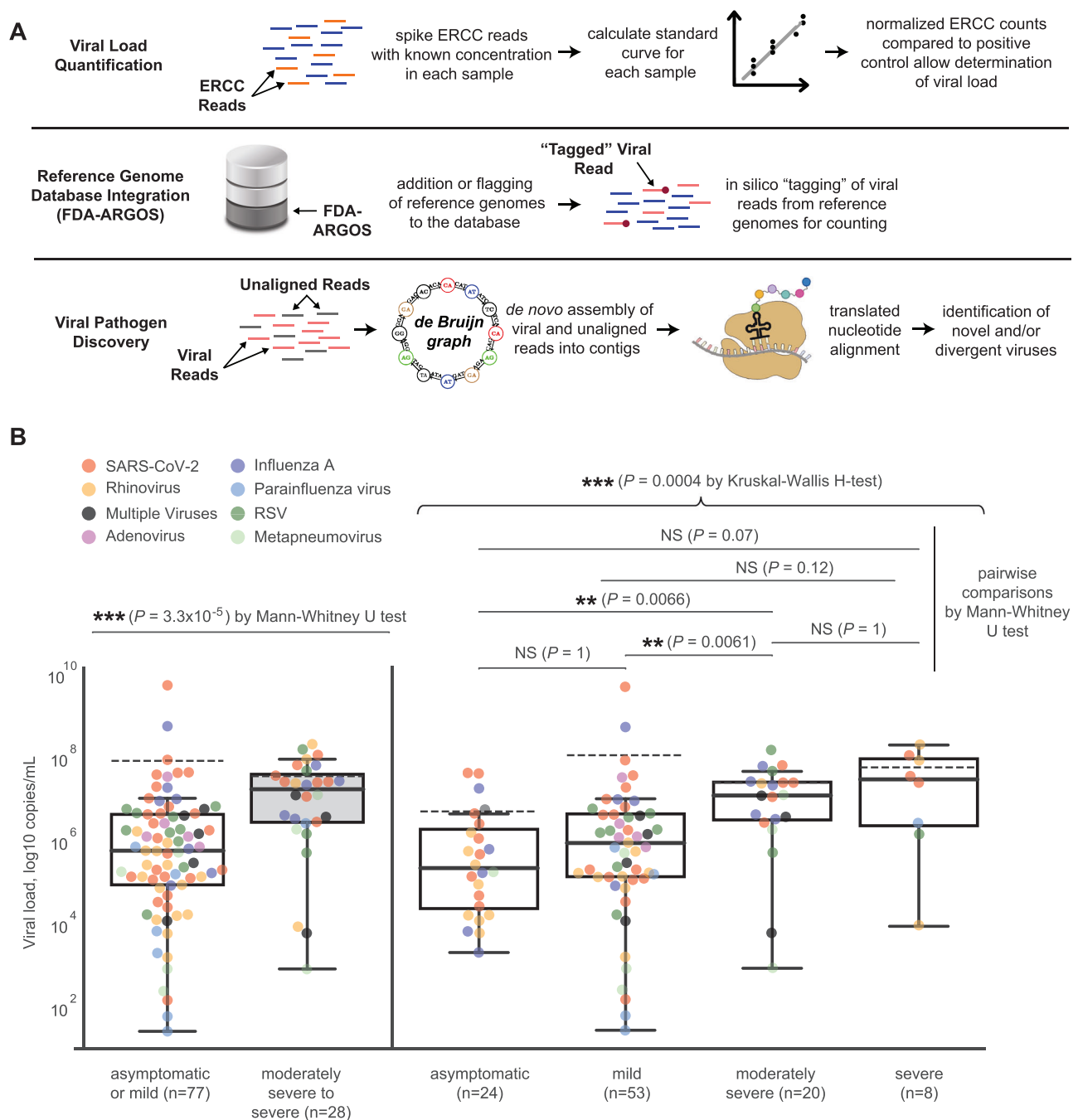
### Precision

We measured intra-assay precision by testing two PC and two NC samples within the same run using different barcodes across 20 runs and inter-assay precision by testing 20 PC and 20 NC samples using

different barcodes across 20 separate runs. Essential agreement (EA) was 100% and intra- and inter-assay precision were within our a priori established limits of  $<10\%$  and  $<30\%$  log-transformed coefficients of variation in reads per million, respectively (Table 1).

### Inclusivity and exclusivity

To evaluate the ability of the mNGS assay to detect a wide range of targets (inclusivity), we obtained commercially available culture supernatants from 17 respiratory viruses representing different sub-lineages and subspecies. Viruses were spiked into negative control matrix at concentrations ranging from  $1.3 \times 10^3$  to  $1.2 \times 10^7$  50% tissue culture infective dose (TCID<sub>50</sub>) per mL in a 1:10 ratio (Table 2). All 17 (100%) of 17 viruses in these contrived samples were correctly



**Fig. 2 | Enhancements to the SURPI+ bioinformatics pipeline for pathogen identification.** **A** Schematic diagram of modifications made to the SURPI+ bioinformatics pipeline to enhance its pathogen detection capabilities. The modifications include (1) calculation of the estimated viral load for each detected virus in the sample using a quantitative internal spiked ERCC control (top row), (2) incorporation of reference-grade databases such as the FDA-ARGOS database by “tagging” of GenBank accession numbers in the SURPI+ database (middle row), and (3) identification of novel, sequence-divergent viruses using de novo viral genome assembly and translated nucleotide (amino acid) alignments to a viral protein database (bottom row). **B** Pairwise and overall comparisons of viral load medians

among groups stratified by severity: asymptomatic ( $n = 24$ ), mild ( $n = 53$ ), moderately ( $n = 20$ ), or severe ( $n = 8$ ) respiratory infection. For the box and whiskers plots, the solid line within each box represents the median log viral load, while the dashed line indicates the mean log viral load. The interquartile range (IQR) is shown by the height of the box, with whiskers extending to the minimum and maximum values within 1.5 times the IQR. Each point corresponds to a detected virus, with different colors representing different virus species or genera. Mann-Whitney U and Kruskal-Wallis H tests were used for pairwise and overall significance testing, respectively. All tests were two-sided with Bonferroni correction for multiple comparisons, and the significance level was set at 0.05.

identified by mNGS assay at the sublineage or subspecies level. Additionally, we identified subtypes of rhinovirus and enterovirus from PCR-positive clinical samples that were not differentiated by multiplex RT-PCR (Fig. 5A). We also evaluated the ability of the mNGS assay to identify uncommon or rare viral pathogens associated with respiratory

infections ( $n = 8$  virus-positive tracheal aspirate samples) or central nervous system (CNS) infections ( $n = 4$  cerebrospinal fluid samples) in severely ill hospitalized patients (Table 2 and Fig. 5B). The assay detected 11 (100%) of 11 viruses in these samples. To assess the exclusivity of the mNGS assay, we spiked two mixtures of

**Table 1 | Performance characteristics of the UCSF viral respiratory mNGS assay**

Metrics	Method	Expected Target		Results	
Limit of detection (LoD)	Detection of PC dilution by probit analysis	<1000 copies/mL		Target SARS-CoV-2 Influenza A Influenza B RSV	LoD 439 copies/mL 706 copies/mL 493 copies/mL 563 copies/mL
Linearity	Correlation of PC with assay quantification	$R^2 > 90\%$		$R^2 = 100\%$	
Precision	Intra-Assay: PC and NC within the same run across 20 runs.	Concordance 100% EA	Log-transformed CV <10%	Concordance 100% EA	Log-transformed CV <10%
	Inter-Assay: PC and NC across 20 separate runs	100% EA	<30%	100% EA	<30%
Inclusivity	Detection of viruses from diluted culture supernatant	100% detection		100% detection (17/17)	
	Detection of viruses in positive BAL/CSF diluted samples	100% detection		100% detection (11/11)	
Exclusivity	Detection of viruses in known organism mixtures <sup>a</sup>	No false-positive		No false-positive	
Contamination	Detection of cross-contamination on the sample wells	No carryover contamination		Cross-contamination of 0.1% between adjacent wells but no carryover contamination	
Interference	Detection of PC spiked with hemolytic blood	Detection at all concentrations		Detection at all concentrations	
	Detection of PC spiked with Human RNA	Detection at all concentrations		Detection at all concentrations	
	Detection of PC spiked with bacterial DNA/RNA	Detection at concentration $\leq 10^7$ cells/mL		Detection at concentration $\leq 10^7$ cells/mL	
	Detection of virus-positive overtly mucoid BAL samples	Detection in all BAL samples		Target detected in 13/14 (92.9%) valid sample runs	
Stability	Detection of targets in samples held at 4 °C for 7 days or after 3 freeze-thaw cycles	100% concordance		100% concordance	
Accuracy	Detection in virus positive and negative samples ( $n = 191$ )	Sensitivity > 90% Specificity > 90% Accuracy > 90% PPA > 90% NPA > 90%		Original testing Sensitivity: 93.6% Specificity: 93.8 % Accuracy: 93.7 %	After discrepancy testing and clinical adjudication PPA: 98.7% NPA: 98.1% Overall: 97.9%
Detection of divergent viruses	Detection by an in silico analysis of divergent viruses ( $n = 70$ )	Sensitivity > 95% Specificity > 95%		Sensitivity: 98.6% Specificity: 100%	

PC positive control consisting of 4 respiratory viruses spiked into pooled nasopharyngeal swab matrix, IC spiked internal control consisting of a RNA MS2 phage, NC negative control, EA essential agreement, CV coefficient of variation, PPA positive percent agreement, NPA negative percent agreement.

<sup>a</sup>Two mixtures were assessed. The first mixture included detectable concentrations of CMV, HIV, *Klebsiella pneumoniae*, *Streptococcus agalactiae*, *Aspergillus niger*, *Cryptococcus neoformans* and *Toxoplasma gondii*, and corresponds to positive control material from a previously validated CSF assay<sup>21</sup>. The second mixture was a commercial reference panel, the ZymoBIOMICS Microbial Community Standard (Zymo Research, Tustin, CA), and consisted of 10 bacterial and fungal pathogens at varying concentrations (*Listeria monocytogenes*—12%, *Pseudomonas aeruginosa*—12%, *Bacillus subtilis*—12%, *Escherichia coli*—12%, *Salmonella enterica*—12%, *Lactobacillus fermentum*—12%, *Enterococcus faecalis*—12%, *Staphylococcus aureus*—12%, *Saccharomyces cerevisiae*—2%, and *Cryptococcus neoformans*—2%) that were spiked into negative nasopharyngeal swab matrix.

microorganisms, including a previously reported positive control mNGS panel consisting of 7 representative pathogens<sup>21</sup> and a commercial reference panel consisting of 10 bacterial and fungal species, into negative nasopharyngeal swab matrix and analyzed multiple aliquots (Table 1 and Supplementary Table 4). Detected reads from non-viral pathogenic organisms did not result in any false-positive detections for viral pathogens.

### Contamination, matrix effect, and stability

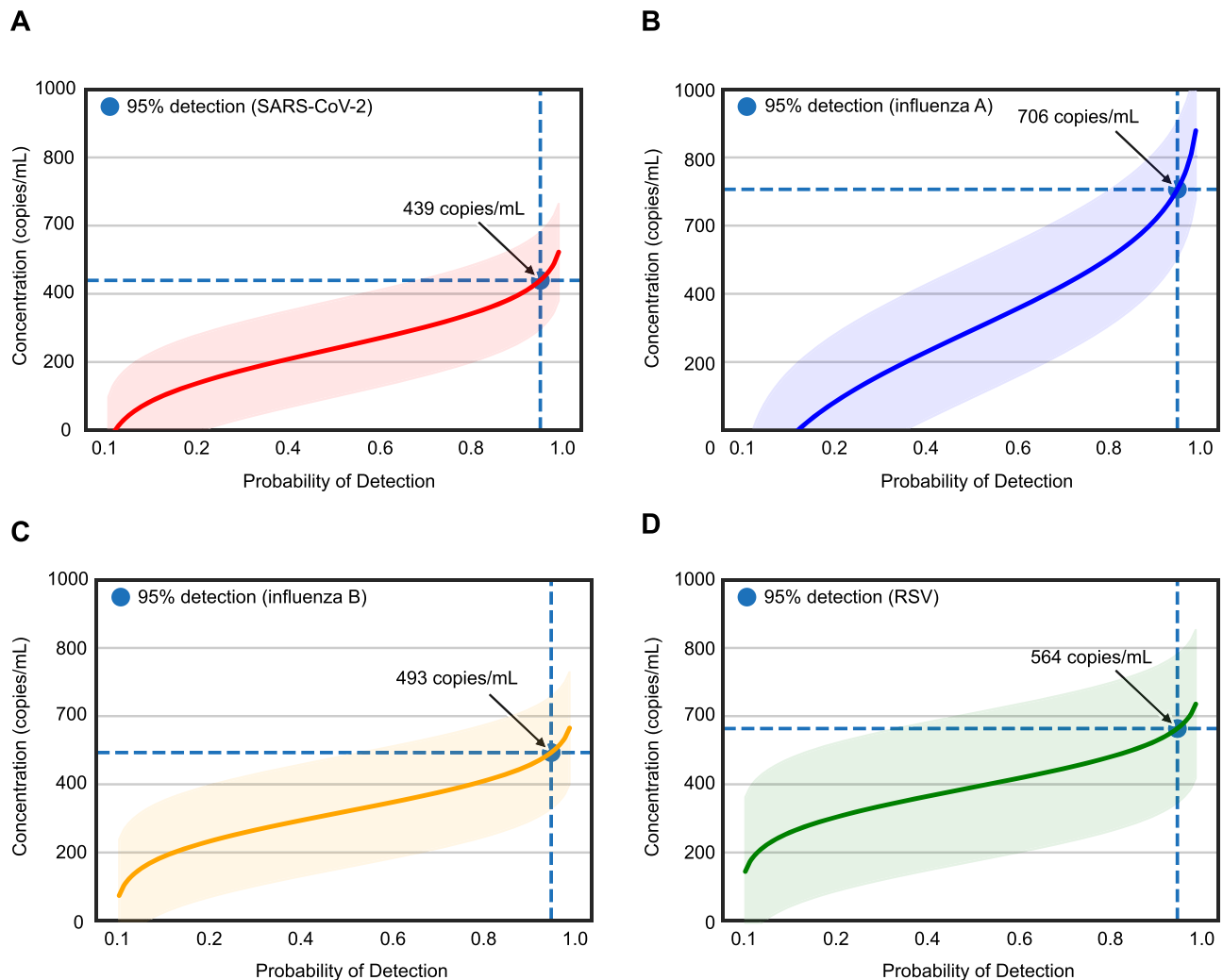
We evaluated potential cross-contamination between nearby sample wells and carryover contamination across successive runs from 10 SARS-CoV-2 high-titer clinical samples and 24 controls (cycle threshold, or  $C_t = 16$ –20) loaded in a modified checkerboard pattern (with at least one space between samples) on a 96-well plate, to mimic a single run on the Illumina NextSeq instrument. Only one possible cross-contamination event was observed, with a single SARS-CoV-2 read detected in one of the negative control wells at a subthreshold reporting level. We also evaluated the effects of interference from potential interfering substances, human RNA, and bacterial DNA/RNA on mNGS assay performance. Hemolysis, lipids, bilirubin, and human genomic RNA spiked into PC matrix at concentrations of 0.1–100  $\mu\text{g}/\text{mL}$  did not interfere with respiratory virus detection, but bacterial

DNA/RNA spiked into PC matrix at concentrations  $\geq 1 \times 10^7$  cells/mL resulted in failure to detect viruses due to high background. To evaluate the potential matrix effect from samples with high host background, we analyzed 14 PCR-positive highly mucoid bronchoalveolar lavage (BAL) samples obtained from lung transplant or cystic fibrosis patients undergoing surveillance bronchoscopy (Supplementary Table 5). All 14 samples had high host background, and 13 (92.9%) of 14 samples had very high host background. As a result, 6 (42.9%) of 14 samples had neither detection of the internal spiked MS2 phage control nor of a respiratory virus, and thus excluded from further analysis, as they not pass sequencing quality control criteria (Supplementary Table 1). The respiratory viral pathogen was detected in all (100%) of the remaining 8 samples. We concluded that highly mucoid samples can inhibit the assay due to high host background. Finally, we evaluated mNGS assay stability; qualitative detection was not affected by keeping samples for up to 7 days at 4 °C or subjecting the samples to 3 freeze/thaw cycles.

### Accuracy

To evaluate accuracy, 191 residual samples after routine clinical testing were obtained from the UCSF Clinical Microbiology Laboratory, including 110 virus-positive samples (104 upper respiratory swab





**Fig. 3 | Limits of detection (LoD) study.** Probit regression analysis curves plotting the viral titer in copies/mL (y-axis) against the calculated detection probability (x-axis) of (A) SARS-CoV-2, (B) influenza A, (C) influenza B and (D) respiratory syncytial virus (RSV). The regression curves and error bands (surrounding shaded areas) representing 95% confidence intervals for each curve as determined by probit

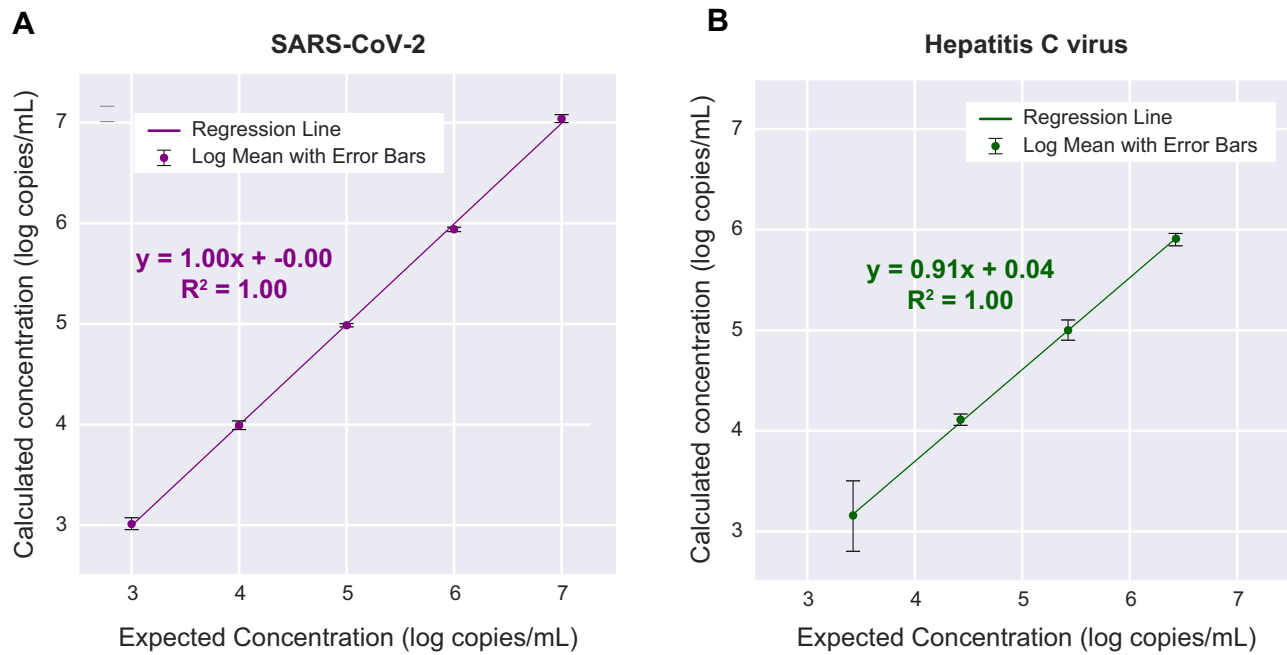
regression analysis are shaded in a different color for each virus. The probability of detection corresponding to 95% is denoted with a blue circle for each virus. Probit analyses were done using Python software (version 3.7.12). Results show a LoD ranging from 439 to 706 copies/mL for the 4 respiratory viruses in the positive control.

samples and 6 BAL fluids) from patients with acute respiratory infection (Supplementary Data 1), along with 81 virus-negative samples (52 upper respiratory swab samples and 29 BAL fluids) (Fig. 6). As more than one target may be positive with mNGS and respiratory viral multiplex panel (RVP) testing using FDA-approved in vitro diagnostic assays, sensitivity/specificity analyses were performed by assessing each result independently to assign true/false-positive/negative calls (see Methods for details). Compared to results from RVP RT-PCR testing, the mNGS assay exhibited 93.6% (103 of 110) sensitivity, 93.8% (76 of 81) specificity, and 93.7% (179 of 191) accuracy.

Discrepancy testing and clinical adjudication (DTCA) of 14 mNGS positive-RVP negative samples using blinded chart review by two board-certified infectious diseases physician (PB and CYC) and orthogonal assays run by the California Department of Public Health Viral and Rickettsial Disease Laboratory confirmed the presence of 9 respiratory viruses missed by RVP, allowing them to be reclassified as true positives (Supplementary Table 6). Viruses detected by mNGS but not targeted by RVP were not considered false-positive results. In one case, while the original RVP and orthogonal PCR testing returned negative results, mNGS identified rhinovirus C with high confidence. A

review of the viral sequences revealed 12 non-overlapping reads across the human rhinovirus C genome (Fig. 7A, B). Cross-contamination was ruled out, as no other sample in the sequencing batch tested positive for rhinovirus. A nucleotide BLAST (blastn) search confirmed sequences with high homology (95–98% identity) to known rhinovirus C strains. Although the exact primer binding sites for the clinical RT-PCR assays used in the current study are unknown, we identified, for the rhinovirus C sample, the presence of mismatches in primer and probe regions from previously reported RT-PCR assays targeting the 5'-untranslated region (UTR)<sup>28,29</sup> (Fig. 7C), which explained the detection by mNGS despite negative RT-PCR results.

Similarly, DTCA was performed on the 7 mNGS negative/RVP positive samples along with repeating the RVP assay (if possible, on a different instrument). This reassessment resulted in 5.5 samples being reclassified as true negatives (1 sample harbored two organisms adjudicated as one true negative and one false negative) (Supplementary Table 7). Compared to a composite standard that incorporates discrepancy testing and clinical adjudication, positive, negative, and overall predictive agreements of the mNGS assay were 98.7% (110.5 of 113), 98.1% (76.5 of 78), and 97.9% (187 of 191), respectively.



**Fig. 4 | Evaluation of linearity and viral load quantification for the mNGS assay.** **A** A quantified SARS-CoV-2 PCR-positive nasopharyngeal swab from a patient with COVID-19 was serially diluted in donor nasal swab matrix and tested across 4 log<sub>10</sub> dilutions. **B** A quantified HCV PCR-positive plasma sample from a patient with

hepatitis C infection was serially diluted in donor plasma and tested across 4 log<sub>10</sub> dilutions. At each dilution, the calculated mean concentration from three replicates is plotted against the expected concentration on a log scale, and the  $R^2$  correlation coefficient is determined by linear regression.

**Detection of novel, sequence-divergent viruses**

To benchmark the capability of the modified SURPI+ pipeline for detection of novel, highly divergent viruses in silico, we created a simulated sequencing output file containing many known human viral pathogens of clinical and public health significance, including those with pandemic potential (Fig. 8A). We then removed all viral reference

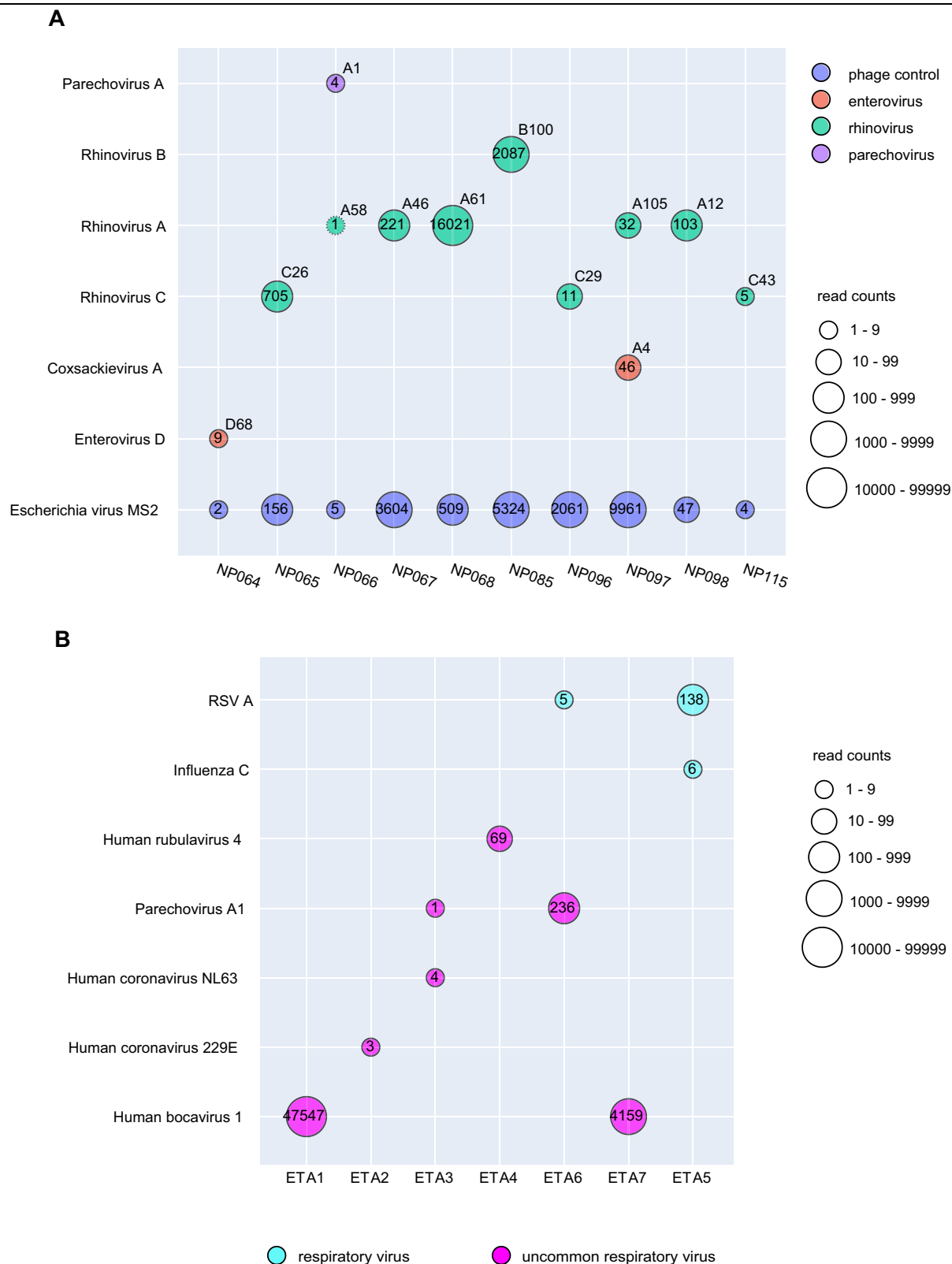
sequences of the same type (for example, all human polyomaviruses, coronaviruses, or parainfluenza viruses) or corresponding to the same genus or species from the SURPI+ 2019 reference database. Next, we used the SURPI+ pipeline to analyze the simulated sequencing file against both the original and “filtered” reference databases. In this analysis, 98.6% (69 of 70) of human viruses were detected at a sequencing depth of 100 reads per million (RPM) and 100% (70 of 70) at 1000 RPM based on homology to known animal or plant viruses (Fig. 8B). Of note, bunyaviruses pathogenic to humans, which are among the most divergent viruses, were still identified by translated nucleotide (amino acid) alignment to plant viruses (for example, detection of Venezuelan equine encephalitis virus based on homology to vanilla latent virus).

**Table 2 | Detection of a broad range of viruses in contrived samples**

Contrived Sample Type	Correctly Identified Virus by mNGS Assay	
Positive cerebrospinal fluid (CSF) spiked in negative matrix	Lymphocytic choriomeningitis virus (LCMV)	
	Herpes simplex virus 2 (HSV-2)	
	Varicella-zoster virus (VZV)	
	Herpes simplex virus 1 (HSV-1) and Epstein-Barr virus (EBV)	
Positive bronchoalveolar lavage (BAL) spiked in negative matrix	Parainfluenza virus Type 4	Parechovirus A
	Influenza C virus	Human bocavirus
	Primate bocaparvovirus 1	Coronavirus 229E
	Coronavirus NL63	
Viral culture fluid spiked in negative control matrix (1:10)	Adenovirus Type 1	Coronavirus 229E
	Coronavirus NL63	Coxsackie virus Type A1
	Echovirus	Human metapneumovirus 16
	Influenza B virus	Measles virus
	Mumps virus	Parainfluenza virus type 2
	Parainfluenza virus type 3	Parainfluenza virus type 4A
	Parechovirus type 1	Rhinovirus A16
	Rhinovirus B14	Rubella virus
	Influenza B virus	

**Discussion**

We validated a clinical mNGS assay in a CLIA laboratory as a Laboratory Developed Test (LDT) for agnostic viral respiratory pathogen detection intended to aid in patient diagnosis and public health surveillance. Our main goal was to develop, optimize, and streamline a protocol for respiratory viral mNGS testing that could be deployed and run routinely in clinical or public health laboratories. The mNGS assay developed here has favorable performance characteristics compared to clinical RVP testing, including a limit of detection of ~500 copies/mL, viral load quantification with 100% linearity, and sensitivity, specificity, and accuracy ranging from 93.6–93.8%. However, in contrast to targeted assays such as RVP, the mNGS assay is capable of detecting, in principle, all known as well as novel viral pathogens in respiratory samples. In addition, mNGS assay performance was found to be superior to RVP (97.9% versus 95.0% overall agreement) after discrepancy testing and clinical adjudication. The correlations we observed between viral load and disease severity highlight the potential for complementary quantitative viral load measurements to aid in distinguishing between asymptomatic infection or colonization and overt severe respiratory disease, thereby informing clinical management and treatment, as has been previously demonstrated for certain non-respiratory viruses such as CMV<sup>30</sup>. Following completion

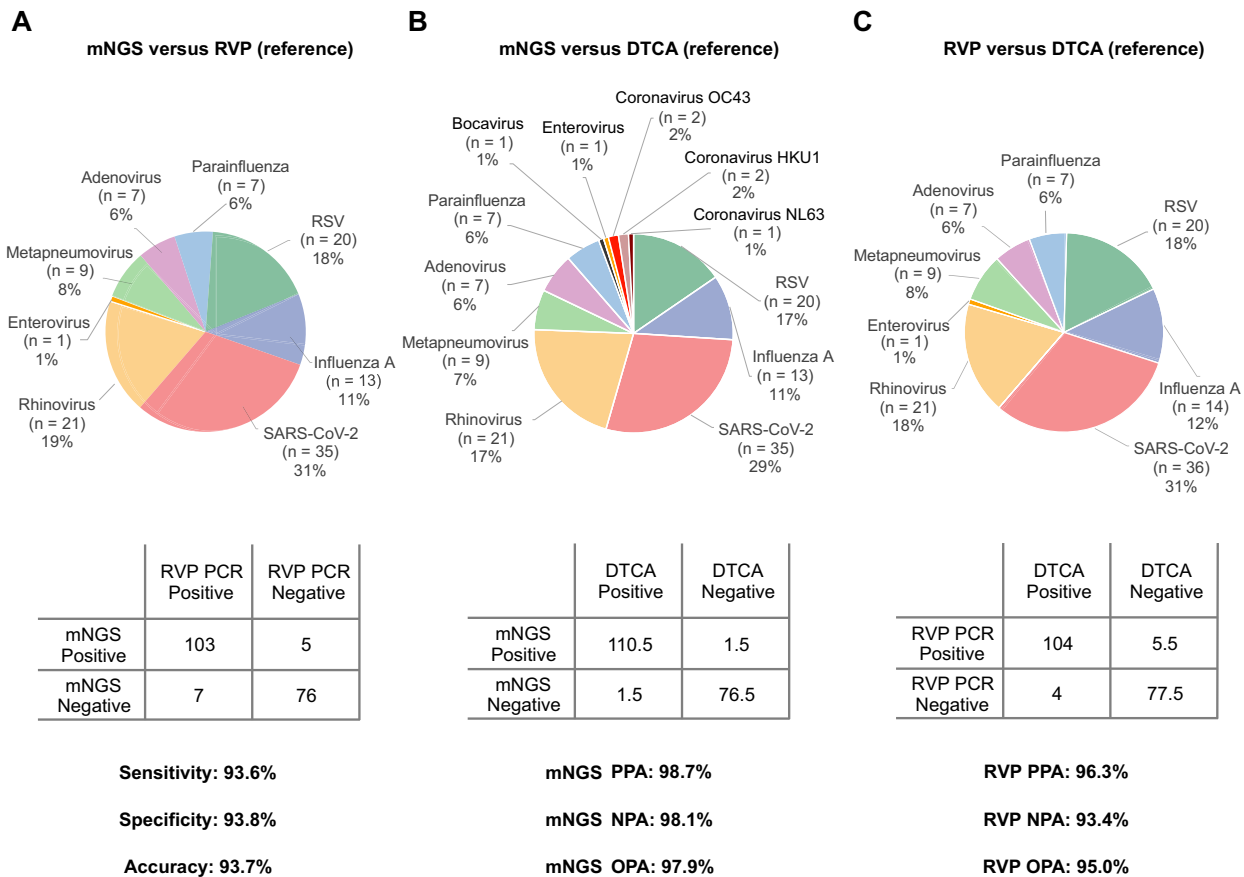


**Fig. 5 | Demonstration of inclusivity and clinical use cases for the mNGS assay.**

**A** Genotyping of rhinovirus and enterovirus subtypes from PCR-positive nasal swab samples. Conventional clinical multiplex RT-PCR tests do not distinguish between rhinoviruses and enteroviruses, nor are they able to subtype more pathogenic strains such as rhinovirus C or enterovirus D68 in association with acute flaccid myelitis<sup>34,35</sup>. **B** Detection of uncommon or rare viral pathogens causing respiratory

infections in critically ill mechanically ventilated hospitalized patients. The circles correspond to detected viruses and are color-coded by virus and scaled by read counts. For each detected virus, the read count is shown in the circle, while the identified genotype after SURPI+ pipeline is shown in the upper right quadrant. Abbreviations: ETA endotracheal aspirate.





**Fig. 6 | Accuracy evaluation for the mNGS assay.** Pie charts and 2 × 2 contingency tables showing the distribution of detected viruses and performance metrics.

**A** mNGS against RVP testing, **(B)** mNGS testing against DTCA, and **(C)** RVP testing against DTCA. RVP testing using FDA IVD assays includes detection of respiratory syncytial virus, parainfluenza viruses 1–3, metapneumovirus, rhinovirus/enterovirus, influenza A virus, influenza B virus, and adenovirus. Discrepant samples that were mNGS-positive/RVP-negative or mNGS-negative/RVP-positive underwent

orthogonal testing by targeted virus-specific PCR at the state public health laboratory and medical chart review for the most likely diagnosis by clinical adjudication. Abbreviations: mNGS metagenomic next-generation Sequencing, PCR polymerase chain reaction, RVP viral respiratory panel, DTCA discrepancy testing and clinical adjudication, PPA positive percent agreement, NPA negative percent agreement, OPA overall percent agreement, RSV respiratory syncytial virus, FDA Food and Drug Administration, IVD in vitro diagnostic.

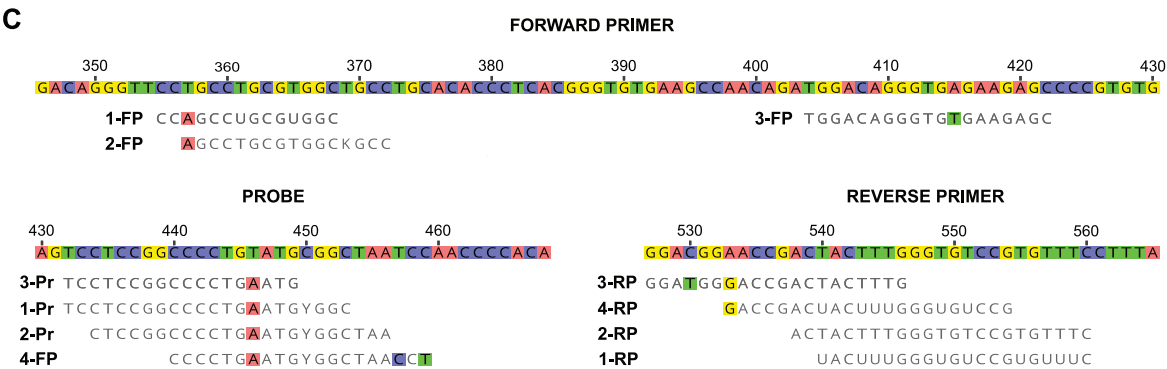
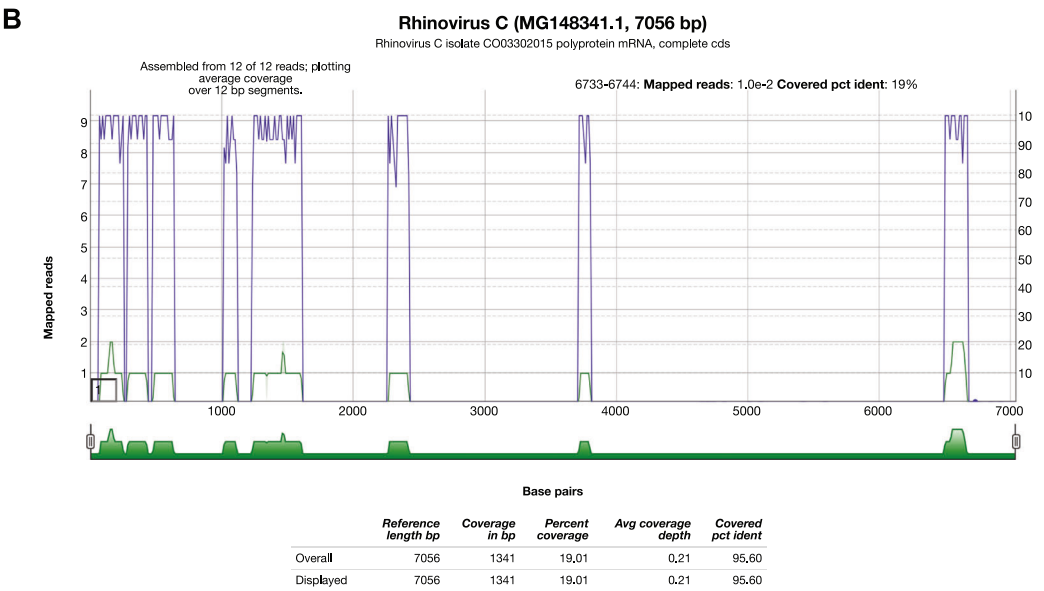
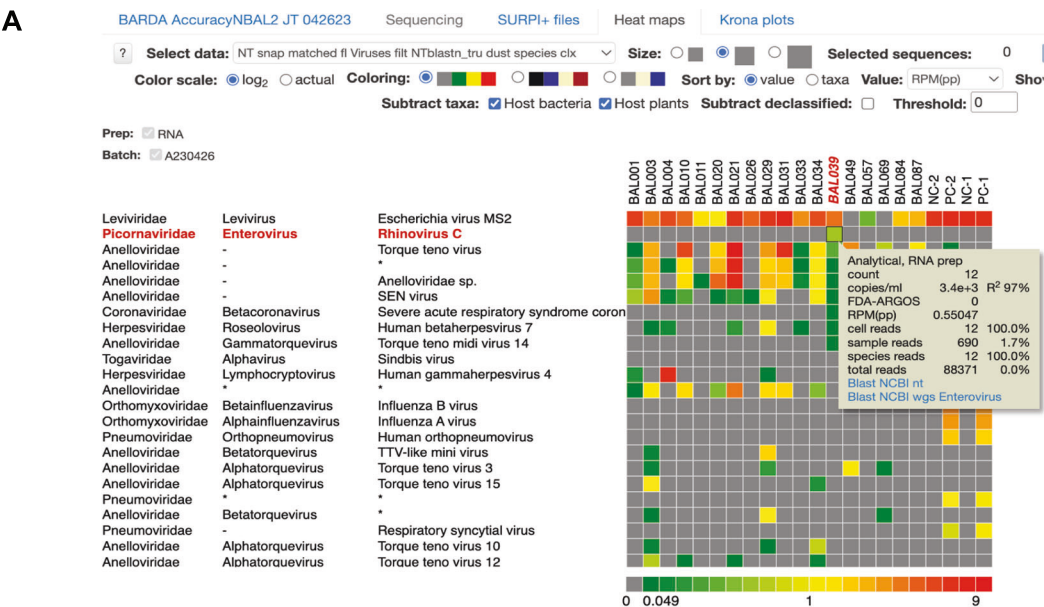
of the validation, our assay received breakthrough device designation from the US Food and Drug Administration (FDA) in August of 2023. Widespread implementation of highly accurate, rapid mNGS assays such as this, with enhanced capacity to detect novel viruses, will support robust preparation for and rapid responses to the next viral pandemic.

Speed is a critical factor for diagnosis of respiratory infections, especially in critically ill patients with lower respiratory involvement and in outbreak investigations of novel or emerging viruses with pandemic potential. Here we also aimed to develop an assay that could be deployable widely in clinical and public health laboratories. Thus, we optimized many of the steps of the mNGS assay and moved the key RNA/cDNA library preparation step to an automated platform, the MagicPrep NGS system (Tecan Genomics, Inc., Männedorf, Switzerland). We further demonstrated that sequencing can be performed on the Illumina MiniSeq using the Rapid Reagent Kit for a faster 5-h turnaround time or on the Illumina NextSeq 550Dx using the Mid-Output Reagent Kit for a 13-h turnaround time, depending on laboratory needs and priorities. All together, these modifications resulted in an assay with a turnaround time of 14–24 h and <2 h of hands-on technician time.

Orthogonal testing and clinical adjudication performed on discordant results demonstrated that the RVP assay is an imperfect gold standard with which to judge mNGS performance. The mNGS assay was able to not only detect uncommon infections from viruses not

covered on existing RVP panels, but also, in multiple cases, detect viruses that are detectable by RVP in principle but tested negative. Unlike RVP, mNGS does not rely on specific primers or probes and is hence less susceptible to primer failure due to viral evolution, as evidenced by the mNGS positive and RVP negative rhinovirus case presented here. Thus, RVP assay sensitivity will likely decrease over time by continual viral mutations, which is an inevitable feature of SARS-CoV-2 and many other RNA viruses<sup>31</sup>. Notably, a previous study evaluating the usefulness of published PCR primers in detecting rhinovirus infection reported that none of the published rhinovirus-specific PCR primer pairs could detect all human rhinoviruses in 101 genotyped clinical specimens<sup>32</sup>. In addition, broader sampling of the viral genome by mNGS may result in increased sensitivity of virus detection compared to RVP due to increased robustness to variability in the relative levels of viral gene expression in infected cells<sup>33</sup>. Most of the false-negative mNGS samples were confirmed as true negative after chart review and repeating the RVP assay. Most likely, these represented false-negative results during the original RVP run, either due to low viral titers associated with high cycle thresholds (>36) or degradation of samples over time and/or after multiple freeze-thaw cycles.

In the study, we used several approaches to demonstrate the capacity of the mNGS assay to identify novel and/or emerging viruses with divergent genomes. The assay was successful in detecting uncommon and unusual viral pathogens associated with both severe respiratory infections from bronchoalveolar lavage fluid and central



nervous infections from contrived CSF samples. mNGS testing also enabled subtyping of specific viral strains with increased virulence, such as enterovirus D68, which has been linked to acute flaccid myelitis in children<sup>34,35</sup>, and rhinovirus C, which has been associated with invasive pulmonary and bloodstream infection in immunocompromised patients<sup>36,37</sup>. Importantly, the mNGS assay was also able to detect DNA viruses, such as adenovirus and bocavirus, in both clinical and

contrived samples, despite the incorporation of DNase treatment in the protocol. Detection of DNA viruses is presumably based on detection of transcribed viral mRNA in infected cells, although may also be enabled by incomplete DNA digestion from the DNase enzyme. To evaluate the capacity for mNGS testing to identify novel viruses using a modified SURPI+ computational pipeline, we performed an in silico analysis of a contrived metagenomic dataset consisting of

**Fig. 7 | In-depth analysis of a rhinovirus C detection by mNGS that was discrepant with RT-PCR. A** A heat map generated from SURPI+ analysis shows 12 reads aligning to rhinovirus C from a single sample, excluding the possibility of cross-contamination. Each column denotes a clinical sample, while each row corresponds to a taxonomic identification at the species, genus, or family level. The asterisks refer to “declassification” of reads from one level to the next higher taxonomic level (for example, from species to genus). **B** A coverage map shows that the 12 reads span the genome of the most closely matched rhinovirus C genome in the reference database identified by SURPI+ (accession number MG148341.1)

reads from the genomes of human viruses of pandemic potential spiked into background using a reference database depleted of all known human viral sequences. This analysis was done to simulate whether “novel” human viruses with pandemic potential could be identified based on homology to known plant and animal viruses. All 70 of the human viral pathogens tested were successfully identified, including those with distant homology to other viruses. Indeed, chikungunya virus, in the *Alphavirus* genus of the *Togaviridae* family, was only identified after removal of all alphavirus sequences because of distant homology to vanilla latent virus in the family *Alphaflexiviridae*. Notably, alphaflexiviruses contain a distinct lineage of alphavirus-like replication proteins that lack a recognized protease domain<sup>38</sup>. These in silico results demonstrate that the pipeline is able to detect highly diverse viruses from families that are known to be potentially pathogenic to humans and that emerge from animal reservoirs (for example, *Bunyaviridae*, *Flaviviridae*, and *Adenoviridae*). If a novel, highly divergent virus from an uncharacterized family were detected, with little to no homology to any viral reference sequence, much more work would be needed to ascertain its clinical significance, or whether it is even capable of infecting humans, including formal assessment of Koch’s postulates with modifications by Rivers for causality<sup>39</sup>.

Our validation study has limitations. First, we tested very few bronchoalveolar lavage fluid samples from patients with acute respiratory infection ( $n=6$ ) and very few clinical samples harboring rare or unusual respiratory viruses ( $n=7$ ). Further validation of assay performance with these kinds of samples is needed. Second, mNGS testing was performed exclusively on samples from US patients, so viral pathogen diversity may not be representative of all populations globally. Third, we did not formally prove that the mNGS assay would be able to detect a novel, sequence-divergent virus, but instead demonstrated the ability of the test to detect such a virus using an in silico analysis, an approach which nonetheless has been used in previous studies to benchmark mNGS bioinformatic pipelines for viral pathogen discovery<sup>40,41</sup>. Finally, we did not address the utility of the mNGS assay for routine diagnosis in patients with unexplained infections or for outbreak surveillance in public health. Both efforts will likely require future prospective clinical and/or epidemiologic investigation.

In our study, the raw materials and labor costs for running the mNGS validation samples were ~\$300 USD per sample (Supplementary Table 8). However, this represents a lower limit for costs and does not account for costs related to assay implementation, bioinformatics analysis and director review, proficiency testing, quality and regulatory management, incomplete batch testing, the use of different sequencers (for example, NextSeq versus MiniSeq), and sample accessioning/reporting, among others. Thus, the actual costs for running the assay in clinical and/or commercial laboratories are much higher. In contrast, the estimated costs for running RVP assays in our clinical laboratory range from \$100–\$150 USD per sample. Nevertheless, the benefits for mNGS testing of greatly expanded scope of detection, capability to identify novel emerging viruses, and comparable performance likely outweigh the costs under certain clinical and public health scenarios. Further investigations that include cost-benefit analyses are needed to identify clinical use cases and indications for viral respiratory mNGS testing.

without overlap, with coverage of 19% of the ~7000 base pair (bp) genome.

**C** Several mismatches in the primer and probe sequences from published RT-PCR assays targeting the 5′-untranslated region (5′-UTR) are observed when compared to the viral mNGS reads, providing a likely explanation for the discrepant mNGS and RT-PCR results. The four assays are labeled 1 through 4 and correspond to Lu et al.<sup>29</sup> (1), Tapparel et al.<sup>47</sup> (2), Gunson et al.<sup>48</sup> (3), and Steininger et al.<sup>49</sup> (4). The mismatched nucleotides are highlighted with a background colour. Note that the assay from Steininger, et al. does not include a probe. Abbreviations: FP forward primer, Pr probe, RP reverse primer.

Even though the mNGS assay described here has exhibited high performance characteristics for sensitivity and specificity for the detection of viral pathogens, it is currently unlikely to replace multiplex RVP assays as a first-line test, as these panels are inexpensive and have more rapid turnaround times than mNGS. In addition, RVP assays easy to perform, with self-contained instrumentation that does not require batching and some platforms being CLIA-waived for use in point of care settings. However, mNGS testing could be particularly useful in public health laboratories that are more likely to receive and test samples from patients infected with unusual or novel viruses that are not part of the standard RVP testing panels. Of note, a modified protocol based on the assay was used to identify adeno-associated virus 2 in co-infections with adenoviruses and herpesviruses in cases of acute severe hepatitis in children as part of a nationwide US outbreak<sup>42</sup>. The respiratory mNGS assay developed here could also be implemented as a second-line test in clinical laboratories for patients with presumed viral bronchiolitis and pneumonia when RVP testing is negative. This strategy would be useful for diagnosis of rare and/or unexpected infections in immunocompromised patients or returning travelers, for whom there is a wider differential diagnosis.

## Methods

### Human sample collection

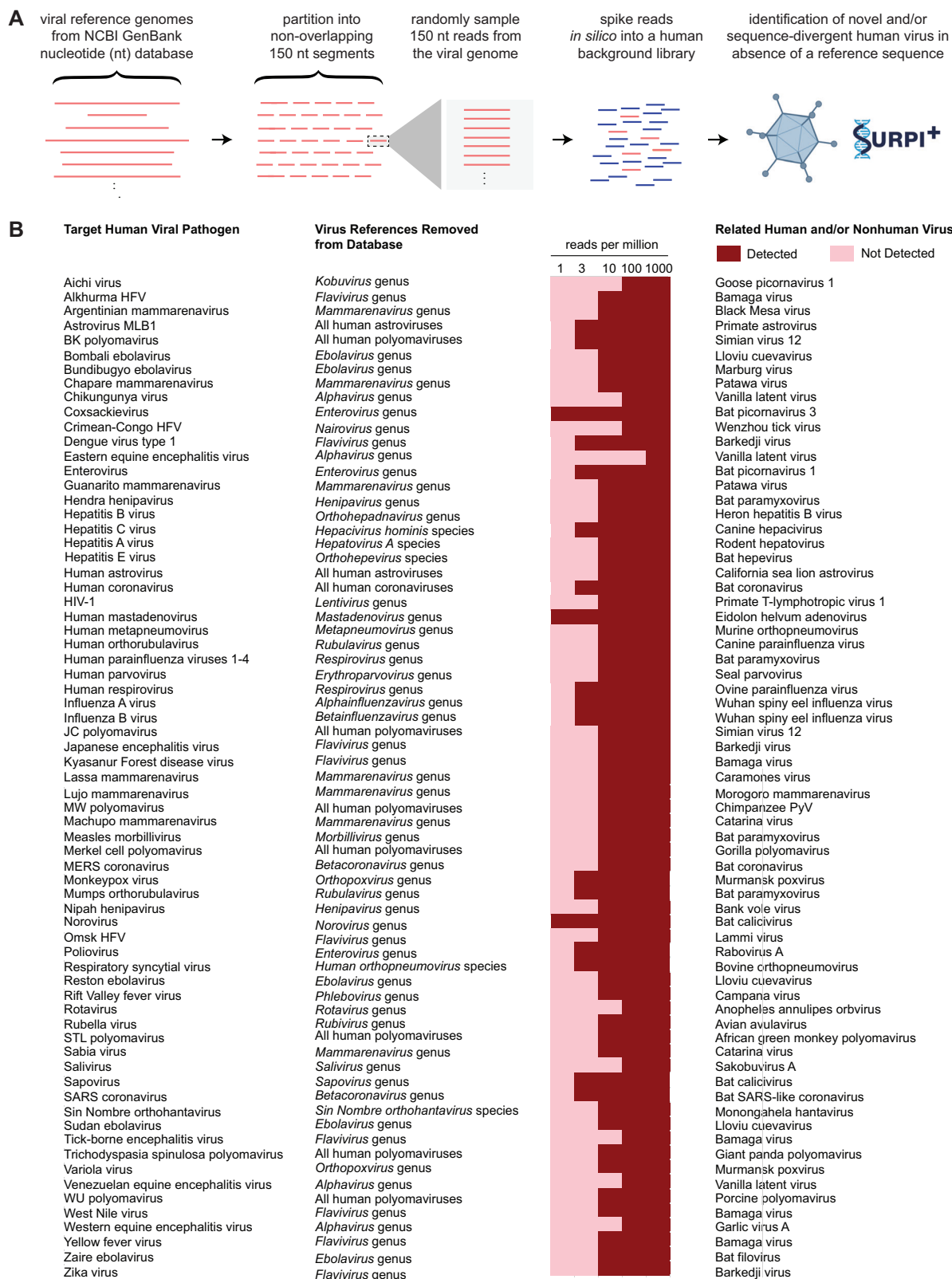
Residual laboratory-confirmed virus-positive upper respiratory swab or BAL samples from clinical patient testing were retrieved from the UCSF Clinical Microbiology Laboratory. Acceptable upper respiratory swab samples included (1) bilateral nasopharyngeal swabs, (2) bilateral anterior nares swabs, (3) oropharyngeal swabs, (4) combined nasopharyngeal and oropharyngeal swabs, and (5) combined oropharyngeal/mid-turbinate nasal swabs. All samples were required to meet minimal sample handling, storage, and volume requirements for inclusion in our study. Samples were stored at 4 °C for <24 h before being de-identified, aliquoted, and stored in -80 °C freezer prior to mNGS processing, thus undergoing one freeze-thaw cycle.

### Inclusion and ethics

All samples meeting minimal volume ( $\geq 450 \mu\text{L}$ ), sample handling (at most one freeze-thaw step), and storage (kept frozen at -80 °C) requirements were included in this study. Samples along with clinical and laboratory metadata were collected according to a biobanking protocol with waiver of consent approved by the UCSF Institutional Review Board (protocol no. 11-05519)

### External controls preparation

The external positive control (PC) was prepared by spiking a pooled negative nasal swab matrix with a commercially available reference material, the Accuplex Verification Panel (SeraCare, Milford, MA). This panel consisted of a mixture of non-infectious SARS-CoV-2, influenza A, influenza B, and RSV genomes encapsidated in a synthetic protein coat to mimic the structure of a viral capsid. This PC material was “spiked in” at a titer of  $10^4$  copies/mL for each virus control, 1–2 logs higher than the estimated limit of detection of the assay (~500 copies/mL). The negative matrix was prepared by pooling nasopharyngeal swab samples from asymptomatic individuals and was used as an external negative control (NC).



### Nucleic acid extraction

500  $\mu$ L of upper respiratory swab or BAL fluid was centrifuged at 16,000  $\times g$  for 10 min. The MagMAX<sup>™</sup> Viral/Pathogen II (MVP II) Nucleic Acid Isolation Kit (cat # A48383, Thermo Fisher Scientific, Waltham, MA) and the KingFisher<sup>™</sup> Flex Purification System with a 96 deep-well head (Thermo Fisher Scientific, Waltham, MA) were used for

total nucleic acid extraction. This protocol was modified to include DNase treatment using TURBO<sup>™</sup> DNase (cat # AM2238, Thermo Fisher Scientific, Waltham, MA) as a host depletion step during extraction. Bacteriophage MS2 (cat # 22-156-880, Zeptomatrix, Buffalo, NY) was added to all samples including the negative control as an internal qualitative control.



**Fig. 8 | In silico demonstration of novel, sequence-divergent virus detection using the mNGS assay.** **A** Representative viral reference genomes corresponding to outbreak viruses of clinical and public health significance with pandemic potential are retrieved from the NCBI GenBank database, partitioned into non-overlapping segments, and then randomly sampled and spiked in silico into a negative nasal swab matrix sequencing library. A higher-level set of taxonomic identifiers (species, genus, and/or family) corresponding to these viruses is removed from the SURPI+ reference dataset and the simulated sequencing file is analyzed using both the original and “restricted reference” databases. **B** Viruses can

be detected using the modified SURPI+ pipeline despite lacking a taxonomic reference at levels down to 10–100 reads per million (RPM). Abbreviations: EEEV Eastern equine encephalitis virus, ERCC External RNA Controls Consortium, FDA-ARGOS FDA dAtabase for Reference Grade microBial Sequences, HFV hemorrhagic fever virus, HIV human immunodeficiency virus, JCPyV JC polyomavirus, PC positive control, PyV polyomavirus, TSPyV trichodysplasia spinulosa polyomavirus, SURPI+ sequence-based ultrarapid pathogen identification, VEEV Venezuelan equine encephalitis virus, WEEV Western equine encephalitis virus.

### Library preparation and sequencing

Simultaneous reverse transcription of purified RNA, spiked in with ERCC RNA controls (cat # 4456740, Invitrogen, Waltham, MA), and ribosomal RNA (rRNA) depletion were carried out using NEBNext® Ultra™ II RNA First Strand Synthesis Module (cat #s E7771S/ E7771L, New England Biolabs, Ipswich, MA) and QIAseq FastSelect-rRNA HMR Kit (cat # 334385, Qiagen, Germantown, MD), respectively, followed by second strand cDNA synthesis using Sequenase™ Version 2.0 DNA Polymerase (cat # 70775Z1000UN, Thermo Fisher Scientific, Waltham, MA). Complementary DNA (cDNA) was purified using AMPure XP beads (cat # A63881, Beckman Coulter, Brea, CA) and loaded on the MagicPrep NGS instrument (Tecan Genomics, Inc., Männedorf, Switzerland) to undergo end-repair, adapter ligation, and barcoding, amplification (25 cycles) and purification using the DNA-Seq Mech kit (cat #s 30186627/30186628/30186629, Tecan Genomics, Inc., Männedorf, Switzerland). Libraries were quantified and normalized using the Qubit dsDNA HS Assay (cat # Q32854, Thermo Fisher Scientific, Waltham, MA) on the Qubit Flex (Thermo Fisher Scientific, Waltham, MA). Final pooled libraries were sequenced as single-end reads on either the Illumina (San Diego, CA) MiniSeq using the Rapid Reagent Kit (100 cycles) or on the Illumina NextSeq 550 using the Mid-Output or High-Output Kit (150 cycles).

### Bioinformatics

The SURPI+ computational pipeline, run as a container (v1.0.0) on either a secure server or cloud infrastructure, was used for identification of respiratory viral pathogens from mNGS data. Reads were pre-processed by trimming of adapters and removal of low-complexity and low-quality sequences, followed by computational subtraction of human reads. The Scalable Nucleotide Alignment Program<sup>43</sup> nucleotide aligner was run using an edit distance of 16 against the National Center for Biotechnology Information (NCBI) nucleotide (NT) database (March 2019, with inclusion of the SARS-CoV-2 WuHan-Hu-1 genome accession number NC\_045512), which was pre-filtered to retain only viral reads. The pipeline was modified to include “tagging”, or annotation, of entries from reference sequences that constitute a subset of the NCBI NT database, such as FDA-ARGOS<sup>23</sup>. Note that the FDA-ARGOS database, while quality controlled and regulated, contains only 1428 microbial strains, the majority of which are bacterial. It had also not been updated with recent viruses such as SARS-CoV-2; thus, this study did not detect any reads matching to viral genomes in FDA-ARGOS. The pipeline was modified to accommodate additional reference databases as needed such as GISAID<sup>44</sup>. The pipeline was also modified to use SPAdes (v3.15.4)<sup>45</sup> and DIAMOND (v2.0.15)<sup>46</sup>, respectively, for optional de novo assembly of reads into contiguous sequences (contigs) and translated nucleotide sequence alignment for identification of sequence-divergent viruses. Viral reads were identified using DIAMOND at a *e*-value cutoff of  $10^{-5}$ . Coverage maps were automatically generated by mapping SURPI+ -classified viral reads to the most likely reference genome.

Quality control metrics for the assay were based on those previously established for cerebrospinal fluid<sup>21</sup>, and include a minimum of 5 million preprocessed reads per sample, >75% of data with quality score >30 ( $Q > 30$ ), and successful detection of the 4 respiratory

viruses in the PC and the internal spiked MS2 phage control. A criterion of  $\geq 3$  non-overlapping viral reads or contigs aligning to the target viral genome was considered a positive detection.

### Evaluation of mNGS analytical performance characteristics

The automated standard operating procedures and sequencing runs for these clinical validation studies were performed by a California state-licensed clinical laboratory scientist. LoD was determined for each of the four representative organisms in the PC by probit analysis using a series of dilutions ranging from 100 to 5,000 copies/mL, with 10 to 40 replicates at each concentration. Linearity was demonstrated by plotting the standard curve. To validate the quantification using the ERCC and the positive control, we serially diluted an HCV positive plasma to known concentration ranging from  $4 \times 10^6$  to  $4 \times 10^3$  copies/mL in triplicate. We then compared the quantitative measure to the known measure. Precision was determined using repeat analysis of two PC and two NC samples across 20 runs (intra-assay reproducibility) and by testing 20 PC and 20 NC samples across 20 separate runs (inter-assay reproducibility). To assess inclusivity, commercially available cultured supernatants were obtained to assess the assay's ability to detect the intended targets. Each of the 17 respiratory viruses, with titers ranging from  $1.3 \times 10^4$  to  $1.2 \times 10^8$  TCID<sub>50</sub>/mL, were spiked into the negative control matrix at a 1:10 dilution. These viruses represented known sublineages and subspecies and we evaluated the ability of the assay to detect the virus. We also tested samples of confirmed virus-positive BAL ( $n=7$ ) and CSF samples ( $n=4$ ) spiked into negative matrix to evaluate assay performance with respect to detection of unusual viruses. To assess the exclusivity of the mNGS assay, we spiked a previously established mixture of seven representative pathogenic organisms to determine the false positive detection rate for viral pathogens. We evaluated cross-contamination between adjacent sample wells and carryover contamination across successive runs from samples with high viral loads. Interference was determined using PC spiked with known amounts of hemolytic blood, lipids, bilirubin, human RNA, and bacterial DNA/RNA. The effect of mucus in BAL positive fluids was also assessed. Stability was determined by keeping samples for up to 7 days at 4 °C or subjecting the samples to 3 freeze/thaw cycles. Accuracy was determined using 191 clinical samples comprising 110 virus-positive samples (103 upper respiratory swab samples and 7 BAL fluids) from patients with acute respiratory infection, along with 81 virus-negative samples (52 upper respiratory swab samples and 29 BAL fluids). Samples were obtained from patients at the University of California, San Francisco (UCSF). The viral RT-PCR comparator assays that were used include the Genmark ePlex (Carlsbad, CA), Luminex NxTAG (Austin, TX), and/or Luminex Verigene RP Flex Respiratory Pathogen Panels. mNGS results were compared with original clinical testing and then with a composite reference standard including discrepancy testing and clinical adjudication. In the second comparison, when results were discordant, orthogonal testing was performed using a different instrument or an independent CLIA laboratory (the California Department of Public Health) in addition to clinical adjudication to reclassify mNGS results. The second comparison was reported as positive percent agreement (PPA) and negative percent agreement (NPA), as selective discrepancy testing can bias sensitivity and specificity results.



## Orthogonal discrepancy testing at the California Department of Public Health

Specimens were tested by real-time PCR based on CDC protocols using a viral respiratory panel, an unpublished CDPH laboratory-developed test (LDT). Viruses that can be detected by this panel included human metapneumovirus, respiratory syncytial virus, adenovirus, parainfluenza virus (types 1, 2, 3, and 4), enterovirus/rhinovirus, and human coronaviruses 229E, OC43, NL63, and HKU1.

## In silico analysis for identification of novel, sequence-divergent viruses using the SURPI+ pipeline

To assess detection capability for novel, sequence-divergent viruses, an in silico analysis was performed. Representative viral reference genomes corresponding to outbreak viruses of clinical and public health significance with pandemic potential were retrieved from the NCBI GenBank database, partitioned into non-overlapping segments, and then randomly sampled and spiked in silico into a negative nasal swab matrix sequencing library. We then took a higher-level set of taxonomic identifiers (species, genus, and/or family) corresponding to these viruses and removed all entries with these taxonomic identifiers from the SURPI+ reference dataset. Next, we used the SURPI+ pipeline to analyze the simulated sequencing file against both the original and “restricted reference” databases and evaluated the performance of the pipeline in detecting “simulated” novel and/or divergent viruses that lacked a reference sequence.

## Statistical analyses

Statistical analyses were performed using scipy (version 1.5.3) and rstatix (version 0.7.0) packages as implemented in Python (version 3.7.12) and R (version 4.0.3), respectively. The non-parametric Mann-Whitney U test was used for pairwise comparisons of viral load medians, while the Kruskal-Wallis H test was used for comparisons of medians across all severity groups. Probit regression analyses were done using scipy (version 1.5.3), numpy (version 1.19.1), and statsmodels (version 0.12.2) as implemented in Python software (version 3.7.12).

Sensitivity and specificity analyses were performed as follows: as more than one target may be positive with mNGS and RVP, each result was independently assessed in every sample and true/false-negative/positive were accordingly assigned to each result. However, the total number of observations was kept constant (one sample = one observation = 1). For instance, in the case a test detected two organisms, namely the culprit pathogen and a contaminant, the former was assigned 0.5 true-positive and the latter 0.5 false-positive, such that their sum was always equal to 1. In addition, as we used RVP as a comparator which included only a limited number of targets, mNGS positive-RVP negative results that were not a target for the RVP were not considered as false-positive results.

## Reporting summary

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

## Data availability

Human-subtracted raw sequence data were submitted to the Sequence Read Archive (SRA) database. (BioProject accession number PRJNA1084017 and umbrella BioProject accession number PRJNA171119). Source data are provided as a Source Data file. Sequence metadata is available in a Zenodo data repository (<https://zenodo.org/doi/10.5281/zenodo.10553378>). Source data are provided with this paper.

## Code availability

Custom scripts and code for data analyses and visualization are available in a Zenodo data repository (<https://zenodo.org/doi/10.5281/zenodo.10553378>). The SURPI+ bioinformatics pipeline is described in

prior publications<sup>21,22</sup>. The code for SURPI+ includes proprietary algorithms for taxonomic classification, filtering, and pathogen software that have been filed under US patent 11380421, “Pathogen detection using next generation sequencing”. Please contact the University of California Office of Technology Management regarding access to and use of the software.

## References

- DALYs, G. B. D. et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. *Lancet* **386**, 2145–2191 (2015).
- Jain, S. et al. Community-acquired pneumonia requiring hospitalization among U.S. adults. *N. Engl. J. Med.* **373**, 415–427 (2015).
- Jain, S. et al. Community-acquired pneumonia requiring hospitalization among U.S. children. *N. Engl. J. Med.* **372**, 835–845 (2015).
- Musher, D. M. & Thorner, A. R. Community-acquired pneumonia. *N. Engl. J. Med.* **371**, 1619–1628 (2014).
- Charlton, C. L. et al. Practical guidance for clinical microbiology laboratories: viruses causing acute respiratory tract infections. *Clin. Microbiol. Rev.* **32**, <https://doi.org/10.1128/CMR.00042-18> (2019).
- Evans, S. E. et al. Nucleic Acid-based Testing For Noninfluenza Viral Pathogens In Adults With Suspected Community-acquired Pneumonia. An Official American Thoracic Society Clinical Practice Guideline. *Am. J. Respir. Crit. Care Med.* **203**, 1070–1087 (2021).
- Jain, S. Epidemiology of viral pneumonia. *Clin. Chest Med.* **38**, 1–9 (2017).
- Schlaberg, R. et al. Viral pathogen detection by metagenomics and pan-viral group polymerase chain reaction in children with pneumonia lacking identifiable etiology. *J. Infect. Dis.* **215**, 1407–1415 (2017).
- Jones, K. E. et al. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Lu, R. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
- Chiu, C. Y. & Miller, S. A. Clinical metagenomics. *Nat. Rev. Genet.* **20**, 341–355 (2019).
- Simner, P. J., Miller, S. & Carroll, K. C. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* **66**, 778–788 (2018).
- Blauwkamp, T. A. et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* **4**, 663–674 (2019).
- Gaston, D. C. et al. Evaluation of metagenomic and targeted next-generation sequencing workflows for detection of respiratory pathogens from bronchoalveolar lavage fluid specimens. *J. Clin. Microbiol.* **60**, e0052622 (2022).
- Wilson, M. R. et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* **380**, 2327–2340 (2019).
- Lee, R. A., Al Dhaheri, F., Pollock, N. R. & Sharma, T. S. Assessment of the clinical utility of plasma metagenomic next-generation sequencing in a pediatric hospital population. *J. Clin. Microbiol.* **58**, <https://doi.org/10.1128/JCM.00419-20> (2020).
- Han, D. et al. The real-world clinical impact of plasma mNGS testing: an observational study. *Microbiol. Spectr.* **11**, e0398322 (2023).
- Miller, S. & Chiu, C. The role of metagenomics and next-generation sequencing in infectious disease diagnosis. *Clin. Chem.* **68**, 115–124 (2021).
- Benoit, P. et al. Metagenomic next-generation sequencing of cerebrospinal fluid for diagnosis of central nervous system infections:

- 7-year performance of a clinically validated test. *medRxiv*, <https://doi.org/10.1101/2024.03.14.24304139> (2024).
21. Miller, S. et al. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res* **29**, 831–842 (2019).
  22. Naccache, S. N. et al. A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res* **24**, 1180–1192 (2014).
  23. Sichtig, H. et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. *Nat. Commun.* **10**, 3313 (2019).
  24. Clinical Laboratory Standards Institute. *Molecular Methods for Genotyping and Strain Typing of Infectious Organisms* 1st edn, Vol. 24 (ed C. A. L. S. Institute) (Clinical and Laboratory Standards Institute, 2021).
  25. Clinical Laboratory Standards Institute. *Validation and Verification of Multiplex Nucleic Acid Assays* 2nd edn, Vol. 9 (ed C. A. L. S. Institute) (Clinical and Laboratory Standards Institute, 2018).
  26. Espy, M. J. et al. Real-time PCR in clinical microbiology: applications for routine laboratory testing. *Clin. Microbiol. Rev.* **19**, 165–256 (2006).
  27. Hayden, R. T. et al. Progress in quantitative viral load testing: variability and impact of the WHO quantitative international standards. *J. Clin. Microbiol.* **55**, 423–430 (2017).
  28. Andeweg, A. C., Bestebroer, T. M., Huybregths, M., Kimman, T. G. & de Jong, J. C. Improved detection of rhinoviruses in clinical samples by using a newly developed nested reverse transcription-PCR assay. *J. Clin. Microbiol.* **37**, 524–530 (1999).
  29. Lu, X. et al. Real-time reverse transcription-PCR assay for comprehensive detection of human rhinoviruses. *J. Clin. Microbiol.* **46**, 533–539 (2008).
  30. Razonable, R. R. & Hayden, R. T. Clinical utility of viral load in management of cytomegalovirus infection after solid organ transplantation. *Clin. Microbiol. Rev.* **26**, 703–727 (2013).
  31. Clark, C., Schrecker, J., Hardison, M. & Taitel, M. S. Validation of reduced S-gene target performance and failure for rapid surveillance of SARS-CoV-2 variants. *PLoS ONE* **17**, e0275150 (2022).
  32. Faux, C. E. et al. Usefulness of published PCR primers in detecting human rhinovirus infection. *Emerg. Infect. Dis.* **17**, 296–298 (2011).
  33. Russell, A. B., Trapnell, C. & Bloom, J. D. Extreme heterogeneity of influenza virus infection in single cells. *Elife* **7**, <https://doi.org/10.7554/eLife.32303> (2018).
  34. Greninger, A. L. et al. A novel outbreak enterovirus D68 strain associated with acute flaccid myelitis cases in the USA (2012–14): a retrospective cohort study. *Lancet Infect. Dis.* **15**, 671–682 (2015).
  35. Messacar, K. et al. Enterovirus D68 and acute flaccid myelitis—evaluating the evidence for causality. *Lancet Infect. Dis.* **18**, e239–e247 (2018).
  36. Lupo, J. et al. Disseminated rhinovirus C8 infection with infectious virus in blood and fatal outcome in a child with repeated episodes of bronchiolitis. *J. Clin. Microbiol.* **53**, 1775–1777 (2015).
  37. Sayama, A. et al. Comparison of rhinovirus A-, B-, and C-associated respiratory tract illness severity based on the 5'-untranslated region among children younger than 5 years. *Open Forum Infect. Dis.* **9**, ofac387 (2022).
  38. Kreuze, J. F. et al. ICTV virus taxonomy profile: alphaflexiviridae. *J. Gen. Virol.* **101**, 699–700 (2020).
  39. Guo, C. & Wu, J. Y. Pathogen discovery in the post-COVID era. *Pathogens* **13**, <https://doi.org/10.3390/pathogens13010051> (2024).
  40. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).
  41. Flygare, S. et al. Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol.* **17**, 111 (2016).
  42. Servellita, V. et al. Adeno-associated virus type 2 in US children with acute severe hepatitis. *Nature* **617**, 574–580 (2023).
  43. Zaharia, M. et al. Alignment in a SNAP: cancer diagnosis in the genomic age. *Lab. Investig.* **92**, 458a–458a (2012).
  44. Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data—from vision to reality. *Euro Surveill* **22**, <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494> (2017).
  45. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
  46. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
  47. Tapparel, C. et al. New respiratory enterovirus and recombinant rhinoviruses among circulating picornaviruses. *Emerg. Infect. Dis.* **15**, 719–726 (2009).
  48. Gunson, R. N., Collins, T. C. & Carman, W. F. Real-time RT-PCR detection of 12 respiratory viral infections in four triplex reactions. *J. Clin. Virol.* **33**, 341–344 (2005).
  49. Steininger, C., Aberle, S. W. & Popow-Kraupp, T. Early detection of acute rhinovirus infections by a rapid reverse transcription-PCR assay. *J. Clin. Microbiol.* **39**, 129–133 (2001).

## Acknowledgements

We thank the staff at the UCSF Clinical Microbiology Laboratory for help in collecting nasopharyngeal swab and bronchoalveolar lavage fluid samples. This work was financially supported in part by BARDA EZ-BAA award 75A50122C00022 (C.Y.C.), US CDC grants 75D30122C15360 and 75D30121C12641 (C.Y.C.), Abbott Laboratories (C.Y.C.), and the Chan-Zuckerberg Biohub (C.Y.C.). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication. The content of this paper is solely the responsibility of the authors and does not represent the official views or opinions of the National Institutes of Health, Kaiser Permanente, California Department of Public Health or the California Health and Human Services Agency. Use of trade names and commercial sources is for identification only and does not imply endorsement by the California Department of Public Health or the California Health and Human Services Agency. Figures 1A, B, 2A, and 8A were created in part using images from BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International (CC-BY-NC-ND) license.

## Author contributions

C.Y.C. conceived of and designed the study. J.K.T., V.S., D.S., J.S., N.S., A.F., H.J.H., J.N., M.O., N.B., J.T., D.I., B.F., H.R., M.H., C.M., D.A.W. and C.Y.C. coordinated the sequencing efforts and laboratory studies. J.K.T., A.C., H.G. and S.Y. processed samples. J.K.T., V.S., D.S., E.K., A.C., H.G., S.Y., M.D.L., P.B. and C.Y.C. analyzed data. J.K.T., J.S., N.S., A.F., J.N., M.O., P.M.M. and C.R.L. collected samples. J.K.T., V.S., E.K., P.B., M.D.L. and C.Y.C. wrote the manuscript. J.K.T., V.S., E.K., P.B., and C.Y.C. prepared the figures. J.K.T., V.S., D.S., E.K., N.S., A.F., H.J.H., J.N., M.O., N.B., J.T., D.I., B.F., H.R., M.H., D.A.W., P.M.M., C.R.L., M.D.L., P.B. and C.Y.C. edited the manuscript. J.K.T., V.S., E.K., M.D.L., P.B. and C.Y.C. revised the manuscript. All authors read the manuscript and agree to its contents.

## Competing interests

C.Y.C. is a founder of Delve Bio and on the scientific advisory board for Delve Bio, Flightpath Biosciences, Biomeme, Mammoth Biosciences, BiomeSense and Poppy Health. He is also an inventor on US patent 11380421, “Pathogen detection using next generation sequencing”, under which algorithms for taxonomic classification, filtering, and pathogen detection are used by SURPI+ software. C.Y.C. receives research support from Delve Bio and Abbott Laboratories, Inc. The other authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51470-y>.

**Correspondence** and requests for materials should be addressed to Charles Y. Chiu.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons 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) 2024

Jessica Karielle Tan<sup>1,2,8</sup>, Venice Servellita<sup>1,2,8</sup>, Doug Stryke<sup>1,2,8</sup>, Emily Kelly<sup>1</sup>, Jessica Streithorst<sup>1</sup>, Nanami Sumimoto<sup>1,2</sup>, Abiodun Foresythe<sup>1,2</sup>, Hee Jae Huh<sup>1,2,3</sup>, Jenny Nguyen<sup>1,2</sup>, Miriam Oseguera<sup>1,2</sup>, Noah Brazer<sup>1,2</sup>, Jack Tang<sup>1,2</sup>, Danielle Ingebrigtsen<sup>1</sup>, Becky Fung<sup>1</sup>, Helen Reyes<sup>1</sup>, Melissa Hillberg<sup>1</sup>, Alice Chen<sup>4</sup>, Hugo Guevara<sup>4</sup>, Shigeo Yagi<sup>4</sup>, Christina Morales<sup>4</sup>, Debra A. Wadford<sup>4</sup>, Peter M. Mourani<sup>5</sup>, Charles R. Langelier<sup>6,7</sup>, Mikael de Lorenzi-Tognon<sup>1,2</sup>, Patrick Benoit<sup>1,2</sup> & Charles Y. Chiu<sup>1,2,6,7</sup> ✉

<sup>1</sup>Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, USA. <sup>2</sup>Abbott Pandemic Defense Coalition, Abbott Park, IL, USA. <sup>3</sup>Department of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. <sup>4</sup>Viral and Rickettsial Disease Laboratory, Center for Laboratory Sciences, California Department of Public Health, Richmond, CA, USA. <sup>5</sup>Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, AR, USA. <sup>6</sup>Division of Infectious Diseases, Department of Medicine, University of California San Francisco, San Francisco, CA, USA. <sup>7</sup>Chan-Zuckerberg Biohub, San Francisco, CA, USA. <sup>8</sup>These authors contributed equally: Jessica Karielle Tan, Venice Servellita, Doug Stryke. ✉ e-mail: [charles.chiu@ucsf.edu](mailto:charles.chiu@ucsf.edu)