

1  
2  
3  
4  
5 **COV-ID: A LAMP sequencing approach for high-throughput co-**  
6 **detection of SARS-CoV-2 and influenza virus in human saliva**  
7  
8  
9

10  
11 Robert Warneford-Thomson<sup>1,2,3</sup>, Parisha P. Shah<sup>3,4</sup>, Patrick Lundgren<sup>5</sup>, Jonathan Lerner<sup>3</sup>,  
12 Benjamin S. Abella<sup>6</sup>, Kenneth S. Zaret<sup>2,3</sup>, Jonathan Schug<sup>7</sup>, Rajan Jain<sup>3,4</sup>, Christoph A. Thaiss<sup>5</sup>,  
13 and Roberto Bonasio<sup>2, 3\*</sup>.

14  
15  
16 <sup>1</sup>Graduate Group in Biochemistry and Biophysics, University of Pennsylvania Perelman School of  
17 Medicine, Philadelphia, PA, USA

18 <sup>2</sup>Epigenetics Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA,  
19 USA

20 <sup>3</sup>Department of Cell and Developmental Biology, University of Pennsylvania Perelman School of  
21 Medicine, Philadelphia, PA, USA

22 <sup>4</sup>Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia,  
23 PA, USA

24 <sup>5</sup>Department of Microbiology, University of Pennsylvania Perelman School of Medicine,  
25 Philadelphia, PA, USA

26 <sup>6</sup>Department of Emergency Medicine and Penn Acute Research Collaboration, University of  
27 Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

28 <sup>7</sup>Next-Generation Sequencing Core, Department of Genetics, University of Pennsylvania  
29 Perelman School of Medicine, Philadelphia, PA, USA

30  
31  
32 \*Lead contact, to whom correspondence should be addressed: [roberto@bonasiolab.org](mailto:roberto@bonasiolab.org).

## 33 **ABSTRACT**

34 The COVID-19 pandemic has created an urgent need for rapid, effective, and low-cost SARS-  
35 CoV-2 diagnostic testing. Here, we describe COV-ID, an approach that combines RT-LAMP with  
36 deep sequencing to detect SARS-CoV-2 in unprocessed human saliva with high sensitivity (5–10  
37 virions). Based on a multi-dimensional barcoding strategy, COV-ID can be used to test thousands  
38 of samples overnight in a single sequencing run with limited labor and laboratory equipment. The  
39 sequencing-based readout allows COV-ID to detect multiple amplicons simultaneously, including  
40 key controls such as host transcripts and artificial spike-ins, as well as multiple pathogens. Here  
41 we demonstrate this flexibility by simultaneous detection of 4 amplicons in contrived saliva  
42 samples: SARS-CoV-2, influenza A, human *STATHERIN*, and an artificial SARS spike-in. The  
43 approach was validated on clinical saliva samples, where it showed 100% agreement with RT-  
44 qPCR. COV-ID can also be performed directly on saliva adsorbed on filter paper, simplifying  
45 collection logistics and sample handling.

## 46 **INTRODUCTION**

47 Within the first year of the COVID-19 pandemic SARS-CoV-2 has swept across the world, leading  
48 to more than 130 million infections and over 2.8 million deaths worldwide (as of April 2021). In  
49 many countries, non-pharmaceutical interventions, such as school closures and national  
50 lockdowns, have proven to be effective, but could not be sustained due to economic and social  
51 impact<sup>1, 2</sup>. Regularly performed population-level diagnostic testing is an attractive solution<sup>3</sup>,  
52 particularly as asymptomatic individuals are implicated in rapid disease transmission, with a  
53 strong overdispersion in secondary transmission<sup>4</sup>. Maintenance of population-level testing can be  
54 successful in isolating asymptomatic individuals and preventing sustained transmission<sup>5, 6</sup>;  
55 however, considerable barriers exist to the adoption of such massive testing strategies. Two such  
56 barriers are cost and supply constraints for commercial testing reagents, both of which make it  
57 impractical to test large numbers of asymptomatic individuals on a recurrent basis. A third major  
58 barrier is the lack of “user-friendly” protocols that can be rapidly adopted by public and private  
59 organizations to establish high-throughput surveillance screening. In addition, while COVID-19  
60 testing of symptomatic individuals might be effective during the summer season, when other  
61 respiratory infections are rare, new strategies are needed to facilitate rapid differential diagnosis  
62 between SARS-CoV-2 and other respiratory viruses in winter.

63 Recent adaptations of reverse transcription and polymerase chain reaction (RT-PCR) to amplify  
64 viral sequence and perform next-generation DNA sequencing have opened promising new  
65 avenues for massively parallel SARS-CoV-2 detection. In general, sequencing-based protocols  
66 use libraries of amplification primers to tag reads originating from each individual patient sample  
67 with a unique index that can be identified and deconvoluted after sequencing, thus allowing  
68 pooling of tens of thousands of samples in a single assay. SARSeq, SPAR-Seq, Swab-seq, and  
69 INSIGHT, directly amplify the viral RNA by RT-PCR and simultaneously introduce barcodes<sup>7-10</sup>.  
70 While effective, these methods rely on individual PCR amplification of each patient sample, thus  
71 requiring a large number of thermal cyclers for massive scale-up. An alternative approach,  
72 ApharSeq, addresses this bottleneck by annealing barcoded RT primers to viral RNA and pooling  
73 samples prior to amplification but the need for specialized oligo-dT magnetic beads might  
74 constitute a separate adoption barrier for this method<sup>11</sup>. Finally, several recent methods have  
75 been designed to take advantage of the extreme sensitivity and isothermal conditions of loop-  
76 mediated isothermal amplification (LAMP)<sup>12-14</sup>, but these methods either require additional  
77 manipulation to introduce barcodes<sup>12, 13</sup> or do not allow for convenient multiplexing<sup>14</sup>.

78 In this study, we present COV-ID, a method for SARS-CoV-2 identification based on LAMP, which  
79 enables large-scale diagnostic testing at low cost and with minimal on-site equipment. COV-ID is  
80 a robust method that can be used to test tens of thousands of samples for multiple pathogens  
81 with modest reagent costs and 2–4 laboratory personnel, generating results within 24 hours.  
82 COV-ID uses unpurified saliva or saliva adsorbed on filter paper as input material, thus enabling  
83 the massively parallel, inexpensive testing required for population-level surveillance of the  
84 COVID-19 pandemic (**Fig. 1A**).

## 85 RESULTS

### 86 Two-step amplification and indexing of viral and human sequences via RT-LAMP and PCR

87 The molecular basis for COV-ID is reverse transcription loop-mediated isothermal amplification  
88 (RT-LAMP), an alternative to PCR that has been used extensively for viral DNA or RNA detection  
89 in clinical samples<sup>15-18</sup>, including SARS-CoV-2<sup>19, 20</sup>. RT-LAMP requires 4–6 primers that recognize  
90 different regions of the target sequence<sup>21, 22</sup> and proceeds through a set of primed and self-primed  
91 steps to yield many inverted copies of the target sequence spanning a range of molecular sizes  
92 (**Fig. S1**). The forward inner primer (FIP) and backward inner primer (BIP), which recognize

93 internal sequences, are incorporated in opposite orientation across the target sequence in the  
94 final amplified product (**Fig. S1**).

95 Previous studies have shown that the FIP and BIP tolerate insertion of exogenous sequence  
96 between their different target homology regions<sup>23</sup>. We exploited this flexibility and introduced 1)  
97 patient-specific barcodes as shown previously<sup>12, 14, 23</sup> and 2) artificial sequences that allowed for  
98 PCR amplification of a small product compatible with Illumina sequencing library construction  
99 (**Fig. 1, Fig. S1**). These innovations allow us to pool individually barcoded RT-LAMP reactions  
100 and amplify them in batch via PCR, while introducing unique P5 and P7 dual indexes in different  
101 pools, thus enabling two-dimensional barcoding and dramatically increasing method throughput  
102 (see **Table S1** for PCR primer sequences). To minimize pool variability, PCR primers can be  
103 titrated to 100 nM and pooled PCRs carried out to completion, resulting in each pool being  
104 amplified to the same approximate concentration. Uniquely amplified and barcoded pools are  
105 mixed into a single “super-pool” that can be sequenced on an Illumina NextSeq or similar  
106 instrument (**Fig. 1A**). Combining individual barcodes embedded in the product at the RT-LAMP  
107 step with dual indexes introduced at the pool level during the PCR step allows for deconvolution  
108 of thousands or tens of thousands of samples in a single sequencing run.

109 To determine whether introduction of these exogenous sequences into the primers inhibited the  
110 isothermal amplification step, we performed RT-LAMP on inactivated SARS-CoV-2 virus using an  
111 extensively validated primer set against the N2 region of the nucleocapsid protein<sup>24</sup> including  
112 either the conventional BIP and FIP primers or their modified version re-engineered for the COV-  
113 ID workflow (**Fig. 1B**). Although the appearance of the amplified viral product was slightly delayed  
114 when using COV-ID primers, all reactions reached saturation rapidly and without detectable  
115 amplification of negative controls (**Fig. 1C**). Next, we tested whether COV-ID was compatible with  
116 RT-LAMP using newly designed primers against a host (human) transcript and whether the  
117 second step of COV-ID, direct library construction and indexing via PCR amplification (**Fig. 1D**),  
118 yields the desired product. For this, we designed RT-LAMP primers against the human beta-actin  
119 (*ACTB*) transcript that included sequences necessary for COV-ID. After RT-LAMP, reactions were  
120 diluted 100-fold before PCR with barcoded Illumina adapters. A PCR product of the expected size  
121 was visible in reactions containing total HeLa RNA, whereas no PCR product was observed in  
122 the absence of template (**Fig. 1E**). Sanger sequencing of the PCR product confirmed that RT-  
123 LAMP followed by PCR generated the product expected by the COV-ID method design, including  
124 the sample barcode introduced during the RT-LAMP step.

125 Thus, our data show that RT-LAMP is tolerant of sequence insertions in the BIP and FIP primers  
126 that allow introduction of LAMP-level barcodes as well as sequences homologous to Illumina  
127 adapters for direct amplification, indexing, and library construction via PCR.

## 128 **Sequencing-based detection of SARS-CoV-2 RNA from saliva using COV-ID**

129 We next evaluated the utility of COV-ID to detect viral RNA in saliva. Saliva is an attractive sample  
130 material for COVID-19 diagnostics with potential for early detection<sup>25</sup>, and has been shown to be  
131 a viable template for nucleic acid amplification via RT-PCR<sup>26</sup>, recombinase polymerase  
132 amplification (RPA)<sup>27</sup> as well as RT-LAMP<sup>28, 29</sup>. We prepared human saliva for RT-LAMP using a  
133 previously described treatment that inactivates SARS-CoV-2 virions, saliva-borne RNases and  
134 LAMP inhibitors (**Fig. 2A**)<sup>29</sup>. We performed RT-LAMP followed by PCR on inactivated saliva  
135 spiked with water or 1,000 total copies of inactivated SARS-CoV-2 virus. We observed a single  
136 band of the expected size in reactions performed on saliva spiked with virus but not in control  
137 reactions (**Fig. 2B**). The sequence of the amplified and barcoded viral product was confirmed by  
138 Sanger sequencing (**Fig. S2A**). Next, we subjected the libraries to deep sequencing. Reads  
139 aligned uniformly to the *N* gene, the region targeted by the N2 primer set, in COV-ID libraries  
140 constructed from viral samples but not in control libraries (**Fig. 2C**).

141 In several SARS-CoV-2 FDA approved tests, parallel amplification of a host (human) amplicon is  
142 implemented as a metric for sample integrity and quality. That is, if no human RNA is amplified  
143 from a clinical sample, no conclusion can be drawn from a negative test result<sup>30</sup>. However, in most  
144 tests, viral and human amplicons must be detected separately, resulting in a multiplication of the  
145 number of reactions to be performed. We reasoned that the deep sequencing nature of COV-ID  
146 would allow for simultaneous detection of viral, human, and other control amplicons, without  
147 increasing the number of necessary reactions. In fact, given that the PCR handles inserted in the  
148 BIP and FIP are the same for all RT-LAMP amplicons (**Fig. 1B**), the same P5 and P7 Illumina  
149 primers allow the simultaneous amplification of all RT-LAMP products obtained with COV-ID-  
150 modified primer sets (**Fig. 1D**). To identify a suitable human control, we compared conventional  
151 RT-LAMP primers for the mRNA of *ACTB*<sup>24</sup> or *STATHERIN* (*STATH*), a gene expressed  
152 specifically in saliva<sup>31</sup>. To determine which of the two RT-LAMP primer sets was a better proxy to  
153 measure RNA integrity in saliva samples, we assayed for amplification of the respective products  
154 in presence or absence of RNase. Whereas addition of RNase A abolished the *STATH* signal, it  
155 was ineffectual for *ACTB* (**Fig. S2B**), suggesting that amplification of genomic DNA made  
156 considerable contributions to the RT-LAMP signal observed for the latter. Therefore, we utilized  
157 *STATH* mRNA as a human control in subsequent experiments.

158 We used COV-ID-adapted primer sets for *N2* and *STATH* (**Table S1**) in multiplex on inactivated  
159 saliva spiked with a range of SARS-CoV-2 from 5 to 10,000 virions/ $\mu$ L. Subsequently, each RT-  
160 LAMP reaction was separately amplified via PCR using a unique P5 and P7 index combination,  
161 pooled, quantified, and deep-sequenced to an average depth of 6,000 reads per sample. After  
162 read trimming, alignment, and filtering (see Methods), 76% of reads from saliva COV-ID reactions  
163 were informative (**Fig. S2C**). In order to differentiate SARS-CoV-2 positive and negative samples,  
164 we calculated the ratio between *N2* reads and reads mapping to the human *STATH* control. Using  
165 the highest *N2/STATH* read ratio in control (SARS-CoV-2 negative saliva) as a threshold, 95%  
166 (19/20) of samples with spiked-in virus were correctly classified as positives (**Fig. 2D**). Using  
167 COV-ID, we consistently detected SARS-CoV-2 in saliva samples containing as low as 5 virions  
168 per  $\mu$ L, a sensitivity comparable and in some cases superior to those of established testing  
169 protocols<sup>32</sup>.

170 Scaling COV-ID to handle higher sample numbers requires pooling samples immediately  
171 following RT-LAMP, prior to the PCR step (**Fig. 1A**). We designed 32 unique 5-nucleotide  
172 barcodes for several target LAMP amplicons (**Fig. S2D** and **Table S2**). We first individually  
173 validated each barcode and primer combination by real-time fluorescence and PCR efficiency.  
174 Certain barcodes inhibited the RT-LAMP reaction, possibly due to internal micro-homology and  
175 primer self-hybridization<sup>33</sup>. Nonetheless, out of 32 barcodes tested in 3 separate RT-LAMP  
176 reactions (*N2*, *ACTB*, and *STATH*), 25 successfully amplified all three target RNAs (**Fig. S2D**).  
177 Saliva samples spiked with various concentrations of inactivated SARS-CoV-2 were amplified via  
178 barcoded RT-LAMP, then optionally pooled prior to PCR and sequencing (**Fig. S2E**). CoV-  
179 2/*STATH* ratios demonstrated no loss of sensitivity or specificity in the pooled samples compared  
180 to the individual PCRs.

181 To test the potential of COV-ID on patient samples, we tested saliva specimens, collected and  
182 previously analyzed at the Hospital of the University of Pennsylvania (see Methods). We carried  
183 out multiplex barcoded RT-LAMPs on each sample (COV-ID step I, **Fig. 1B**), pooled the reactions  
184 and then constructed libraries via PCR (COV-ID step II, **Fig. 1D**). After deep sequencing, analysis  
185 of *N2/STATH* ratios showed 100% (8/8) concordance with viral copy numbers generated by a  
186 standard clinical test (RNA purification followed by RT-qPCR) (**Fig. 2E**), demonstrating the  
187 effectiveness of the COV-ID approach.

188 Taken together, our data show that COV-ID can be utilized to detect viral and human amplicons  
189 in multiplex directly from saliva. The samples that can be batch amplified and deconvoluted after  
190 deep sequencing.

## 191 Calibration of COV-ID using an artificial spike-in

192 Existing deep sequencing approaches for massively parallel COVID-19 testing based on RT-PCR  
193 incorporate artificial spike-ins, which serve as an internal calibration controls and allow for better  
194 estimates of viral loads by end-point PCR<sup>7,8</sup>. At the same time, adding to the reactions an artificial  
195 substrate for amplification helps minimizing spurious signals as it can “scavenge” viral  
196 amplification primers in negative samples. Finally, by providing a baseline amplification even in  
197 empty samples, a properly designed spike-in strategy can reduce variance in total amounts of  
198 final amplified products across samples, which compresses the dynamic-range of sequence  
199 coverage for each patient in a complex pool and, therefore, reduces the risk of inconclusive  
200 samples due to low sequencing coverage<sup>8</sup>.

201 We reasoned that a spike-in approach for LAMP-based quantification would provide similar  
202 benefits in the context of COV-ID. To generate a SARS-CoV-2 spike-in, we synthesized a  
203 fragment of the *N2* RNA that retained all primer-binding regions for RT-LAMP and contained a  
204 divergent 7-nt stretch of sequence to distinguish reads originating from the spike-in from those  
205 originating from the natural virus (**Fig. S3A**). After confirming that the spike-in template was  
206 efficiently amplified via RT-LAMP with the *N2* primer set (**Fig. S3B**), we performed pooled COV-  
207 ID on virus-containing saliva in the presence of 20 fg of *N2* spike-in RNA. As expected<sup>8</sup>, addition  
208 of a constant amount of viral spike-in across reactions reduced the variability in total read numbers  
209 for individual samples in the final pool (**Fig. S3C**). As discussed above, a narrower range in  
210 sequencing output across samples in a pool optimizes the utilization of sequencing reads, and  
211 ultimately lowers the cost per sample. Because the spike-in provides an internal calibration that  
212 is independent of the RNA quality found in saliva, in several cases normalization against the  
213 spike-in resulted in lower levels of false positive signal from negative samples (**Fig. S3D**). This is  
214 likely because in cases where very few *STATH* reads were obtained, possibly due to degradation  
215 of host RNA in the saliva sample, the resulting small denominator inflated the *N2/STATH* ratio  
216 even for SARS-CoV-2 signal that was low in absolute terms and likely spurious.

217 Thus, these data show that spike-in strategies are compatible with the COV-ID workflow and  
218 provide a means to stabilize total amplification and read allocation per sample while also offering  
219 an additional calibration control to better estimate the viral load in samples where the endogenous  
220 *STATH* mRNA might be below detection due to improper collection or handling.

221

222

## 223 Simultaneous detection of SARS-CoV-2 and influenza A by COV-ID

224 Given the challenge of distinguishing early symptoms of COVID-19 from other respiratory  
225 infections, we evaluated COV-ID for the simultaneous detection of more than one viral pathogen.  
226 Multiple distinct products can be simultaneously amplified by RT-LAMP in the same tube by  
227 providing the appropriate primer sets in multiplex, as we demonstrated above by co-amplifying  
228 *N2* and *STATH* in the same COV-ID reaction (see **Fig. 2**). In fact, simultaneous detection of  
229 SARS-CoV-2 and influenza virus by RT-LAMP was previously achieved, albeit in a fluorescent-  
230 based, low-throughput type of assay<sup>34</sup>. We reasoned that the sequencing-based readout of COV-  
231 ID would allow extending this approach to the simultaneous detection of multiple pathogens as  
232 well as endogenous (host mRNA) and artificial (spike-in) calibration standards, all in a single  
233 reaction.

234 To test the ability of COV-ID to simultaneously detect multiple viral templates, we selected and  
235 validated a generic “flu” RT-LAMP primer set that recognizes several strains, including influenza  
236 A virus (IAV) and influenza B<sup>34, 35</sup>, and modified the BIP and FIP sequence to introduce the COV-  
237 ID barcodes and handles for PCR (**Fig. S2D** and **Table S1**). We added inactivated SARS-CoV-2  
238 virus (BEI resources) and IAV strain H1N1 RNA (Twist Biosciences) to saliva according to a 3 x  
239 4 matrix of ( $10^4$ ,  $10^3$ , or 0 copies per  $\mu\text{L}$ ) SARS-CoV-2 RNA against H1N1 RNA ( $10^5$ ,  $10^4$ ,  $10^3$ , or  
240 0 copies per  $\mu\text{L}$ ) (**Fig. 3A**), as well as the *N2* spike-in control. We performed multiplex COV-ID on  
241 these samples using primers sets for *STATH*, *N2* (to detect SARS-CoV-2), and IAV (to detect  
242 H1N1) and sequenced to an average depth of 21,000 reads per sample. Both H1N1 and SARS-  
243 CoV-2 were detected above background and the signal correlated with the amount of the  
244 respective template added to saliva (**Fig. 3B–C**). Overall, multiplex COV-ID correctly identified  
245 samples that contained only SARS-CoV-2 (7/8) or H1N1 (6/8). For samples that contained both  
246 pathogens we observed reduced sensitivity (11/16 identification of both pathogens), which was  
247 also observed in a previous multiplexing attempt<sup>34</sup>. However, in practice individuals who are  
248 simultaneously infected with both viruses presumably would be rare, and for these cases the  
249 ability to detect at least one virus successfully would allow to follow up with further diagnostic  
250 testing. We found that of the samples containing both viruses, 16/16 showed positive detection of  
251 at least one pathogen (SARS-CoV-2 or H1N1), suggesting the reduced sensitivity of the multiplex  
252 assay is due to interference between amplification of both viral templates. This also demonstrates  
253 that COV-ID can be used as an effective screening approach for multiple viral templates.

254



## 255 Paper-based saliva sampling for COV-ID

256 As an additional step toward increasing the throughput of the COV-ID approach, we explored  
257 avenues to simplify collection, lower costs, and expedite processing time. Absorbent paper is an  
258 attractive alternative to sample vials for collection, given its low cost, wide availability, and smaller  
259 environmental footprint. In fact, paper has been used as a means to isolate nucleic acid from  
260 biological samples for direct RT-PCR testing<sup>36</sup> as well as RT-LAMP<sup>37, 38</sup>.

261 We sought to determine whether the COV-ID workflow would be compatible with saliva collection  
262 on absorbent paper. First, we immersed a small square of Whatman filter paper into water  
263 containing various dilutions of inactivated SARS-CoV-2. After 2 min, the paper was removed and  
264 transferred to PCR strip tubes followed by heating at 95°C for 5 minutes to air-dry the sample  
265 (**Fig. 4A**). Next, we added the RT-LAMP mix containing the *N2* COV-ID primer set directly to the  
266 tubes containing the paper squares and let the reaction proceed in the usual conditions. COV-ID  
267 PCR products of the correct size were evident in all samples containing viral RNA, with sensitivity  
268 of at least 100 virions /  $\mu\text{L}$  (**Fig. 4B**) and in none of the controls, demonstrating that the presence  
269 of paper does not interfere with the RT-LAMP reaction and subsequent PCR amplification with  
270 Illumina adapters.

271 To assay direct COV-ID detection from saliva on paper, we saturated Whatman filter paper  
272 squares with saliva containing different amounts of inactivated SARS-CoV-2 virus, which, we  
273 reasoned, would be equivalent to a patient collecting their own saliva by chewing on a small piece  
274 of absorbent paper. Next, we placed the paper squares into reaction tubes containing  
275 TCEP/EDTA inactivation buffer (see Methods) similar to that used for the in-solution samples  
276 used in our previous experiments (see **Fig. 1A**). We dried the paper at 95°C and performed RT-  
277 LAMP followed by PCR (**Fig. 4C**), which resulted in the appearance of COV-ID products of the  
278 correct size starting from saliva spiked with as few as 50 virions /  $\mu\text{L}$  (**Fig. 4D**). We then performed  
279 COV-ID sequencing on saliva collected on paper using primers *N2* and *STATH* in the presence  
280 of the *N2* spike-in RNA. The sequence data showed more variability and limited coverage of the  
281 control amplicons compared to in-solution COV-ID likely due to the more challenging reaction  
282 conditions; therefore, we normalized viral reads using both *STATH* and spike-in. This paper-  
283 based COV-ID proof-of-principle experiment detected the presence of viral RNA in samples with  
284 as little as 320 copies /  $\mu\text{L}$  (**Fig. 4E**), a lower sensitivity compared to that of in-solution COV-ID  
285 but still well within the useful range<sup>39</sup> to detect infections.

286 Taken together, these data show that the RT-LAMP step of COV-ID is compatible with the  
287 presence of paper in the reaction tube and suggest that self-collection of saliva by patients directly

288 on absorbent paper could provide a simple and cost-effective strategy to collect and test  
289 thousands of saliva samples for multiple pathogens (**Fig. 4F**).

## 290 **DISCUSSION**

291 Testing strategies are vital to an effective public health response to the COVID-19 pandemic,  
292 particularly with the spread of the disease by asymptomatic individuals. An ongoing challenge to  
293 COVID-19 testing is the need for massive testing strategies for population-level surveillance that  
294 are needed for efficient contact tracing and isolation. Most FDA-approved clinical SARS-CoV-2  
295 diagnostic tests are based on time-consuming and expensive protocols that include RNA  
296 purifications and RT-PCR<sup>32</sup> and must be performed by trained personnel in well-equipped  
297 laboratories. Point-of-care antigen tests provide a much faster turnaround time and require little  
298 manipulation, but there remains limited data on their specificity in real-world applications<sup>40</sup>.  
299 Because of reagent limitations and diagnostic testing bottlenecks, prioritization of COVID  
300 diagnostic testing continues to be for symptomatic individuals and individuals who are particularly  
301 vulnerable for infection after exposure<sup>41</sup>. Private organizations, including colleges and  
302 universities, have circumvented some of these challenges by contracting with private laboratories  
303 to establish asymptomatic surveillance testing protocols; this is a costly option for population-level  
304 surveilling of asymptomatic SARS-CoV-2 infections.

305 Several effective COVID-19 vaccines have been developed and there is a concerted ongoing  
306 global vaccination effort, providing a concrete means to end the pandemic. Despite this progress  
307 there are several potential risks that require vigilance: possible COVID-19 transmission in  
308 vaccinated individuals, emergence of vaccine-resistant viral variants, and public skepticism of  
309 vaccines or faltering compliance with social distancing guidelines<sup>42</sup>. For these reasons ongoing  
310 testing and surveillance efforts will remain important for the foreseeable future, both to monitor  
311 the progress of vaccination in reducing symptomatic cases and to detect emerging variants.

312 In order to scale testing to an effective volume and frequency, surveillance tests must  
313 demonstrate the following qualities: 1) sensitivity, to identify both asymptomatic and symptomatic  
314 carriers; 2) simplicity in methodology, to be performed in a number of traditional diagnostic  
315 laboratories, without specialized equipment; 3) low cost and easily accessible reagents; 4) ease  
316 of collection method; 5) rapid turnaround time to allow for isolation and contact tracing; and 6)  
317 ability to co-detect multiple respiratory viruses, given the overlap in patient symptoms. To this  
318 end, we have developed COV-ID, an RT-LAMP-based parallel sequencing SARS-CoV-2

319 detection method that can provide results from tens of thousands of samples per day at relatively  
320 low cost to simultaneously detect multiple respiratory viruses.

321 COV-ID features several key innovations that make it well-suited to high-throughput testing. First,  
322 COV-ID uses a two-dimensional barcoding strategy<sup>8</sup>, where the same 96 barcodes are used in  
323 each RT-LAMP plate, making it possible to pre-aliquot barcodes in 96-well plates ahead of time  
324 and store them at -20°C, simplifying execution of the assay and shortening turnaround times.  
325 Second, since RT-LAMP does not require thermal cycling, tens of thousands of samples can be  
326 run simultaneously in a standard benchtop-sized incubator or hybridization oven held at 65°C.  
327 Third, individual samples are pooled immediately following RT-LAMP; therefore, a single  
328 thermocycler has the potential to process up to 96 or 384 RT-LAMP plates, generating 9,216 or  
329 36,864 individually barcoded samples, respectively (**Fig. 1A, 4F, 4G**). Only 96 unique FIP  
330 barcodes are required for this scaling; here, we show that 28 out of 32 LAMP barcodes tested  
331 were functional for both *N2* and *STATH*. This proof-of-principle experiment demonstrates the  
332 feasibility of generating the library of barcodes required to apply COV-ID to a large population. An  
333 additional advantage of sequencing-based approaches, such as COV-ID is that with carefully  
334 designed primers it would be possible to recover information about viral variants directly from the  
335 sequencing reads<sup>43</sup>. Finally, COV-ID can generate ready-to-sequence libraries directly from saliva  
336 absorbed onto filter paper, which would allow for major streamlining of the often-challenging  
337 logistical process of sample collection (**Fig. 4**). Thus, COV-ID libraries for thousands and tens of  
338 thousands of samples can be generated with relatively minimum effort in biological laboratories  
339 with basic equipment and easily accessible reagents.

340 With the average throughput of an Illumina NextSeq 500/550, a relatively affordable next-  
341 generation sequencer up to 9,216 (96 RT-LAMPs x 96 pools) can be sequenced at a depth of  
342 ~48,000 reads per sample, and up to 36,864 (96 RT-LAMPs x 384 pools) can be sequenced at a  
343 depth of ~12,000 reads, which, we showed, is more than sufficient to obtain information about  
344 multiple viral and control amplicons. Considering that reagents for one NextSeq run cost ~1,500  
345 U.S. dollars, the theoretical sequencing cost per sample could be as low as \$0.04 (**Fig. 4G**). While  
346 sequencing instruments are relatively specialized and not ubiquitous, amplified COV-ID DNA  
347 libraries could be shipped to remote facilities for sequencing in a cost-effective manner as  
348 previously proposed by the inventors of LAMP-seq<sup>14</sup>. Finally, because of the limited sequence  
349 space against which reads must be aligned, computational analysis of the resulting data can be  
350 performed in a matter of minutes with optimized pipelines, providing results shortly after the  
351 sequencing run has completed.

352 COV-ID has sensitivity of 5–10 virions of SARS-CoV-2 per  $\mu\text{L}$  in contrived saliva samples (**Fig.**  
353 **2D**) and at least 300 virions /  $\mu\text{L}$  in saliva collected from patients in a clinical setting (**Fig. 2E**).  
354 However, this was sufficient to properly classify 100% of the clinical samples analyzed, given that  
355 all positive samples had an estimated viral load  $> 300$  virions /  $\mu\text{L}$ . Importantly, this was also the  
356 apparent limit of sensitivity of paper-based COV-ID (**Fig. 4E**), suggesting that even in these  
357 settings COV-ID would be capable of accurately classifying the majority of patient-derived  
358 samples.

359 In conclusion, COV-ID is a flexible platform that can be executed at varying levels of scale with  
360 additional flexibility in sample input, making it an attractive platform for surveillance testing.  
361 Population-level monitoring of SARS-CoV-2 infections will be critical while vaccines are being  
362 distributed to the global population, and continued surveillance will likely remain an effective  
363 strategy to protect immune-compromised and unvaccinated members in society and within  
364 entities and organizations where regular monitoring is critical to social isolation strategies. To that  
365 end, effective, low-cost, multiplexed, and readily-implementable strategies for surveillance  
366 testing, such as COV-ID, are important to mitigate the effects of the current and future pandemics.

## 367 **METHODS**

### 368 **RT-LAMP primer design**

369 Primers against *ACTB* were designed using PrimerExplorerV5 (<https://primerexplorer.jp/e/>) using  
370 default parameters and including loop primers (**Table S1**).

371 For COV-ID, priming sequences for PCR were inserted in FIP and BIP primers between the target  
372 homology regions (F1c and F2, and B1c and B2, respectively, see **Fig. S1**). After testing, we  
373 determined that 12 nts and 11 nts were most effective for the P5 and P7 binding regions,  
374 respectively, being the shortest insertion that allowed reliable PCR amplification from LAMP  
375 products without impacting LAMP efficiency. In addition a 5 nt barcode sequence was inserted at  
376 the immediate 3' end of the P5-binding region of the FIP primer.

### 377 **LAMP barcode design**

378 Starting from the total possible 1,024 unique 5-nt barcodes, we removed those that matched any  
379 sequence within the RT-LAMP primers used in this study (**Table S1**) in either sense or anti-sense  
380 orientation. From the remaining pool, we selected 32 barcodes with hamming distance of at least  
381 2 between all candidates. We tested FIPs incorporating candidate barcodes for *ACTB*, *STATH*,

382 N2, and IAV primer sets on saliva RT-LAMP with 1,000 copies target amplicon. Primers that failed  
383 to show LAMP signal by real time fluorescence monitoring or generate expected PCR product  
384 were discarded. Final usable barcodes are provided in (**Table S2**).

### 385 **Saliva preparation**

386 We prepared 100x TCEP/EDTA buffer (250 mM TCEP, 100 mM EDTA, 1.15 N NaOH)<sup>29</sup>.  
387 TCEP/EDTA buffer was added to human saliva at 1:100 volume, then samples were capped,  
388 vortexed to mix and heated in a thermocycler (95°C 5 min, 4°C hold) until ready to use for RT-  
389 LAMP. When indicated, heat-inactivated SARS-CoV-2 (BEI Resources Cat. NR-52286) or H1N1  
390 genomic RNA (Twist Biosciences Cat. 103001) was added to inactivated saliva prior to RT-LAMP.

### 391 **N2 spike-in synthesis**

392 To prepare the *in vitro* transcription template for SARS-CoV-2 N2 spike-in RNA, we performed  
393 RT-PCR using Power SYBR RNA-to-Ct kit (Thermo Cat. 4389986) of heat inactivated SARS-  
394 CoV-2 (BEI Resources Cat. NR-52286) using the following primers: N2-B3 and N2-spike-T7 S.  
395 PCR product was purified and used as a template for *in vitro* transcription using HiScribe T7  
396 transcription kit (NEB Cat. E2040S). RNA was purified with Trizol (Thermo Cat. 15596026),  
397 quantified via A<sub>260</sub>, then aliquoted in BTE buffer (10 mM bis-tris pH 6.7, 1 mM EDTA) and stored  
398 at -80°C. Primers used and final spike-in sequence are provided in **Table S1**.

### 399 **RT-LAMP**

400 All RT-LAMP reactions were set up in clean laminar flow hoods and all steps before and after  
401 LAMP were carried out in separate lab spaces to avoid contamination. RT-LAMP reactions were  
402 set up on ice as follow: for each amplicon 5 or 6 LAMP primers were combined into 10x working  
403 stock at established concentrations: 16 µM FIP, 16 µM BIP, 4 µM LF, 4 µM LB, 2 µM F3, 2 µM  
404 B3. For multiplexed COV-ID reactions 10x working primer mixes for each amplicon were either  
405 added proportionally so that the total primer content remained constant, or mixed so that BIP and  
406 FIP primers were scaled down depending on amplicon number while remaining primers (LF  
407 and/or LB, F3, B3) were kept at same concentration as in single reactions.

408 Each 10 µL RT-LAMP reaction mix consisted of 1x Warmstart LAMP 2x Master Mix (NEB Cat.  
409 E1700S), 0.7 µM dUTP (Promega Cat. U1191), 1 µM SYTO-9 (Thermo Cat. S34854), 0.1 µL  
410 Thermolabile UDG (Enzymatics Cat. G5020L), 1 µL of saliva template and optionally 20 fg of N2  
411 Spike RNA. Reactions were prepared in qPCR plates or 8-well strip tubes, sealed, vortexed and  
412 centrifuged briefly, then incubated in either a QuantStudio Flex 7 or StepOnePlus instrument

413 (Thermo) for 65°C 1 hr. Real-time fluorescence measurements were recorded every 30 sec to  
414 monitor reaction progress but were not used for data analysis. Following LAMP the reactions were  
415 heated at 95°C 5 min to inactivate LAMP enzymes.

#### 416 **Library construction by PCR amplification**

417 All post-LAMP steps were carried out on a clean bench separate from LAMP reagents and  
418 workspace. For individual LAMP samples, LAMP amplicons were diluted either 1:100 or 1:1,000  
419 in water. For pooling of individually barcoded LAMP reactions, equal amounts of all LAMP  
420 reactions were combined and then either diluted 1:1000 or purified via SPRIselect beads  
421 (Beckman Coulter Cat. B23317) using a bead-to-reaction ratio of 0.1x. Purified material was  
422 diluted to final 100-fold dilution relative to LAMP.

423 1 µL of diluted LAMP material was used as a template for PCR using OneTaq DNA polymerase  
424 (NEB Cat. M0480L) with 100 nM each of custom dual-indexed Illumina P5 and P7 primers in  
425 either 10 or 25 µL reaction (**Table S1**). PCR reactions were incubated as follows: (25 cycles of  
426 stage 1 [94°C x 15 sec, 45°C x 15 sec, 68°C x 10 sec], 10 cycles of Stage 2 [ 94°C x 15 sec, 68°C  
427 x 10 sec], 68°C x 1 min, 4°C x ∞). Note, for initial pilot COV-ID and clinical sample experiments  
428 (**Fig. 2D–E, Fig. S2C**) PCR incubation was performed as above with modification: [Stage 1 x 10  
429 cycles, Stage 2 x 25 cycles].

430 PCR products were resolved on 2% agarose gel to confirm library size, then all were pooled and  
431 purified via MinElute PCR purification kit (Qiagen Cat. 28004) and quantified using either Qubit  
432 dsDNA High Sensitivity kit (Thermo Cat. Q32851) or Kapa Library Quantification Kit for Illumina  
433 (Kapa Cat. 07960140001).

#### 434 **Human samples**

435 Clinical saliva samples used for **Fig. 2E** were obtained and characterized as part of a separate  
436 study at the University of Pennsylvania<sup>44</sup> and collected under Institutional Review Board (IRB)-  
437 approved protocols (IRB protocol #842613 and #813913). Briefly, salivary samples were collected  
438 from possible SARS-CoV-2 positive patients at one of three locations: (1) Penn Presbyterian  
439 Medical Center Emergency Department, (2) Hospital of the University of Pennsylvania  
440 Emergency Department, and (3) Penn Medicine COVID-19 ambulatory testing center. Inclusion  
441 criteria including any adult (age > 17 years) who underwent SARS-CoV-2 testing via standard  
442 nasopharyngeal swab at the same visit. Patients with known COVID-19 disease who previously  
443 tested positive previously were excluded. After verbal consent was obtained by a trained research

444 coordinator, patients were instructed to self-collect saliva into a sterile specimen container which  
445 was then placed on ice until further processing for analysis.

446 The saliva used in the remaining experiments was donated by one of the authors. Because it was  
447 only used for protocol optimization the Penn IRB has determined that it did not constitute human  
448 subjects research and therefore approval was not required.

#### 449 **Paper COV-ID**

450 Squares of Whatman no. 1 filter paper (2 mm x 2 mm) were cut using a scalpel on a clean surface  
451 under a laminar flow hood and stored at room temperature until used. Using ethanol-sterilized  
452 fine-nosed tweezers a single square was dipped twice into unprocessed, freshly collected saliva  
453 with or without added SARS-CoV-2 (BEI Resources Cat. NR-52286) until saliva was saturated on  
454 paper by eye. Paper was then transferred to well of 96-well plate containing 10 ul of 1x  
455 TCEP/EDTA buffer (2.5 mM TCEP, 1 mM EDTA, 1.15 NaOH). Plate was placed on heat block  
456 inside laminar flow hood or inside open thermocycler and incubated at 95°C x 10 min.

457 10 ul RT-LAMP mixture was prepared as described above in the absence of the N2 Spike RNA.  
458 10 ul of RT-LAMP reaction mixture was added to each paper strip, then plate was sealed and  
459 incubated 65°C x 1 hr, 95°C x 5 min in QuantStudio Flex 7 (Thermo). 1 ul of each reaction was  
460 either diluted 1:100 or purified via SPRIselect beads and PCR amplified as described above.

#### 461 **Sequencing**

462 Libraries were sequenced on one of the following Illumina instruments: MiSeq, NextSeq 500,  
463 NextSeq 550, NovaSeq 6000 and sequenced using single end programs with a minimum of 40  
464 cycles on Read 1 and 8 cycles for index 1 (on P7) and index 2 (on P5).

#### 465 **Sequence Analysis**

466 Reads were filtered for optical quality using FASTX-toolkit utility fastq\_quality\_filter  
467 ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)), then cutadapt<sup>45</sup> was used to remove adapters and  
468 demultiplex LAMP barcodes. Reads were aligned to a custom index containing SARS-CoV-2  
469 genome (NC\_045512.2), Influenza H1N1 coding sequences (NC\_026431.1, NC\_026432.1,  
470 NC\_026433.1, NC\_026434.1, NC\_026435.1, NC\_026436.1, NC\_026437.1, NC\_026438.1),  
471 STATH coding sequence (NM\_003154.3), and custom N2 spike sequence (**Table S3**), target  
472 sequences using bowtie2<sup>46</sup> with options --no-unal and --end-to-end. Alignments with greater than  
473 1 mismatch were removed and the number of reads mapping to each target for all barcodes were  
474 extracted and output in a matrix. Barcodes with fewer than 25 total mapped reads were discarded.

475 **Data Availability**

476 Next generation sequencing data generated for this study are available at the NCBI GEO with  
477 accession GSE172118.

478 **ACKNOWLEDGMENTS**

479 The authors thank E. Shields for careful proofreading of analysis scripts; B. Morris and R. Collman  
480 for the collection and distribution of clinical saliva samples; F. Bushman, S. Sherril-Mix, and  
481 Abigail Glascock for sharing RT-qPCR data on the clinical samples; the UPenn rapid assay task  
482 force for project feedback; and the gLAMP weekly forum for advice and guidance.

483 **DISCLOSURE**

484 R.W-T., C.A.T. and R.B. are inventors on a patent application filed by the University of  
485 Pennsylvania related to this work.



## 486 REFERENCES

- 487 1. Haug, N. et al. Ranking the effectiveness of worldwide COVID-19 government  
488 interventions. *Nat Hum Behav* **4**, 1303-1312 (2020).
- 489 2. Tian, H. et al. An investigation of transmission control measures during the first 50 days  
490 of the COVID-19 epidemic in China. *Science* **368**, 638-642 (2020).
- 491 3. Taipale, J., Romer, P. & Linnarsson, S. Population-scale testing can suppress the spread  
492 of COVID-19. *medRxiv*, 2020.2004.2027.20078329 (2020).
- 493 4. Endo, A., Centre for the Mathematical Modelling of Infectious Diseases, C.-W.G., Abbott,  
494 S., Kucharski, A.J. & Funk, S. Estimating the overdispersion in COVID-19 transmission  
495 using outbreak sizes outside China. *Wellcome Open Res* **5**, 67 (2020).
- 496 5. Holt, E. Slovakia to test all adults for SARS-CoV-2. *Lancet* **396**, 1386-1387 (2020).
- 497 6. Larremore, D.B. et al. Test sensitivity is secondary to frequency and turnaround time for  
498 COVID-19 screening. *Sci Adv* (2020).
- 499 7. Bloom, J.S. et al. Swab-Seq: A high-throughput platform for massively scaled up SARS-  
500 CoV-2 testing. *medRxiv*, 2020.2008.2004.20167874 (2020).
- 501 8. Yelagandula, R. et al. SARSeq, a robust and highly multiplexed NGS assay for parallel  
502 detection of SARS-CoV2 and other respiratory infections. *medRxiv*,  
503 2020.2010.2028.20217778 (2020).
- 504 9. Aynaud, M.-M. et al. A Multiplexed, Next Generation Sequencing Platform for High-  
505 Throughput Detection of SARS-CoV-2. *medRxiv*, 2020.2010.2015.20212712 (2020).
- 506 10. Wu, Q. et al. INSIGHT: A population-scale COVID-19 testing strategy combining point-of-  
507 care diagnosis with centralized high-throughput sequencing. *Sci Adv* **7** (2021).
- 508 11. Chappleboim, A. et al. ApharSeq: An Extraction-free Early-Pooling Protocol for Massively  
509 Multiplexed SARS-CoV-2 Detection. *medRxiv*, 2020.2008.2008.20170746 (2020).
- 510 12. James, P. et al. LamPORE: rapid, accurate and highly scalable molecular screening for  
511 SARS-CoV-2 infection, based on nanopore sequencing. *medRxiv*,  
512 2020.2008.2007.20161737 (2020).
- 513 13. Dao Thi, V.L. et al. A colorimetric RT-LAMP assay and LAMP-sequencing for detecting  
514 SARS-CoV-2 RNA in clinical samples. *Sci Transl Med* **12** (2020).
- 515 14. Schmid-Burgk, J.L. et al. LAMP-Seq: Population-Scale COVID-19 Diagnostics Using a  
516 Compressed Barcode Space. *bioRxiv*, 2020.2004.2006.025635 (2020).
- 517 15. Li, S. et al. Simultaneous detection and differentiation of dengue virus serotypes 1-4,  
518 Japanese encephalitis virus, and West Nile virus by a combined reverse-transcription  
519 loop-mediated isothermal amplification assay. *Virology* **8**, 360 (2011).
- 520 16. Shirato, K. et al. Detection of Middle East respiratory syndrome coronavirus using reverse  
521 transcription loop-mediated isothermal amplification (RT-LAMP). *Virology* **11**, 139 (2014).

- 522 17. Calvert, A.E., Biggerstaff, B.J., Tanner, N.A., Lauterbach, M. & Lanciotti, R.S. Rapid  
523 colorimetric detection of Zika virus from serum and urine specimens by reverse  
524 transcription loop-mediated isothermal amplification (RT-LAMP). *PLoS One* **12**, e0185340  
525 (2017).
- 526 18. Enomoto, Y. et al. Rapid diagnosis of herpes simplex virus infection by a loop-mediated  
527 isothermal amplification method. *J Clin Microbiol* **43**, 951-955 (2005).
- 528 19. Augustine, R. et al. Loop-Mediated Isothermal Amplification (LAMP): A Rapid, Sensitive,  
529 Specific, and Cost-Effective Point-of-Care Test for Coronaviruses in the Context of  
530 COVID-19 Pandemic. *Biology (Basel)* **9** (2020).
- 531 20. United States Food and Drug Administration. Color Genomics SARS-CoV-2 RT-LAMP  
532 Diagnostic Assay - EUA Summary. <https://www.fda.gov/media/138249/download>.  
533 Published: November 2, 2020, Accessed: December 22, 2020.
- 534 21. Nagamine, K., Hase, T. & Notomi, T. Accelerated reaction by loop-mediated isothermal  
535 amplification using loop primers. *Mol Cell Probes* **16**, 223-229 (2002).
- 536 22. Notomi, T. et al. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res* **28**,  
537 E63 (2000).
- 538 23. Yamagishi, J. et al. Serotyping dengue virus with isothermal amplification and a portable  
539 sequencer. *Sci Rep* **7**, 3510 (2017).
- 540 24. Butler, D.J. et al. Shotgun Transcriptome and Isothermal Profiling of SARS-CoV-2  
541 Infection Reveals Unique Host Responses, Viral Diversification, and Drug Interactions.  
542 *bioRxiv*, 2020.2004.2020.048066 (2020).
- 543 25. Savela, E.S. et al. SARS-CoV-2 is detectable using sensitive RNA saliva testing days  
544 before viral load reaches detection range of low-sensitivity nasal swab tests. *medRxiv*,  
545 2021.2004.2002.21254771 (2021).
- 546 26. Ranoa, D.R.E. et al. Saliva-Based Molecular Testing for SARS-CoV-2 that Bypasses RNA  
547 Extraction. *bioRxiv*, 2020.2006.2018.159434 (2020).
- 548 27. Myhrvold, C. et al. Field-deployable viral diagnostics using CRISPR-Cas13. *Science* **360**,  
549 444-448 (2018).
- 550 28. Lalli, M.A. et al. Rapid and extraction-free detection of SARS-CoV-2 from saliva with  
551 colorimetric LAMP. *medRxiv*, 2020.2005.2007.20093542 (2020).
- 552 29. Rabe, B.A. & Cepko, C. SARS-CoV-2 detection using isothermal amplification and a rapid,  
553 inexpensive protocol for sample inactivation and purification. *Proc Natl Acad Sci U S A*  
554 **117**, 24450-24458 (2020).
- 555 30. Babiker, A., Myers, C.W., Hill, C.E. & Guarner, J. SARS-CoV-2 Testing. *Am J Clin Pathol*  
556 **153**, 706-708 (2020).

- 557 31. Satoh, T. et al. Development of mRNA-based body fluid identification using reverse  
558 transcription loop-mediated isothermal amplification. *Anal Bioanal Chem* **410**, 4371-4378  
559 (2018).
- 560 32. MacKay, M.J. et al. The COVID-19 XPRIZE and the need for scalable, fast, and  
561 widespread testing. *Nat Biotechnol* **38**, 1021-1024 (2020).
- 562 33. Torres, C. et al. LAVA: an open-source approach to designing LAMP (loop-mediated  
563 isothermal amplification) DNA signatures. *BMC Bioinformatics* **12**, 240 (2011).
- 564 34. Zhang, Y. & Tanner, N.A. Development of Multiplexed RT-LAMP for Detection of SARS-  
565 CoV-2 and Influenza Viral RNA. *medRxiv*, 2020.2010.2026.20219972 (2020).
- 566 35. Takayama, I. et al. Development of real-time fluorescent reverse transcription loop-  
567 mediated isothermal amplification assay with quenching primer for influenza virus and  
568 respiratory syncytial virus. *J Virol Methods* **267**, 53-58 (2019).
- 569 36. Glushakova, L.G. et al. Detection of chikungunya viral RNA in mosquito bodies on cationic  
570 (Q) paper based on innovations in synthetic biology. *J Virol Methods* **246**, 104-111 (2017).
- 571 37. Kellner, M.J. et al. A rapid, highly sensitive and open-access SARS-CoV-2 detection assay  
572 for laboratory and home testing. *bioRxiv*, 2020.2006.2023.166397 (2020).
- 573 38. Yaren, O. et al. Ultra-rapid detection of SARS-CoV-2 in public workspace environments.  
574 *medRxiv*, 2020.2009.2029.20204131 (2020).
- 575 39. Winnett, A. et al. SARS-CoV-2 Viral Load in Saliva Rises Gradually and to Moderate  
576 Levels in Some Humans. *medRxiv* (2020).
- 577 40. Pettengill, M.A. & McAdam, A.J. Can We Test Our Way Out of the COVID-19 Pandemic?  
578 *J Clin Microbiol* **58** (2020).
- 579 41. Schuetz, A.N. et al. When Should Asymptomatic Persons Be Tested for COVID-19? *J Clin*  
580 *Microbiol* **59** (2020).
- 581 42. Aschwanden, C. Five reasons why COVID herd immunity is probably impossible. *Nature*  
582 **591**, 520-522 (2021).
- 583 43. Everett, J. et al. SARS-CoV-2 Genomic Variation in Space and Time in Hospitalized  
584 Patients in Philadelphia. *mBio* **12** (2021).
- 585 44. Sherrill-Mix, S. et al. LAMP-BEAC: Detection of SARS-CoV-2 RNA Using RT-LAMP and  
586 Molecular Beacons. *medRxiv*, 2020.2008.2013.20173757 (2020).
- 587 45. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.  
588 *2011* **17**, 3 (2011).
- 589 46. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods*  
590 **9**, 357-359 (2012).
- 591

## 592 **FIGURE LEGENDS**

### 593 **Figure 1. Barcoding and PCR amplification of RT-LAMP products**

594 (A) Overview of COV-ID. Saliva is collected and inactivated prior to RT-LAMP performed with up  
595 to 96 individual sample barcoded primers. LAMP reactions are pooled and further amplified via  
596 PCR to introduce Illumina adapter sequences and pool-level dual indexes. A single thermal cyclers  
597 can amplify 96 or 384 such pools and the resulting “super-pool” can be sequenced overnight to  
598 detect multiple amplicons from 9,216 or 36,864 individual patient samples (number of reads in  
599 parenthesis assume an output of ~450M reads from a NextSeq 500).

600 (B) Schematic of the RT-LAMP (step I) of COV-ID. Selected numbered intermediates of RT-LAMP  
601 reaction are shown to illustrate how the LAMP barcode, shown in yellow, and the P5 and P7  
602 homology sequences (blue and pink, respectively) are introduced in the final LAMP product. Upon  
603 generating the dumb-bell intermediate the reaction proceeds through rapid primed and self-  
604 primed extensions to form a mixture of various DNA amplicons containing sequences for PCR  
605 amplification. A more detailed version of the LAMP phase of COV-ID, including specific  
606 sequences, is illustrated in Fig. S1.

607 (C) Conventional RT-LAMP primers (solid lines) or primers modified for COV-ID (dotted lines)  
608 were used for RT-LAMP of SARS-CoV-2. The numbers of inactivated SARS-CoV-2 virions per  
609  $\mu\text{L}$  is indicated in the color legend.

610 (D) Schematic of the PCR (step II) of COV-ID. Following RT-LAMP, up to 96 reactions are pooled  
611 and purified and Illumina libraries are generated directly by PCR with dual-indexed P5 and P7  
612 adapters in preparation for sequencing.

613 (E) COV-ID primers targeting ACTB mRNA were used for RT-LAMP with HeLa total RNA. LAMP  
614 was diluted 1:100, amplified via PCR and resolved on 2% agarose gel.

### 615 **Figure 2. Sequencing-based detection of SARS-CoV-2 in saliva samples**

616 (A) Saliva preparation. Crude saliva was inactivated via TCEP/EDTA addition and 95°C  
617 incubation prior to RT-LAMP.

618 (B) RT-LAMP followed by COV-ID PCR performed directly on saliva. Saliva with and without  
619 addition of 1,000 copies of inactivated SARS-COV-2 templates was inactivated as described in  
620 (A), then used as template.

621 (C) Alignment of sequenced reads against SARS-COV-2 genome from COV-ID of inactivated  
622 saliva spiked with without 1,280 virions SARS-COV-2 per  $\mu\text{L}$ . All SARS-COV-2 reads align  
623 exclusively to expected region of the N gene. Open reading frames of viral genome are depicted  
624 via gray boxes below alignment. Inset: scale shows reads per 1,000.

625 (D) Scatter plot for the ratio of SARS-CoV-2 / (*STATH* + 1) reads obtained by COV-ID (y axis)  
626 versus the number of virions per  $\mu\text{L}$  spiked in human saliva (x axis). The threshold was set above  
627 the highest values scored in a negative control (dashed line).

628 (E) COV-ID performed on clinical saliva samples. The scatter plot shows the SARS-CoV-2 /  
629 (*STATH* + 1) read ratio (y axis) versus the viral load in the sample estimated by a clinically  
630 approved, qPCR-based diagnostic test. The threshold was set based on the negative controls  
631 shown in (D).

### 632 **Figure 3. COV-ID multiplex detection of SARS-COV-2 and Influenza A**

633 (A) TCEP/EDTA treated saliva was spiked with indicated amounts of BEI heat-inactivated SARS-  
634 CoV-2 or H1N1 influenza A RNA to the indicated concentration of virions/genomes per  $\mu\text{L}$ . 1  $\mu\text{L}$  of  
635 saliva was used for COV-ID reactions.

636 (B) COV-ID was performed in two independent experiments on saliva samples from the matrix  
637 shown in (A) in the presence of 20 femtograms synthetic N2 spike-in using N2, influenza34 and  
638 *STATH* COV-ID primers. N2/(N2 Spike + 1) and influenza/(*STATH* + 1) read ratios are displayed  
639 with bars showing median  $\pm$  interquartile range. Samples were labeled as positive for a given  
640 sequence if the associated read ratio was greater than 2x the maximum value in the control saliva  
641 samples.

642 (C) Heatmaps of SARS-CoV-2 (left) or H1N1 (right) COV-ID signal in multiplex reaction.  
643 Heatmaps are colored according by percentage of viral reads observed.

### 644 **Figure 4. Application of COV-ID with paper assay**

645 (A) Scheme for COV-ID on viral RNA absorbed on paper.

646 (B) PCR reactions from paper samples immersed in water with indicated viral concentrations then  
647 amplified with N2 COV-ID primers.

648 (C) Scheme for COV-ID on saliva spiked with viral and RNA and absorbed on paper.

649 (D) Same as (B) but on saliva absorbed on paper.

650 (E) SARS-CoV-2 virus was added to saliva and prepared as in (C). RT-LAMP and sequencing  
651 was carried out in presence of SARS spike-in RNA. Viral reads are presented as ratio against the  
652 sum of *STATH* and *N2* spike-in reads. Positive threshold was set as 2x maximum value in  
653 negative saliva and indicated by dashed horizontal line.

654 (F–G): Paper-based COV-ID workflow (F) and cost calculations (G). Saliva is collected orally on  
655 a precut strip of paper, from which a 2 mm square would be cut out and added to a reaction vessel  
656 containing TCEP/EDTA inactivation buffer and processed as shown in (C).

657  
658

## 659 SUPPLEMENTARY FIGURE LEGENDS

### 660 **Figure S1. Detailed COV-ID mechanism**

661 Steps of COV-ID protocol are depicted, showing RT-LAMP mechanism and ultimate amplicon  
662 that is sequenced. For clarity only selected steps of RT-LAMP reaction are shown and loop primer  
663 intermediates are not depicted. For full LAMP mechanism see <sup>21</sup>.

### 664 **Figure S2. Optimization of COV-ID in human saliva**

665 (A) Saliva COV-ID sequence validation. Single saliva COV-ID reaction using *N2* primers was  
666 sequenced by the Sanger method.

667 (B) Validation of control human amplicons for RT-LAMP on saliva. RT-LAMP of TCEP/EDTA  
668 inactivated saliva was performed with conventional RT-LAMP primer sets for *ACTB* and *STATH*  
669 in the presence or absence of RNase A.

670 (C) Characterization of COV-ID sequencing libraries. Breakdown of reads for sequence data  
671 presented in Fig. 2D. Samples without added template consist of predominantly adapter dimers.

672 (D) Validation of COV-ID LAMP barcodes. 32 potential barcodes were tested for LAMP primer  
673 sets indicated, incompatible barcodes are marked in red.

674 (E) Validation of pooled PCR. COV-ID was performed on saliva samples using unique LAMP  
675 barcodes. The RT-LAMP reactions were then amplified either by individual PCR or by first pooling  
676 and then performing a single PCR on the pool.

### 677 **Figure S3. RT-LAMP amplification of SARS-CoV-2 spike-in RNA**

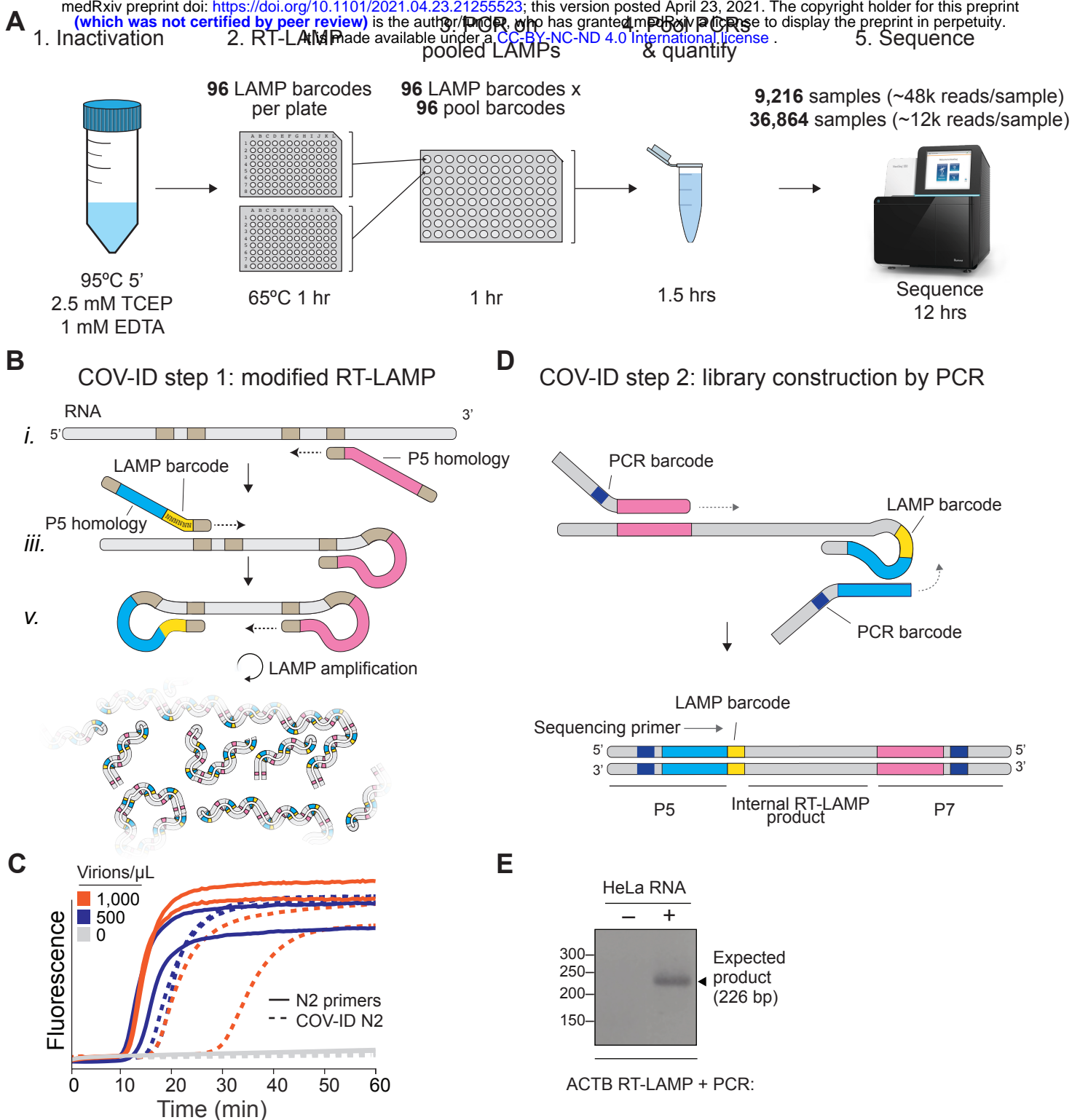
678 (A) Synthetic *N2* Spike RNA. SARS-CoV-2 *N2* RNA fragment was synthesized including 7 nt  
679 divergent sequence inside the forward loop primer-binding site, maintaining all other LAMP primer  
680 binding sites and identical GC content.

681 (B) RT-LAMP using COV-ID *N2* primers was carried out on indicated amounts of spike-in RNA,  
682 showing rapid amplification down to picogram quantities of added template.

683 (C) Total number of reads per barcode in COV-ID pool obtained by including (+) or omitting (-)  
684 the *N2* spike-in.

685 (D) Spurious COV-ID signal for the *N2* amplicon in negative control samples after normalization  
686 either to the *STATH* control in absence of spike-in (left) or to the *N2* spike-in control.

687



**Figure 1. Barcoding and PCR amplification of RT-LAMP products**

(A) Overview of COV-ID. Saliva is collected and inactivated prior to RT-LAMP performed with up to 96 individual sample barcoded primers. LAMP reactions are pooled and further amplified via PCR to introduce Illumina adapter sequences and pool-level dual indexes. A single thermal cycler can amplify 96 or 384 such pools and the resulting “super-pool” can be sequenced overnight to detect multiple amplicons from 9,216 or 36,864 individual patient samples (number of reads in parenthesis assume an output of ~450M reads from a NextSeq 500).

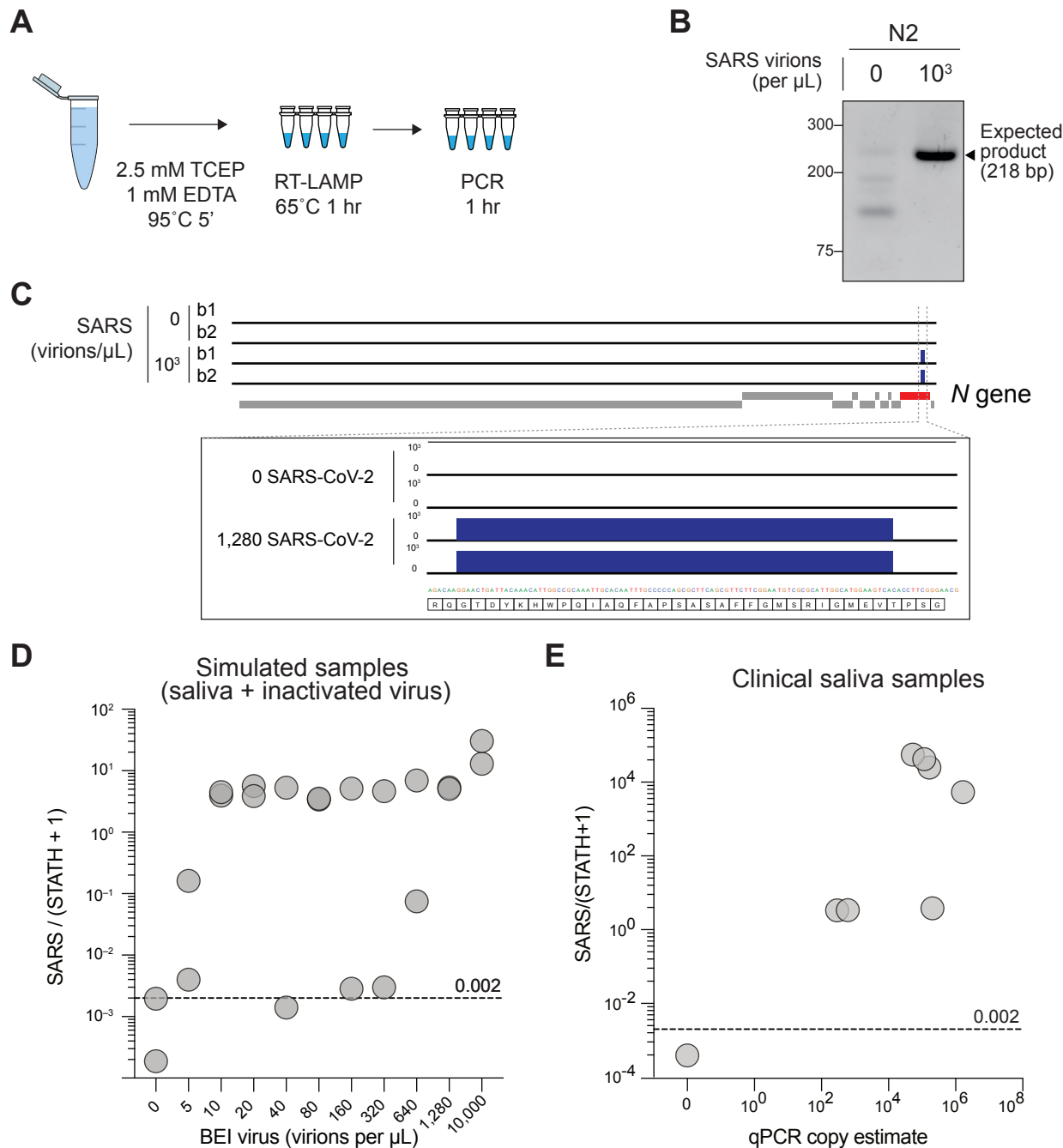
(B) Schematic of the RT-LAMP (step I) of COV-ID. Selected numbered intermediates of RT-LAMP reaction are shown to illustrate how the LAMP barcode, shown in yellow, and the P5 and P7 homology sequences (blue and pink, respectively) are introduced in the final LAMP product. Upon generating the dumb-bell intermediate the reaction proceeds through rapid primed and self-primed extensions to form mixture of various DNA amplicons containing sequences for PCR amplification. A more detailed version of the LAMP phase of COV-ID, including specific sequences, is illustrated in Fig. S1.

(C) Conventional RT-LAMP primers (solid lines) or primers modified for COV-ID (dotted lines) were used for RT-LAMP of SARS-CoV-2. The numbers of inactivated SARS-CoV-2 virions per  $\mu\text{L}$  is indicated in the color legend.

(D) Schematic of the PCR (step II) of COV-ID. Following RT-LAMP, up to 96 reactions are pooled and purified and Illumina libraries are generated directly by PCR with dual-indexed P5 and P7 adapters in preparation for sequencing.

(E) COV-ID primers targeting ACTB mRNA were used for RT-LAMP with HeLa total RNA. LAMP was diluted 1:100, amplified via PCR and resolved on 2% agarose gel.





## Figure 2. Sequencing-based detection of SARS-CoV-2 in saliva samples

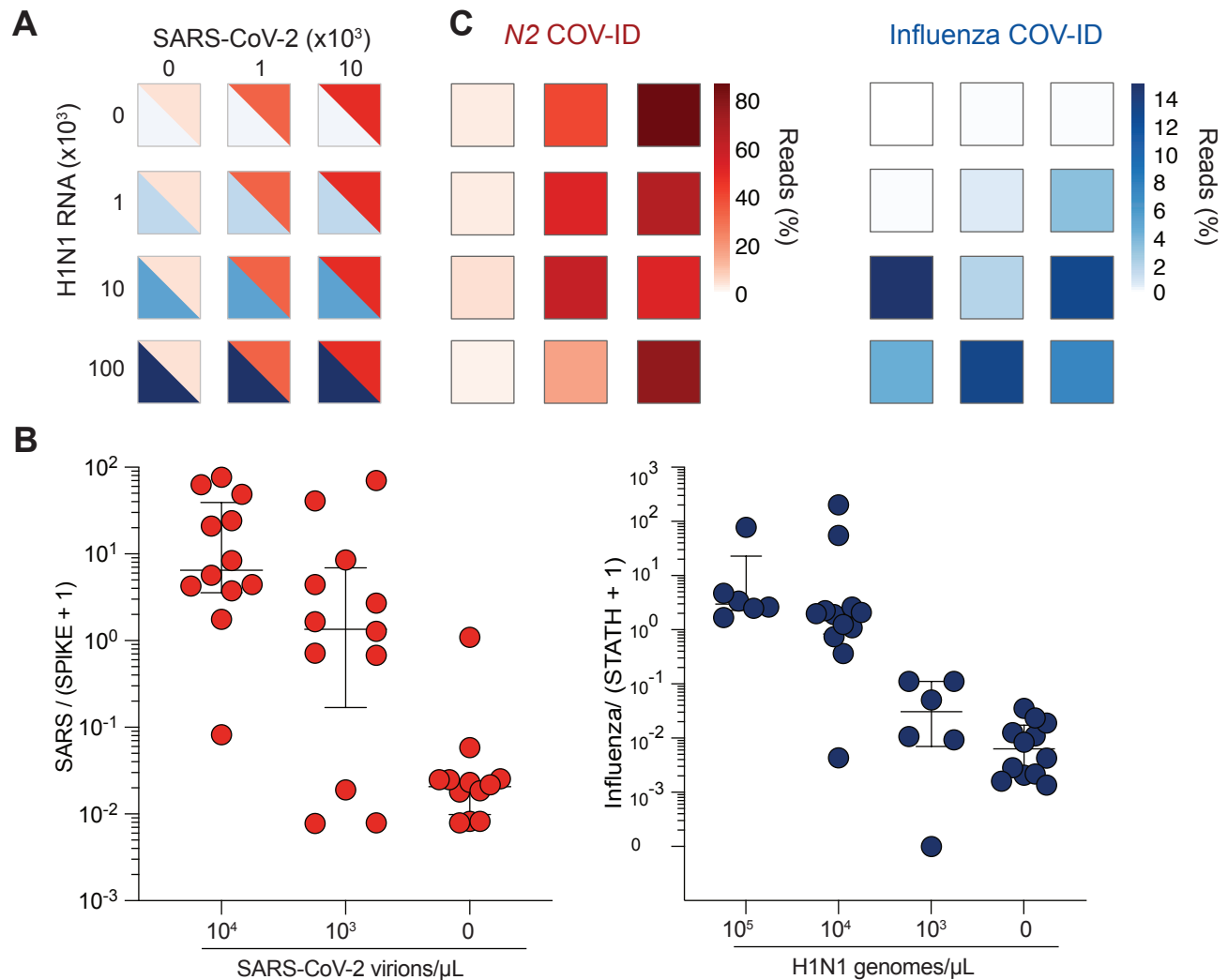
(A) Saliva preparation. Crude saliva was inactivated via TCEP/EDTA addition and 95°C incubation prior to RT-LAMP.

(B) RT-LAMP followed by COV-ID PCR performed directly on saliva. Saliva with and without addition of 1,000 copies of inactivated SARS-COV-2 templates was inactivated as described in (A), then used as template.

(C) Alignment of sequenced reads against SARS-COV-2 genome from COV-ID of inactivated saliva spiked with without 1,280 virions SARS-COV-2 per  $\mu\text{L}$ . All SARS-COV-2 reads align exclusively to expected region of the N gene. Open reading frames of viral genome are depicted via gray boxes below alignment. Inset: scale shows reads per 1,000.

(D) Scatter plot for the ratio of SARS-CoV-2 / (STATH + 1) reads obtained by COV-ID (y axis) versus the number of virions per  $\mu\text{L}$  spiked in human saliva (x axis). The threshold was set above the highest values scored in a negative control (dashed line).

(E) COV-ID performed on clinical saliva samples. The scatter plot shows the SARS-CoV-2 / (STATH + 1) read ratio (y axis) versus the viral load in the sample estimated by a clinically approved, qPCR-based diagnostic test. The threshold was set based on the negative controls shown in (D).

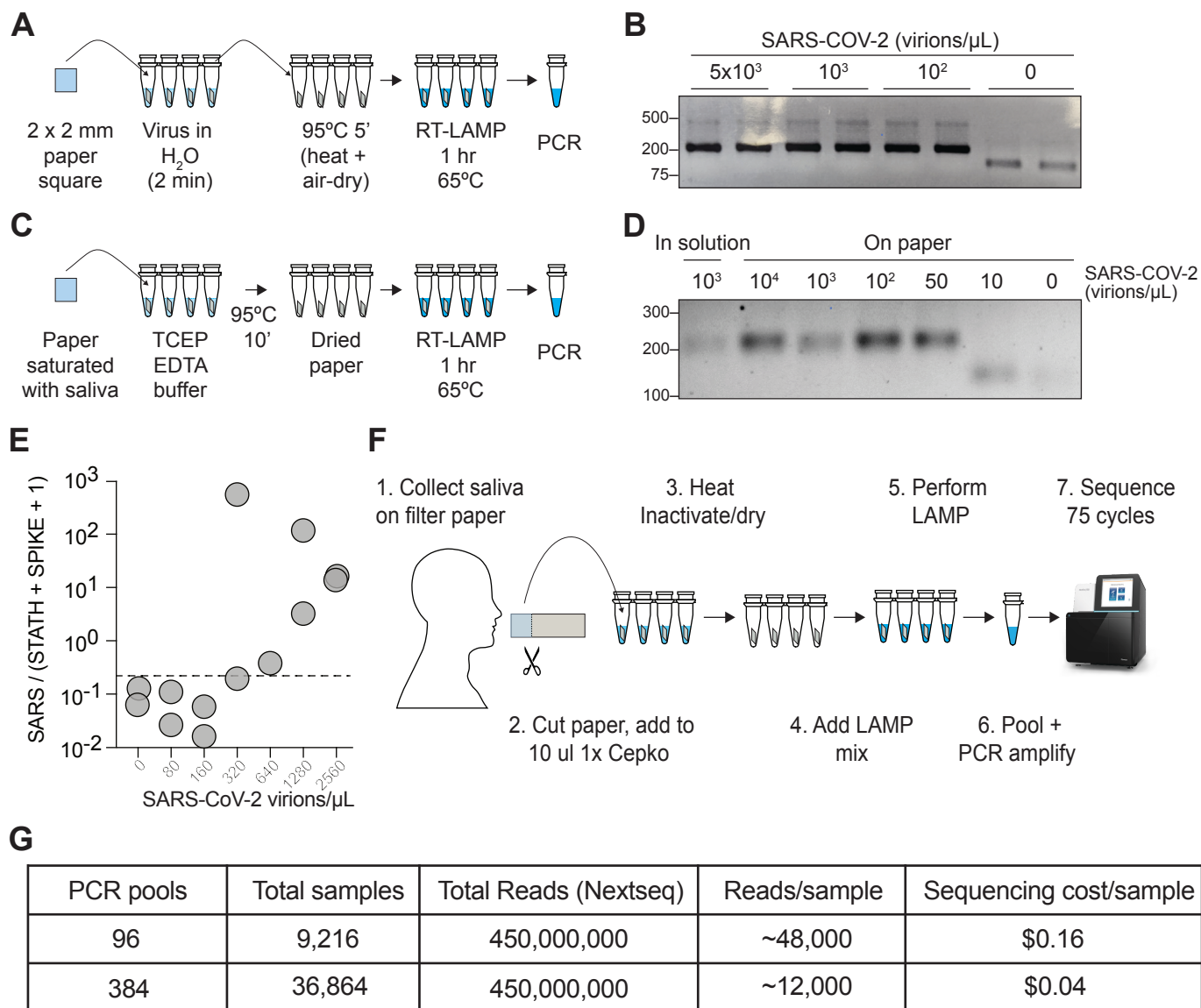


**Figure 3. COV-ID multiplex detection of SARS-COV-2 and Influenza A**

(A) TCEP/EDTA treated saliva was spiked with indicated amounts of BEI heat-inactivated SARS-CoV-2 or H1N1 influenza A RNA to the indicated concentration of virions/genomes per  $\mu$ L. 1  $\mu$ L of saliva was used for COV-ID reactions.

(B) COV-ID was performed in two independent experiments on saliva samples from the matrix shown in (A) in the presence of 20 femtograms synthetic N2 spike-in using N2, influenza (Zhang and Tanner, 2020) and STATH COV-ID primers. N2/(N2 Spike + 1) and influenza/(STATH + 1) read ratios are displayed with bars showing median  $\pm$  interquartile range.

(C) Heatmaps of SARS-CoV-2 (left) or H1N1 (right) COV-ID signal in multiplex reaction. Heatmaps are colored according by percentage of viral reads observed.



**Figure 4. COV-ID on saliva collected on paper**

(A) Scheme for COV-ID on viral RNA absorbed on paper.

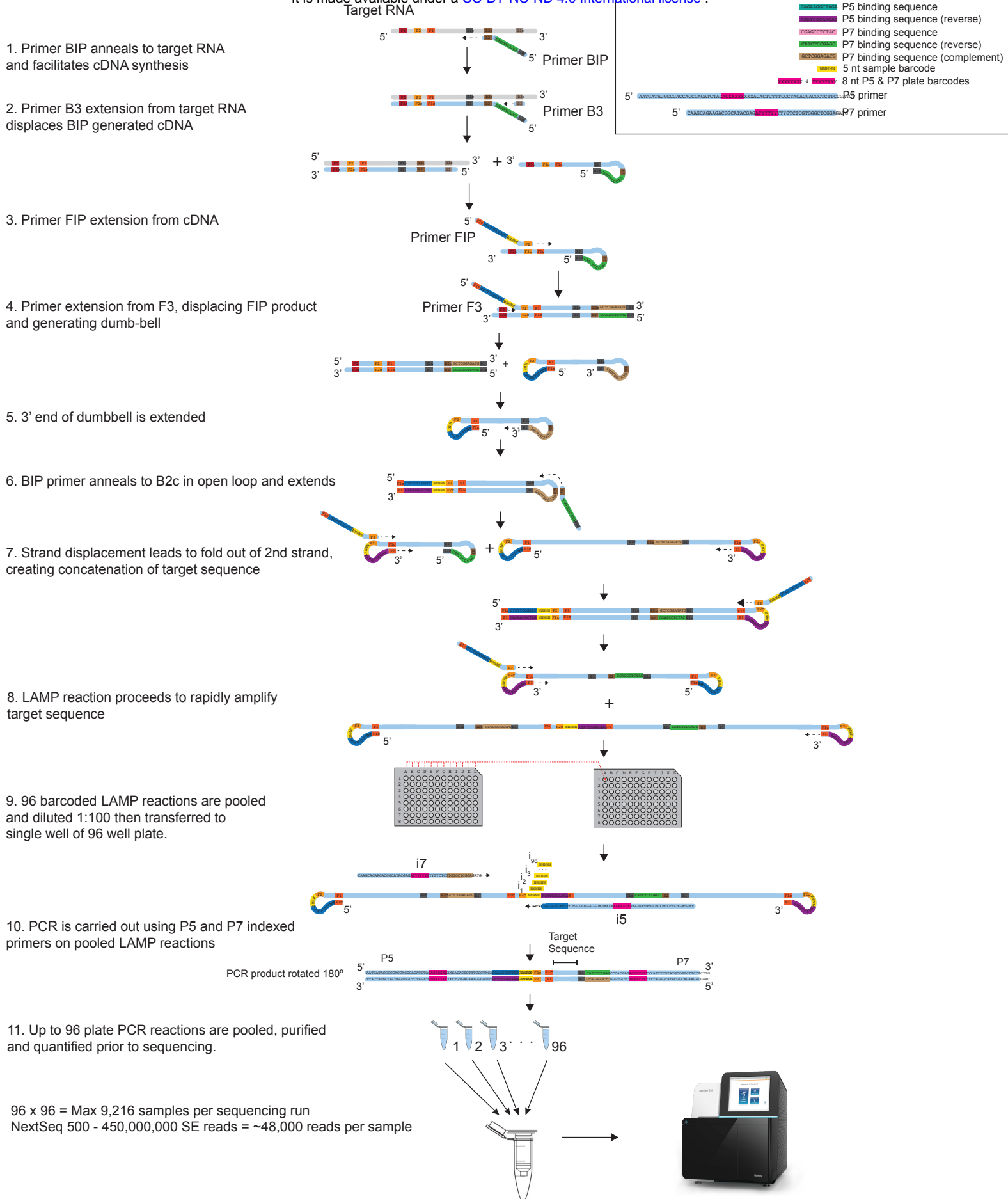
(B) PCR reactions from paper samples immersed in water with indicated viral concentrations then amplified with N2 COV-ID primers.

(C) Scheme for COV-ID on saliva spiked with viral and RNA and absorbed on paper.

(D) Same as (B) but on saliva absorbed on paper.

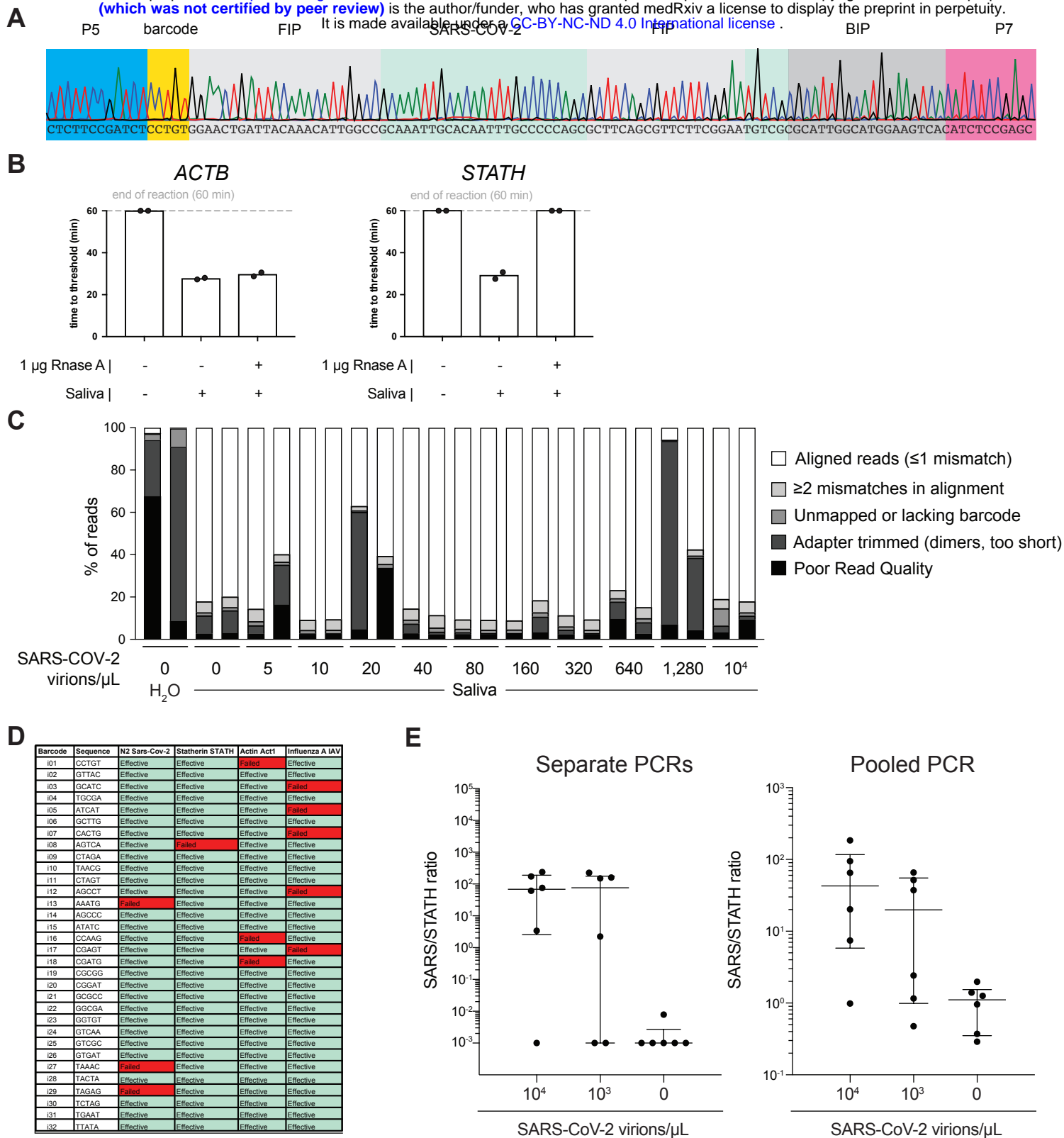
(E) SARS-CoV-2 virus was added to saliva and prepared as in (C). RT-LAMP and sequencing was carried out in presence of SARS spike-in RNA. Viral reads are presented as ratio against the sum of STATH and N2 spike-in reads. Positive threshold was set as 2x maximum value in negative saliva and indicated by dashed horizontal line.

(F–G): Paper-based COV-ID workflow (F) and cost calculations (G). Saliva is collected orally on a pre-cut strip of paper, from which a 2 mm square would be cut out and added to a reaction vessel containing TCEP/EDTA inactivation buffer and processed as shown in (C).



**Figure S1. Detailed COV-ID mechanism**

The steps of the COV-ID protocol are depicted, showing RT-LAMP mechanism and the final barcoded amplicon that is sequenced. For clarity only, selected steps of RT-LAMP reaction are shown and loop primer intermediates are not depicted. For full LAMP mechanism see (Nagamine et al., 2002).



**Figure S2. Optimization of COV-ID in human saliva**

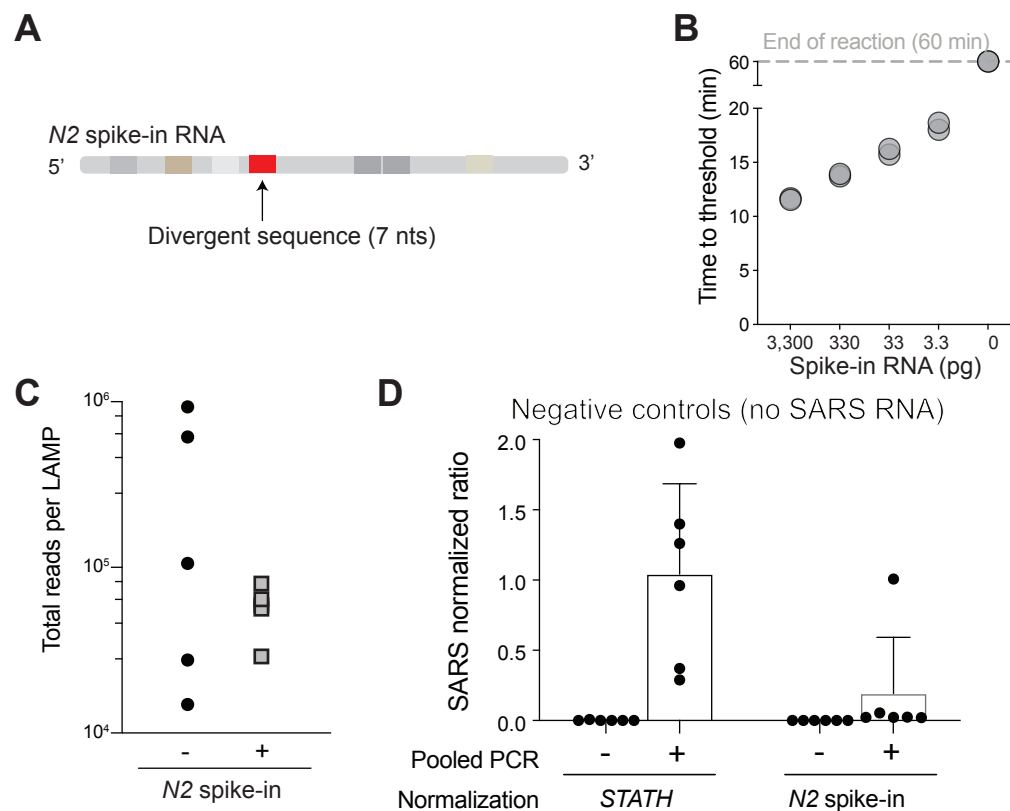
(A) Saliva COV-ID sequence validation. Single saliva COV-ID reaction using N2 primers was sequenced by the Sanger method.

(B) Validation of control human amplicons for RT-LAMP on saliva. RT-LAMP of TCEP/EDTA inactivated saliva was performed with conventional RT-LAMP primer sets for *ACTB* and *STATH* in the presence or absence of RNase A.

(C) Characterization of COV-ID sequencing libraries. Breakdown of reads for sequence data presented in Fig. 2D. Samples without added template consist of predominantly adapter dimers.

(D) Validation of COV-ID LAMP barcodes. 32 potential barcodes were tested for LAMP primer sets indicated, incompatible barcodes are marked in red.

(E) Validation of pooled PCR. COV-ID was performed on saliva samples using unique LAMP barcodes. The RT-LAMP reactions were then amplified either by individual PCR or by first pooling and then performing a single PCR on the pool.



**Figure S3. Spike-in strategy for COVID-19`**

(A) Synthetic N2 Spike RNA. SARS-CoV-2 N2 RNA fragment was synthesized including 7 nt divergent sequence inside the forward loop primer-binding site, maintaining all other LAMP primer binding sites and identical GC content.

(B) RT-LAMP using COVID-19 N2 primers was carried out on indicated amounts of spike-in RNA, showing rapid amplification down to picogram quantities of added template.

(C) Total number of reads per barcode in COVID-19 pool obtained by including (+) or omitting (-) the N2 spike-in.

(D) Spurious COVID-19 signal for the N2 amplicon in negative control samples after normalization either to the STATH control in absence of spike-in (left) or to the N2 spike-in control.