Swab-Seq: A high-throughput platform for massively scaled up SARS-CoV-2 testing

Joshua S. Bloom[1,2,3#], Eric M. Jones[3], Molly Gasperini[3], Nathan B. Lubock[3], Laila Sathe[4], Chetan Munugala[1,2], A. Sina Booeshaghi[5], Oliver F. Brandenberg[1,2,6], Longhua Guo[1,2,6], James Boocock[1,2,6], Scott W. Simpkins[3], Isabella Lin[1,4], Nathan LaPierre[7], Duke Hong[8], Yi Zhang[1], Gabriel Oland[9], Bianca Judy Choe[10], Sukantha Chandrasekaran[4], Evann E. Hilt[4], Manish J. Butte[11,12], Robert Damoiseaux[13,14,15], Clifford Kravit[15], Aaron R. Cooper[3], Yi Yin[1], Lior Pachter[17], Omai B. Garner[4], Jonathan Flint[1,18], Eleazar Eskin[1,7,8], Chongyuan Luo[1], Sriram Kosuri[3,19#], Leonid Kruglyak[1,2,6#], Valerie A. Arboleda[1,4#]

1 Department of Human Genetics, David Geffen School of Medicine, UCLA
2 Howard Hughes Medical Institute, HHMI
3 Octant, Inc.
4 Department of Pathology & Laboratory Medicine, David Geffen School of Medicine, UCLA
5 Department of Mechanical Engineering, Caltech
6 Department of Biological Chemistry, David Geffen School of Medicine, UCLA
7 Department of Computer Science, Samueli School of Engineering, UCLA
8 Department of Computational Medicine, David Geffen School of Medicine, UCLA
9 Department of Surgery, David Geffen School of Medicine, UCLA
10 Department of Emergency Medicine, David Geffen School of Medicine, UCLA
11 Department of Pediatrics, David Geffen School of Medicine, UCLA
12 Department of Microbiology, Immunology & Molecular Genetics, David Geffen School of Medicine, UCLA
13 California NanoSystems Institute, UCLA
14 Department of Bioengineering, Samueli School of Engineering, UCLA
15 David Geffen School of Medicine, Research Information Technology
16 Department of Medical and Molecular Pharmacology, David Geffen School of Medicine, UCLA
17 Division of Biology and Bioengineering & Department of Computing and Mathematical Sciences, Caltech
18 Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, UCLA
19 Department of Chemistry and Biochemistry, UCLA
Correspondence should be addressed to jbloom@mednet.ucla.edu, sri@octant.bio, lkruglyak@mednet.ucla.edu, or varboleda@mednet.ucla.edu

ABSTRACT

The rapid spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is due to the high rates of transmission by individuals who are asymptomatic at the time of transmission[1,2]. Frequent, widespread testing of the asymptomatic population for SARS-CoV-2 is essential to suppress viral transmission and key for safely reopening society. Despite increases in testing capacity, multiple challenges remain in deploying traditional reverse transcription and quantitative PCR (RT-qPCR) tests at the scale required for population screening of asymptomatic individuals. We have developed SwabSeq, a high-throughput testing platform for SARS-CoV-2 that uses next-generation sequencing as a readout. SwabSeq employs sample-specific molecular barcodes to enable thousands of samples to be combined and simultaneously analyzed for the presence or absence of SARS-CoV-2 in a single run. Importantly, SwabSeq incorporates an *in vitro* RNA standard that mimics the viral amplicon, but can be distinguished by sequencing. This standard allows for end-point rather than quantitative PCR, improves quantitation, reduces requirements for automation and sample-to-sample normalization, enables purification-free detection, and gives better ability to call true negatives. We show that SwabSeq can test nasal and oral specimens for SARS-CoV-2 with or without RNA extraction while maintaining analytical sensitivity better than or comparable to that of fluorescence-based RT-qPCR tests. We have tested 1215 clinical samples: purified RNA from nasal swabs (n=380), and extraction free protocols for nasal swabs (n=298) and extraction free saliva (n=537) that combined show 96.6% agreement, high sensitivity and specificity when compared to high sensitivity nasal swab qPCR assays. SwabSeq is simple, sensitive, flexible, rapidly scalable, inexpensive enough to test widely and frequently, and, in clinical operation, provides a turn around time of 12 to 18 hours.

INTRODUCTION

In the absence of an effective vaccine or prophylactic treatment, public health strategies remain the only tools for controlling the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the cause of COVID-19. In contrast to SARS-CoV-1, for which infectivity is associated with symptoms[3,4], infectivity of SARS-CoV-2 is high during the asymptomatic/presymptomatic phase[5,6]. As a consequence, containing transmission based solely on symptoms is impossible, which makes molecular screening for SARS-CoV-2 essential for pandemic control.

As regional lockdowns have been lifted and people have returned to work and resumed other activities, rates of infection have started to rise again[7]. In many parts of the United States, the rise in cases has overwhelmed the capacity of quantitative RT-PCR (qRT-PCR) tests that make up the majority of FDA-authorized tests for COVID-19. Delays in obtaining test results, which are due to capacity constraints rather than assay times[8], render testing ineffective for the public health aims of preventing viral transmission and suppressing local outbreaks. Even where expanded capacity exists, the ~$100 price of tests (current Medicare reimbursement rates[9]) prohibits the widespread adoption by large employers and schools on a regular basis for effective viral suppression[10,11]. Frequent, low-cost population testing, combined with contact tracing and isolation of infected individuals, would help to halt the spread of COVID-19 and reopen society[12,13]**.** Here we describe SwabSeq, a SARS-CoV-2 testing platform that leverages next-generation sequencing to massively scale up testing capacity[14,15].

SwabSeq improves on one-step reverse transcription and polymerase chain reaction (RT-PCR) approaches in several key areas. Like other sequencing approaches, SwabSeq utilizes molecular barcodes that are embedded in the RT-PCR primers to uniquely label each sample and allow for simultaneous sequencing of hundreds to thousands of samples in a single run (LampSeq[16], Illumina CovidSeq[17], DxSeq[18]). SwabSeq uses very short reads, reducing sequencing times so that results can be returned in less than 24 hours.

To deliver robust and reliable results at scale, SwabSeq adds to every sample a synthetic *in vitro* RNA standard that is almost the same as that of the virus, but can be distinguished easily by sequencing. SARS-CoV-2 detection is based on the ratio of the counts of true viral sequencing reads to those from the *in vitro* viral standard. Since every sample contains the synthetic RNA, SwabSeq controls for failure of amplification: samples with no SARS-CoV-2 detected are those in which only *in vitro* viral standard reads are observed, while those without viral or *in vitro* viral standard reads are inconclusive.

The RNA control confers a number of additional advantages on the SwabSeq assay. Since we are only interested in the ratio of real virus to *in vitro* standard, the PCR can be run to the endpoint, where all primers are consumed, rather than for a set number of cycles. By driving the reaction to endpoint, we overcome the presence of varying amounts of RT and PCR inhibitors and effectively force each sample to have similar amounts of final product. Using *in vitro* standard RNA with end-point PCR has two important consequences. First, by overcoming the heterogeneity that inevitably occurs with clinical samples, we can pool reaction products after PCR, without the need to normalize each sample individually. Second, it enables direct processing of extraction-free samples. Inhibitors of RT and PCR present in mucosal tissue or saliva should affect both the virus and the *in vitro* standard equally. Endpoint PCR overcomes the effect of inhibition, while keeping the ratio of reads between the two RNA species approximately constant, and so avoids the need for extraction.

Here we show that SwabSeq has extremely high sensitivity and specificity for the detection of viral RNA in purified samples. We also demonstrate a low limit of detection in extraction-free lysates from mid-nasal swabs and oral fluids. These results demonstrate the potential of SwabSeq to be used for SARS-CoV-2 testing on an unprecedented scale, offering a potential solution to the need for population-wide testing to stem the pandemic.

RESULTS

SwabSeq is a simple and scalable protocol, consisting of 5 steps (**Figure 1A**): (1) sample collection, (2) reverse transcription and PCR using primers that contain unique molecular indices at the i7 and i5 positions (**Figure 1B, Figure S1**) as well as *in vitro* standards, (3) a simple pooling (no normalization) and cleanup of the uniquely barcoded samples for library preparation, (4) sequencing of the pooled library, and (5) computational assignment of barcoded sequencing reads to each sample for counting and viral detection.

Our assay consists of two primer sets that amplify two genes: the S2 gene of SARS-CoV-2 and the human *Ribonuclease P/MRP Subunit P30* (*RPP30*). We include a synthetic *in vitro* RNA standard that is identical to the viral sequence targeted for amplification, except for the most upstream 6 bp (**Figure 1C**), which allows us to distinguish sequencing reads corresponding to the *in vitro* standard from those corresponding to the target sequence**.** The primers amplify both the viral and the synthetic sequences with equal efficiency (**Figure S2**). We have also added a second RNA standard for RPP30 with a similar design. The ratio of the number of native reads to the number of *in vitro* standard reads provides a more accurate and quantitative measure of the number of viral genomes in the sample than native read counts alone (**Figure 1D**,**E**). The *in vitro* standard also allows us to retain linearity over a large range of viral input despite the use of endpoint PCR (**Figure S3**). With this approach, the final amount of DNA in each well is largely defined by the total primer concentration rather than by the viral input—negative samples have high amounts of *in vitro* S2 standard (abbreviated as S2-spike) and low/zero amounts of viral reads, and positive samples have low amounts of *in vitro* S2 standard and high amounts of viral reads (**Figure S3**). In addition to viral S2, we reverse-transcribe and amplify a human housekeeping gene to control for specimen quality, as in traditional qPCR assays (**Figure 1F**). The i5/i7 barcodes used are designed to be at least several base-pair edits away from one another, allowing for assignment even in the face of sequencing errors.

After RT-PCR, samples are combined at equal volumes, purified, and used to generate one sequencing library. We have used both the Illumina MiSeq and the Illumina NextSeq 550 to sequence these libraries (**Figure S4**). We minimize instrument sequencing time by sequencing only the minimum required 26 base pairs (**Methods**). Each read is classified as deriving from native or *in vitro* standard S2, or RPP30, and assigned to a sample based on the associated index sequences (barcodes). To maximize specificity and avoid false-positive signals arising from incorrect classification or assignment, conservative edit distance thresholds are used for this matching operation (**Methods** and **Supplemental Results**). A sequencing read is discarded if it does not match one of the expected sequences. Counts for native and *in vitro* standard S2 and RPP30 reads are obtained for each sample and used for downstream analyses (**Methods**).

We have estimated that approximately 5,000 reads per well are sufficient to detect the presence or absence of viral RNA in a sample (**Methods**). This translates to at least 1,500 samples per run on a MiSeq or MiniSeq, 20,000 samples per run on a NextSeq 550, and up to 150,000 samples per run on a NovaSeq S2 flow cell. Computational analysis takes only minutes per run[19]. The MiniSeq newly released rapid output flow cell has a turn around time of 2.5 hours, which significantly improves our throughput and turn-around time. Our optimized protocols allow for a single person to process 6-384 well plates in a single hour; which is equivalent to 2304 samples an hour. This process cloned over multiple liquid handlers can rapidly scale to 10,000 samples per hour with a staff of 5 people.

We have optimized the SwabSeq protocol by identifying and eliminating multiple sources of noise (**Supplemental Results**) to create a streamlined and scalable protocol for SARS-CoV-2 testing.

Validation of SwabSeq as a diagnostic platform

We first validated SwabSeq on purified RNA nasopharyngeal (NP) samples that were previously tested by the UCLA Clinical Microbiology Laboratory with a standard RT-qPCR assay (ThermoFisher Taqpath COVID19 Combo Kit). To determine our analytical limit of detection, we diluted inactivated virus with pooled, remnant clinical NP swab specimens. The remnant samples were all confirmed to be negative for SARS-CoV-2. In these remnant samples, we performed a serial, 2-fold dilution of heat-inactivated SARS-CoV-2 *(*ATCC® VR-1986HK), from 8,000 to 125 genome copy equivalents (GCE) per mL. We detected SARS-CoV-2 in 34/34 samples down to 250 GCE per mL, and in 28/34 samples down to 125 GCE per mL **(Figure 2A**). These results established that SwabSeq is highly sensitive, with an analytical limit of detection (LOD) of 250 GCE per mL for purified RNA from nasal swabs. This limit of detection is lower than those of many currently FDA authorized and highly sensitive RT-qPCR assays for SARS-CoV-2.

SwabSeq detects the SARS-CoV-2 genome with high clinical sensitivity and specificity. We retested 668 RNA-purified nasopharyngeal samples from the UCLA Clinical Microbiology Laboratory in which SARS-CoV-2 positive (n=94) and negative (n=574). We observed 99.2% agreement with RT-qPCR results for all samples (**Figure 2B and S5D**) with 100% concordance in samples that were positive for SARS-CoV-2. We sequenced these samples on either a MiSeq, MiniSeq or on a NextSeq550 (**Figure S5**), with 100% concordance between the different sequencing instruments.

One of the major bottlenecks in scaling up RT-qPCR diagnostic tests is the RNA purification step. RNA extraction is challenging to automate, and supply chains have not been able to keep up with the demand for necessary reagents during the course of the pandemic. Thus, we explored the ability of SwabSeq to detect SARS-CoV-2 directly from a variety of extraction-free sample types. There are several types of media that are recommended by the CDC for nasal swab collection: viral transport medium (VTM)[20], Amies transport medium[21], and normal saline[21]. A main technical challenge arises from RT or PCR inhibition by ingredients in

the collection buffers. We found that dilution of specimens with water overcame the RT and PCR inhibition and allowed us to detect viral RNA in contrived and positive clinical patient samples, at limits of detections between 4000 and 6000 GCE/mL (**Figure S6**). We also tested nasal swabs that were collected directly into Tris-EDTA (TE) buffer, diluted 1:1 with water. This approach yielded a limit of detection of 560 GCE/mL (**Figure 2C**). We next performed a comparison between traditional purified RT-qPCR protocols and our extraction free-protocol for nasopharyngeal samples collected into either normal saline (n=128) or TE (n=170). We showed 96.0% overall agreement, 92.0% positive percent agreement and 97.6% negative percent agreement for all samples and a high correlation (r^2=0.68) between SwabSeq signal and RT-qPCR (**Figure 2D, Figures S7-S8**).

We also tested extraction-free saliva protocols in which saliva is collected directly into a matrix tube using a funnel-like collection device **(Figure S9)**. The main technical challenges in demonstrating the detection of virus in saliva samples have been preventing degradation of the inactivated SARS-CoV-2 virus that is added to saliva and ensuring accurate pipetting of this heterogeneous and viscous sample type. We found that heating the saliva samples to 95 □ for 30 minutes[22] reduced PCR inhibition and improved detection of the S2 amplicon (**Figure S10**). After heating, we diluted samples at a 1:1 volume with 2x TBE with 1% Tween-20[22]. Using this method, we obtained a LoD of 2000 GCE/mL (**Figure 2E)**. In a study performed in the UCLA Emergency Department, we collected paired saliva and nasopharyngeal swab samples in patients with COVID19-like symptoms. In a comparison between our extraction-free, saliva SwabSeq protocol and purified nasopharyngeal swab samples run with high-sensitivity RT-qPCR in the UCLA Clinical Microbiology lab, we showed 95.3% overall agreement, 90.2% positive percent agreement, and 96.3% negative percent agreement **(Figure S11)**.

With a highly scalable diagnostic platform, such as SwabSeq, specimen accessioning, securing supply chains, and processing samples all present challenges. We have redesigned the pre-analytic processes to prioritize self-registration, rapid sample collection, and leverage

simple automation once samples are received in the lab (**Figure S9**). We have developed a web app that allows us to push the registration of samples to the individual. The second innovation is sample collection into tubes that can be uncapped in a racked, 96-tube format. This allows us to pipet batches of 96-samples at a time directly into a RT-PCR mix **(Figure S9)**.

The development of extraction free protocols prioritized “low-touch”. Our inactivation protocol requires a simple heat step of the specimen tube, minimizing exposure and pipetting of the sample. In addition, we only use only a 1 liquid-handler tip per sample thereby minimizing supply chain challenges that continue to emerge.

We have calculated that a single person can process 6 384-well plates per hour using an automated liquid handler. In our COVID-19 testing laboratory, we have scaled this to four-liquid handlers capable of processing up to 24 plates/hour, which is equivalent to a capacity of 9,216 samples per hour. A laboratory running 24 hours a day could process well over 100,000 samples per day. Furthermore, the post-PCR multiplexing and library generation is designed such that a single person can handle thousands of samples concurrently, the number of samples is dependent only on the capacity of the sequencer. For smaller scale university settings, having an automated decapper, single liquid handler and MiniSeq, and investment of less than $150,000 in equipment, would be able to diagnose up to ~3000 samples a day.

These optimizations allow for rapid scaling of the SwabSeq diagnostic platform. We tested sample collection and processing workflows in a variety of settings, including the emergency room, return-to-school testing and university-wide population screening. These and other optimizations demonstrate a path to rapid scaling in population-dense settings such as university campuses.

DISCUSSION

Swabseq has the potential to alleviate existing bottlenecks in diagnostic clinical testing. We believe that it has even greater potential to enable testing on a scale necessary

for pandemic suppression via population surveillance. The technology represents a novel use of massively parallel next-generation sequencing for infectious disease surveillance and diagnostics. We have demonstrated that SwabSeq can detect SARS-CoV-2 RNA in clinical specimens from both purified RNA and extraction-free lysates, with clinical and analytical sensitivity and specificity comparable to RT-qPCR performed in a clinical diagnostic laboratory. We have optimized SwabSeq to prioritize scale and low cost, as these are the key factors missing from current COVID-19 diagnostic platforms.

Methods for surveillance testing, such as SwabSeq, should be evaluated differently than those for clinical testing. Clinical testing informs medical decision-making, and thus requires high sensitivity and specificity. For surveillance testing, the most important factors are the breadth and frequency of testing and the turn-around-time[12]. Sufficiently broad and frequent testing with rapid return of results, contact tracing, and quarantining of infectious individuals can effectively contain viral outbreaks, avoiding blanket stay-at-home orders. Epidemiological modeling of surveillance testing on university campuses has shown that diagnostic tests with only 70% sensitivity, performed frequently with a short turn-around time, can suppress transmission[13]. However, there remain major challenges for practical implementation of frequent testing, including the cost of testing and the logistics of collecting and processing thousands of samples per day.

The use of next generation sequencing in diagnostic testing has garnered concern about turn-around-time and cost. SwabSeq uses short sequencing runs that read out the molecular indexes and 26 base pairs of the target sequence in as little as 3 hours, followed by computational analysis that can be performed on a desktop computer in 5 minutes. The cost of 1,000 samples analyzed in one MiSeq run is less than $1 per sample for sequencing reagents. Running 10,000 samples on a NextSeq550, which generates 13 times more reads per flow cell, can reduce this sequencing cost approximately 10-fold. We estimate that the total consumable cost, inclusive of the collection kit and informatics infrastructure, ranges from 4 to 6 dollars per

test. The total cost of operations depends on a wide range of factors, including the costs of setting up CLIA-certified laboratory, certified personnel, logistics of sample collection, and result reporting. Ongoing optimization to decrease reaction volumes and to use less expensive RT-PCR reagents can further decrease the total cost per test.

Finally, scaling up testing for SARS-CoV-2 requires high-throughput sample collection and processing workflows. Manual processes, common in most academic clinical laboratories, are not easily compatible with simple automation. The current protocols with nasopharyngeal swabs into viral transport media, Amies buffer or normal saline are collection methods that date back to the pre-molecular-genetics era, when live viral culture was used to identify cytopathic effects on cell lines. A fresh perspective on collection methods that are easily scalable would be enormously beneficial to scaling up centralized laboratory testing approaches.

Several groups, including ours, have piloted “lightweight” sample collection approaches, which push sample registration and patient information collection directly onto the individual tested via a smartphone app. Much of the labor of sample acquisition is due to a lack of interoperability between electronic health systems, with laboratory professionals manually entering information for every sample by hand. By developing a HIPAA-compliant registration process, we aim to streamline labor-intensive sample accessioning. To promote scalability, we have also started to develop sample collection protocols that use smaller-volume tubes that are compatible with simple automation, such as automated capper-decapper and 96-head liquid handlers[23,24]. These approaches decrease the amount of hands-on work required in the laboratory to process and perform tests leading to higher reproducibility, faster turn-around time and decreasing exposure risk to laboratory workers.

The SwabSeq diagnostic platform complements traditional clinical diagnostics tests[25], as well as the growing arsenal of point-of-care rapid diagnostic platforms[26] emerging for COVID-19, by increasing test capacity to meet the needs of both diagnostic and widespread surveillance testing. Looking forward, SwabSeq is easily extensible to accommodate additional pathogens

and viral targets. This would be particularly useful during the winter cold and flu season, when multiple respiratory pathogens circulate in the population and cannot be easily differentiated based on symptoms alone. Surveillance testing is likely to become a part of the new normal as we aim to safely reopen the educational, business and recreational sectors of our society.

SOFTWARE AND DATA

https://github.com/joshsbloom/swabseq

https://github.com/octantbio/SwabSeq

The core technology has been made available under the Open Covid Pledge, and software and data under the MIT license (UCLA) and Apache 2.0 license (Octant Inc.).

Acknowledgments. We thank Jane Semel. Without her support this work would not have been possible. We also thank the Held Foundation for all their support of this project. We thank the UCLA David Geffen School of Medicine's Dean's Office for their support, the Fast Grants, Inc for funding of this work. We also thank Lea Starita, Beth Martin, Jase Gehring, Sanjay Srivatsan, Jay Shendure, and the members of the Covid Testing Scaleup Slack for their input, guidance and openness in sharing their processes. This work was supported by funding from the Howard Hughes Medical Institute (to LK) and DP5OD024579 (to VA). IL is supported by T32GM008042. We thank Marlene Berro for her guidance with the FDA EUA201963. We also thank the clinical lab specialists in the UCLA Clinical Microbiology lab for their assistance in collecting and processing the remnant specimens and data. Biorender.com was used to generate figures with cartoons. We thank all the video games and video game makers that have helped keep our loved ones sane as we spent all our time on SwabSeq.

E.M.J, M.G., N.B.L., S.W.S. and S.K. are employed by and hold equity, J.S.B. consults for and holds equity, and A.R.C holds equity in Octant Inc. which initially developed SwabSeq, and has filed for patents for some of the work here, though they have been made available under the Open Covid License:

https://www.notion.so/Octant-COVID-License-816b04b442674433a2a58bff2d8288df .

Data Availability

Source data for all figures are available on https://github.com/joshsbloom/swabseq**.** All protocols and primers are available under an Open Covid License https://www.notion.so/Octant-COVID-License-816b04b442674433a2a58bff2d8288df .

Code Availability

All code can be accessed on https://github.com/joshsbloom/swabseq

Author Contributions

JSB and VA wrote the manuscript with assistance from CL, JF, LK, EE, EJ, AC, NL, MG, SK. EJ, AC, NL, MG, SS, JSB, SK designed barcodes and performed early testing and analysis of protocols and reagents. CL, YY, YZ, LG, RD, MB provided early guidance and key automation resources. EE, DH, NLP, KC and CK developed the registration webapp and IT infrastructure. LS, CM, MG, EJ, NL, SK, IL, OFB, VA, JSB performed and analyzed experiments. ASB and LP analyzed mis-assignment of index barcodes. VA, OG, SC, EH, GO, BJC collected and processed clinical samples. EE, LK, JF, CL, YY, YZ, JB provided helpful insight into protocols, software, and development and optimization of our specimen collection and handling.

METHODS

Sample Collection. All patient samples used in our study were deidentified. All samples were obtained with UCLA IRB approval. Nasopharyngeal samples were collected by health care providers from individuals whom physicians suspected to have COVID19.

Creation of Contrived Specimens. For the clinical limit of detection experiments, we pooled confirmed, COVID-19 negative remnant nasopharyngeal swab specimens collected by the UCLA Clinical Microbiology Laboratory. Pooled clinical samples were then spiked with ATCC Inactivated Virus (ATCC 1986-HK) at specified concentrations and extracted as described below. For the clinical purified RNA samples, they were collected as nasopharyngeal swabs and purified using the KingFisherFlex (Thermofisher Scientific) instrument using the MagMax bead extraction. All extractions were performed according to manufacturer's protocols. For extraction-free samples, we first contrived samples at specified concentrations into pooled, confirmed negative clinical samples and diluted samples in TE buffer or water prior to adding to the RT-PCR master mix.

Processing of Extraction-Free Saliva Specimens

Direct saliva is collected into a Matrix tube (Thermofisher, 3741-BR) using a small funnel (TWDRer 6565). The saliva samples were collected into a matrix tube and heated to 95□ for 30 minutes. Samples were then either frozen at -80□ or processed by dilution with 2X TBE with 1% Tween-20, for final concentration of 1x TBE and 0.5% Tween-20[22]. We also tested 1x Tween with Qiagen Protease and RNA Secure (ThermoFisher), which also works but resulted in more sample-to-sample variability and required additional incubation steps.

Processing of Extraction-Free Nasal Swab Lysates

All extraction-free lysates were inactivated using a heat inactivation at 56□ for 30 minutes. Samples were then diluted with water at a ratio of 1:4 and directly added to mastermix. Dilution amounts varied depending on the liquid media that was used. We found that of the CDC recommended media, normal saline performed the most robustly. Viral Transport Media and Amies Buffer showed significant PCR inhibition that was difficult to overcome, even with dilution in water. We recommend placing the swab directly into the diluted TE buffer, which has little PCR inhibition.

Barcode Primer Design

Barcode primers were chosen from a set of 1,536 unique 10bp i5 barcodes and a set of 1,536 unique 10bp i7 barcodes. These 10 bp barcodes satisfied the criteria that there is a minimum Levenshtein[27] distance of 3 between any two indices (within the i5 and i7 sets) and that the barcodes contain no homopolymer repeats greater than 2 nucleotides. Additionally, barcodes were chosen to minimize homo- and hetero-dimerization using helper functions in the python API to Primer3[28]. Additional details and code for primer design can be found at https://github.com/octantbio/SwabSeq.

Construction of S2 and RPP30 *in vitro* standard

S2_FP	TAATACGACTCACTATAGGGCTGGTGCTGCAGCTTATTATGTGGGTATAG AACAACCTAGGACTTTTCTATTAA
S2_RP	AACGTACACTTTGTTTCTGAGAGAGG
RPP30_FP_1	CTGACCTGAAGGCTGACGCCGGACTTGTGGAGACAGC
RPP30_FP_2	TAATACGACTCACTATAGGGAGATTTGGACCTGCGAGCGGGTTCTGACC TGAAGGCTGA
RPP30_R	GGTTTTTCAATTTCCTGTTTCTTTTCCTTAAAGTCAACG

RT-PCR was performed using primers shown above on gRNA of SARS-CoV-2 (Twist BioSciences, #1) for construction of a *in vitro* S2 standard DNA template. RT PCR (FP_1, R) and a second round of PCR (FP_2, R) was performed on HEK293T lysate for construction of an *in vitro* RPP30 standard DNA template. Products were run on a gel to identify specific products at ~150 bp. DNA was purified using Ampure beads (Axygen) using a 1.8 ratio of beads:sample volume. The mixture was vortexed and incubated for 5 minutes at room temperature. A magnet was used to bind beads for 1 minute, washed twice with 70% EtOH, beads were air-dried for 5 minutes, and then removed from the magnet and eluted in 100 µL of IDTE Buffer. The bead solution was placed back on the magnet and the eluate was removed after 1 minute. DNA was quantified by nanodrop (Denovix).

This prepared DNA template was used for standard HiScribe T7 in vitro transcription (NEB). IVT reactions prepared according to the manufacturer's instructions using 300 ng of template DNA per 20 uL reaction with a 16 hour incubation at 37□. IVT reactions were treated with DNAseI according to the manufacturer's instructions. RNA was purified with an RNA Clean & Concentrator-25 kit (Zymo Research) according to the manufacturer's instructions and eluted

into water. RNA standard was quantified both by nanodrop and with a RNA screen tape kit for the TapeStation according to the manufacturer's instructions (Agilent) to verify the RNA was the correct size (~133 nt).

One-Step RT-PCR

RT-PCR were performed using either the Luna® Universal One-Step RT-qPCR Kit (New England BioSciences E3005) or the TaqPath™ 1-Step RT-qPCR Master Mix (Thermofisher Scientific, A15300) with a reaction volume of 20 µL. Both kits were used according to the manufacturer's protocol. The final concentration of primers in our mastermix was 50 nM for RPP30 F and R primers and 400 nM for S2 F and R primers. Synthetic S2 RNA was added directly to the mastermix at a copy number of 500 copies per reaction. Sample was loaded into a 20 µL reaction. All reactions were run on a 96- or 384-well format and thermocycler conditions were run according to the manufacturer's protocol. We observe significant differences between the amplification of samples from purified RNA versus extraction-free (unpurified swab) samples (**Figure S12**). For purified RNA samples we performed 40 cycles of PCR. For extraction-free samples, we performed endpoint PCR for 50 cycles.

Multiplex Library Preparation

After the RT-PCR reaction, samples were pooled using a multichannel pipet or Integra Viaflow Benchtop liquid handler. 6 µL from each well were combined in a sterile reservoir and transferred into a 15 mL conical tube and vortexed. 100 uL of the pool was transferred to a 1.7 mL eppendorf tube for a double-sided SPRI cleanup [29]. Briefly, 50 µL of AmpureXP beads (Beckman Coulter A63880) were added to 100 µL of the pooled PCR volume and vortexed. After 5 minutes, a magnet was used to collect beads for 1 minute and supernatant transferred to a new eppendorf tube. An additional 130 µL of Ampure XP beads were added to the 150 µL of supernatant and vortexed. After an additional 5 minutes, the magnet was used to collect beads for 1 minute and the beads were washed twice with fresh 70% EtOH. DNA was eluted off the beads in 40 µL of Qiagen EB buffer. The magnet was used to collect beads for 1 minute and 33 µL of supernatant was transferred to a new tube. Purified RNA was quantified and library quality was assessed using the Agilent TapeStation. We observe some differences in non-specific peaks in our TapeStation analysis of the final library preparation, particularly when sequencing unpurified samples out to 50 cycles **(Figure S13).** The presence of non-specific reads affects the quantification of the library, loading concentration, and cluster density. Therefore, we suggest quantifying the final library based on the proportion of the desired peaks.

439

440 Sequencing Protocol

441 Libraries were sequenced on either an Illumina MiSeq (2012), NextSeq550, or MiniSeq. Prior to
442 each MiSeq run, a bleach wash was performed using a sodium hypochlorite solution (Sigma
443 Aldrich, 239305) according to Illumina protocols. We also perform a maintenance wash between
444 each run. On the NextSeq550 and MiniSeq, the post-run wash was performed automatically by
445 the instruments, and no human intervention was required. The pooled and quantitated library
446 was diluted to a concentration of 6 nM (based on Qubit 4 Fluorometer and Illumina's formula for
447 conversion between ng/µl and nM) and was loaded on the sequencer at either 25 pM (MiSeq),
448 1.35 pM (NextSeq), or 1.6 pM (MiniSeq). PhiX Control v3 (Illumina, FC-110-3001) was spiked
449 into the library at an estimated 30-40% of the library. PhiX provides additional sequence
450 diversity to Read 1, which assists with template registration and improves run and base quality.
451 For this application, the MiSeq and MiniSeq (Rapid Run Kit) require 2 custom
452 sequencing primer mixes, the Read1 primer mix and the i7 primer mix. Both mixes have a final
453 concentration of 20 µM of primers (10 µM of each amplicon's sequencing primer). The NextSeq
454 requires an additional sequencing primer mix, the i5 primer mix, which also has a final
455 concentration of 20 µM. The MiSeq Reagent Kit v3 (150-cycle; MS-102-3001) is loaded with 30
456 µL of Read1 sequencing primer mix into reservoir 12 and 30 µL of the i7 sequencing primer
457 primer mix into reservoir 13. The Miniseq Rapid High Output Reagent Kit (100-cycle; 20044338)
458 is loaded with 17 µl of Read1 sequencing primer mix into reservoir 24 and 26 µl of i7
459 sequencing primer mix into reservoir 28. The NextSeq 500/550 Mid Output Kit (150 cycles;
460 20024904) is loaded with 52 µl of Read1 sequencing primer mix into reservoir 20, 85 µl of i7
461 sequencing primer mix into reservoir 22, and 85 µl of i5 sequencing primer mix into reservoir 22.
462 Index 1 and 2 are each 10 bp, and Read 1 is 26 bp.

463

464

465

466 Analysis

467 The bioinformatic analysis consists of standard conversion of BCL files into FASTQ sequencing
468 files using Illumina's bcl2fastq software (v2.20.0.422). Demultiplexing and read counts per
469 sample are performed using our custom software. Here read1 is matched to one of the three
470 expected amplicons allowing for the possibility of a single nucleotide error in the amplicon
471 sequence. The *hamming distance* is the number of positions at which the corresponding
472 sequences are different from each other and is a commonly used measure of distance between

sequences. Samples are demultiplexed using the two index reads in order to identify which sample the read originated from. Observed index reads are matched to the expected index sequences allowing for the possibility of a single nucleotide error in one or both of the index sequences. The set of three reads are discarded if both index1 and index2 have hamming distances greater than 1 from the expected index sequences. The count of reads for each amplicon and each sample is calculated. In this analysis we make use of a few custom scripts written in R that rely on the ShortRead [30] and stringdist [31] packages for processing fastq files and calculating hamming distances between observed and expected amplicons and indices. This approach was conservative and gave us very low level control of the sequencing analysis. However, we anticipate that continued development of the kallisto and bustools SwabSeq analysis tools [19] will be a more user-friendly and computationally efficient solution for SwabSeq.

Criteria for Classification of Purified Patient Samples

For our analytic pipeline, we developed QC metrics for each type of specimen. For purified RNA, we require each sample to have at least 10 reads detected for RPP30 and that the sum of S2 and *in vitro* S2 standard reads exceeds 2,000 reads. If these conditions are not met, the sample is rerun one time and if there is a second fail we request a resample. To determine if SARS-CoV-2 is present, we calculate if the ratio of S2 to *in vitro* S2 standard exceeds 0.003. (We note that we add 1 count to both S2 and *in vitro* S2 standard before calculating this ratio to facilitate plotting the results on a logarithmic scale.) If the ratio is greater than 0.003 we concluded that SARS-CoV-2 is detected for that sample and if it is less than or equal to 0.003 we conclude that SARS-CoV-2 is not detected **(Figure S5C).**

The same pair of primers will amplify both the S2 and *in vitro* S2 standard amplicons. Because we run an endpoint assay, the primers will be the limiting reagent to continued amplification. In developing this assay, we observed that as S2 counts increase for a sample, the *in vitro* S2 standard counts decrease **(Figure S3)**. We found that at very high viral levels, *in vitro* S2 standard read counts decreased to less than 1000 reads. Therefore, analysis of S2 and *in vitro* S2 standard together allowed our QC to call SARS-CoV-2 even at extremely high viral levels.

Since the S2 and *in vitro* S2 standard are derived from the same primer pair, to account for the scenario where *in vitro* S2 standard counts are low because S2 amplicon counts are very high and the sample contains large amounts of SARS-CoV-2 RNA **(Figure S3)** in the QC we require that the sum of S2 and *in vitro* S2 standard counts together exceeds 2000. For example,

if we detected greater than 2000 S2 counts and 0 *in vitro* S2 standard counts this would certainly be a SARS-CoV-2 positive sample and we would result: SARS-CoV-2 detected.

Criteria for Classification of Extraction Free Patient Samples

We require that the sum of S2 and S2 synthetic spike-in reads exceeds 500 reads, or the results are considered inconclusive. With inhibitory lysates we have observed that this lower total is acceptable for maintaining sensitivity and specificity while limiting the number of tests that are considered inconclusive. If the sum of S2 and S2 synthetic spike-in reads exceeds 500, we determine if SARS-CoV-2 is detected in a sample by seeing if the ratio of S2 to S2 spike exceeds 0.003. (We note that we add 1 count to both S2 and S2 spike before calculating this ratio to facilitate plotting the results on a logarithmic scale.) If the ratio is greater than 0.003 we concluded that SARS-CoV-2 is detected for that sample and if it is less than or equal to 0.003 we conclude that SARS-CoV-2 is not detected as long as 10 reads are detected for RPP30 for that sample. This serves as a control that sample collection took place properly and contains a human specimen. If fewer than 10 reads are detected for RPP30 and the ratio of S2 to S2 spike is less than or equal to 0.003 the results are considered inconclusive. We have modified this criteria such that only samples without SARS-CoV-2 signal are considered inconclusive if RPP30 reads are fewer than 10. This ensures that we do not miss SARS-CoV-2 positive samples that may have fewer RPP30 reads **(Figure S7)**.

Downsampling analysis

Reads were downsampled from the results for the NP purified confirmatory LoD shown in **Figure S5B**. We observed that downsampling down to 5,000 reads per well resulted in no instances of mis-classification of SARS-CoV-2 presence or absence. At 5,000 reads per well approximately 3% of wells would no longer pass the filter that the sum of S2 and S2 spike reads exceeded 1,000 reads and would result in a sample being classified as 'Inconclusive'. A logistic regression classifier described elsewhere[19] should robustly tolerate a small fraction of outlier samples with slightly lower read depth.

Analysis of index mis-assignment

Unique dual indices and amplicon specific indices were used to study index mis-assignment. In this scheme, each sample was assigned two unique indices for the S2 or Spike amplicon and two unique indices for the RPP30 amplicon for a total of four unique indices per sample. A count

matrix with all possible pairwise combinations, i.e. a “matching matrix”, was generated for each index pair (one i7 and one i5) using kallisto and bustools [32]. The counts on the diagonal of the matching matrix correspond to input samples and counts off of the diagonal correspond to index swapping events. The extent of index mis-assignment for the i7 and i5 index was determined by computing the row and column sums, respectively, of the off-diagonal elements of the matching matrix. The observed rate of index swapping to wells with a known zero amount of viral RNA was determined by computing the mean of the viral S counts to spike ratio for those wells.

Supplemental Results

Improving Limit of Detection Requires Minimizing Sources of Noise

One of the major challenges in running a highly sensitive molecular diagnostic assay is that even a single contaminant or source of noise can decrease the test's analytical sensitivity. In the process of developing SwabSeq, we observed S2 reads from control samples in which no SARS-CoV-2 RNA was present (**Figure 1D**). We subsequently refer to these reads as "no template control" (NTC) reads. A key part of SwabSeq optimization has been understanding and minimizing the sources of NTC reads in order to improve the limit of detection (LoD) of the assay. We identified two important sources of NTC reads: molecular contamination and mis-assignment of sequencing reads.

To minimize molecular contamination, we followed protocols and procedures that are commonly used in molecular genetic diagnostic laboratories[33]. To limit molecular contamination, we use a dedicated hood for making dilutions of the synthetic RNA controls and master mix. At the start of each new run, we sterilize the pipettes, dilution solutions, and PCR plates with 10% bleach, followed by UV-light treatment for 15 minutes.

To prevent post-PCR products that are at high concentration from contaminating our pre-PCR processes, we physically separated pre- and post-PCR steps of our protocol into two separate rooms, where any post-PCR plates were never opened within the pre-PCR laboratory space. To further protect from post-PCR contamination, we compared RT-PCR mastermixes with or without Uracil-N-glycosylase (UNG). The presence of UNG in the TaqPath™ 1-Step RT-qPCR Master Mix (Thermofisher Scientific) showed a significant improvement reducing post-PCR contamination of S2 reads present in the negative patient samples as compared with the Luna One Step RT-PCR Mix (New England Biosciences) (**Figure S14**). The RT-PCR mastermix contains a mix of dTTP and dUTP such that post-PCR amplicons are uracil containing DNA. These post-PCR that are remnants of previously run SwabSeq experiments therefore can be

572 selectively eliminated by UNG. Importantly, this addition does not interfere with downstream
573 sequencing.

574 A third source of molecular contamination was carryover contamination on the
575 sequencer template line of the Illumina MiSeq[34]. Without a bleach maintenance wash, we found
576 that indices from the previous sequencing run were identified in a subsequent experiment where
577 those indices were not included. While the number of reads for some indices were present at a
578 number of S2 reads, the presence of carryover contamination affects the sensitivity and
579 specificity of our assay. After an extra maintenance and bleach wash, we substantially reduced
580 the amount of carryover reads present to less than 10 reads **(Figure S15)**.

581 Another source of NTC reads is mis-assignment of amplicons. Mis-assignment of
582 amplicons occurs when sequencing (and perhaps at a lower rate, oligo synthesis) errors result
583 in an amplicon sequence that originates from the *in vitro* S2 standard but is mistakenly assigned
584 to the S2 sequence within a given sample. Only 6 bp distinguishes S2 from *in vitro* S2 standard
585 at the beginning of read 1. Sequencing errors can result in *in vitro* S2 standard reads being
586 misclassified as S2 reads as error rates appear to be higher in the beginning of the read (**Figure**
587 **S16A**). If computational error correction of the amplicon reads is too tolerant, these reads may
588 be inadvertently counted to the wrong category. To reduce this source of S2 read
589 misassignment, we use a more conservative thresholding on edit distance (**Figure S16B**).
590 Future redesigns or extensions to additional viral amplicons should consider engineering longer
591 regions of sequence diversity here.

592 An additional source of NTC reads is when S2 amplicon reads are mis-assigned to the
593 wrong sample based on the indexing strategy. In our assay, individual samples are identified by
594 pairs of index reads (**Figure 1B**). Mis-assignment of samples to the wrong index could occur if
595 there is contamination of index primer sequences, synthesis errors in the index sequence,
596 sequencing errors in the index sequences, or “index hopping” [35].

We leveraged multiple indexing strategies in our development of SwabSeq, from fully combinatorial indexing (where each possible combination of i5 and i7 indices was used to tag samples in the assay) to unique-dual indexing (UDI) where each sample has distinct and unrelated i7 and i5 indices (**Figure S14**). However, the ability to scale can be limited due to the substantial upfront cost of developing that many unique primers. Fully combinatorial indexing approaches significantly expand the number of unique primer combinations. We have also explored a compromise strategy between fully combinatorial indexing and UDI where sets of indices are only shared between small subsets of samples. Such designs reduce the effect of sample mis-assignment while facilitating scaling to tens of thousands of patient samples (**Figure S17**). With a fully combinatorial indexing (**Figure S17**) we observed that NTC read depth was correlated with the total number of S2 reads summed across all samples that shared the same i7 sequence (**Figure S18A**). This is consistent with the effect of index hopping from samples with high S2 viral reads to samples that share the same indices. It is possible to computationally correct for this effect, for example using a linear mixed model (**Figures S18B**).

Finally, the challenges associated with combinatorial and semi-combinatorial indexing strategies can be mitigated by using unique dual indexing (UDI), a known strategy to reduce the number of index-hopped reads by two orders of magnitude[36]. We have observed consistently lower S2 viral reads for negative control samples with this strategy. It also enables quantification of index mis-assignment by counting reads for index combinations that should not occur in our assay (**Figure S19 A and B**). The number of index hopping events is correlated with the total number of S2 + S2 spike reads (**Figure S19 C and D**), indicating that hopped reads are more likely to come from wells where the expected index has strong viral signal. We quantify the overall rate of hopping as 1-2% on a MiSeq, and expect this rate may be higher on patterned flow cell instruments.

There are many sources of noise in amplicon-based sequencing, from environmental contamination in the RT-PCR and sequencing steps to misassignment of reads based on

623 computational correction and “index-hopping” on the Illumina flow cells. Preventing and
624 correcting these sources of error considerably improves the limit of detection of the SwabSeq
625 assay.

626
627
628
629 **Supplementary Documents**
630 OptimizedProtocol.docx
631 EquipmentList.xlsx
632 SampleExperimentalDesignSetup.xlsx
633 SampleTrackingWorkbook.xlsx
634
635 **Supplementary Figures**
636 SupplementaryFigures.pdf
637
638 **Supplementary Videos**
639 1_PrePCR.mov
640 2_PostPCR.mov
641

REFERENCES

1. Furukawa, N. W., Brooks, J. T. & Sobel, J. Evidence Supporting Transmission of Severe Acute Respiratory Syndrome Coronavirus 2 While Presymptomatic or Asymptomatic. *Emerg. Infect. Dis.* **26**, (2020).
2. Lavezzo, E. *et al.* Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* (2020) doi:10.1038/s41586-020-2488-1.
3. Peiris, J. S. M., Yuen, K. Y., Osterhaus, A. D. M. E. & Stöhr, K. The severe acute respiratory syndrome. *N. Engl. J. Med.* **349**, 2431–2441 (2003).
4. Cheng, C., Wong, W.-M. & Tsang, K. W. Perception of benefits and costs during SARS outbreak: An 18-month prospective study. *J. Consult. Clin. Psychol.* **74**, 870–879 (2006).
5. Gandhi, M., Yokoe, D. S. & Havlir, D. V. Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19. *The New England journal of medicine* vol. 382 2158–2160 (2020).
6. Kimball, A. *et al.* Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 377–381 (2020).
7. Fernandez, M. & Mervosh, S. Texas Pauses Reopening as Virus Cases Soar Across the South and West. *The New York Times* (2020).
8. News Division. HHS Details Multiple COVID-19 Testing Statistics as National Test. https://www.hhs.gov/about/news/2020/07/31/hhs-details-multiple-covid-19-testing-statistics-as-national-test-volume-surges.html.
9. CMS Increases Medicare Payment for High-Production Coronavirus Lab Tests. https://www.cms.gov/newsroom/press-releases/cms-increases-medicare-payment-high-production-coronavirus-lab-tests-0.
10. Kliff, S. Most Coronavirus Tests Cost About 100. Why Did One Cost 2,315? *The New York*

Times (2020).

11. Pollitz, K. Free Coronavirus Testing for Privately Insured Patients? *Kasier Family Foundation* https://www.kff.org/coronavirus-policy-watch/free-coronavirus-testing-for-privately-insured-patients/ (2020).
12. Larremore, D. B. *et al.* Surveillance testing of SARS-CoV-2. *Infectious Diseases (except HIV/AIDS)* (2020) doi:10.1101/2020.06.22.20136309.
13. Paltiel, A. D., David Paltiel, A., Zheng, A. & Walensky, R. P. Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States. *JAMA Network Open* vol. 3 e2016818 (2020).
14. Eric M Jones, Aaron R Cooper, Joshua S Bloom, Nathan B. Lubock, Scott W. Simpkins, Molly Gasperini, Sriram Kosuri. Octant SwabSeq Testing. https://www.notion.so/Octant-SwabSeq-Testing-9eb80e793d7e46348038aa80a5a901fd (2020).
15. Jones, E. M. *et al.* A Scalable, Multiplexed Assay for Decoding GPCR-Ligand Interactions with RNA Sequencing. *Cell Syst* **8**, 254–260.e6 (2019).
16. Schmid-Burgk, J. L. *et al.* LAMP-Seq: Population-Scale COVID-19 Diagnostics Using a Compressed Barcode Space. *bioRxiv* 2020.04.06.025635 (2020) doi:10.1101/2020.04.06.025635.
17. COVIDSeq Test (RUO and PEO Versions). https://www.illumina.com/products/by-type/clinical-research-products/covidseq.html.
18. rapidmicrobiology Bioinnovation's DxSeq™ Sequences Filoviruses. https://www.rapidmicrobiology.com/news/bioinnovations-dxseq-sequences-filoviruses.
19. Booeshaghi, A. S. *et al.* Fast and accurate diagnostics from highly multiplexed sequencing assays. *Health Informatics* (2020) doi:10.1101/2020.05.13.20100131.
20. Relich, R. F. *PREPARATION OF VIRAL TRANSPORT MEDIUM*. https://www.cdc.gov/coronavirus/2019-ncov/downloads/Viral-Transport-Medium.pdf (2020).
21. CDC. Information for Laboratories about Coronavirus (COVID-19). *Centers for Disease*

Control and Prevention https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html (2020).

22. Ranoa, D. R. E. *et al.* Saliva-Based Molecular Testing for SARS-CoV-2 that Bypasses RNA Extraction. *bioRxiv* 2020.06.18.159434 (2020) doi:10.1101/2020.06.18.159434.
23. COVID-19 Testing at Broad. https://covid-19-test-info.broadinstitute.org/.
24. iSWAB Rack Format - Mawi DNA Technologies. *Mawi DNA Technologies* https://mawidna.com/our-products/iswab-rack-format/.
25. Covid, T. Multiplex Diagnostic Solution| Thermo Fisher Scientific-US. (19AD).
26. Clark, T. W. *et al.* Diagnostic accuracy of the FebriDx host response point-of-care test in patients hospitalised with suspected COVID-19. *J. Infect.* (2020) doi:10.1016/j.jinf.2020.06.051.
27. Yujian, L. & Bo, L. A normalized Levenshtein distance metric. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1091–1095 (2007).
28. Untergasser, A. *et al.* Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
29. Quail, M. A., Swerdlow, H. & Turner, D. J. Improved protocols for the illumina genome analyzer sequencing system. *Curr. Protoc. Hum. Genet.* **Chapter 18**, Unit 18.2 (2009).
30. Morgan, M. *et al.* ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* **25**, 2607–2608 (2009).
31. Van der Loo, M. P. J. The stringdist package for approximate string matching. *R J.* **6**, 111–122 (2014).
32. Melsted, P. *et al.* Modular and efficient pre-processing of single-cell RNA-seq. *bioRxiv* 673285 (2019) doi:10.1101/673285.
33. Furtado, L. V. *et al.* The 2013 AMP Clinical Practice Committee consisted of Matthew J. Bankowski, Milena Cankovic, Jennifer Dunlap.
34. Nelson, M. C., Morrison, H. G., Benjamino, J., Grim, S. L. & Graf, J. Analysis, optimization

and verification of Illumina-generated 16S rRNA gene amplicon surveys. *PLoS One* **9**, e94249 (2014).

35. Valk, T. van der *et al.* Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. *Molecular Ecology Resources* (2019) doi:10.1111/1755-0998.13009.

36. MacConaill, L. E. *et al.* Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics* **19**, 30 (2018).

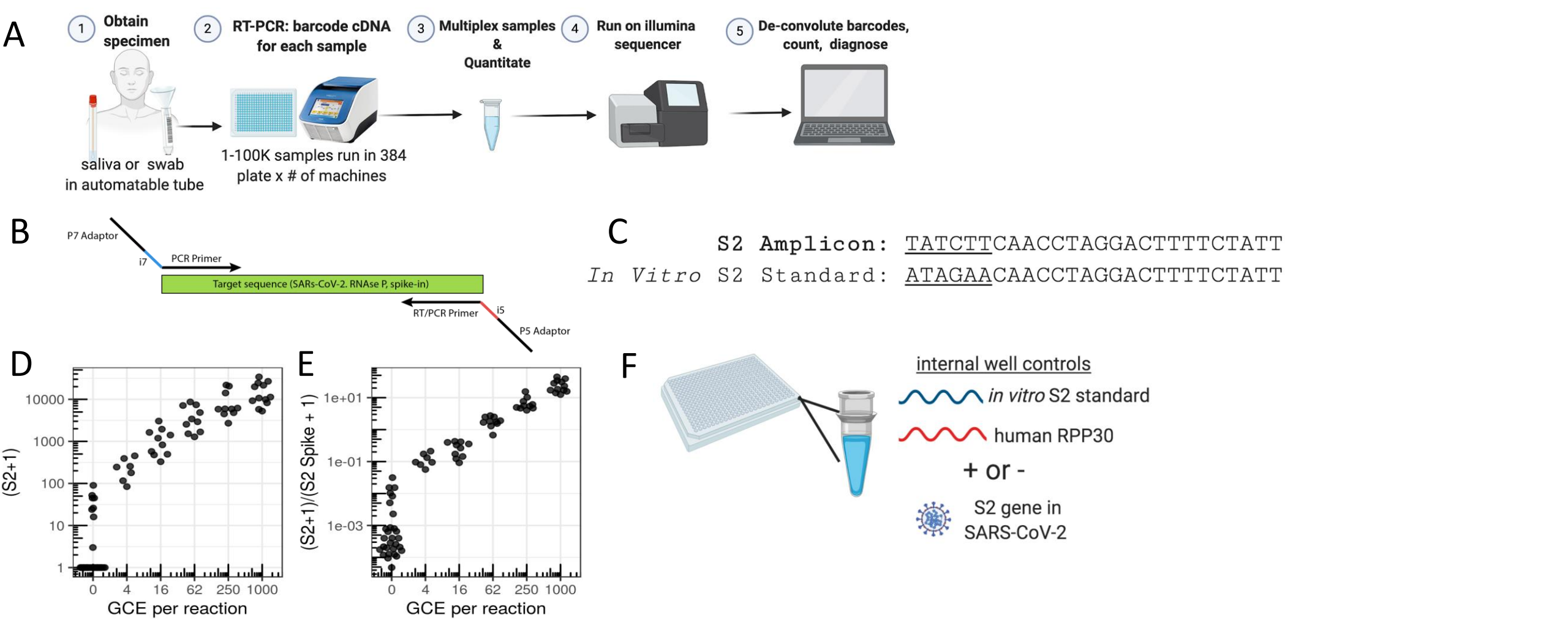


Figure 1. SwabSeq Diagnostic Testing Platform for COVID19. A) The workflow for SwabSeq is a five step process that takes approximately 12 hours from start to finish. B) In each well, we perform RT-PCR on clinical samples. Each well has two sets of indexed primers that generate cDNA and amplicons for SARS-CoV-2 S2 gene and the human RPP30 gene. Each primer is synthesized with the P5 and P7 adaptors for Illumina sequencing, a unique i7 and i5 molecular barcodes, and the unique primer pair. Importantly, every well has a synthetic *in vitro* S2 standard that is key to allowing the method to work at scale. C) The *in vitro* S2 standard (abbreviated as S2-Spike) differs from the virus S2 gene by 6 base pairs that are complemented (underlined). (D) Read count at various viral concentrations (E) Ratiometric normalization allow for in-well normalization for each amplicon (F) Every well has two internal well controls for amplification, the *in vitro* S2 standard and the human RPP30. The RPP30 amplicon serves as a control for specimen collection. The *in vitro* S2 standard is critical to SwabSeq's ability to distinguish true negatives.

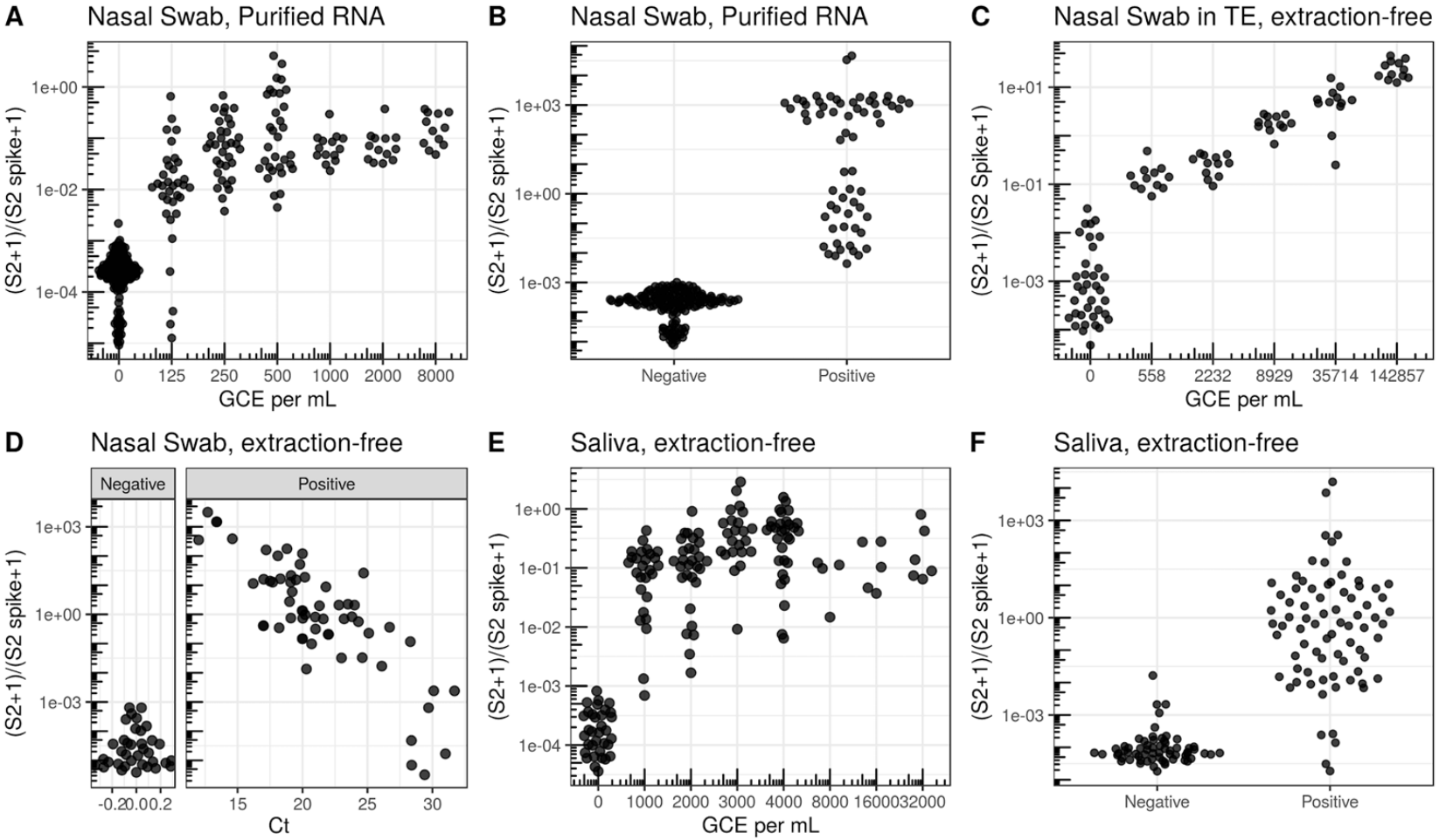


Figure 2. Validation in clinical specimens demonstrate a limit of detection equivalent to sensitive RT-qPCR reactions. A) Limit of Detection in nasal swab samples with no SARS-CoV2 were pooled and ATCC inactivated virus was added at different concentrations. Nasal Swab sample was RNA purified and using SwabSeq showed a limit of detection of 250 genome copy equivalents (GCE) per mL. B) RNA-purified clinical nasal swab specimens obtained through the UCLA Health Clinical Microbiology Laboratory were tested based on clinical protocols using FDA authorized platforms and then also tested using SwabSeq. We show 100% agreement with samples that tested positive for SARS-CoV-2 (n=63) and negative for SARS-CoV-2 (n=159). C) We also tested RNA purified samples from extraction-free nasopharyngeal swab and showed a limit of detection of 558 GCE/mL. D) Relationship between Ct from RT-qPCR targetting the S gene (x-axis) and SwabSeq ratio for extract-free swabs into normal saline or Tris-EDTA (y-axis) for patient samples classified as testing positive or negative for SARS-CoV-2 by the UCLA Clinical Microbiology Laboratory. Samples with no virus detected were assigned a Ct of 0 for this visualization. E) Extractio- free processing of saliva specimens show a limit or detection down to 1000 GCE per mL. F) Extraction-free processing of saliva clinical specimens using swabseq (y-axis) compared to classification of SARS-CoV-2 status from RNA-purified clinical nasal swab specimens for matched samples (x-axis).

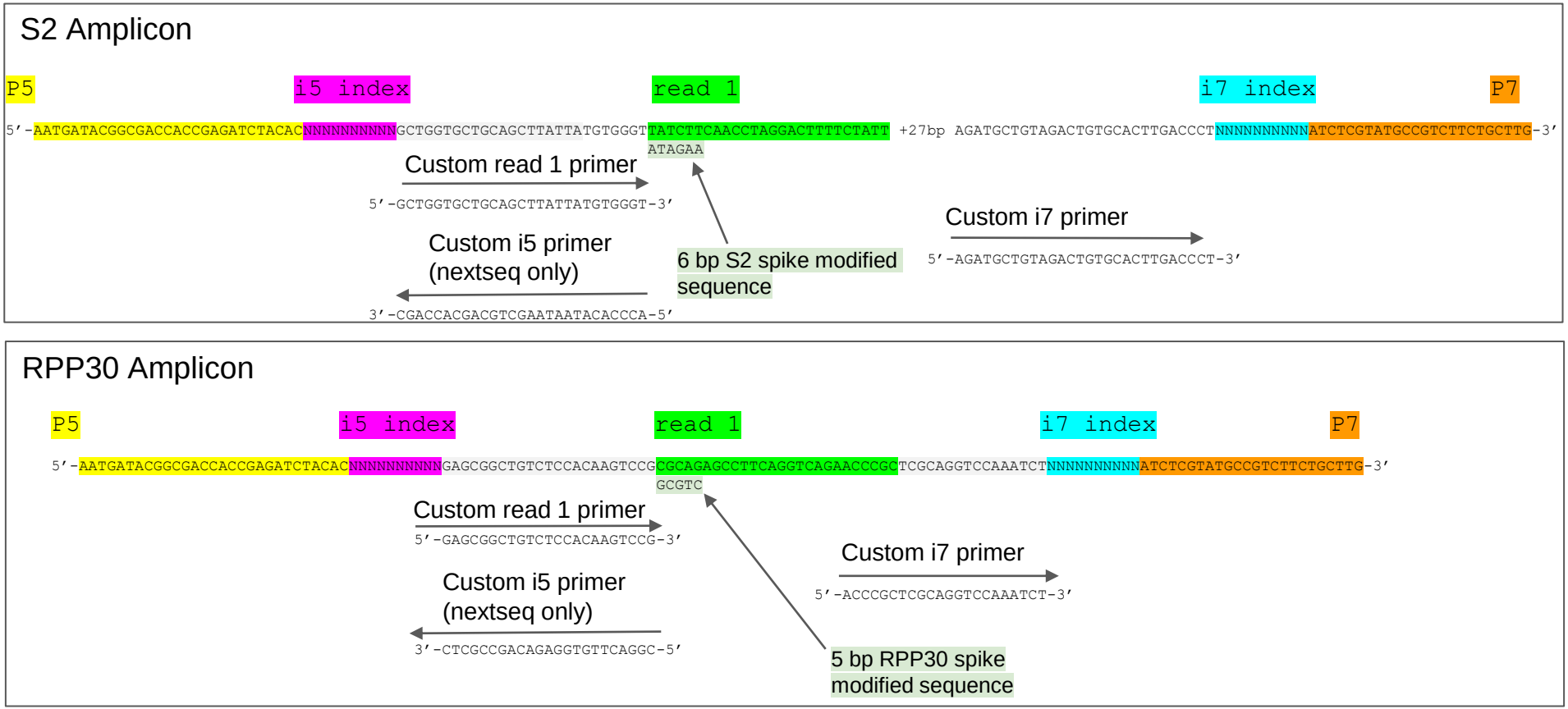


Figure S1. Sequencing library design. The amplicon designs are shown for the S2 (top) and RPP30 (bottom) amplicons. Amplicons were designed such that the i5 and i7 molecular indexes uniquely identify each sample. SwabSeq was designed to be compatible with all Illumina platforms.

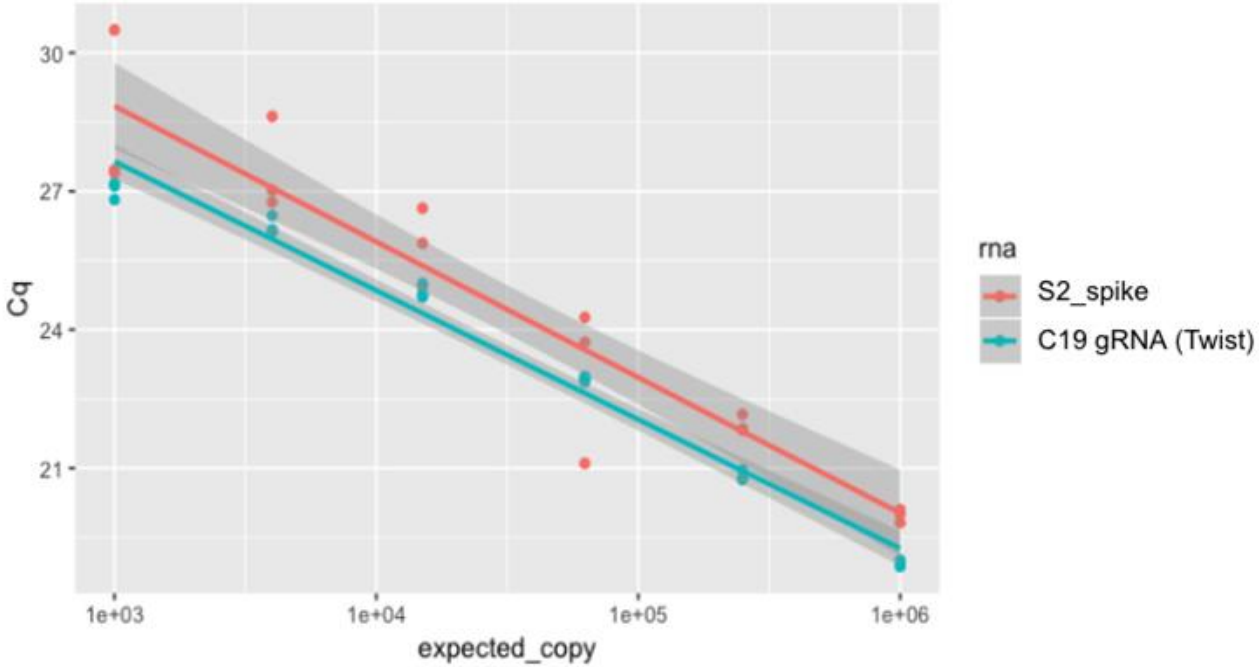


Figure S2. The S2 primers show equivalent PCR efficiency when amplifying the COVID-19 S2 gene amplicon and the synthetic *in vitro* S2 standard. Slope of PCR efficiency of the primers with either the S2_spike (labeled in red) or the SARS-CoV-2 viral (labeled in green as C19gRNA) input are as follows: S2_spike slope = -6.68e-6 and C19gRNA (Twist Control) slope = -6.74e-6. The slopes are expected to equivalent (parallel) if the primers do not show preferential amplification of the S2 spike RNA versus the C19gRNA. This shows that the S2 spike and C19gRNA have equivalent amplification efficiencies using the S2 primer pair. The bands represent 95% confidence intervals for predicted values, are non-overlapping due to different intercepts, and are not relevant for this analysis of slopes.

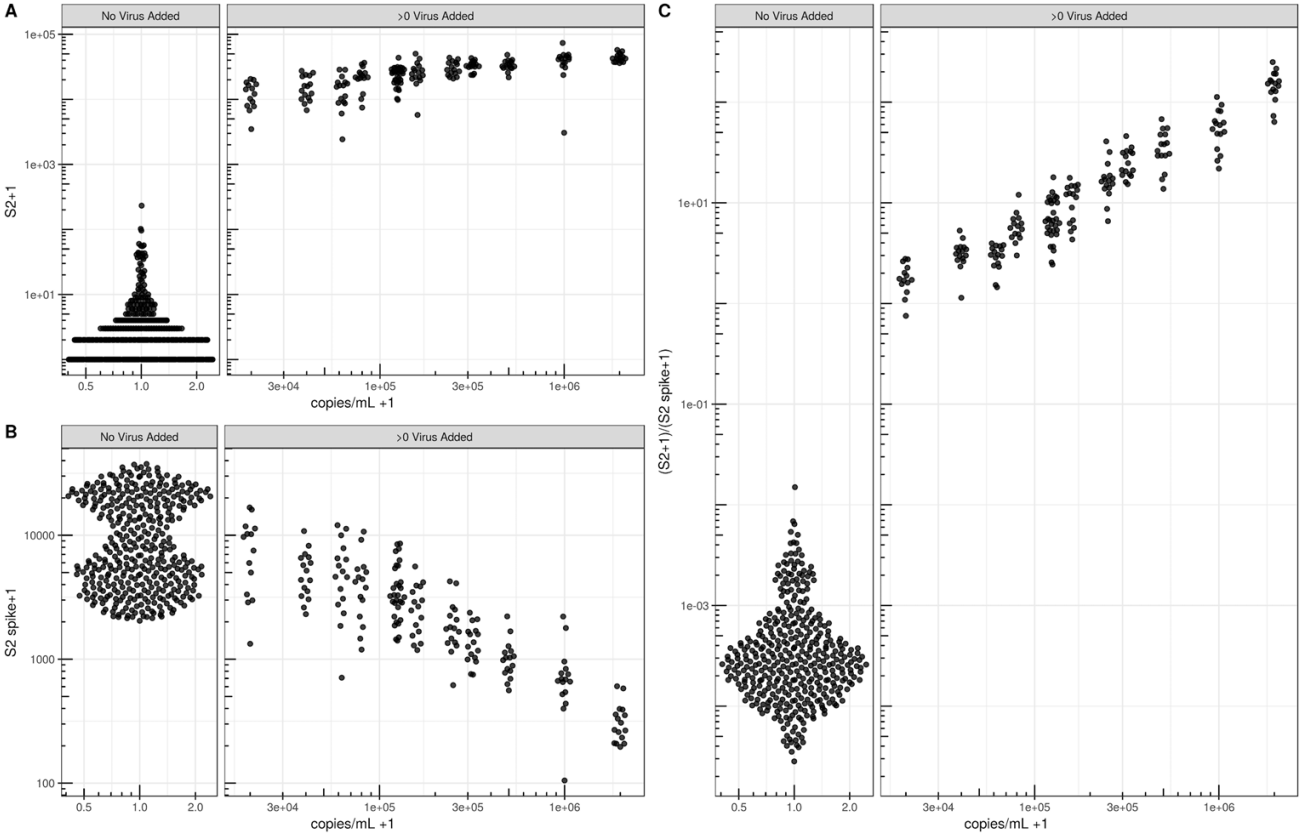


Figure S3. At very high viral concentrations SwabSeq maintains linearity. We include an internal well control, the S2 Spike, to enable us to call negative samples, even in the presence of heterogeneous sample types and PCR inhibition. (A) As virus concentration increases, we observe increased reads attributed to S2 and (B) decreased reads attributed to the S2 Spike. (C) The ratio between the S2 and S2 Spike provides an additional level of ratiometric normalization and exhibits linearity up to at least 2 million copies/mL of lysate. Note that ticks on both axes are spaced on a log10 scale.

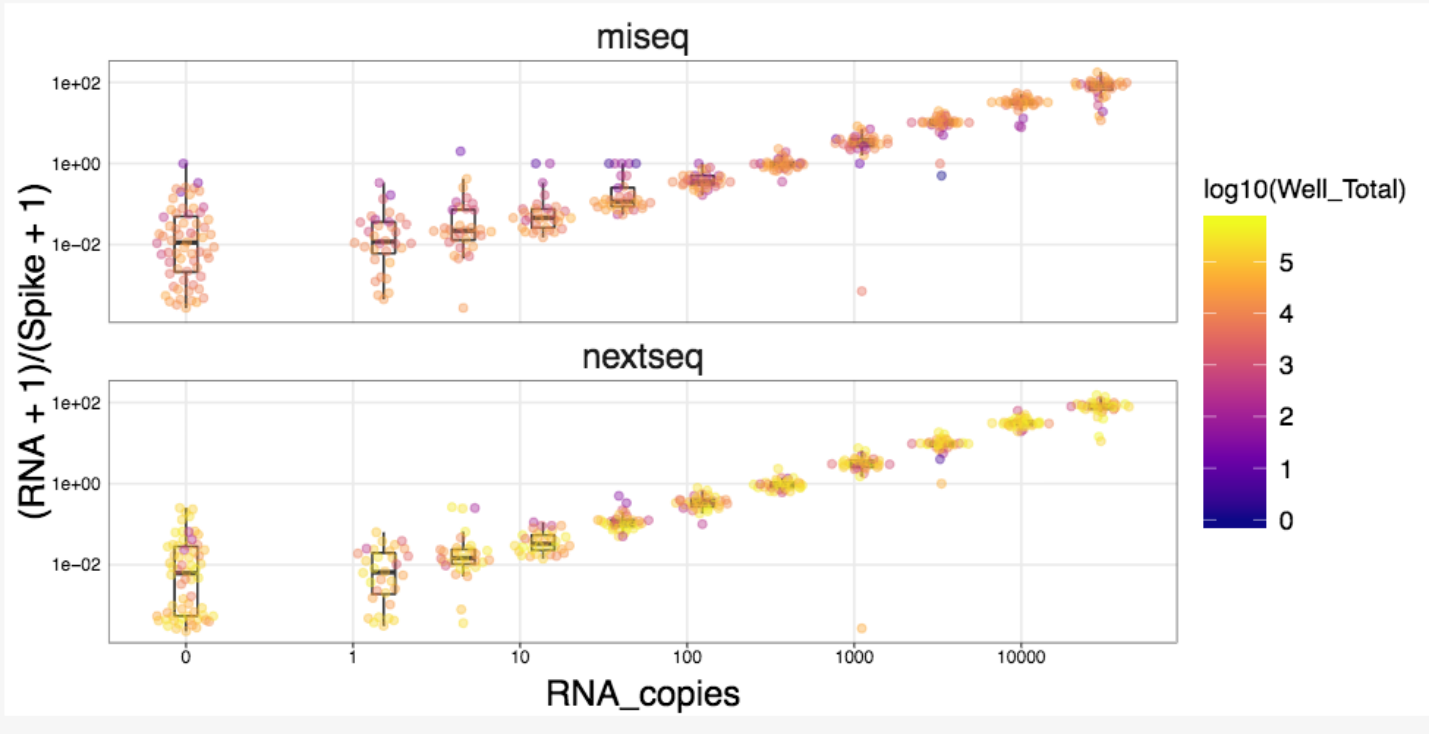
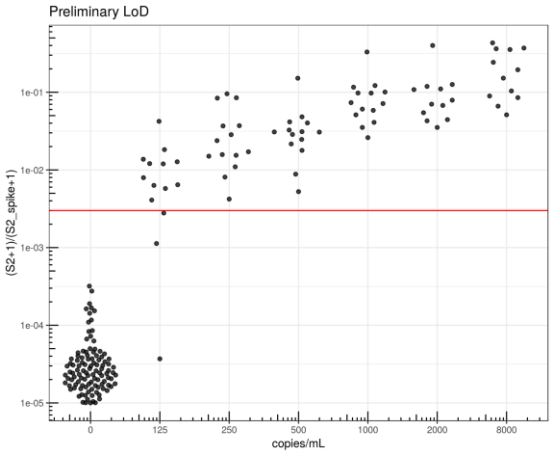
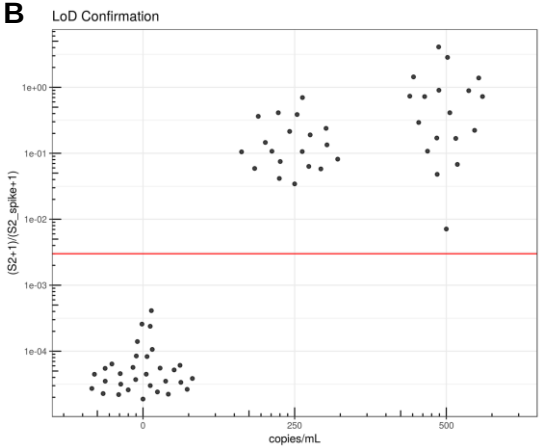


Figure S4. Sequencing is performed on MiSeq or NextSeq Machine with similar sensitivity. Multiplexed libraries run on both MiSeq and NextSeq showed linearity across a wide range of SARS-CoV2 virus copies in a purified RNA background.

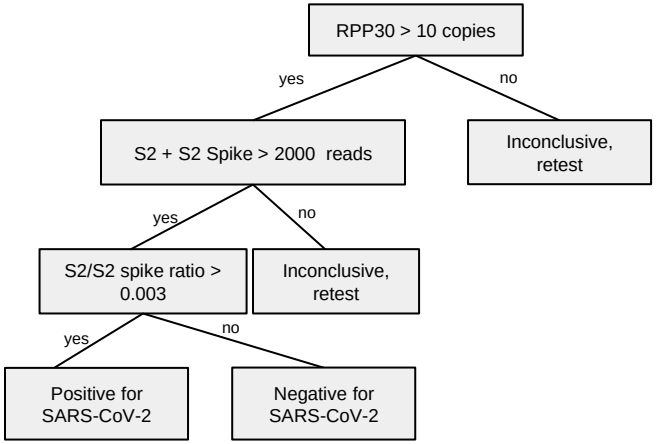
A



B



C



D

		UCLA RT-qPCR		
		Positive	Negative	
SwabSeq	Positive	94	0	94
	Negative	0	282	282
	Inconclusive	0	4	4
	Total	94	286	380

Positive Percent Agreement	1
Negative Percent Agreement	98.6%
Overall Agreement	98.9%

Figure S5. Preliminary and Confirmatory Limit of Detection Data for RNA purified Samples A) Our preliminary LOD data identified a LOD of 250 copies/mL using the NextSeq550 B) Confirmatory studies showed an LOD of 250 copies/mL using the NextSeq550. C) Our result interpretation guidelines for purified RNA. D) Concordance of the 380 clinical samples that were run during our validation process. Concordance is 98.6%, with 100% positive percent agreement (PPA) and 100% negative percent agreement (NPA).

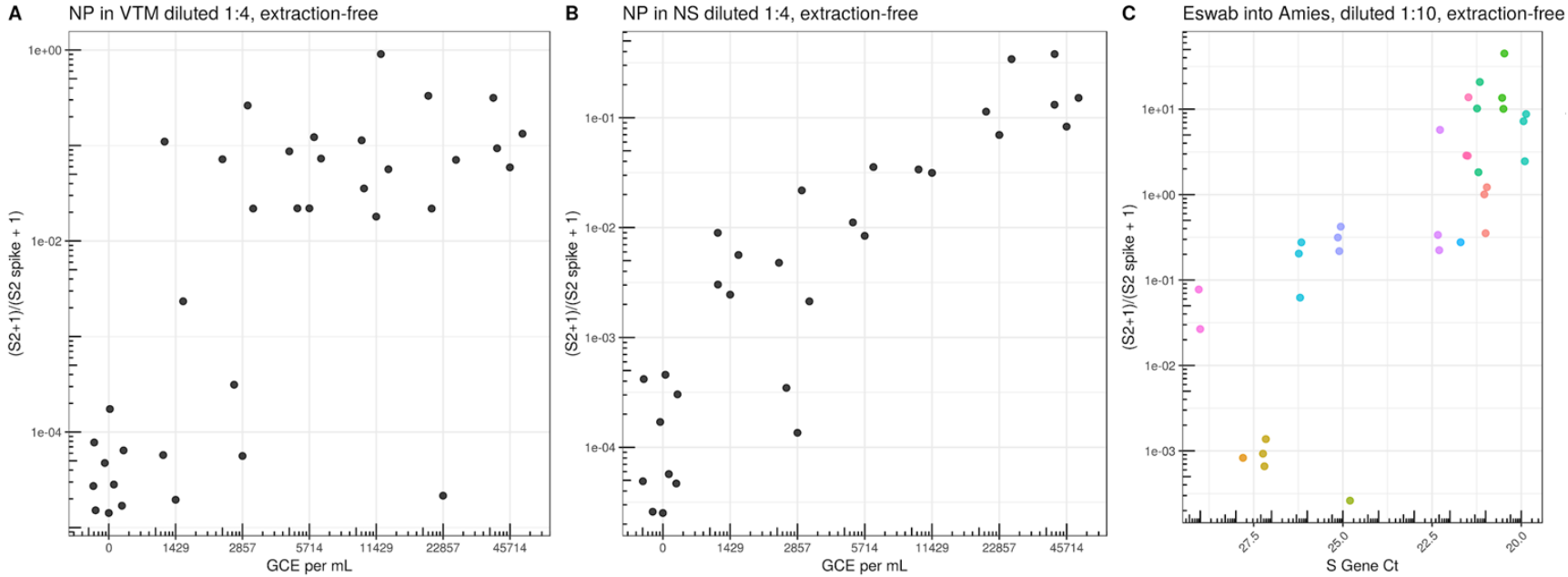


Figure S6. Extraction-Free protocols into traditional collection medias and buffers require dilution to overcome effects of RT and PCR inhibition. A) We tested extraction free protocols for nasopharyngeal swabs that were placed into viral transport media (VTM). We spiked ATCC live inactivated virus at varying concentrations into pooled VTM and then diluted samples 1:4 with water before adding to the RT-PCR reaction. We observed a limit of detection of 5714 copies per mL. B) We also tested nasopharyngeal swabs that were collected in normal saline, pooled and then spiked with ATCC live inactivated virus at varying concentrations. Contrived samples were diluted 1:4 in water. Here, our early studies show a similar similar limit of detection between 2857 and 5714 copies per mL. C) We tested natural clinical samples that were collected into Amies Buffer (ESwab). Here we compare S gene Ct count (x-axis) from positive samples to the SwabSeq S2 to S2 spike ratio (y-axis). Samples were run in triplicate (colors). We observed high concordance for Ct counts of 27 and lower but more variability for Ct counts greater than 27 suggesting that RT and PCR inhibition were affecting our limit of detection. Based on these data we opted only to further explore extraction free protocols into normal saline or tris-EDTA buffer.

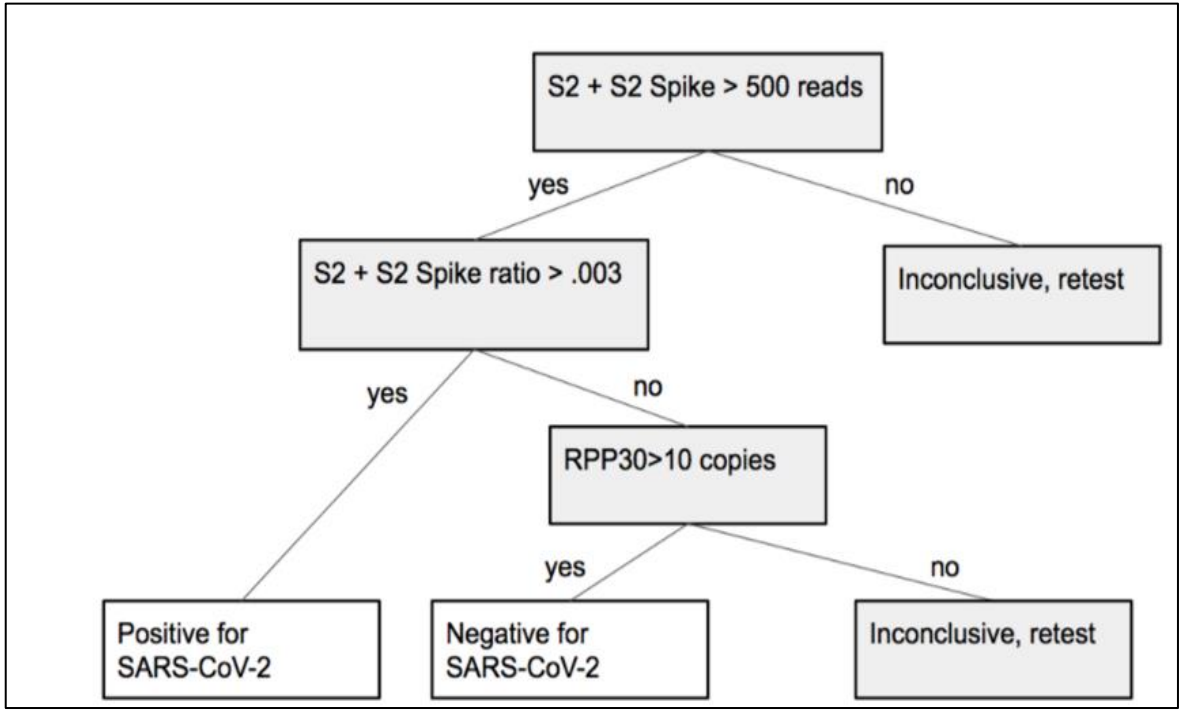
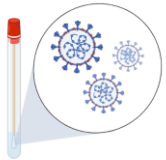
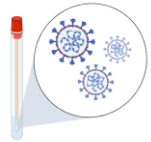


Figure S7. Per sample decision tree for extraction-free samples. Given the slight modifications for our extraction free protocols, we have modified our decision tree to reflect the differences in extraction-free sample types. Our result interpretation guidelines for extraction-free samples relax threshold for S2 + S2 spike to 500 reads due to the increased PCR inhibition observed in extraction free sample types.





		UCLA RT-qPCR		
		Positive	**Negative**	
SwabSeq	**Positive**	80	0	80
	Negative	6	206	212
	Inconclusive	1	5	6
		87	211	298

	Positive Percent Agreement	92.0%
	Negative Percent Agreement	97.6%
	Overall Agreement	96.0%

Figure S8. Comparison of extraction-free NP samples run on SwabSeq to NP Swab samples processed to Clinical pathway using RNA purification and RT-qPCR. Evaluation of extraction free nasal swabs processed into normal saline or Tris-EDTA ph 8.0 that have previously tested positive or negative in the UCLA Clinical Microbiology Laboratory. We have explored the sources of false negatives in our data set. Three of the four false negatives stem from differences in the limit of detection, where we do not always detect samples with Ct > 30.

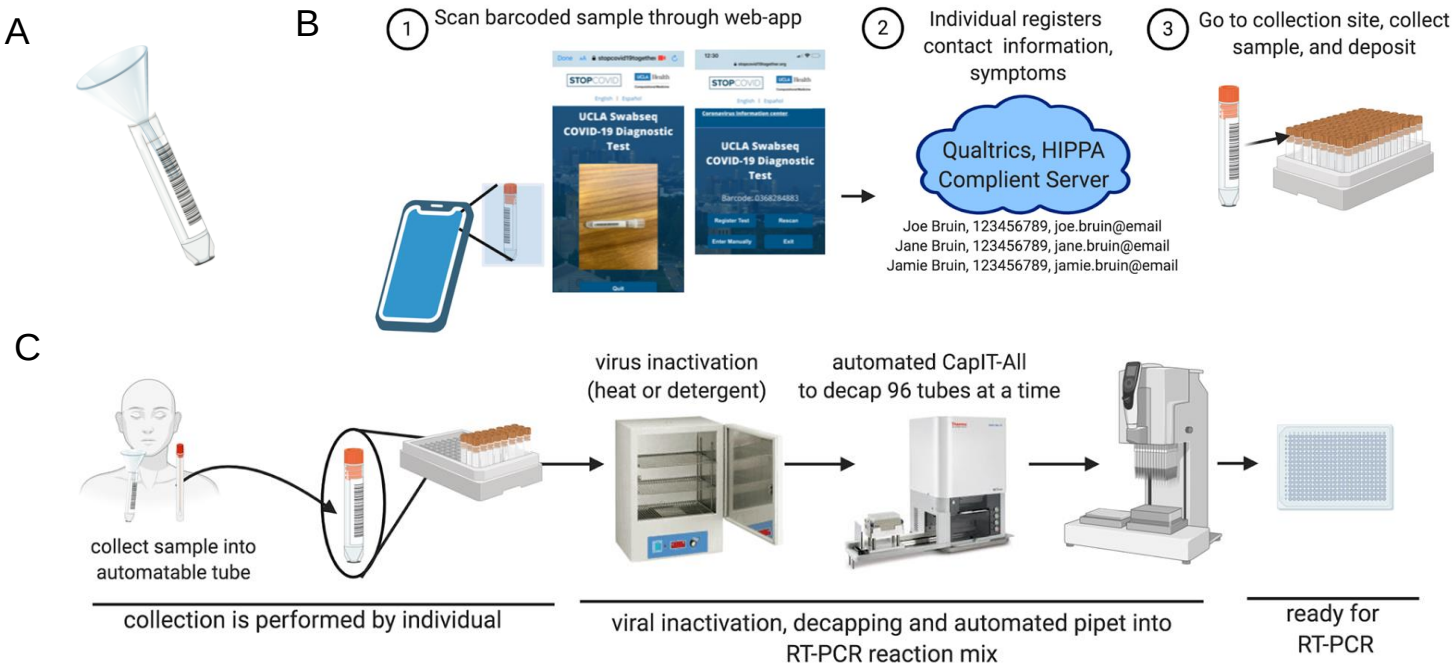


Figure S9. Developing a lightweight sample accessioning, collection and processing to allow for scalable testing into the thousands of samples per day. A) To address the challenge of sample collection, we have developed lightweight collection methods that collect sample directly into an automatable tube. Here a funnel is used for an individual to deposit a small sample of saliva (0.25 mL into the funnel and tube). This setup can accommodate multiple sample types. B) To facilitate the sample accessioning and collection, we developed a web-based app for individuals to register their sample tube using a barcode reader and send their identifying information into a secure instance of Qualtrics. Individuals then collect their sample and then place the tube in the rack. This low-touch pre-analytic process allows us to process thousands of samples a day without heavy administrative burden. C) The overall workflow streamlines processing in the lab. First, individuals collect samples into an automatable tube and place them into a 96-tube rack. Samples arrive in the lab in a 96-rack format allowing us to efficiently inactivate and process the samples, drastically increasing the flow of samples through our platform. Optimized throughput by this approach allows for a single person to rapidly process 6 384-well plates per hour (2,304 samples / hour / technician).

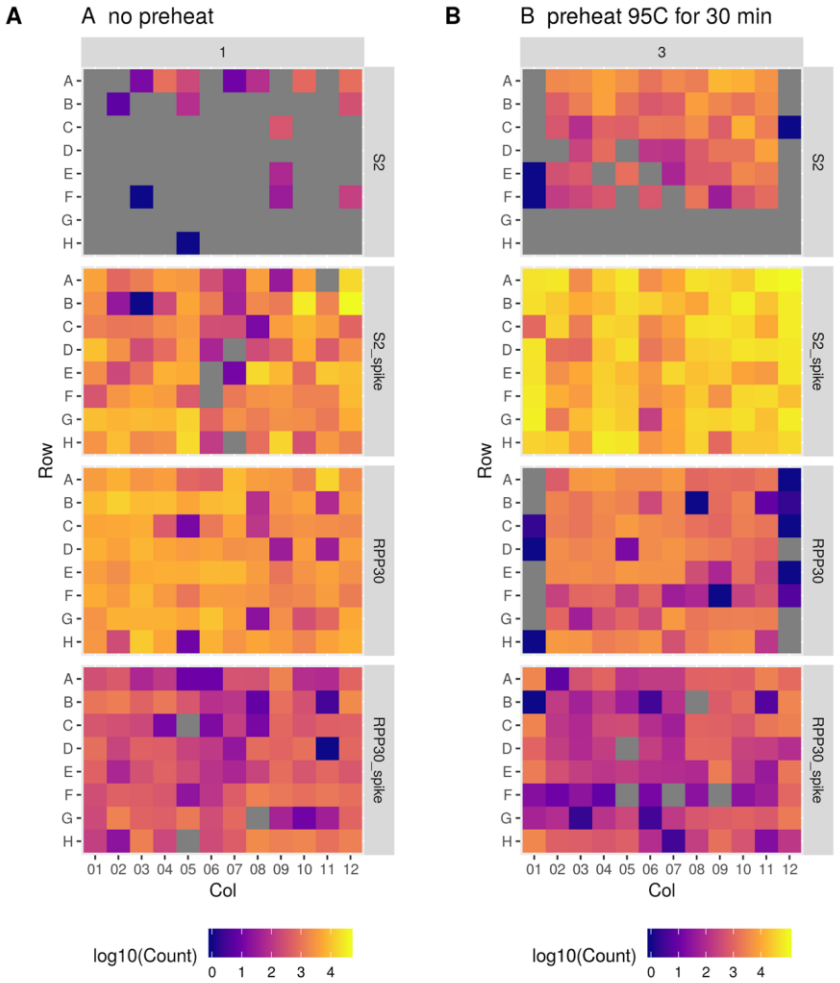
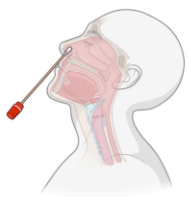
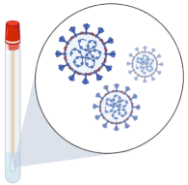
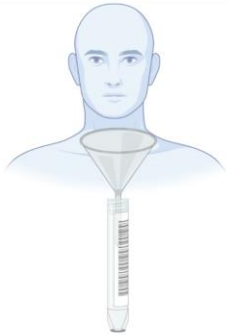


Figure S10. Preheating Saliva to 95C for 30 minutes drastically improves RT-PCR. Detection of viral genome and shows improved robustness in detection of our controls. A) Without preheating, detection of S2 spike is minimal and there are lower counts for the control amplicons. B) with a 95C preheating step for 30 minutes, we observe robust detection of the S2 amplicon and synthetic S2 Spike.







SwabSeq Saliva		UCLA RT-qPCR NP Swab		
		Positive	Negative	
	Positive	74	2	76
	Negative	6	438	444
	Inconclusive	2	15	17
		82	455	537

Positive Percent Agreement	90.2%
Negative Percent Agreement	96.3%
Overall Agreement	95.3%

Figure S11. Comparison of extraction-free saliva samples run on SwabSeq to NP Swab samples processed to Clinical pathway using RNA purification and RT-qPCR.

We performed a series of studies to compare the concordance of Saliva and NP swab performed within 2 hours of each other. These collections were obtained in the UCLA ED and UCLA Student Health Center over the course of several months.

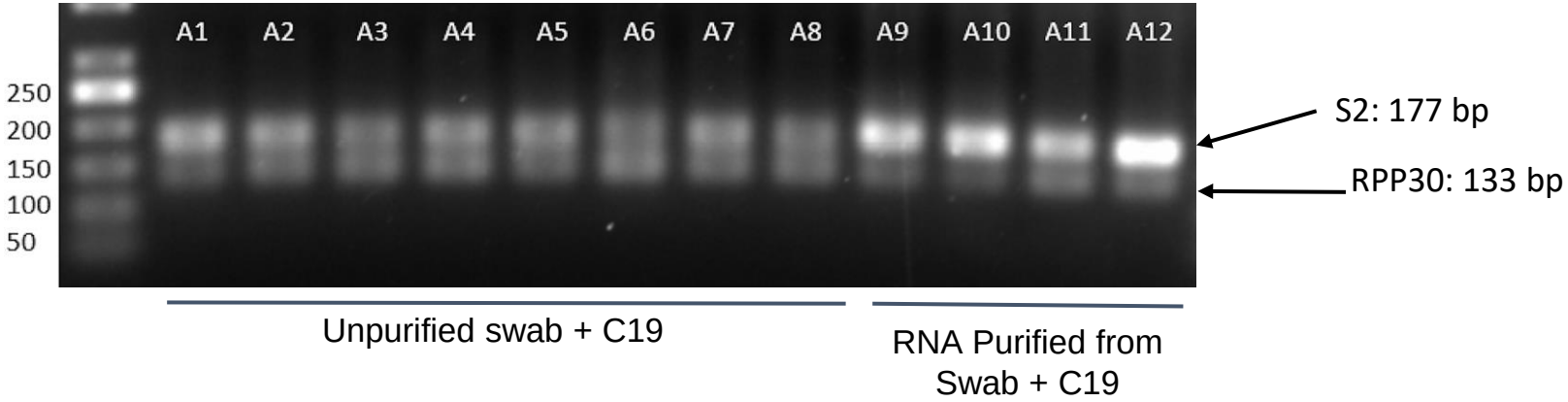


Figure S12. PCR inhibition has significant effect on amplification products. A) 2% Agarose gene was run for a subset of wells from our Rt-PCR reactions. We observe RT-PCR inhibition from swabs in unpurified lysate (A1-A8) as compared to purified RNA (A9-A12). We observe two bands in this subset of wells representing 2 amplicons for the S2 or S2 spike (177bp) and RPP30 (133 bp) primer pairs.

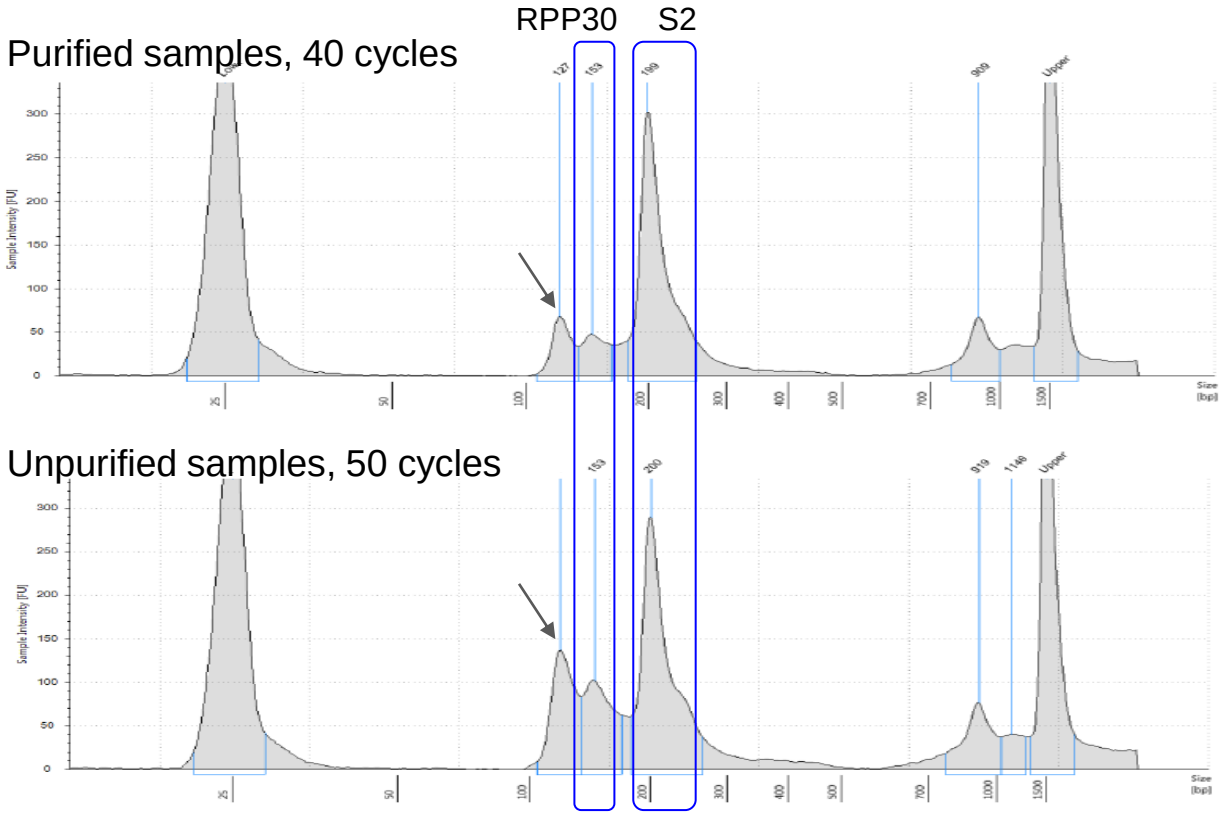
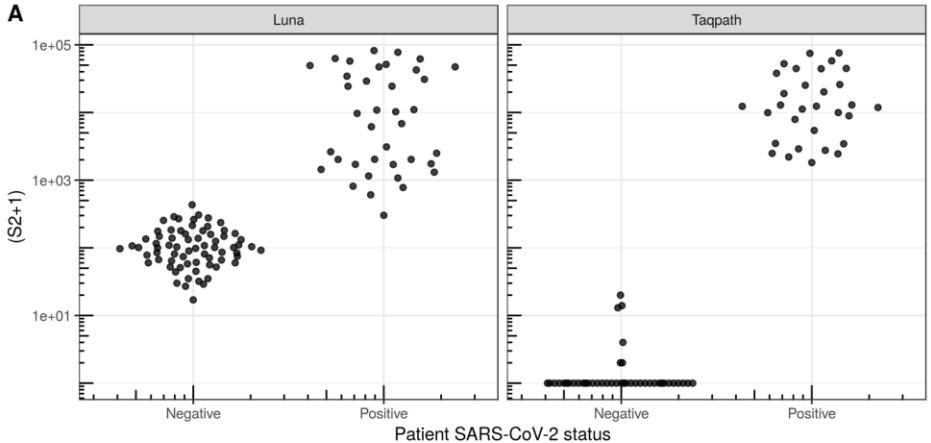


Figure S13. Tapestation Increasing the number of PCR cycles and working with unpurified or inhibitory samples types (eg. Saliva) was seen to increase the size of a nonspecific peak in our library preparation. Representative result from Agilent TapeStation for our purified amplicon libraries. We observe a nonspecific peak slightly above 100bp (arrow) in both library traces, but this peak increases in size with unpurified samples and an increased number of PCR cycles. While we have not confirmed the identity of this peak, we believe this peak may be the result of adapter dimers or unsequenceable PCR artifacts. Importantly, we observe that an increase in the size of this nonspecific peak leads to inaccurate library quantification. Therefore, in order to optimize cluster density on Illumina sequencers, we suggest quantifying the loading concentration of the final library based on the proportion of the desired peaks (RPP30 and S2).



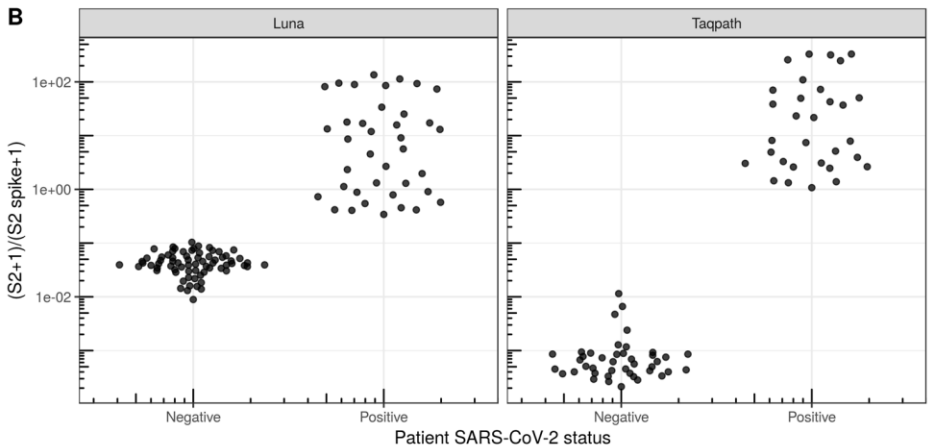


Figure S14. TaqPath decreases the number of S2 reads in SARS-CoV2-negative samples relative to NEB Luna. We compared Luna One Step RT-PCR Mix (New England Biosciences) to TaqPath™ 1-Step RT-qPCR Master Mix (Thermofisher Scientific). It is likely that the presence of UNG in the TaqPath Mastermix significantly reduced the number of S2 reads in the SARS-CoV-2-negative samples allowing us to more accurately distinguish SARS-CoV-2-positive and SARS-CoV-2-negative samples.

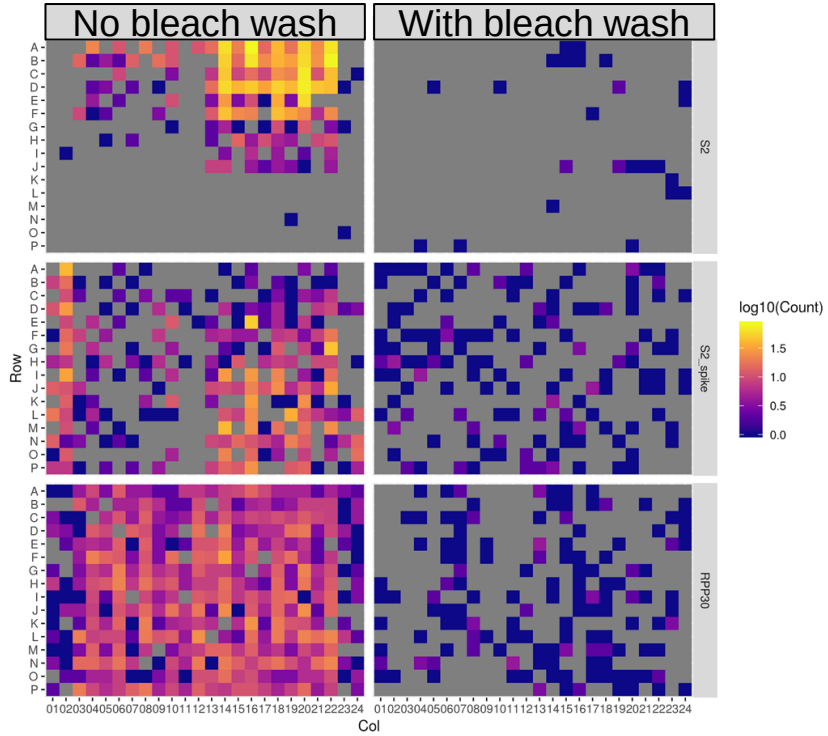
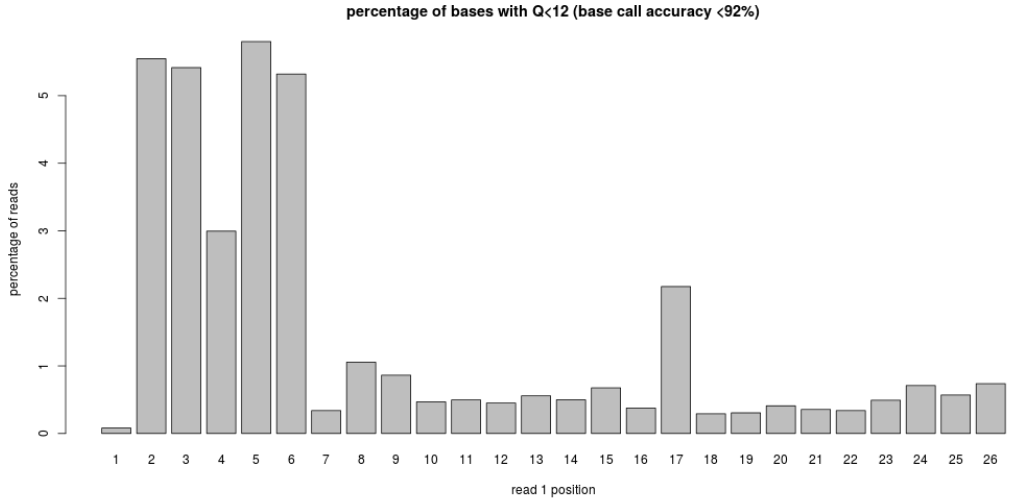


Figure S15. Carryover contamination from template line in a MiSeq contributes to cross contamination. In this experiment we did RT-PCR on four 384-well plates but only pooled three plates. On the left are observed counts of each of the amplicons for each sample for the 384-well plate not included in our run (but for which the indices were used in the previous run). Amplicon reads for indices used in the previous run are present at a low level (0-150 reads). We then performed a bleach wash in addition to regular wash prior to the subsequent run. In this subsequent run, we pooled three different plates and left out the fourth 384 well plate. On the right are observed counts of each of the amplicons for sample indices corresponding to the left-out plate (again, for which the indices were used in the previous run). We observe a remarkable decrease in the amount of carryover contamination, where carryover reads are <10 per sample.

A



B

Hamming distance from S2 Spike

Hamming distance from S2	0	1	2	3	4	5	6
0	0	0	0	0	0	0	981607
1	0	0	0	0	0	49010	235315
2	0	0	0	0	22578	19024	111529
3	0	0	0	25516	7089	15226	59012
4	0	0	74519	8036	7576	10021	36688
5	0	522655	21671	9469	3292	5824	13217
6	9062990	222210	48991	2270	1337	1669	2140
7	0	928836	19700	2439	277	253	348

Figure S16. Sequencing errors in amplicon read and potential amplicon mis-assignment. In experiment v18 we loaded less PhiX than usual (11%) and the overall quality of read1 was lower. Trends noticed here persist in other runs but this run more clearly highlights issues that can occur due to sequencing errors and overly tolerant error-correction. A) The percentage of reads with base quality scores less than 12 for each position in read 1. Note that the first 6 bases of read1 distinguish S2 from S2 spike and have the highest percentage of low quality base calls. B) The hamming distance between each read1 sequence and either the expected S2 sequence (rows) or S2 spike sequence (columns), In yellow are perfect match and edit distance 1 sequences that can be clearly identified as S2 or S2 spike. In red are sequences with errors that may be mis-assigned (S2 spike assigned as S2 is most problematic for this assay.)

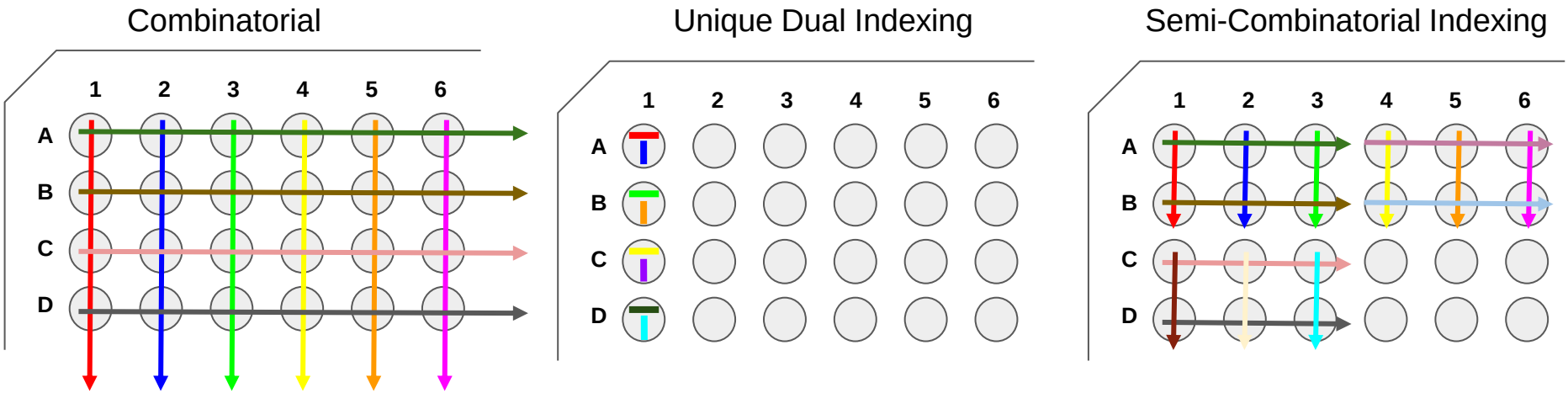


Figure S17. Visualization of different indexing strategies. Here i5 indices are depicted as horizontal lines, i7 indices are depicted as vertical lines, and colors represent unique indices. In combinatorial (or fully-combinatorial) indexing, the i5 and i7 indices are combined to make unique combinations, but each i5 and i7 index may be used multiple times within a plate, and all possible i5 and i7 . For unique dual indexing, each i5 and i7 index are only used 1 time per plate. This requires many more oligos to be synthesized. For Semi-combinatorial indexing, the combinations used are more limited, such that indices are only repeated for a subset of wells and many possible combinations are not used. In practice (not depicted here), we've used a design where the i7 index is unique but the i5 index can be repeated up to four times across a 384-well plate. For the majority of our Swabseq development, we used either semi-combinatorial indexing (384x96) that allowed for 1536 combinations or samples to be run or unique dual indexing (384 UDI)

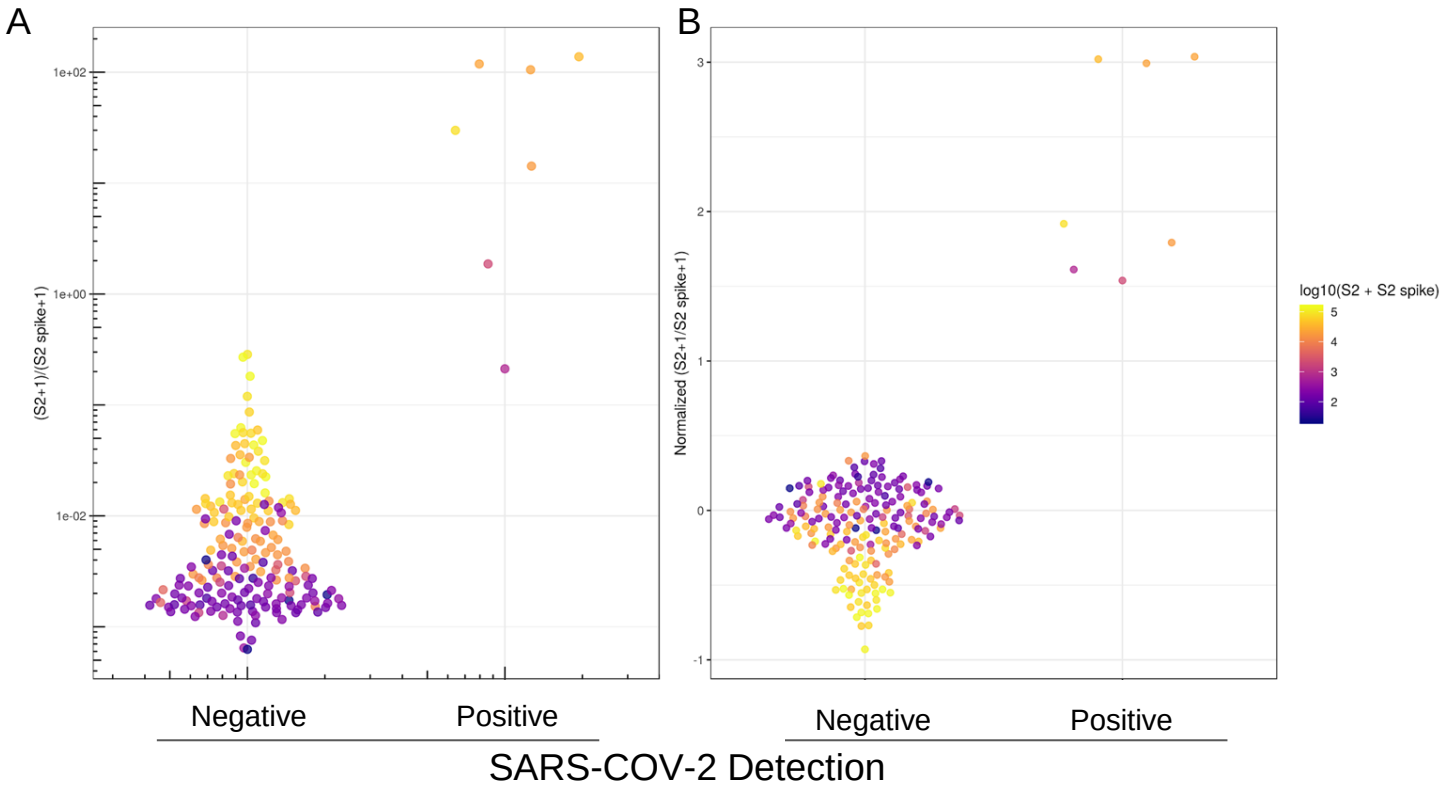


Figure S18. Computational correction for index mis-assignment using a mixed-model. To expand the number of samples we are capable of testing, we can use a combinatorial indexing strategy. In this experiment we used a single index on i5 to uniquely identify a plate and 96 i7 indices to identify wells. (A) The ratio of S2 to S2 spike (y-axis) is plotted for clinical samples based on whether Covid was detected by RT-qPCR (x-axis). SARS-CoV-2 positive samples were filtered to have Ct<32. The effects of index mis-assignment across plates can be observed as i7 indices that have high a sum of S2 and S2 spike across all samples that share the same i7 barcode across plates (colors). (B) Best linear unbiased predictor residuals are plotted (y-axis) for data in A, after computational correction of the log10(S2+1/S2_spike+1) ratio by treating the identity of the i7 barcode as a random effect.

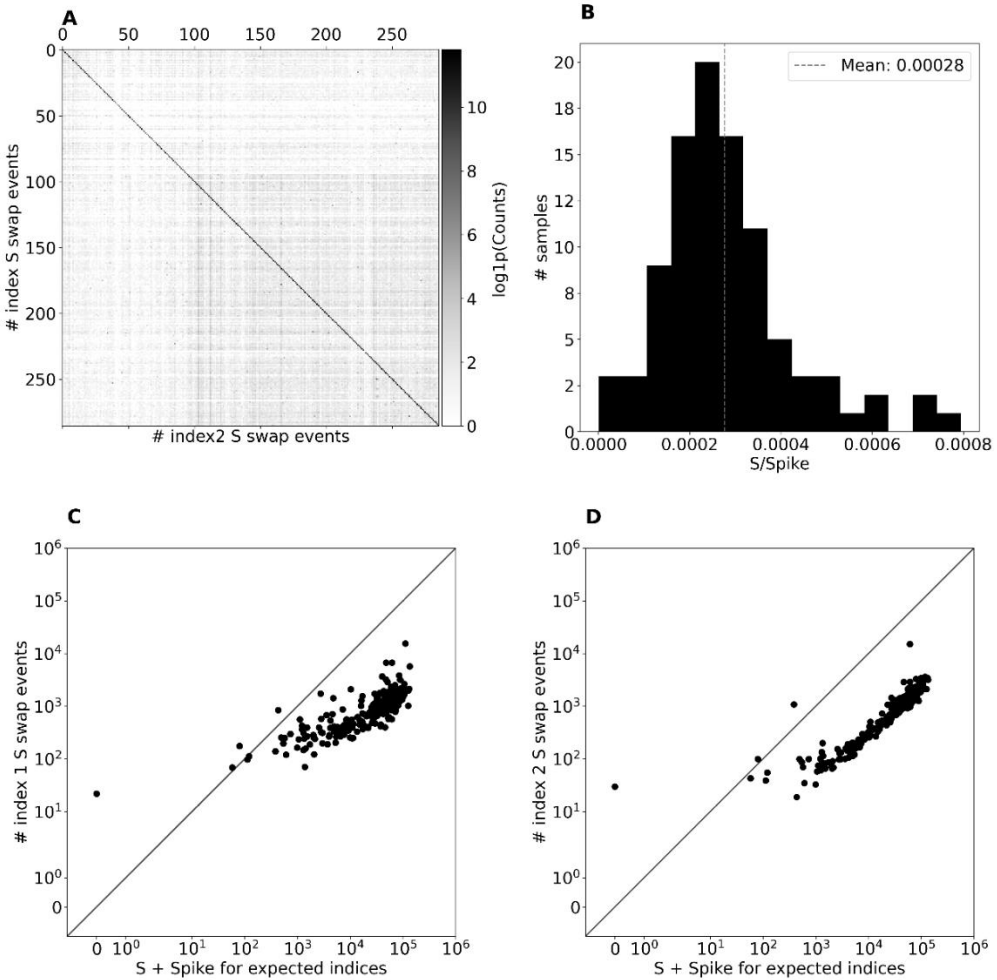


Figure S19. Quantifying the role of index mis-assignment as a source of noise in the S2 reads. A) A matching matrix for the viral S2 + S2 spike count for each pair of i5 and i7 index pairs from run v19 that used a unique dual index design. The index pairs along the diagonal correspond to expected index pairs for samples present in the experiment (expected matching indices) and the index pairs off of the diagonal correspond to index mis-assignment events. B) The distribution of ratios of viral S counts to Spike counts for samples with known zero amount of viral RNA. The mean ratio is 0.00028. C) The number of i7 mis-assignment events vs the number of viral S2 + S2 Spike counts for each sample. D) The number of i5 mis-assignment events vs the number of viral S2 + S2 Spike counts for each sample.