HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

in the subject of Biological and Biomedical Sciences

have examined a dissertation entitled

Development of nucleic acid detection methods for object provenance and viral diagnostics

presented by Jason Qian

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: Peng Yin
Peng Yin (May 5, 2021 11:57 EDT)

Typed Name: Dr. Peng Yin

Signature: Constance Cepko
Constance Cepko (May 25, 2021 08:32 EDT)

Typed Name: Dr. Constance Cepko

Signature: Pardis Sabeti

Typed Name: Dr. Pardis Sabeti

Signature: Cameron Myhrvold
Cameron Myhrvold (May 5, 2021 3:26 EDT)

Typed Name: Dr. Cameron Myhrvold

Date: April 28, 2021

Development of nucleic acid detection methods for object provenance and viral diagnostics

A dissertation presented

by

Jason Qian

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

April 2021

**Development of nucleic acid detection methods for
object provenance and viral diagnostics**

Abstract

Starting with the discovery of DNA structure, advances in molecular and systems biology in the past few decades have propelled nucleic acid amplification and detection technologies with far-reaching applications including diagnostics for healthcare to food industries, such as disease prognosis and surveillance during a global pandemic, or object provenance to determine locations of food-borne illness outbreaks. Many of these technologies require extensive instrumentation and infrastructure, restricting access in low-resource environments; therefore, recent efforts have been directed towards developing isothermal nucleic acid amplification methods amenable to field-deployable applications. In this thesis, we used isothermal nucleic acid detection method to develop field-deployable systems in two areas: object provenance and viral diagnostics. First, determining the location history of an object is a fundamental challenge for human health, commerce, and food safety. We leveraged location-specific barcoded microbial spores to create a system that can determine object provenance that is rapid, fieldable, sensitive, cheap, and can be safely introduced into and recovered from the environment. This technology has broad potential as it can help solve object provenance challenges in a wide range of applications. Second, the global SARS-CoV-2 pandemic has highlighted in many countries the limitations of inadequate testing infrastructure and the pressing need for different testing modalities. Rapid, inexpensive, and sensitive testing is essential for disease surveillance and containment strategies to be effective. We optimized the isothermal nucleic acid detection method used for microbial spores to develop a SARS-CoV-2 point-of-care diagnostics that can detect down to 5 viral RNA molecules and can provide quantitative information in crude patient samples. This method can be quickly and easily adapted to future infectious disease outbreaks. Taken together, this thesis provides some groundwork for bringing isothermal nucleic acid detection methods closer to a future where fully field-deployable technologies may become a reality.

Acknowledgements

It is difficult to imagine I would be sitting here writing this thesis without the people who have continuously supported me scientifically and personally through this entire journey. I would like to express my deep and sincere gratitude for:

My research advisor, Dr. Michael Springer, who took the chance on me as a graduate student. Mike has provided great mentorship and has taught me how to ask and critically think about biological questions and carefully design solutions to solve them. He has instilled in me the importance of collaboration and working with others. He also taught me the importance of science communication and distilling complex science into simple concepts. Mike will always serve as someone I deeply respect; he is not only brilliant scientifically, but also an incredible human being.

My tech dev teammates – Dr. Zhi-xiang Lu, Dr. Sarah Boswell and Mary Pettit, for being great collaborators who made all the projects we worked together on enjoyable and enriching experiences, even during the SARS-CoV-2 pandemic mayhem. I especially enjoyed the constant discussion and flow of ideas that made research intellectually stimulating. Together, we devised projects, performed scientifically rigorous experiments, wrote papers, developed and taught each other new skills and technologies, all the while encouraging each other to achieve our fullest potentials.

The rest of the Springer lab, both current and past members, for being a great lab. I thank them for productive scientific discussions and giving feedback on various projects and during lab meetings. I also thank LSP and the Lab Ops support team in the Systems Biology department, especially Sibylle Kellner, Samantha Jalbert, Jennie Epp, Tenzin Phulchung, Nicole Anderson who helped with all the orderings, bookings, and made our research run smoothly. Dr. Becky Ward for not only her masterful writing skills and polishing our papers, but also for bringing the best out of everyone which in turn makes the Systems Biology department a great environment to work in.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Overview

Rapid detection of nucleic acids plays a primordial role in healthcare (*1*), biothreat (*2*) or food industry (*3*, *4*), such as disease prognosis and surveillance, or monitoring and testing contaminated human food (*4*, *5*). Viral outbreaks in the past several years (e.g. Zika) (*6*) and the current worldwide SARS-CoV-2 pandemic (*7*) highlight the urgent need and challenges for diagnosing viral infection rapidly, sensitively, and specifically in order to contain the spread of pathogenic agents. Failure to control the viral SARS-CoV-2 infections has amounted to millions of deaths worldwide and put the majority of the world under lockdown which has disrupted means of livelihood and costed trillions of dollars in economic damage (*8*). Food industry is another example where there is a need for rapid DNA detection methods to identify the source of food contamination, as it often takes weeks with current approaches due to the complex modern market chain (*9*). Foodborne illness is a global health issue with 48 million cases reported every year in the US alone and costing the economy more than $15 billions (*10*). In recent years, there has been a massive expansion of isothermal nucleic acid amplification technologies (INAATs) to offer highly sensitive and rapid nucleic acid detection. These INAATS can be harnessed for detection of infectious diseases and other field-deployable applications. This chapter provides an overview of the current INAATs and their applications.

Isothermal nucleic acid amplification technologies (INAATs)

Amplification and detection of nucleic acids are two key steps for developing accurate diagnostics; trace amounts of specific DNA sequence need to be amplified in a specific manner to levels that are then detectable by a readout method (*1*). Numerous DNA amplification methods have been developed which can be divided into two categories, 1) those that require thermal cycling, and 2) those that operate at constant temperature (isothermal).

For nucleic acid amplification methods that require thermal cycling, polymerase chain reaction (PCR) or quantitative PCR (qPCR) remain the gold standard in diagnostic assays. In research and clinical settings, qPCR is used for sensitive amplification, detection, and quantitation of nucleic acids in a given sample. These methods can be adapted for multiplex detection of different targets at once (*11–13*). Nevertheless, in settings others than research laboratories, PCR-based methods are often unfeasible due to the necessity of expensive thermocycling instrument, reliable power supply, technical expertise, and other resources that may not be available in low-resource settings (*14*). To circumvent the need of specialized equipment, INAATs have been developed that often use enzymes to replace the thermocycling steps of PCR to continuously replicate targeted DNA strands (*15*). In recent years, numerous INAATs have been developed, including recombinase polymerase amplification (RPA), loop-mediated isothermal amplification (LAMP), helicase-dependent amplification (HDA), nucleic acid sequence-based amplification (NASBA), and rolling circle amplification (RCA). For each method, there are trade-offs amongst sensitivity, specificity, simplicity, cost, and speed (*15*), which are summarized in **Table 1**. The ASSURED (Affordable, Sensitive, Specific User-friendly, Robust and Rapid, Equipment-free, and Deliverable) criteria was set by the World Health Organization (WHO) as a general guidance to work towards when developing field-deployable diagnostic assays (*16*).

Recombinase Polymerase Amplification (RPA)

RPA is one of the most promising INAATs due to the low temperature requirement, the simple primer design, the rapid amplification time. RPA uses a combination of recombinases (T4 uvsX, T4 uvsY), strand displacing DNA polymerase (*Bsu*) from *Bacillus subtilis*, and single-stranded DNA-binding (SSB) (T4 gp32) proteins to achieve rapid DNA amplification at near constant ambient temperature (*17*). RPA achieves exponential nucleic acid amplification by just using two target-specific primers, one forward and one reverse as in PCR; however, there is no need for melting temperature since primer annealing and elongation steps are mediated by enzymes and not driven by temperatures. The recombinases form complexes with the primers to scan the double stranded DNA template and facilitate strand exchange at matching sequences. Typical RPA primers are 30-35 bases long to allow the formation of recombinase/primer filaments. Once the primer is bound to the DNA template, the recombinases disengage and are replaced by SSB proteins that coat to the displaced DNA strand and stabilize the resulting D-loop, which prevents the ejection of the primer by branch migration. DNA amplification by the *Bsu* polymerase can then be initiated by primer extension (**Figure 1-1**). Once initiated, the amplification reaction progresses exponentially by cyclic repetition of this process at constant 37-42°C and reaches saturation within minutes with just a few target copies of DNA (*17*). RPA was demonstrated to amplify nucleic acid using just body heat (*18*).

One of the key reagents unique to RPA is polyethylene glycol (PEG), a high-molecular weight crowding agent, that helps to stimulate interaction of enzymes with nucleic acids (*17*). In RPA, the ATP to power the enzymes is generated by creatine kinase and phosphocreatine. Once the enzymes are activated by magnesium acetate, the amplification of DNA template continues until exhaustion of the phosphocreatine pool. RPA can amplify DNA template up to 1.5 kb but in general shorter amplicons are recommended (100 – 200 bp) (*17*).
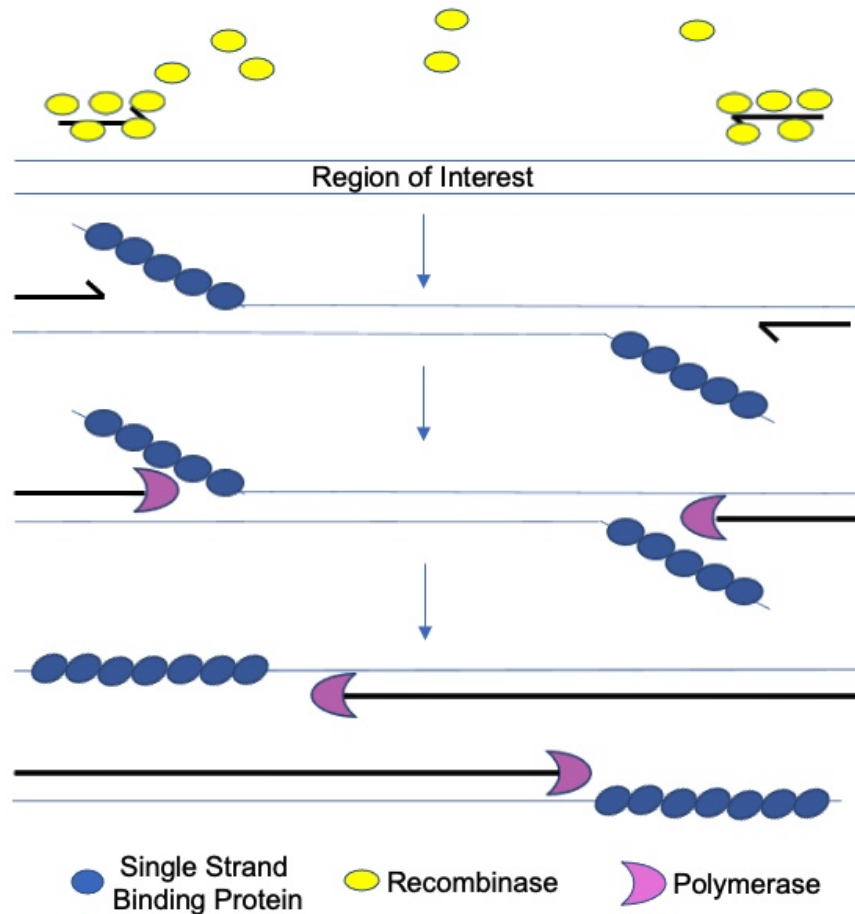
Figure 1-1 Schematic of RPA amplification.

Recombinases (yellow circle) form complexes with forward and reverse primers to scan template and aid strand exchange at homologous sequences. Concurrently, primers are inserted at cognate sites and SSB (blue circle) stabilize D-loops. Then recombinases disengage, and *Bsu* polymerase (pink crescent) initiate primer extension. Exponential amplification happens thru cyclic repetition of this process.

RPA is versatile and can be paired with different detection methods - fluorescent probes (*19*), lateral flow strip assay (*19*, *20*), and CRISPR detection readout (*21*) - making it suitable for field-deployable applications. Fluorescent probes (exo and fpg) have been developed to detect RPA reaction in real-time. Since traditional Taq-man probes are not compatible with RPA due to the 5 → 3' exonuclease activity of Taq polymerase, the exo or fpg fluorescent RPA probes rely on having an abasic base, located between a flutophone and a quencher, that can be cleaves by an exonuclease or a DNA glycosylase (*17*, *19*). Post RPA amplification detection by lateral flow

strip assay is achieved by the LF-probe which usually has a 5'-FAM tag, an abasic residue, and is 3' blocked. The reverse RPA primer is 5' biotinylated and forms a doubly labeled amplicon (FAM/Biotin), which can be captured by gold nanoparticles conjugated with anti-FAM antibodies (*19*). Nevertheless, these RPA probes are expensive and time-consuming to manufacture, making them less adaptable during rapidly evolving pathogenic disease outbreaks when mutated variants can quickly arise. More recently, CRISPR detection methods such as those using programmable Cas12a (*22*) and Cas13a (*21*) endonucleases have been developed and paired with RPA to sensitively detect amplified nucleic acids. Upon binding to a cognate DNA target via the guide RNA, the CRISPR enzyme unleashes its collateral cleavage activity cutting an initial quenched fluorescent reporter, which produces a fluorescent signal (*23*).

RPA carries robust nucleic acid amplification with input from various sample types with minimal sample preparation and is tolerant to common PCR inhibitors, including urine (*24*), soil samples (*25*), milk (*26*), serum or crude swab samples only requiring heat lysis and lysis buffer (*27*). While each sample type and target should be optimized with RPA, the possibility of minimizing sample preparation makes it an ideal method to portable assay development.

A shortcoming of RPA and other INAATs is the amplification of non-specific amplicons, arising from primer dimers and off-target hybrids. The non-specific side reactions are a common issue found for INAATs since in these systems, the Watson-Crick base-pairings are not formed and reformed thermodynamically but rather mediated enzymatically. The formation of these non-specific amplicons is not well understood, but it is well documented that in the absence of DNA template, a reaction containing just the four nucleotide bases (dNTPs), primers, and enzymes generates self-replicating mixture of DNA species (*28*, *29*). This leads to high amounts of non-specific species that compete with specific products and causes lower sensitivity or even false-positive results. Primer designs with modified bases have been explored as potential solutions to prevent non-specific products. Self-avoiding molecular recognition system (SAMRS) minimizes non-specific products in PCR (*30*), RPA (*31*), and other INAATs (*32*, *33*). The SAMRS nucleotides

are variants of the natural bases, and they are capable of binding to natural nucleotides but not to themselves. While placing SAMRS nucleotides in the 3'-end of the primers reduced non-specific amplicons (*31*), these modified nucleotides are expensive and time-consuming to produce. Another primer modification strategy is to insert an RNA base a few nucleotides upstream of the 3'-end of primers. Once the primer is bound its cognate target, the polymerase is blocked from extending the primers until RNAse H2 cuts RNA/DNA hybrid at the RNA base position, releasing the blocked end (*34*). To date, the best strategy to mitigate non-specific amplificons is careful primer design (*35*, *36*) and screening to find ones with best sensitivity for on-target amplification.

RPA and other INAATs currently lack simple quantitation readout. INAATs mostly output a binary signal of either the presence or the absence of a target. A quantitative point-of-care test would be useful to monitor the progress of viral infections or other diseases and potentially guide drug treatment as in the case for HIV titers (*37*). Quantitative real-time RPA using fluorescent probes and competitive amplification has been reported (*14*); however, probes with internal modified bases are difficult and expensive to manufacture and require a sensitive instrument for readout. A quantitative end-point LAMP method has been reported that enables order-of-magnitude quantitation (*38*), but as mentioned in the next section, LAMP requires higher temperature and a more complex primer design than RPA. Other attempts of making RPA quantitative involve compartmentalization of reagents and require more complex setup and instrumentation, including digital plasma separation (*39*), centrifugal step emulsification (*40*), picowell-based technology (*41*), and microfluidics technology (*42*)

Loop-mediated Amplification (LAMP) and other methods

LAMP is another promising INAAT as it requires a set of 4 to 6 primers and a strand displacing polymerase (*Bst* polymerase) and works optimally at 60-65C (*43*). Typically, there are two outer primers (F3 and B3) and two inner primers (forward internal FIP and backward internal BIP) (**Figure 1-2**). Two additional primers (loop F and loop B) are often used to accelerate the

LAMP reaction (*44*). As depicted in **Figure 1-2**, the inner and outer primers and strand displacing polymerase help generate dumbbell intermediary products since both the 5' and 3' ends are self-complementary to sequences further inwards creating the stem-loop structure. This intermediary product and subsequent binding of inner, outer and loop primers generate very long cauliflower-shaped DNA structures (>20 kbp) (*43*). Many methods exist to detect LAMP amplification in real-time including fluorescent dye, turbidity, colorimetric pH dye (*45*). Post-amplification detection methods also exist including oligo strand displacement probes, molecular beacons, lateral flow strips (*45*), and more recently CRISPR (*46–48*). LAMP's requirement of using only 1 enzyme, the polymerase, to amplify DNA makes it an attractive INAAT; however, its 4-6 primer requirement is more complicated than other INAATs and reduces the user's target sequence space. LAMP primer design tools have been reported (*49*).

Helicase-dependent amplification (HDA), like RPA and PCR, uses 2 primers for amplification. HDA relies the helicase enzyme to first unwind double stranded DNA, and SSB stabilizes single stranded DNA allowing primers to anneal and a strand displacing polymerase to extension (*50*). This cyclical process achieves exponential amplification at constant temperature between 37-65C depending on the polymerase (**Figure 1-3**). Despite its simplicity HDA also suffers from non-specific amplification as with other INAATs (*51*). In addition to primer modifications, reagents (e.g. DMSO, betaine, sorbitol) that reduce DNA secondary structures to facilitate primer annealing have been used to minimize non-specific amplification (*52*). HDA reaction is slower than RPA and LAMP.

Nucleic acid sequence-based amplification (NASBA) relies on 2 primers and multiple enzymes (reverse transcriptase, RNase H, RNA polymerase) at 42C for amplification (*53*). NASBA system relies on reverse transcription and in-vitro transcription to produce single stranded RNA products. Rolling circle amplification (RCA) relies on a strand displacing polymerase to amplify a long single stranded circular DNA product (*54*). RCA is linear; therefore, it amplifies more slowly and is less sensitive than other INAATs that rely on exponential amplification.
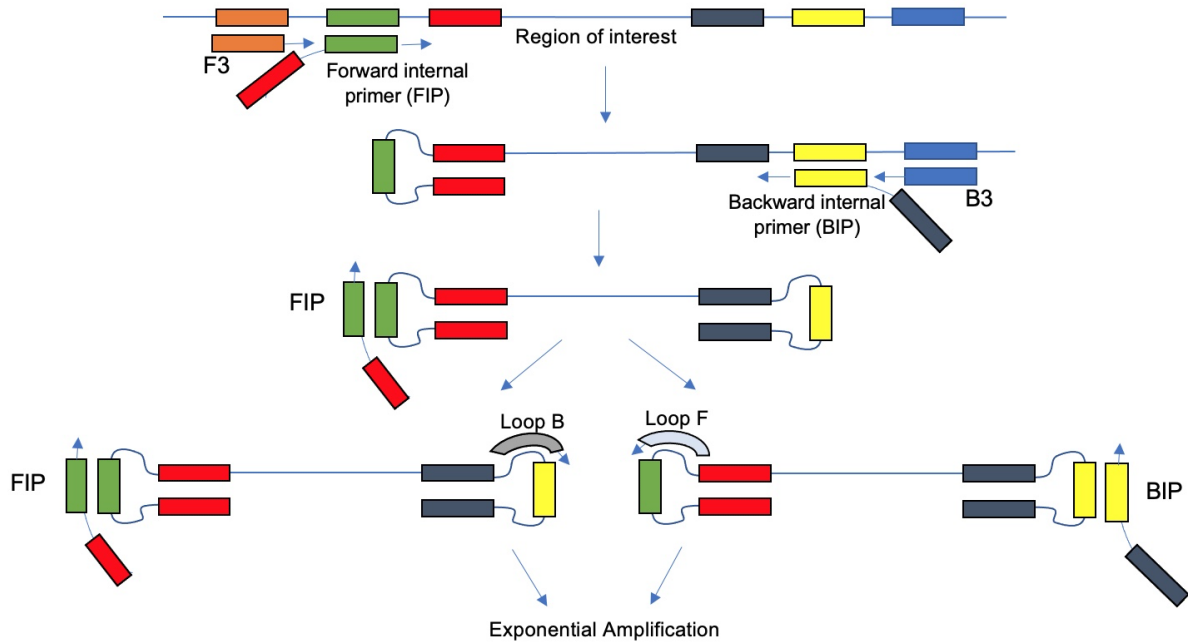
Figure 1-2 Schematic of LAMP amplification

Outer primers (F3 and B3) and inner primers (FIP and BIP) are used to generate stem-loop DNA structures intermediates. Inner primers and loop primers then help to exponentially amplify cauliflower-like products
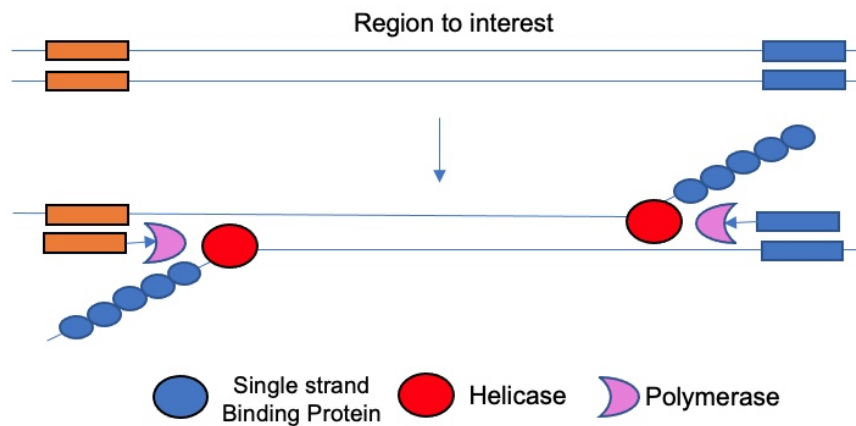
Figure 1-3 Schematic of HDA amplification

Helicase (blue circle) unwinds double stranded DNA. Primers (orange and blue rectangles with arrows) anneal to cognate site (orange and blue rectangles), then primers are extended by polymerase (pink crescent) to form a new dsDNA.

Isothermal Method	Input	Primers needed	Running temperature	Amplification time (min)	Limit of detection (copies)
RPA	DNA/RNA	2	37-42 C	20-40	1
LAMP	DNA/RNA	4-6	60-65 C	60	~5
HDA	DNA/RNA	2	30-65 C	60-120	1
NASBA	RNA	2	42 C	90-180	~5
RCA	DNA/RNA	1	65 C	120 – 240	~10

Table 1 Comparison of INAATs

INAATs applications

INAATs can be widely used in forensics, agriculture and disease diagnostics (*55*). One of most exciting applications of INAATs is in pathogenic disease diagnostics. In recent years, there has been an explosion of studies that have used INAATs for diagnosing infectious diseases including viruses (*56–63*), bacteria (*64–70*), parasites (*71, 72*), and SNP detection in cancer cells (*73, 74*). These studies have primed the scientific community to quickly adapt their methods for the detection of newly emerging disease outbreaks. One of the key steps to consider for developing a sensitive INAAT is to design and screen primer pairs that minimize non-specific reactions. For RPA, there is a lack of streamlined method to screen for primer pairs; current studies rely on running an RPA reaction and using gel electrophoresis as readout, which is low-throughput and inefficient (*75*).

Another exciting application of INAAT is to solving object provenance. Determining the provenance of an object is a major challenge for the food and health industries, as it would permit rapid identification of the source of foodborne illnesses or disease transmission. Foodborne illness cases often take weeks to determine origins of the outbreaks due to the complexities of food supply system (*9*). One way to help track objects through the food supply chain is using microbiome. Recent studies have shown that naturally occurring microbial communities can be transferred from their environment to objects that pass through it (*76*); thus, it has been suggested that the microbial composition of an object could be used as a forensic tool to determine object

provenance (*77, 78*). The potential use of microbiome to derive location origins and distinctive signatures has been reported for samples collected from various locations, including soil (*79, 80*), hair (*81*), skin, and keyboards (*82*). The most common technique to examine the microbiome is analyzing the distribution of microbial species composition in a sample using sequencing of 16S ribosomal RNA gene (*83*). 16S rRNA-based sequencing of office space was shown to be predictive at city-scale resolution. Microbial species varied significantly enough between offices in Tucson, San Francisco and New York City to have predictive power (*84*). While in principle analyzing the transfer of microbiomes between that of an environment and that of objects traversing through it to provide geolocation information is appealing, practically the microbiome of an environment fluctuates too rapidly for this method to be useful for determining the geographical origins of objects at high-resolution. Moreover, existing methods used to analyze microbiome sequencing data also requires expensive instrumentation and technical expertise, including bioinformatics, statistical analysis, and next generation sequencing, which are costly, difficult to scale, and time-consuming (*85*). These current methods set the stage for the development of a novel system which integrates INAATs to make identification of object provenance scalable in a high-throughput manner and easily implementable.

Chapter 2: Barcoded microbial system for high-resolution object provenance

Chapter 2 Contributions

The contents of this chapter were from: <u>J. Qian*</u>, Z. Lu*, C. P. Mancuso*, H.-Y. Jhuang*, R. del C. Barajas-Ornelas*, S. A. Boswell*, F. H. Ramírez-Guadiana, V. Jones, A. Sonti, K. Sedlack, L. Artzi, G. Jung, M. Arammash, M. E. Pettit, M. Melfi, L. Lyon, S. V. Owen, M. Baym, A. S. Khalil, P. A. Silver, D. Z. Rudner, M. Springer, Barcoded microbial system for high-resolution object provenance. *Science*. 368, 1135–1140 (2020). The paper was adapted and formatted for this dissertation chapter.

This work is the result of a fruitful collaboration with members of the following laboratories: MS, DZR, PAS, ASK, MB. JQ, ZL, CPM, HYJ, RCB-O, SAB conceived the study, performed most key experiments and data analyses (supervised by MS). More specifically, JQ and SAB co-led experiments and analyses for detection pipeline (Figure 2-1 thru Figure 2-8). RCB-O and FHR-G co-led *B. subtilis* spore experiments and analyses (Figure 2-3). HYJ performed sporulation validation experiments for *S. cerevisiae* (Figure 2-4). CPM led small-scale test surfaces experiments and metagenomics analyses (Figure 2-9). ZL and JQ co-led in-field testing experiments and analyses (Figure 2-10 thru Figure 2-18). SAB and HYJ co-led initial food provenance experiments and analyses (Figure 2-19, Figure 2-22, Figure 2-23). JQ and ZL co-led food cross-association experiments and analyses (Figure 2-20, Figure 2-21). JQ, ZL, CPM, HYJ, RCB-O, SAB and FHR-G wrote the paper (with MS, DZR). VJ, AS, KS, LA, GJ, MA, MEP, MM, LL, SVO assisted with technical experiments. SAB purified Cas13 proteins used in this study. RCB-O and FHR-G designed and constructed *B. subtilis* strains and purified spores, and microscopy of spores (supervised by DZR). DZR, PAS, ASK, MB provided critical insights used in this study. All authors reviewed the manuscript**.**

Chapter 2 Acknowledgements

Chapter 2 Abstract

Determining where an object has been is a fundamental challenge for human health, commerce, and food safety. Location-specific microbes in principle offer a cheap and sensitive way to determine object provenance. We created a synthetic, scalable microbial spore system that identifies object provenance in under one hour at meter-scale resolution and near single-spore sensitivity, which can be safely introduced into and recovered from the environment. This system solves the key challenges in object provenance: persistence in the environment, scalability, rapid and facile decoding, and biocontainment. Our system is compatible with SHERLOCK, a Cas13a RNA-guided nucleic acid detection assay, facilitating its implementation in a wide range of applications.

Chapter 2 Introduction

Globalization of supply chains has dramatically complicated the process of determining the origins of agricultural products and manufactured goods. Determining the origin of these objects can be critical, for example in cases of foodborne illness, but current labeling technologies are prohibitively labor-intensive and easy to subvert (*86*). Tools that label persons or objects passing through a location of interest could also be useful to law enforcement as a complement to fingerprinting and video surveillance (*87*). Microbial communities offer a potential alternative to standard labeling approaches. Any object gradually adopts the naturally occurring microbes present in its environment (*76*, *78*); thus, it has been suggested that the microbial composition of an object could be used to determine object provenance (*77*). Challenges with this approach include variability of resident microbial community abundance over time, similarities of microbial composition between different locations, and the requirement for extensive, expensive, and time-consuming mapping of natural environments.

To circumvent these challenges, we propose the deliberate introduction and use of synthetic non-viable microbial spores harboring barcodes that uniquely identify locations of interest (e.g. food production areas). These synthetic spores would offer a sensitive, inexpensive, and safe way to map object provenance provided that several important criteria are met, including: 1) the microbes must be compatible with growth at industrial scale, 2) the synthetic spores must be bio-contained and not viable in the wild to prevent adverse ecological effects; 3) the synthetic spores must persist in the environment and reliably label objects that pass through it; and 4) the encoding and decoding of information about object provenance must be rapid, sensitive and specific. Similar barcoding approaches have been explored previously to model pathogen transmission (*88*, *89*), but did not explicitly address those challenges.

Chapter 2 Results and Discussion

Here we report the BMS (Barcoded Microbial Spores) system, a scalable, safe and sensitive system that uses DNA-barcoded microbial spore mixtures to permit the determination of object provenance (**Figure 2-1**).
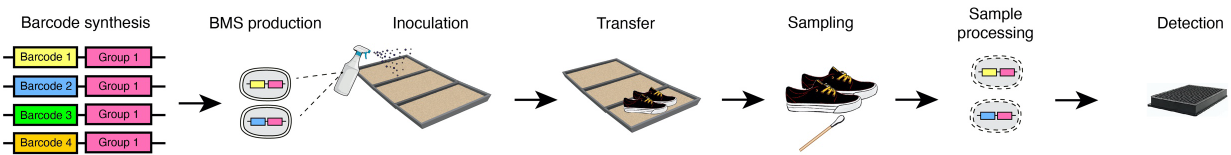
Figure 2-1 Schematic of the Barcoded Microbial Spores (BMS) application and detection pipeline

The BMS system leverages the natural ability of spores to persist for long periods in the environment without growth (*90*). We designed unique DNA barcodes and integrated them into the genomes of *Bacillus subtilis* (*B. subtilis*) and *Saccharomyces cerevisiae* (*S. cerevisiae*) spores, creating a set of BMS that can be used combinatorially to provide a nearly infinite set of unique identification codes. The BMS can be: 1) manufactured at scale using standard cloning and culturing techniques; 2) inoculated to surfaces by spraying; and 3) transferred to objects that come into contact with the inoculated surface. To identify barcodes, BMS sampled from objects are lysed and can be decoded with a range of methods including SHERLOCK, a recombinase polymerase amplification (RPA) method coupled with a Cas13a-based nucleic acid detection assay (*21*), qPCR, and sequencing (**Figure 2-1**and **Figure 2-2**).
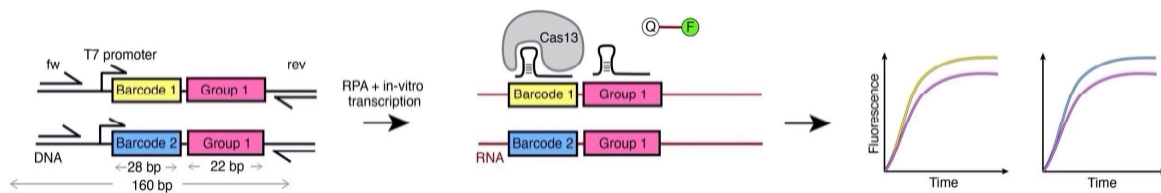
Figure 2-2 Schematic of BMS detection pipeline SHERLOCK.
Schematic of the DNA barcode region design (160 bp) with RPA primers. Specific barcodes (28 bp) are highlighted in yellow and blue. The group barcode (22 bp) is highlighted in pink.

The BMS are designed not to impact the native environment into which they are applied. First, we used auxotrophic strains that require amino acid supplementation for growth. Second, we made the cells germination deficient. For *B. subtilis* spores, we deleted the genes encoding the germinant receptors and the genes that encode the cell wall lytic enzymes required to degrade the specialized spore cell wall. Incubation of $>10^{12}$ spores generated from this mutant strain showed they were unable to form colonies or grow in rich medium, and remained stable and non-germinating at room temperature for >3 months (**Figure 2-3**).
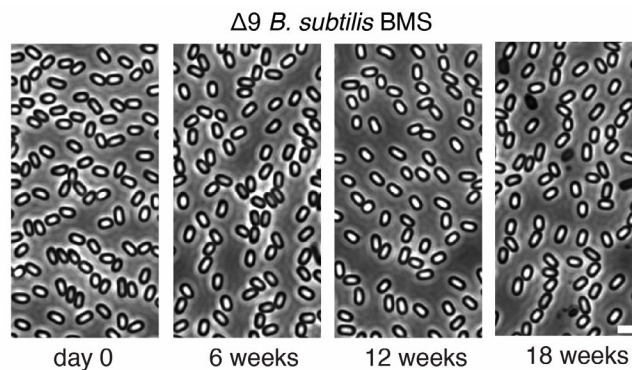
Figure 2-3 *B. subtilis* BMS remain stable and dormant over 4 months
At the indicated time points, spores when stored in PBS at room temperature were analyzed by phase-contrast microscopy. Scale bar indicates 2 µm

For *S. cerevisiae*, we boiled spores for 30 minutes to heat-kill vegetative cells and spores prior to application. Incubation of $>10^8$ boiled spores on rich medium yielded no colonies (**Figure 2-4**).
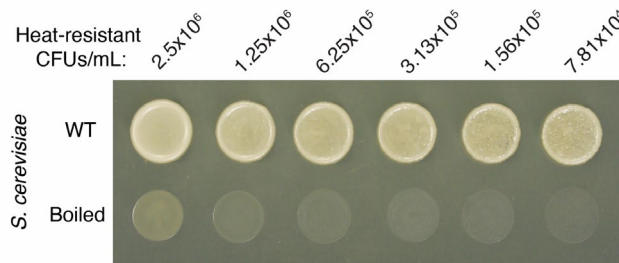
Figure 2-4 S*. cerevisiae* BMS are unable to germinate, outgrow, and form colonies on nutrient rich medium after boiling.
2.5×10^6 cells before and after 1 hour of boiling were each serially diluted at 2-fold and spotted on YPD. Note that the dim circles in the boiled sample on the YPD plate are cell debris and do not indicate growth.

All antibiotic resistance cassettes used to generate the BMS were removed by site-specific recombination to prevent horizontal gene transfer of resistance genes to other organisms in the environment. Finally, the inserted barcode does not encode any gene and should not confer any fitness advantage if horizontally transferred. We tracked the microbiome of soil samples, and found that inoculation with BMS had insignificant effects on the microbiome compared to natural changes over time or in response to watering (**Supp Fig 2-1**). We also note that *B. subtilis* and *S. cerevisiae* are both commonly found in environmental and food samples.

Multiple BMS can be applied and then decoded simultaneously. We designed a series of tandem DNA barcodes, each with a Hamming distance of >5, allowing more than 10^9 possible unique barcodes. To test the specificity of our barcode design in a field-deployable system, we constructed 22 barcodes and their matching crRNAs and assayed all permutations *in vitro* using SHERLOCK. All 22 crRNAs clearly distinguished the correct barcode target (**Figure 2-5**).
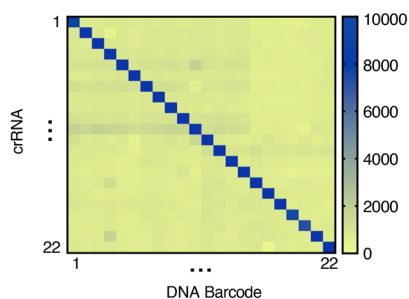
Figure 2-5 SHERLOCK specificity testing.
Heatmap of endpoint fluorescence values from in vitro SHERLOCK reactions. All combinations of 22 barcodes and 22 crRNAs assessing specificity of each barcode-crRNA pair.

To scale the system, we devised a rapid and facile method to screen a large number of barcodes and crRNAs in parallel, to eliminate those with cross-reactivity or background; we validated and performed pooled *n-1 barcode* RPA reactions *in vitro* with corresponding crRNA and water RPA controls testing 94 crRNA-barcode pairs, eliminating 17 for high background and 7 for cross-reactivity (**Figure 2-6**).

19
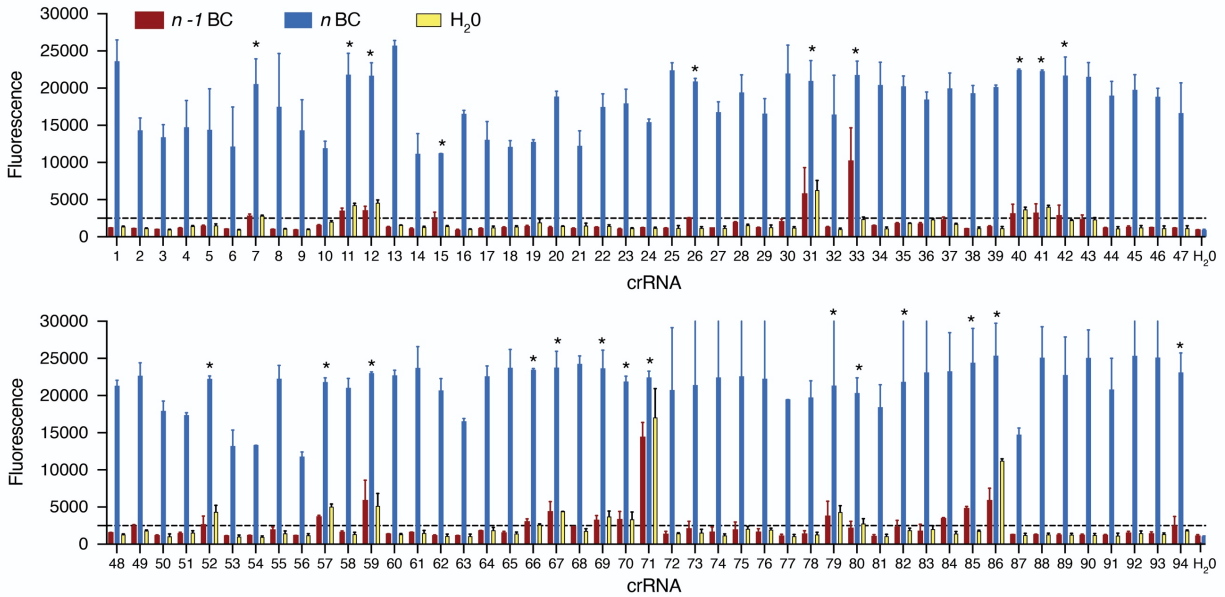
Figure 2-6 In vitro DNA barcode and crRNA cross-reactivity assay.

Bars depict SHERLOCK signal from reactions using different DNA megamer at equimolar template concentrations, with the n-1 barcode reactions in red, barcode specific reactions in blue, and H20 RPA reactions in yellow. * denotes crRNA with high background or cross-reactivity. (n = 3 technical replicates, error bars represent mean + s.e.m.).

To test sensitivity and specificity *in vivo*, we integrated 57 barcodes into *B. subtilis* and 11 into *S. cerevisiae*. We developed an efficient spore lysis protocol using heat and sodium hydroxide (**Supp Fig 2-2**), which allowed us to achieve near single-spore resolution for detection using SHERLOCK (**Figure 2-7**).
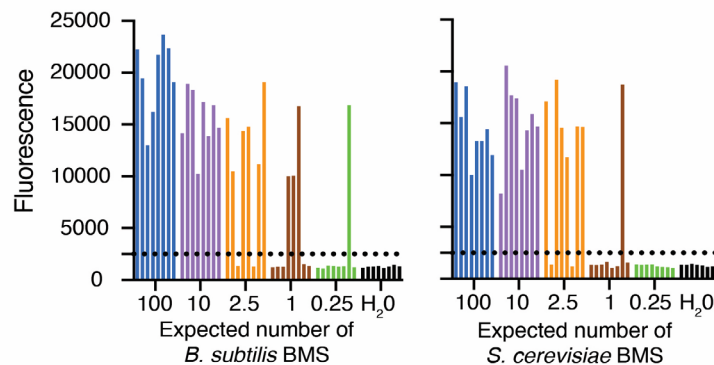
Figure 2-7 Detection limit of B. subtilis and S. cerevisiae BMS by SHERLOCK.

(Each of the 8 biological replicates for each spore concentration are shown). Spore numbers are calculated on a per reaction basis.

In vivo and *in vitro* specificity screenings of crRNA-barcode pairs gave similar results (**Supp Fig 2-3**). In addition, our barcodes are tandemly designed with a unique sequence and a shared group sequence (**Figure 2-2**) to aid in high-throughput detection settings where only a subset of samples contains the BMS of interest. The group sequence is compatible with field-deployable detections and can be used to determine whether a BMS of interest is present before using a second assay to uniquely identify the BMS (**Figure 2-8**). This two-step process solves the throughput limitations of field-deployable detection and lowers the costs of sequencing.
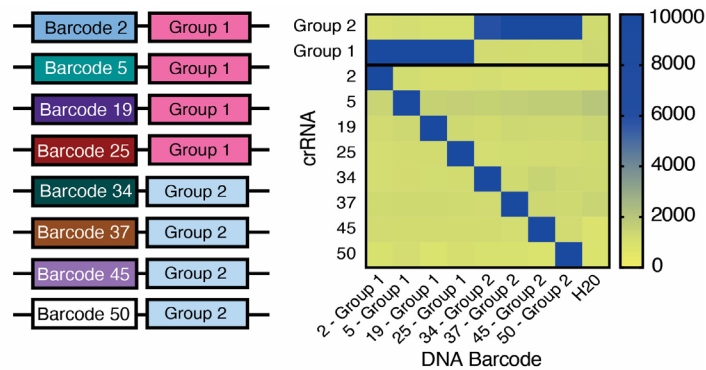
Figure 2-8 Group barcode testing.
Heatmap of endpoint fluorescence values from in vitro SHERLOCK reactions testing specificity of 4 barcodes for group 1 crRNA and 4 barcodes for group 2 crRNA as detected by either unique or group crRNA.
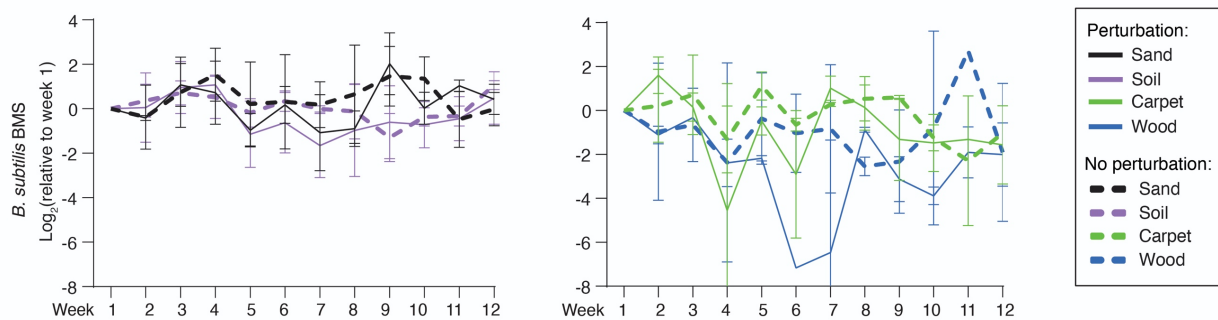
Figure 2-9 BMS persistence on sand, soil, carpet, and wood over 3 months on ~1 m2 test surfaces.
 y-axis: B. subtilis BMS number relative to week 1 levels. (Bars represent standard deviations).
Perturbation: simulated wind, rain, vacuuming, or sweeping in incubators.

The BMS system is robust and can function on different surfaces in simulated real-world environments. First, in ~1 m^2-scale experiments (**Supp Fig 2-4** and **Table S6**), we used qPCR to detect and quantify BMS directly from surface samples or surface swabs. We found that BMS persisted on sand, soil, carpet, and wood surfaces for at least 3 months with little to no loss over time (**Figure 2-9** and **Supp Fig 2-4).**

Notably, multiple tested perturbations (e.g. simulated wind, rain, vacuuming, or sweeping in **Supp Fig 2-4**) did not significantly reduce our ability to detect BMS from the surface. Second, we constructed a ~100 m^2 indoor sandpit (**Figure 2-10** and **Figure 2-12**), inoculated one region with BMS, and were able to readily detect the BMS for 3 months using SHERLOCK (**Figure 2-11** and **Supp Fig 2-5**).
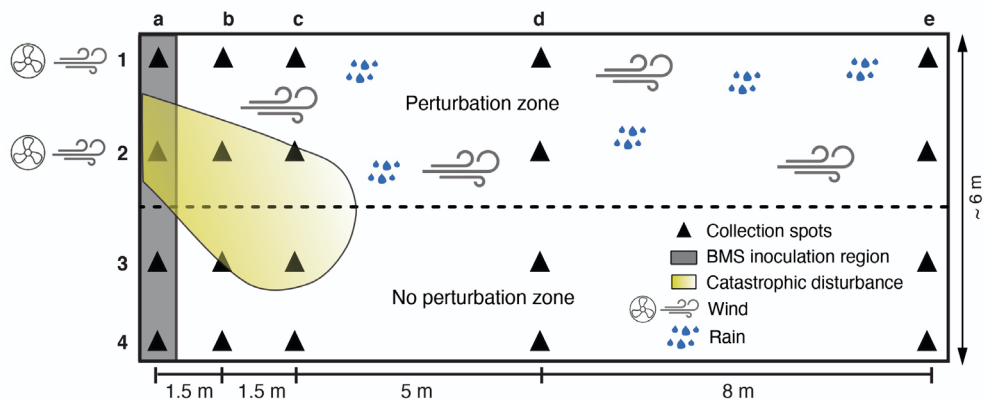
Figure 2-10 Schematic of large-scale (~100 m^2) sandpit.
B. subtilis BC-24 and BC-25 BMS, and *S. cerevisiae* BC-49 and BC-50 BMS were inoculated in the shaded grey region. The yellow shadow represents the area over which the top 2 inches of sand from a 1.5 m^2 area of the inoculation region was redistributed after a large fan fell over.
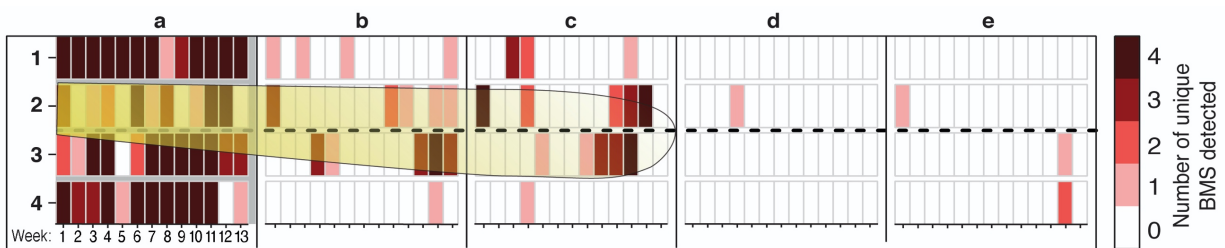
Figure 2-11 BMS persistence in large-scale sandpit.
BMS persist at collection point "a" (within the inoculated region) and do not spread to collection points "d" or "e". Heatmap depicts the number of BMS (out of 4) detected by SHERLOCK at each collection point over 13 weeks.

Importantly, perturbations did not cause appreciable spreading to non-inoculated areas (**Figure 2-11** and **Supp Fig 2-5**); even a catastrophic disturbance where a fan fell into the pit, displacing a significant amount of sand, only spread the BMS several meters (**Figure 2-12**). In an outdoor environment, BMS inoculated on grass was still detectable after 5 months of exposure to natural weather, with minimal spreading outside of the inoculated region (**Figure 2-13**). This is consistent with low levels of re-aerosolization reported for other spore-forming *Bacilli* (*91*)
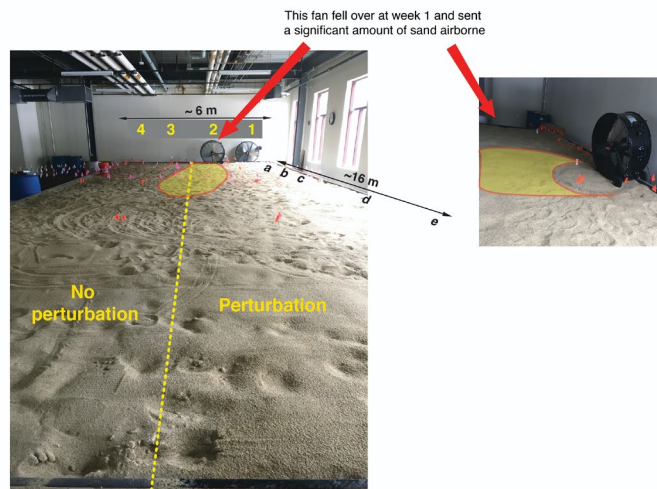
Figure 2-12 Catastrophic disturbance.

Photographs of the large-scale sandpit, with the disturbance area from the fan falling over indicated in yellow.
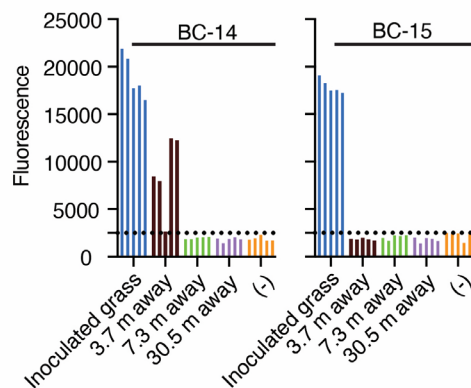
Figure 2-13 BMS persistence on grass in an outdoor environment for at least 5 months.

Grass region was inoculated with B. subtilis BC-14 and 15 BMS. Samples from actual BMS-inoculated region, 3.7 m, 7.3 m and 30.5 m away from the inoculated region were tested by SHERLOCK using crRNA 14 and 15 (each of the 5 biological replicates for each grass region are shown).

The BMS can be transferred onto objects that pass through test environments. In ~1 m^2 scale testing, BMS could be transferred onto rubber or wooden objects simply by placing them on the BMS-inoculated surface for several seconds, yielding up to ~100 spores per microliter of reaction input (**Supp Fig 2-4**). At ~100 m^2 scale, BMS were reliably transferred onto shoes worn in the inoculated sandpit (**Figure 2-14 and Supp Fig 2-6)**.
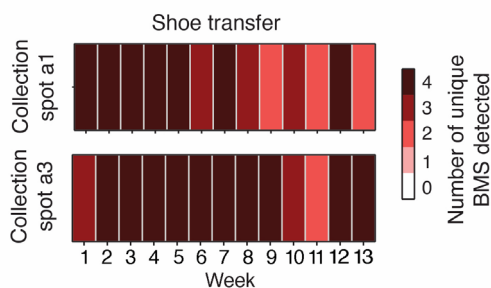
Figure 2-14 BMS transfer onto shoes.
BMS were transferred onto shoes by walking in the inoculated region "a" in the sandpit, and were detected by SHERLOCK.
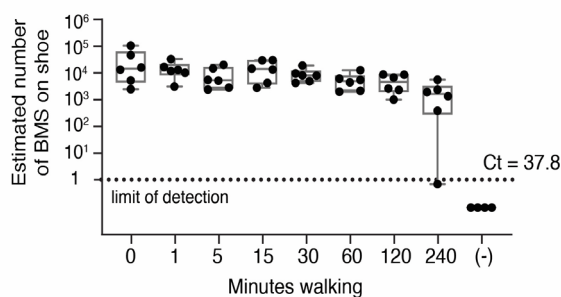
Figure 2-15 BMS maintenance on shoes.
Abundance of BMS on shoes after up to 240 minutes of walking on non-inoculated outdoor areas; y-axis: BMS count based on qPCR standard curve.

Furthermore, the BMS transferred onto shoes could still be detected even after walking on non-inoculated surfaces for several hours, though BMS counts decreased by 2-fold with 2 hours of walking as quantified by qPCR (**Figure 2-15** and **Supp Fig 2-7**). We were unable to detect spores on non-inoculated surfaces after walking on them with shoes that had traveled through BMS-inoculated regions (**Supp Fig 2-8**). We conclude that the BMS can persist in the environment

24

without significant spreading; are transferable onto objects that pass through the environment; are retained on these objects; and can be sensitively and specifically detected using SHERLOCK.

The BMS system can be used to label specific locations of interest to determine whether a person or object has passed through them. We divided different surfaces into grids, inoculated each grid region with 1, 2, or 4 unique BMS (**Figure 2-16**) and traversed them with different test objects (e.g. shoes). To mimic in-field deployment, we used a portable light source, an acrylic filter, and a mobile phone camera to image the SHERLOCK readout (**Figure 2-16**) and determine object provenance (**Figure 2-17**, **Supp Fig 2-9**, **Supp Fig 2-10** and **Supp Fig 2-11**). Notably, provenance could be determined in the field within ~1 hour from sample collection.
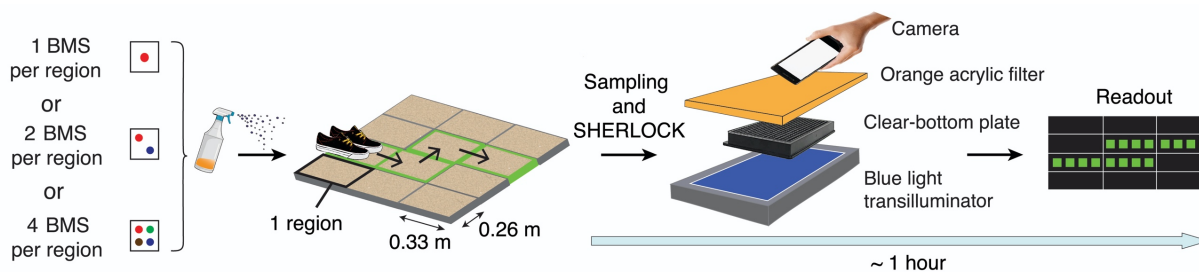
Figure 2-16 Schematic of experimental design for field-deployable detection.
Method to determine previous locations of an object. Each region was inoculated with 1, 2, or 4 BMS. Green outlines indicate the object path through a subset of the regions. SHERLOCK reactions

To evaluate the sensitivity and specificity of our system for determining an object's provenance, we considered different criteria for classification with varying numbers of unique BMS per region. Object provenance could be determined with a 0.6% (1/154) false positive rate and 0% (0/62) false negative rate if regions were inoculated with 4 unique BMS based on the criteria of >=2 positive BMS calls. Inoculating with only 1 or 2 unique BMS per region still permitted object provenance to be determined, albeit at higher error rates (**Figure 2-17, Figure 2-18,** and **Supp Fig 2-12**). Importantly, provenance could be determined on all 4 surface types tested (sand, soil, carpet, or wood) (**Supp Fig 2-11**); further validation will be needed to determine error rates in real-world environments. This experiment demonstrates that the BMS can be used to determine

25

object provenance at meter-scale resolution, which would be extremely difficult to achieve using natural microbiome signatures (*92*).
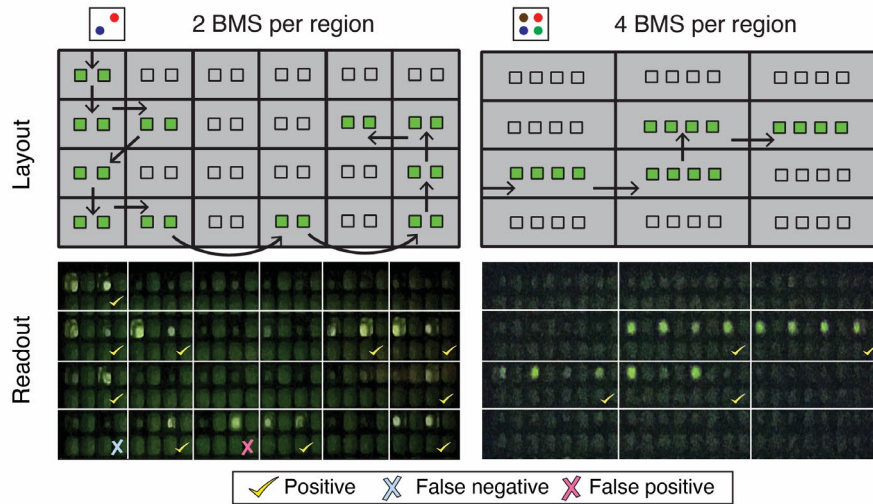
Figure 2-17 Object provenance using 2 and 4 BMS per region.

Left: 2 BMS per region. Right 4 BMS per region. Top: path of an object overlaid on reaction plate. Bottom: photograph of SHERLOCK reaction plate, overlaid with correct or incorrect calls. The call for each region is denoted by color (yellow check: true positive, blue cross: false negative, purple cross: false positive).

BMS per region		Objects (*n*)	False positive rate	False negative rate
1	🔴	32	16.9% (67/397)	19.8% (48/243)
2		20	4.9% (12/245)	9.4% (13/139)
4		22	0.6% (1/154)	0% (0/62)

Figure 2-18 Statistics for SHERLOCK provenance predictions of objects.

Objects traversed regions inoculated with 1, 2, or 4 unique BMS per region. The false positive and false negative rates for 1 and 2 unique BMS per region were based on the criteria of >=1 positive call; the rates for 4 unique BMS per region were based on the criteria of >=2 positive calls.

The BMS system offers a flexible and comprehensive approach to determining food provenance. Foodborne illness is a global health issue with an estimated 48 million cases each year in the US alone (*10*). There is an urgent need for rapid methods for identifying the source of food contamination; current approaches often take weeks and are costly due to the complex modern market chain (*9*). Plants inoculated with *B. subtilis* BMS allowed us to map lab-grown leafy plants

back to the specific pot in which they were grown (**Figure 2-19**). BMS were inoculated 4 times, beginning 1 week after the first set of leaves appeared, to match recommended inoculation protocols for *Bacillus thuringiensis* spores (*Bt*), an FDA-approved biocide that is widely used in agriculture (*93*). One week after the final BMS inoculation, a leaf and soil sample from each pot were harvested and tested using SHERLOCK. All the samples were positively detected except for the 2 plants that had received variant group barcode sequences, demonstrating the specificity of detection (**Figure 2-19**). Using Sanger sequencing, we then identified the pot in which each plant was grown for all 18 BMS (**Supp Fig 2-13**). The process from DNA extraction to sequence identification took less than 24 hours. This time frame could likely be shortened to hours with massively parallelized hybridization-based detection (*94*).
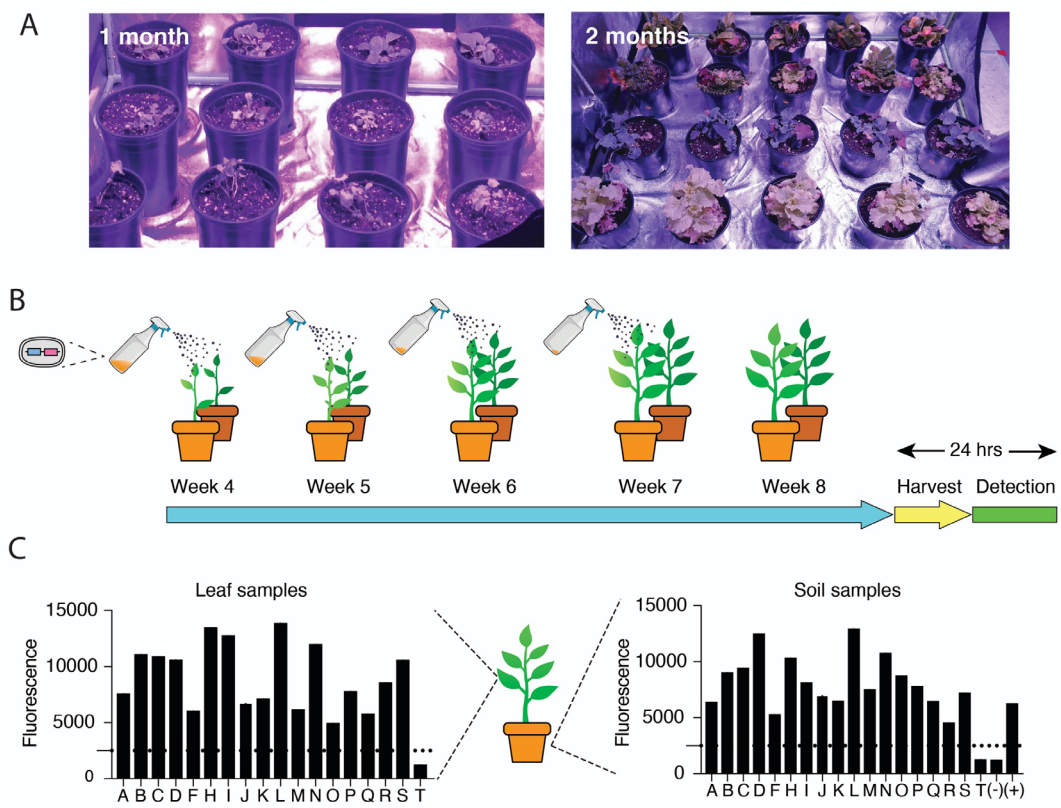
Figure 2-19 Determining the provenance of produce using BMS.

(A) 18 lettuce plants were inoculated with distinct *B. subtilis* BMS, inoculated once a week (4 times total). (B) 18 plants were inoculated with distinct B. subtilis BMS, inoculated once a week (4 times total). (C) Detection of BMS on plants and soil after harvesting by SHERLOCK with a group crRNA; y-axis:

endpoint fluorescence values. Plants A to S were sprayed with BMS; plant T was not sprayed. (-): negative control without DNA template; (+) positive control from DNA. Dashed line is the threshold for positive calls.

Cross-association of BMS-inoculated plants does not compromise the determination of provenance. To simulate cross-association that could occur during food processing, we mixed leaves from plants that were inoculated with unique BMS (**Figure 2-20**). Unlike the other surfaces we inoculated, BMS-inoculated plants did not transfer as easily to objects that came into contact with the plants (**Supp Fig 2-4** vs **Supp Fig 2-14**). While there was detectable transfer between leaves, the amount of transfer still allowed Sanger sequencing to cleanly determine the origin of each leaf (**Figure 2-21** and **Supp Fig 2-14**).
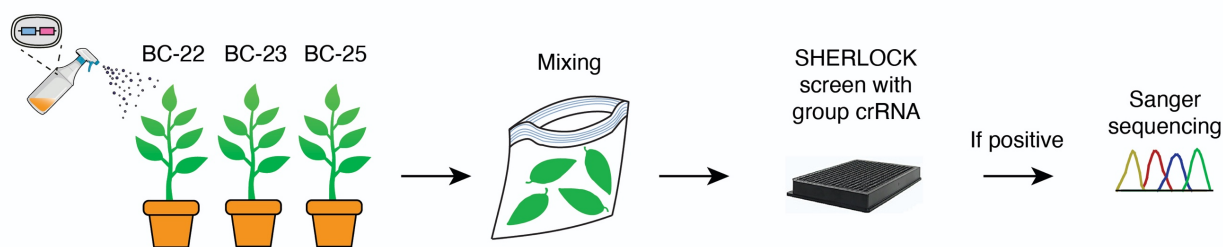
Figure 2-20 Schematic of cross-association experiment.
Leaves from plants inoculated with different BMS were mixed together, SHERLOCK was used to confirm the presence of the BMS, then Sanger sequencing was used to identify the origin of each leaf.
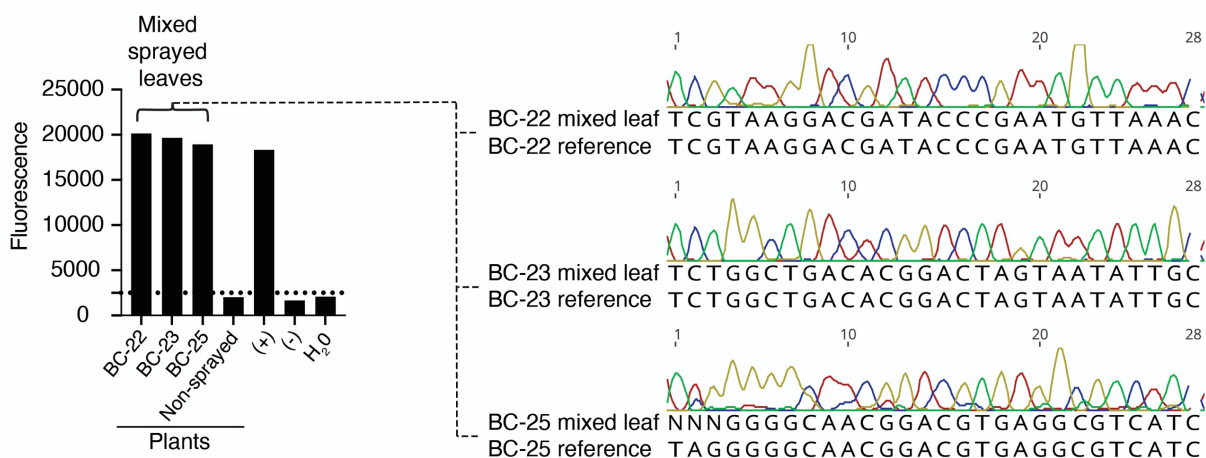
Figure 2-21 BMS detection on cross-associated plants.
(Left) Leaves were screened for presence of BMS by SHERLOCK using a group crRNA, y-axis: endpoint fluorescence values. (+): group 2 positive DNA; (-): group 1 DNA; (H20): water control. (Right) Sanger sequencing identified the plant of origin of the mixed leaves.

Bt could be used to determine food provenance. We used *Bt* spores applied during farming as a surrogate to test whether BMS would persist through conditions of a real-world food supply chain. For plants of known *Bt* inoculation status, we correctly identified all *Bt* positive and negative plants (38 total plants). Furthermore, we detected *Bt* on 10 of 24 store bought produce of *a priori* unknown *Bt* status (**Figure 2-22**).

Test result	Bt positive plants	Bt negative plants	Store-bought produce	Total
Positive	26	0	10	36
Negative	0	12	14	26
Total	26	12	24	62

Figure 2-22 Summary of *Bt* PCR detection results for produce.
Produce known to be inoculated or non-inoculated with *Bt*, and produce purchased from a store with a priori unknown *Bt* status

Strikingly, BMS and *Bt* spores remained detectable even after washing, boiling, frying, and microwaving (**Figure 2-23**), highlighting the potential to determine provenance from cooked foods. These results show the potential to use the BMS system to determine produce provenance.
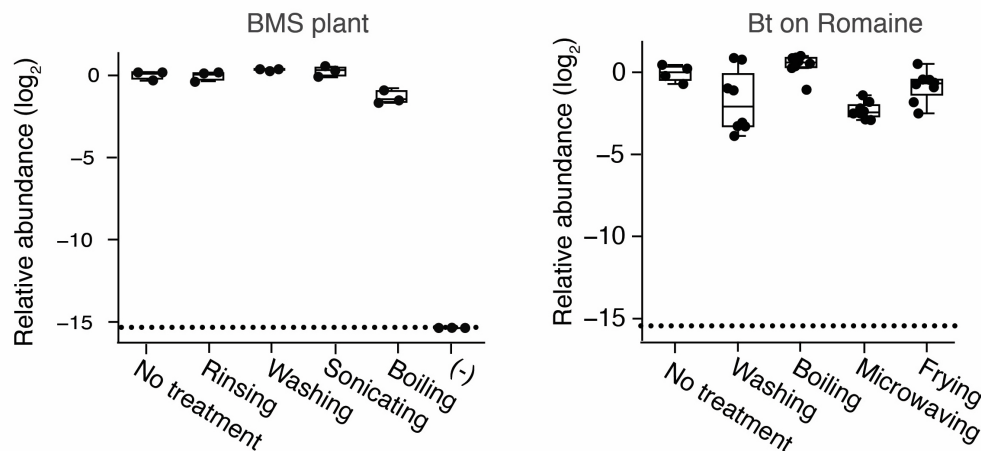
Figure 2-23 BMS and *Bt* spores retention on plants.
(Left) BMS retention was measured by qPCR, normalized to no-treatment. BMS inoculated by spraying on plant samples were retained after a brief rinsing, 1-hour washing, sonicating or boiling (see Methods). (-): H20 negative control. (Right) Romaine lettuce, both positive for *Bt*, were 1-hour washing, boiling, frying, or

microwaving. *Bt* remained detectable after these treatments. (n = 2 technical replicates, n = 4 biological replicates for each treatment, n = 2 biological replicates for non-treated control samples).

Our work shows how rationally engineered microbial spores that can be manufactured in a high-throughput manner provide a new solution to the object provenance problem. We have shown that BMS: 1) persist in the environment, 2) do not spread out of the inoculation area, 3) transfer from soil, sand, wood, and carpet to contacting objects and 4) permit sensitive and rapid readout using laboratory and field-deployable methods. The ability to rapidly label objects and determine their provenance in real-world environments has a broad range of applications across agriculture, commerce, and forensics (*87*). Preliminary data suggests that BMS could work across various environments, though extensive validation in a wider range of real-world conditions is needed. Future iterations of our BMS system could be engineered for limited propagation and actively contained for use in highly trafficked areas. This could also provide time-resolved information about location history, making the BMS system useful for an even wider range of applications.

Chapter 2 Materials and Methods

Barcode generation

A collection of 28-bp DNA barcodes with a hamming distance greater than 5 were bioinformatically generated (**Table S4**). The collection of generated barcodes was screened against GenBank genome data using NCBI BLAST, and any barcodes found to align to genome sequences of *Bacillus subtilis* (*B. subtilis*) or *Saccharomyces cerevisiae* (*S. cerevisiae*) were eliminated from the collection.

Transformation and barcode insertion in bacteria

B. subtilis strains were derived from the wild-type strain 168 and are listed in **Table S1.** Insertion-deletion mutants were from the *Bacillus* knock-out (BKE) collection ((*95*). All BKE mutants were back-crossed twice into *B. subtilis* 168 before assaying and prior to antibiotic cassette removal. Antibiotic cassette removal was performed using a temperature-sensitive plasmid encoding the Cre recombinase (*96*).

DNA barcodes were produced by amplifying 164 bp synthetic megamers (**Table S5**) using oligonucleotide primers oCB034 and oCB035 (**Table S2**) in PCR. The barcode fragments were cloned using standard restriction digest cloning into the plasmid pCB018 (*ycgO*::lox66-kan-lox71), a vector for double-crossover integration at the *ycgO* locus.

Bacterial sporulation

For large scale spore production, *B. subtilis* strains were sporulated in 1 L supplemented Difco Sporulation Medium (DSM) by nutrient exhaustion (*97*) at 37˚C in 4 L flasks. After 36 hours of growth and sporulation, the spores were pellet by centrifugation at 7000 rpm for 30 min, washed 2 times with sterile distilled water, incubated at 80˚C for 40 min to kill non-sporulating cells and then washed 5 times with sterile distilled water. Spores were stored at 4˚C in phosphate-buffered saline.

Evaluating bacterial spore lysis by microscopy

To rapidly assess the efficacy of different spore lysis protocols, we expressed a fluorescent protein in the spore core and monitored release following lysis by fluorescence microscopy. The *mScarlett* gene was PCR amplified with oligonucleotide primers oCB049 and oCB050 (**Table S2**) from plasmid pHCL147 (*98*), and inserted downstream of the strong sporulation promoter P$_{sspB}$ in plasmid pCB137 (*yycR*::P$_{sspB}$-spec), a vector for double-crossover integration at the *yycR* locus.

Spores were immobilized on 2% agarose pads. Fluorescence and phase-contrast microscopy was performed using an Olympus BX61 microscope equipped with an UplanF1 100× phase-contrast objective and a CoolSnapHQ digital camera. Exposure time was 400 ms for mScarlett. Images were analyzed and processed using MetaMorph software.

Transformation and barcode insertion in yeast

Barcodes were introduced into *S. cerevisiae* yeast strain BY4743 with standard lithium acetate chemical transformation with a 15 min heat shock. Following overnight recovery in YPD media (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose), cultures were plated on YPD + G418 to select for transformants. Yeast were transformed with 1 µg of barcode oligos (**Table S5**), and two linearized plasmids: 50 ng of Cas9 plasmid, F48V (2µ-KanR-pRPL18B-Cas9-tPGK1-GapRepair) and 1 µg of gRNA plasmid, F51V, containing a single gRNA targeting HO and 200- to 300-bp sequences homologous to the GapRepair region in F48V. When both are transformed into yeast cells, the two linearized fragments would assemble into a functional plasmid granting G418 resistance. Once the HO locus targeted by the gRNA was replaced with the barcode sequence, the assembled Cas9 + gRNA plasmid was dispensable (*99*). Plasmids were cured by culturing cells in YPD for overnight followed by spreading cells on YPD plates and replica plated to YPD + G418 plates to select for colonies negative for plasmids.

<u>Yeast sporulation</u>

Yeast cells were cultured in 5 mL of YPD medium at 30˚C overnight, then transferred to 1 L of YPD medium and cultured for 24 hours. Cells were pelleted by centrifugation at 3000 g for 3 min and washed twice with sterile distilled water. Finally, cells were resuspended in 500 mL of sporulation medium (10 g/L potassium acetate, 1 g/L yeast extract, and 0.5 g/L dextrose anhydrate) and incubated at room temperature while shaking for 5 days. Presence of spores was confirmed by microscopy at 60x magnification.

Spores were pelleted and the supernatant was carefully removed. The spores were then washed once then resuspended in 25 mL of sterile distilled water and transferred to 50 mL conical tubes. These tubes were boiled at 100˚C for 1 hour to rupture any remaining vegetative cells. After boiling, spores were pelleted, washed twice then resuspended in 25 mL of distilled water.

<u>Production of LsCas13a</u>

LsCas13a was purified as described in (*21*), with some modifications. All buffers were made in UltraPure nuclease-free water and all labware used during purification were cleaned with RNase*Zap* before use. Purification of the expressed LsCas13a protein was performed in batch format using StrepTactin sepharose. The SUMO-protease cleaved LsCas13a was concentrated using an Amicon Ultra-0.5 centrifugal filter with a 100 kDa molecular weight cutoff filter. The protein was concentrated until the sample measured as 2 mg/mL using the BioRad Protein Assay. The LsCas13a was not purified or concentrated further, and instead stored as 2 mg/mL aliquots in lysis buffer supplemented with 1 mM DTT and 5% glycerol. The use of RNase free water for all buffers during the preparation of LaCas13a is critical to achieving low basal activity of LaCas13a. It is suggested that new batches of LaCas13a are tested prior to use to ensure low basal activity in the absence of crRNA.

Recombinase polymerase amplification reaction and primer design

Recombinase Polymerase Amplification (RPA) reactions were performed as described in (*21*). The Twist-Dx Basic kit was used according to the manufacturer's instructions. RPA primers JQ24 and JQ42 (**Table S2**) were used at 480 nM final concentration to amplify a DNA amplicon of 161-bp, containing the T7 promoter and barcode sequences. The T7 promoter sequence was designed in-between the forward RPA primer, JQ42, and the barcode (**Figure 2-2**). RPA reactions were run with 1 µL of template in a total reaction volume of 10 uL for 1-2 hours at 37°C, unless otherwise noted.

SHERLOCK detection reactions and crRNAs

Detection reaction reactions were performed as described (*21*). crRNA preparation was performed as described in (*21*), except the *in-vitro* transcription reaction volume was scaled to 60 µL. All crRNA and barcode sequences used in this study are available in **Table S3,Table S4, Table S5**. BioTek readers were used for measuring fluorescence of reactions (Synergy H1 Plate Reader) at Excitation/Emission = 485/528 nm wavelength for 90 min. Positive threshold cutoff value of 2500 was determined by averaging the values of negative control fluorescence values plus 4σ.

Inoculating surfaces with spores

Spores were diluted to a final concentration ≤1 x 10^8 spores/mL in distilled water in order to reduce the viscosity of the solution. Spores were routinely stored at 4°C for long periods of time, or at room temperature for short periods of time. Diluted spores were sprayed onto surfaces using handheld spray bottles (Fisher Scientific). At this concentration, inoculation had no visible effect on most surfaces tried, though water stains with white residue did appear on hydrophobically-treated wood due to water beading up on the surfaces.

<u>Swab collection and NaOH lysis protocol</u>

Sterile nylon swabs (Becton Dickinson) were dipped into sterile swab solution (0.15 M NaCl + 0.1% Tween-20) and excess liquid was wiped away. The damp swab was rubbed over the object, covering each part of the surface, twice. The tip of the swab was clipped into a microcentrifuge tube, and 200 µL of freshly prepared 200 mM NaOH was pipetted onto the swab. The tube was heated to 95°C for 10 min, then the base was neutralized with 20 µL 2 M HCl, and buffered with 20 µL 10x TE buffer (Tris-HCl 100 mM, EDTA 10 mM, pH 8.0). Lysate samples were optionally purified with 1x AMPure XP bead protocol (Beckman Coulter).

<u>Spore quantitation by qPCR</u>

Quantitative Polymerase Chain Reactions (qPCR) were prepared in 10 µL reactions with SYBR Green I Master Mix (Roche), 1µL of genomic extract as a template, 0.4 mg/mL Bovine Serum Albumin, and 1 µM of each primer (**Table S2**). The reactions were carried out in a LightCycler 480 instrument (Roche) with the following cycling conditions: (i) denaturation, 95°C/10 m (ii) amplification, 45 cycles 95°C/10 s, 60°C/5 s, 72°C/10 s.

<u>Design of ~1 m^2-scale test surfaces and perturbations</u>

We constructed small-scale test surfaces in an incubator to simulate real world conditions in which the Barcoded Microbial Spores (BMS) may be deployed. We assembled 20 test surfaces, from 4 materials (sand, soil, carpet, and wood), divided into control and perturbed conditions (**Table S6**). Each surface was divided into a ~0.2 m x 0.3 m grid, denoting different locations for direct samples and transfer samples each week. The gridded area of each surface was inoculated with different pairs of barcoded strains, using a handheld spray bottles, to a final concentration of ~1.2 x 10^6 *B. subtilis* spores and 3.8 x 10^4 *S. cerevisiae* spores per square cm.

For outdoor conditions, twelve ~0.2 m x 0.3 m trays were filled to a ~2.5 cm depth with either sand or potting mix, and housed on shelved in one of two incubators (Shel Labs) heated to 25˚C. To simulate wind, one incubator was equipped with 140 mm computer fans, one directed at each of the 6 perturbed trays. Simulated rain was applied to perturbed surfaces each week, varying in intensity from handheld spray bottle (200 mL/week) on weeks 1-6 to watering can (500 mL/week) on weeks 7-12. The incubator housing the perturbed surfaces was also programmed to fluctuate in temperature between 25˚C and 35˚C with a period of one week.

For indoor conditions, we cut out 4 sections of carpet and 4 sections of laminate wood flooring and marked out ~0.2 m x 0.3 m sections on each for testing. All 8 surfaces were shelved in an incubator (Shel labs) heated to 25˚C, and humidified to 40-50% RH. To simulate cleaning, perturbed carpet surfaces were cleaned using a handheld vacuum, while perturbed wood surfaces were cleaned with a hand broom.

Sampling from ~1 m^2-scale test surfaces over a three-month period

Each week for a 13-week period, 0.25 g of sand or soil was sampled from each surface from a different adjacent location (~2.5 cm away) on the tray each week using a microcentrifuge tube. Samples were processed using a DNeasy Powersoil kit (Qiagen) to isolate DNA. For carpet and wood samples, we swabbed a different adjacent location on the surface directly using the swab collection and NaOH lysis protocol to generate lysate for qPCR. For all surfaces, 2.5 cm x 5 cm test objects (either rubber or plywood) were used to test transferability of spores; test objects were pressed onto the surface a single time, then processed with the swab collection and NaOH lysis protocol to generate lysate used for qPCR without AMPure XP bead cleanup.

Design of full-scale sandpit and perturbations

A 6 m x 16 m x 0.25 m indoor sandpit was built and equipped with drainage and a slight grade. A 1 m x 6 m section along the top edge was inoculated with *B. subtilis* BC-24 & BC-25

spores with roughly 2.5 x 10^{11} spores each, and *S. cerevisiae* BC-49 & BC-50 spores with roughly 1.25 x 10^{10} spores each. One half of the sand pit was designated for environmental perturbations, with 1 m diameter fans placed at the inoculation end, and hose simulated rain (1.27 cm/week) applied each week.

Sampling from full-scale sandpit over a three-month period

Each week for a 13-week period, 0.25 g of sand was sampled from 20 collection spots in the sandpit (**Figure 2-10**) and processed using the NaOH lysis protocol. Each week, sampling was performed at an adjacent location (<8 cm away) in the region each week using a microcentrifuge tube. For testing transferability, objects (shoes or wood) were pressed onto the surface a single time, then processed with the sample swab and NaOH lysis protocol. In this experiment, 2 μl of lysates were used for RPA reactions followed by SHERLOCK using crRNA 24 and 25 to detect *B. subtilis* BMS*,* and crRNA 49 and 50 to detect *S. cerevisiae* BMS. Threshold was determined as described above, using average values of negative control fluorescence values plus 4σ. A new Cas13a batch was prepared on week 6, so subsequent reactions have different baseline signal and threshold values accordingly.

Sampling from contained outdoor environment

A grass site was inoculated with 2 BMS (BC-14 and BC-15) on August 2019 in a contained outdoor environment at a government research facility. After 5 months of exposure to natural weather (sun, rain, snow, ice, hail, grass-cutting, and wildlife activities), 4 grass samples were sent to us; 1 of which was from the inoculated grass site and the rest were from 12 feet, 24 feet or 100 feet away from the inoculated site. For each grass sample, DNA was isolated using DNeasy PowerSoil Pro kit (Qiagen). 5 samples of DNA were isolated for each of the four grass samples resulting in a total of 20 samples. The extracted DNA was then amplified in 12 μL reactions of RPA using JQ24 and JQ42 primers at 37˚C for 1 hour (**Table S2**). 1 μl of RPA product

was then used for Cas13a detection. Each DNA sample was tested in triplicate for both BC-14 and BC-15 and the reaction was read on BioTek plate reader.

Measuring spore retainment on shoes

24 pairs of shoes were worn in the inoculation region of the ~100 m^2 sandpit (**Figure 2-10**), accumulating spores over a 1-min walking period. These shoes were then divided into 8 groups, and were worn while walking in non-inoculated areas for 0, 1, 5, 15, 30, 60, 120 or 240 min. Non-inoculated areas included city sidewalks, dirt paths, and grassy lawns. Shoes were processed with the swab collection and NaOH lysis protocol, followed by AMPure XP bead purification. 2 µL of purified DNA was used for RPA reactions followed by SHERLOCK using crRNA 24 and 25 to detect *B. subtilis* BMS*,* and crRNA 49 and 50 to detect *S. cerevisiae* BMS. qPCR reactions were prepared in 10 µL volume with PowerUP SYBR Green Master Mix (ThermoFisher Scientific), 2 µL of purified DNA as template, and 2.5 µM of each primer listed in **Table S2**, then ran on a QuantStudio 6 instrument (ThermoFisher Scientific).

Measuring spore re-transfer on shoes

3 sandboxes were inoculated by spraying 2 BMS per sandbox. BMS were transferred to shoes by stepping 5 times with two shoes. One shoe was sampled directly and used as a pre-stepping control; the other shoe was used to walk into 3 different non-inoculated sandboxes. Sand from all the sandboxes were sampled after this initial walk to determine if BMS was transferred to these non-inoculated sandboxes. To simulate re-transfer of spore from sand back onto another shoes, new shoes were stepped five times in the non-inoculated sandboxes. DNA from sand samples were isolated using DNeasy PowerSoil Pro kit (Qiagen). DNA from swabbed shoes were processed with the sample swab and NaOH lysis protocol. BC-1, 2, 23, 25, 90 and 91 qPCR (primers in **Table S2**) were used for qPCR.

Determining object provenance

Grids of varying sizes (averaging 0.25 m^2 per region, averaging 1 m^2 per region) and materials (sand, soil, carpet, wood) were inoculated using approximately 2.5 x 10^{11} *B. subtilis* spores and/or 1.25 x 10^{10} *S. cerevisiae* spores per region. Surfaces were inoculated by spraying then dried for at least 24 hours before shoes and remote-control cars were used as test objects, and were exposed to inoculated surfaces by walking or driving on the test grid surfaces. All surfaces of test objects were swabbed and processed with the NaOH lysis protocol. For 1 unique BMS per sand region, a region was called positive if SHERLOCK reaction was positive for the unique BMS of that region; for 2 unique BMS per sand region, a region was called positive if SHERLOCK reaction was positive for at least 1 of the 2 BMS of that region; for 4 unique BMS per sand region, a region was called positive if SHERLOCK reaction was positive for at least 2 BMS of that region. False positive and false negative rates were calculated using different threshold criteria for a positive call of a region.

Qualitative readout of SHERLOCK using mobile phone camera

To mimic in-field deployment, we set up a portable blue light source, an orange acrylic filter for data collection (**Figure 2-16**). In a dark room, a mobile phone, set at default setting (flash off), was used to photograph the SHERLOCK reaction plate after 30-60 min of reaction.

Barcode identification from a model farm

20 garden pots were filled with potting mix and enclosed in canvas with 12 hours of daily blue light. One seedling was planted in each pot. Temperature was controlled to be around 23˚C. Plants were watered every 2-3 days and exposed to blue light for 12 hours daily. Barcoded *B. subtilis* spores were inoculated by spraying on the plants after the first set of leaves appeared. Inoculation was done for each plant separately once a week for 4 weeks during the growth period. In total, ~10^8-10^9 spores were inoculated onto each plant. One week after the final inoculation,

plant samples were harvested and processed using DNeasy PowerSoil Pro Kit as described above to isolate DNA. Barcode DNA was amplified using BTv2-F and BCv2-R (**Table S2**) using Kapa Biosystems HiFi HotStart ReadyMix, then Sanger sequencing was used to identify the barcode sequences (GENEWIZ).

Barcode identification of co-associated plants

7 leafy plants were bought and each leaf was marked allowing identification of the source plant for each leaf after harvest. For 6 of the 7 plants, each plant was inoculated by spraying once with one unique BMS. One control plant was not inoculated with BMS. In total, ~10^8-10^9 spores were inoculated onto each plant. At 1 week, 4 weeks, and 6 weeks after inoculation, one leaf from each of the 7 plants was harvested, mixed, and shaken together with other leaves for 5 min in a Ziploc bag to simulate co-association of produce during the food supply chain. Mixed leaves were then taken out of the bag and individually processed with DNeasy Powersoil kit (Qiagen) for gDNA extraction. SHERLOCK was used to screen for leaf DNA samples that were positive with group 2 crRNA. For leaf DNA samples that were positive with group 2 crRNA, the amplified DNA was sent for Sanger sequencing for identification of plant provenance.

PCR of *Bacillus thuringiensis* on produce

Around 250 mg of each produce sample was cut into 1-3 mm pieces using a scalpel and then processed with the DNeasy Powersoil Pro kit (Qiagen) to isolate 50 µL of eluted DNA. 1 µL of the DNA was used for PCR with primers BT-1F and BT-1R (**Table S2**) using Phire HotStart II DNA polymerase (ThermoFisher Scientific) with the following cycling conditions: (i) denaturation, 98˚C/30 s (ii) amplification, 36 cycles 98 ˚C/5 s, 60 ˚C/5 s, 72 ˚C/10 s (iii) extension, 72 ˚C/4 min.

Robustness of *Bacillus thuringiensis* on produce

We selected produce that were PCR-positive for *Bacillus thuringiensis*, then treated these samples with various cooking methods: washing, boiling, microwaving, or frying. For washing, the produce pieces were placed in a 50 mL conical tube covered with a screen with tap water running over the sample for 10 min then dried in paper towels. For boiling, a piece of produce was placed in an Eppendorf tube filled with 1 mL of water and placed in a boiling beaker of water for 15 min then dried in paper towels. For microwaving, produce pieces were placed in a petri dish with the cover on at full power for 2 min. For frying, 1 mL of vegetable oil was added to a 250 mL or 400 mL beaker that was pre-heated for 1 min on hotplate set to 350˚C then the piece of produce was added and heated for 1 min with occasional stirring. Around 250 mg of cooked sample was processed with DNeasy PowerSoil Pro Kit (Qiagen). qPCR was performed using primers BT-1F and BT-1R (**Table S2**) and PowerUp SYBR Green Master Mix (ThermoFisher Scientific). The reactions were carried out with the following cycling conditions: (i) denaturation, 95°C/10 m (ii) amplification, 45 cycles 95 °C/10 s, 60°C/5 s, 72°C/10 s.

Sampling and library preparation for microbiome analysis of inoculated surfaces

For incubator-scale experiments simulating outdoor conditions, we sampled 0.25 g of sand or soil each month from each tray, and extracted genomic DNA with a DNeasy Powersoil kit (Qiagen). For each tray, two samples were taken from the same locations each month, one from a region that had been inoculated with BMS the other from a non-inoculated region of the same tray. Sequencing libraries were created in two rounds as described elsewhere (*100*), first targeting the v3-v4 16S rRNA region with primers prCM509 and prCM510 (**Table S2**) using Kapa Biosystems HiFi HotStart ReadyMix with the following cycling conditions: (i) denaturation, 95˚C/5 min (ii) amplification, 20 cycles (or 30 cycles for low-biomass sand samples) 98˚C/20 s, 55˚C/15 s, 72˚C/1 min (iii) extension, 72˚C/5 min. Next, Illumina Nextera XT primers were used to add barcodes using Kapa Biosystems HiFi HotStart ReadyMix with the following cycling conditions:

(i) denaturation, 95˚C/5 min (ii) amplification, 8 cycles 98˚C/20 s, 55˚C/15 s, 72˚C/1 min (iii) extension, 72˚C/10 min. Following each round of PCR, samples were purified using AMPure XP beads (Beckman Coulter). Sequencing was performed using a MiSeq v3 kit to collect 300 bp paired end reads.

Sequencing analysis for 16S metagenomics samples

Sample compositions were determined using QIIME2 v2018.4 (*84*) on the Boston University Shared Computing Cluster. Due to poor read quality, only single end reads (from the v4 region) truncated to 150 bp were analyzed. Sample inference was performed with DADA2 (*101*), then taxonomy was assigned according to the SILVA 132 database to the phylum level for coarse grained analysis, or genus level to determine BMS abundance. We attributed all reads with a genus-level taxonomic assignment of Bacillus to our *B. subtilis* BMS. Weighted UniFrac distance (*102*) calculations for soil samples were calculated from 10000 reads excluding Bacillus reads. The distance between two samples varying only in a single parameter was calculated for all month 2 samples. For example, to determine the effect of inoculation, weighted UniFrac distance was calculated between wet soil +/- inoculation at month 2, and averaged with the distance between dry soil +/- inoculation at month 2, etc.

Supplementary Text

Developing high-throughput methods to screen barcodes and crRNAs

In order to scale the BMS, we devised a facile method to quickly screen a large number of barcodes and crRNAs in parallel to eliminate those with high cross reactivity or background; we performed pooled *n-1 barcode* RPA reactions with a corresponding crRNA and H_2O RPA controls. A pooled *n-1 barcode* RPA reaction denotes a single RPA reaction with all DNA barcodes pooled together except for one barcode that is left out, and this pooled reaction is screened against all crRNAs. Following our initial *in vitro* screen of 22 barcode crRNA pairs, we made 72 additional barcodes and tested all 94 (**Table S5**). Retesting our initial 22 crRNA validated our *n-1 barcode* assay; 19 of the 22 passed while crRNAs 7, 11, and 12 that were just below our cut-off in our pairwise test (see Methods) scored as cross-reactive in the pooled assay. In total, we eliminated 17 of the 94 crRNAs due to high background and 7 additional crRNAs were eliminated due to cross-reactivity with multiple barcodes. Cross-reactivity appeared to be caused by crRNAs instead of barcodes as no barcodes were cross-reactive with all crRNAs (**Figure 2-6**). To perform screening of crRNAs on *n-1* and *n* BMS *in vivo*, one would need to mix spores at equimolar concentration to avoid amplification biases.

Feasibility testing of BMS on incubator-scale test surfaces

Compared to *in vitro* tests, real-world environments often present challenges for enzyme-based analyte detection systems, both by sequestering analyte and inhibiting reactions. Real-world environments also present a challenge to the stability of forensic labels, by degrading or washing away the label over time. We tested the feasibility of our forensic microbial spores by constructing chambers to simulate different real-world environments and perturbations. We chose four material types: sand, soil, carpet, and wood, and housed test surfaces of these materials in modified incubators in order to simulate indoor and outdoor conditions. We also devised perturbations (e.g. rain/wind/cleaning) that might be expected to remove spores from the surface.

Over a 3-month period, qPCR targeted to our forensic microbial spores demonstrated no significant loss of spores over time (**Figure 2-9**). Importantly, the perturbations did not significantly reduce detection compared to control surfaces for any of the 4 materials. Furthermore, in most cases, spores could be transferred to a rubber or wood test object after a single direct exposure to the inoculated surface, and subsequently detected by qPCR.

We note that the diverse range of conditions present in real-world environments is difficult to capture in a simulated incubator experiment. We did not measure or control for all potentially relevant factors that could affect BMS integrity over time: abiotic factors like sunlight radiation, pH, or chemical stresses (e.g. cleaning agents), or biotic factors like consumption or enzymatic degradation by other organisms (*103*). We anticipate that extensive validation in real-world environments will be needed to determine the feasibility of BMS for different applications.

Sensitivity and specificity of PCR-based detection of *Bacillus thuringiensis*

To validate the sensitivity of the primers we used in detecting *Bt*, we first tested a pool of non-*Bt* bacterial gDNA (*Streptomyces Hygroscopicus, S. cerevisiae, B. subtilis, E. coli* and *Pseudomonas*) along with *Bt* gDNA and only *Bt* gave rise to a PCR band. To test the specificity of PCR based *Bt* detection we included 12 negative control produce samples either from personal gardens or a local farm that had not been sprayed with *Bt* and 26 positive controls from a local farm that we know had been sprayed with *Bt*. We then tested produce purchased from the grocery store and found 10 of 24 samples were positive for *Bt* (**Figure 2-22**). To further test the validity of the detected bands from produce, we randomly selected 6 PCR-positive samples and sent their PCR products for Sanger sequencing. The sequences matched the *Bt* cry1A gene and belonged to 3 variants. When the sequences were input into NCBI BLAST against the nr databases, the only organism that showed up in the top 100 hits was *Bt* (not shown). The same primers used in the PCR screens of produce were used in a qPCR assay to assess the level of spores on produce after cooking. To validate our primers used in qPCR, we made a standard cure of purified *Bt*

gDNA, which was sensitive enough to detect the amount of gDNA equivalent to as few as 10 spores.

Outlook on the cost and regulatory challenges to BMS

As BMS are grown in routine bacterial and yeast cell culture techniques, we do not anticipate any major challenges to their production at an industrial scale. A good comparison would be *Bacillus thurigiensis* spores, which are produced industrially for agricultural application at low cost (*104*).

For application in the United States at the time of writing, clearance from the US Environmental Protection Agency was required for release in a contained outdoor environment. Additional regulations may apply in other countries or for certain applications.

Chapter 3: An enhanced isothermal amplification assay for viral detection

Chapter 3 Contributions

The contents of this chapter were from: <u>J. Qian*</u>, S. A. Boswell*, C. Chidley*, Z. Lu*, M. E. Pettit, B. L. Gaudio, J. M. Fajnzylber, R. T. Ingram, R. H. Ward, J. Z. Li, M. Springer, An enhanced isothermal amplification assay for viral detection. *Nat Commun*. 11, 5920 (2020). The paper was adapted and formatted for this dissertation chapter.

This work is the result of a fruitful collaboration with members of the Springer, Sorger and Li labs. JQ, SAB, ZL and MS conceived the study. JQ, SAB, CC, ZL, MEP, and BLG performed most key experiments and data analyses (supervised by MS). More specifically, JQ and CC co-led experiments and analyses for primer screens (Figure 3-1 thru Figure 3-3, Figure 3-5 thru Figure 3-7). CC led experiments and analyses for RT screens (Figure 3-4) JQ, SAB, CC, ZL, MEP, BLG performed limit of detection experiments and analyses (Figure 3-8 thru Figure 3-12). CC and SAB co-led experiments and analyses of eRPA adaptation for VTM and saliva (Figure 3-13 thru Figure 3-21). JQ and SAB co-led experiments and analyses of patients samples (Figure 3-22 thru Figure 3-24). JMF and JZL performed patient collection and initial sample processing. RTI performed bioinformatic analyses. JQ, SAB, CC, ZL, MEP, BLG, RHW, and MS wrote the paper. All authors reviewed the manuscript.

Chapter 3 Acknowledgements

Chapter 3 Abstract

Rapid, inexpensive, robust diagnostics are essential to control the spread of infectious diseases. Current state of the art diagnostics are highly sensitive and specific, but slow, and require expensive equipment. Here we report the development of a molecular diagnostic test for SARS-CoV-2 based on an enhanced recombinase polymerase amplification (eRPA) reaction. eRPA has a detection limit on patient samples down to 5 viral copies, requires minimal instrumentation, and is highly scalable and inexpensive. eRPA does not cross-react with other common coronaviruses, does not require RNA purification, and takes ~45 minutes from sample collection to results. eRPA represents a first step toward at-home SARS-CoV-2 detection and can be adapted to future viruses within days of genomic sequence availability.

Chapter 3 Introduction

SARS-CoV-2 has rapidly spread around the world with serious consequences for human life and the global economy (*105*). In many countries, efforts to contain the virus have been hampered by a lack of adequate testing (*106*). Rapid, inexpensive, and sensitive testing is essential for contact tracing and isolation strategies to be effective (*107*). While numerous different tests exist, the overwhelming global need for testing has led to limitations in both the supplies of reagents, e.g. swabs and purification kits, and instrumentation, e.g. quantitative polymerase chain reaction (qPCR) or ID NOW machines. In most cases, overcoming these limitations would require scaling of supply lines by several orders of magnitude over current production capacities. Therefore, in an effort to avoid overrun health care systems and high death tolls, many countries have resorted to costly lockdowns.

The ability to reopen economies safely depends crucially on the testing capacity available. Efforts to increase testing capacity have included testing from saliva (*108*), using non-standard storage media or dry swabs (*109*), and eliminating the normal RNA purification step from the standard RT-qPCR tests (*110*, *111*). Strategies such as pooling samples followed by detection using traditional or high throughput sequencing approaches have also been proposed as a way to allow significantly more testing at a highly reduced cost (*112*, *113*). In general, such strategies force a trade-off between throughput and sensitivity.

Isothermal amplification technologies have long held promise to offer highly sensitive detection at high throughput, and to allow for widely distributed testing including at-home/point-of-need (PON) tests (*7*, *114*). However, isothermal amplification is plagued by nonspecific amplification events that require secondary amplification and detection steps. These steps add extra complexity to the reactions, removing many of the benefits of the isothermal amplification approach. Many ongoing efforts aim to circumvent these problems for SARS-CoV-2 detection. Most of the approaches developed so far still require an extraction step and/or two amplification

steps to achieve high specificity, or have low sensitivities that give poor concordance with the gold standard RT-qPCR test (*7*).

We set out to determine the underlying reasons for the poor performance of isothermal amplification technologies in viral detection applications. We selected reverse transcription-recombinase polymerase amplification (RT-RPA) as the most promising current technology. RT-RPA is an isothermal amplification method in which the double stranded DNA denaturation and strand invasion that is typically achieved by heat cycling in PCR is instead accomplished by a cocktail of recombinase enzymes, single-stranded binding proteins, and DNA polymerases (*17*). RPA has potential advantages over other isothermal amplification technologies such as loop-mediated isothermal amplification (LAMP) (*43*) as it can be performed near ambient temperature (37-42°C) and is more rapid. While several creative applications of LAMP technologies to diagnose COVID-19, the disease caused by SARS-CoV-2, have recently been developed and show promise (*47*, *115–118*), RT-RPA has been less explored.

Here, we present a method to screen for efficient RPA primers and show that proper RT enzyme selection with the addition of RNase H to a standard RPA reaction enhances detection of viral RNA targets. We show that the enhanced RPA (eRPA) reaction allows for specific detection of SARS-CoV-2 N and S gene down to 5 molecules per reaction. We demonstrate how eRPA can be used to detect SARS-CoV-2 on unextracted samples from saliva or swab transport media. eRPA is validated on clinical samples and gives concordant results with RT-qPCR in all samples above 5 molecules per reaction. These important improvements for RPA represent a step toward at-home SARS-CoV-2 detection using isothermal amplification.

Chapter 3 Results and Discussion

Reverse transcriptase choice can greatly affect RPA efficiency

We designed RPA primers to both the SARS-CoV-2 N gene and S gene (**Figure 3-1**) and quantified the amplification of a RT-RPA assay with ProtoScript II reverse transcriptase by qPCR (**Figure 3-2**). The detection limit of this standard assay was poor, requiring between 100 and 300 RNA molecules for reliable detection (**Figure 3-3**). Some studies have used longer reaction times to partially counteract the poor yield of RT-RPA (*119*), but we set out to determine whether alternative approaches were possible.
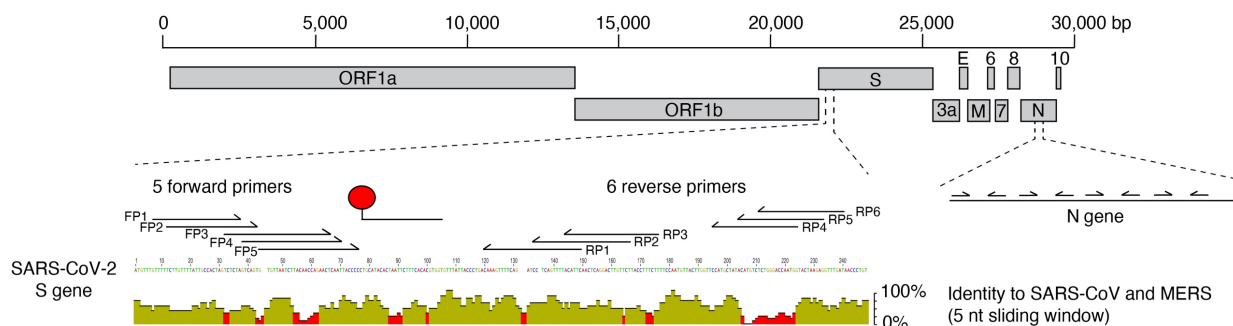
Figure 3-1 SARS-CoV-2 genome and location of regions in the S and N genes targeted by eRPA. Detailed mapping of the binding site of all forward and reverse primers tested in the primer optimization screen and of the biotin hybridization probe was shown for S gene only for display purposes. SARS-CoV-2 was aligned to the closely related SARS-CoV and MERS to identify regions of low homology which were targeted by primers and hybridization probes used in the assay.
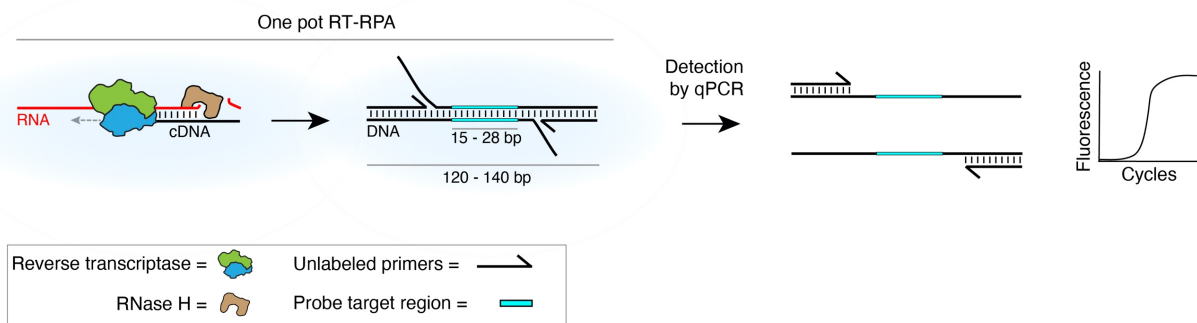
Figure 3-2 Schematic of the workflow used for optimization of eRPA. The cDNA product amplified by recombinase polymerase amplification (RPA) using forward and reverse unlabeled primers was quantified in a subsequent qPCR assay.
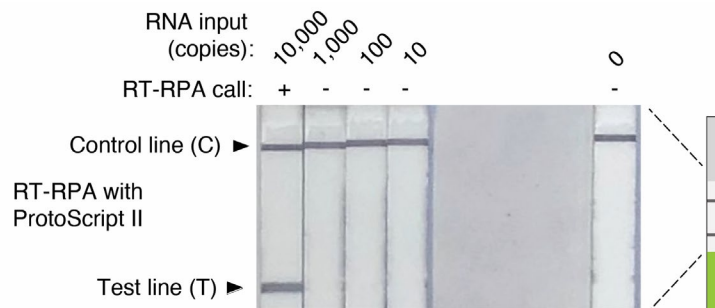
Figure 3-3 Limit of detection of standard RPA with lateral flow strip detection.
In vitro transcribed (IVT) N gene SARS-CoV-2 RNA was amplified by RT-RPA using ProtoScript II and reactions were read out on a lateral flow strip.

We reasoned that the poor performance of RT-RPA could either be due to a specific inhibitor of the RPA reaction from the RT (reverse transcription) reaction or to non-specific primer oligomerization products that could dominate the amplification reaction before the RT reaction occurs. These possibilities are not mutually exclusive. As the RPA reaction is both fast and sensitive when DNA is used as an input (*17*, *120*), we further hypothesized that the product of the RT reaction, i.e. the RNA:DNA hybrid duplex, might inhibit the RPA reaction. We explored methods to circumvent both of these possible problems. To address the problem of kinetic interference by non-specific oligomerization, we screened multiple reverse transcriptases; and to attempt to remove interference from RNA:DNA hybrids, we introduced RNase H, which selectively degrades the RNA strand in these hybrids. Our tests showed that both RT enzyme choice and RNase H addition affected the sensitivity of the RT-RPA reaction, suggesting that both of our hypothesized mechanisms affect RT-RPA efficiency (**Figure 3-4**). The best combination we identified was SuperScript IV reverse transcriptase with RNase H. The magnitude of the effect of the addition of RNase H was correlated with the intrinsic RNase H activity of the RT enzyme. Both SuperScript IV and Maxima H Minus reverse transcriptases are engineered to have minimal RNase H activity in order to improve their processivity, robustness, and synthesis rate (*121*), and we saw the largest effect of RNase H addition in RT-RPA reactions using these enzymes.
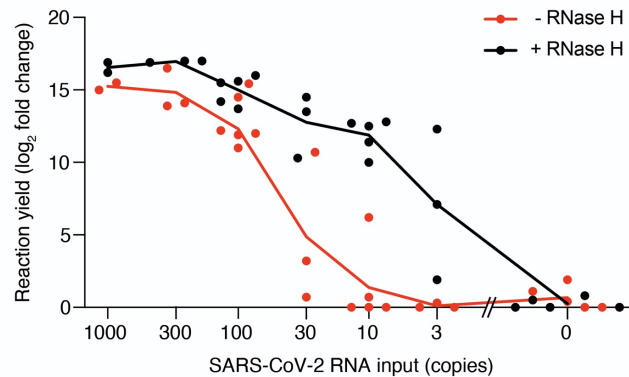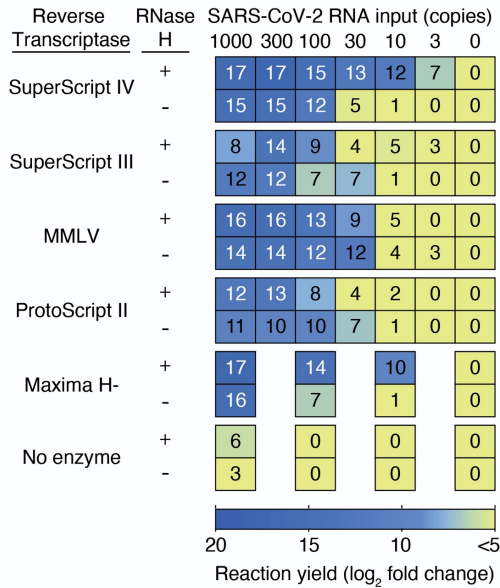
Figure 3-4 Screen for reverse transcriptase (RT) enzyme and effect of RNase H. (Left) SARS-CoV-2 RNA was amplified by RT-recombinase polymerase amplification (RT-RPA) using five different RTs with or without RNase H addition and the yield of each reaction was determined by quantitative PCR (qPCR). At least two biological and two technical replicates were used for each data point; numbers in each square represent mean log2 fold amplification. Samples labeled as zero yielded only non-specific amplification products. (Right) IVT N gene SARS-CoV-2 RNA was amplified by RT-RPA with or without RNase H addition and the yield of each reaction was determined by quantitative PCR. Data represent the average yield of two technical replicates and is staggered on the x axis for visualization purposes.

Reducing non-specific primer reactions increases RT-RPA yield

In addition to the performance issues addressed above, non-specific amplification reactions of primer dimers can greatly inhibit the ability of RPA to amplify the sequence of interest (*114*). To determine whether primer choice affects the importance of these non-specific reactions, we designed forward and reverse primers to both the SARS-CoV-2 N gene and S gene (**Figure 3-1**). Our primer designs avoided regions with strong homology to other coronaviruses including MERS and SARS-CoV, as well as HCoV-229E, HCoV-HKU1, HCoV-NL63, HCoV-OC43, which cause respiratory illnesses such as the common cold. We also avoided regions that have high variability across sequenced SARS-CoV-2 strains. Primer pairs were screened by performing qPCR on diluted RT-RPA products so that both specific and non-specific reaction yield could be

determined, using a modification of a method we previously developed (**Figure 3-5**) (*122*). Many primer pairs gave high levels of amplification at 100 molecules of input RNA, but only a small fraction of those yielded significant amplification products at 10 molecules of input RNA (**Figure 3-5**). We selected primer pairs that gave a high yield of the desired target sequence while minimizing the amount of non-specific amplicons for further assay development.
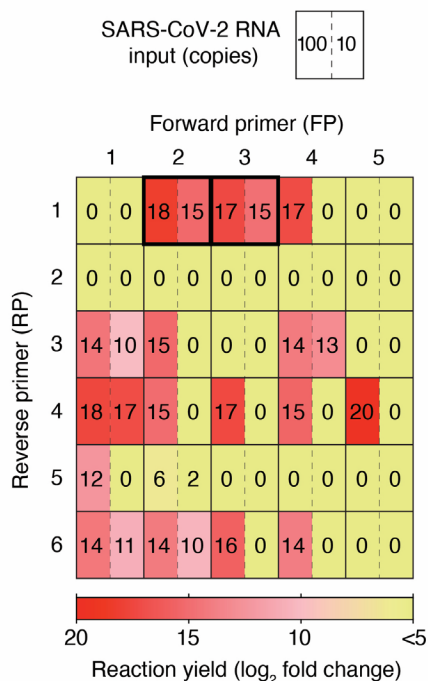
Figure 3-5 Primer optimization screen.
SARS-CoV-2 RNA was amplified by RT-RPA using forward and reverse primers specific to the S gene. The yield of each reaction was determined by qPCR using the same primer pair as for the RT-RPA reaction. Data represent mean log2 fold amplification from 2 technical replicates for each RNA input.

An optimized RT-RPA reaction allows for simple detection

Our optimized RT-RPA assay's product can be hybridized and detected with a commercial lateral flow assay (LFA) without further amplification. LFAs allow accurate read-out by eye by minimally trained personnel, and even opens up the possibility of home-based testing (*123*). We chose to use Milenia Biotec HybriDetect lateral flow test strips that contain a streptavidin band, an anti-Ig band, and carry gold nanoparticle-labeled anti-FAM antibodies for visualization. Based on the results shown in **Figure 3-5**, we selected two primer pairs that amplify part of the S gene,

added a FAM label to the reverse primer, and hybridized the product amplicon to a biotinylated capture probe. Consistent with expectations from qPCR, both primer pairs reproducibly yielded bands with 10 input molecules, and one gave consistent bands with 3 input molecules (**Figure 3-6**).

Primer pair FP2/RP1:

RNA input (copies): 100 30   10   3   1   0

RT-RPA call: + + + + + + + + - - - - - -

C ▶

T ▶

Primer pair FP3/RP1:

RNA input (copies): 100 30 10   3   1   0

RT-RPA call: + + + + + - - - - - -

C ▶

T ▶
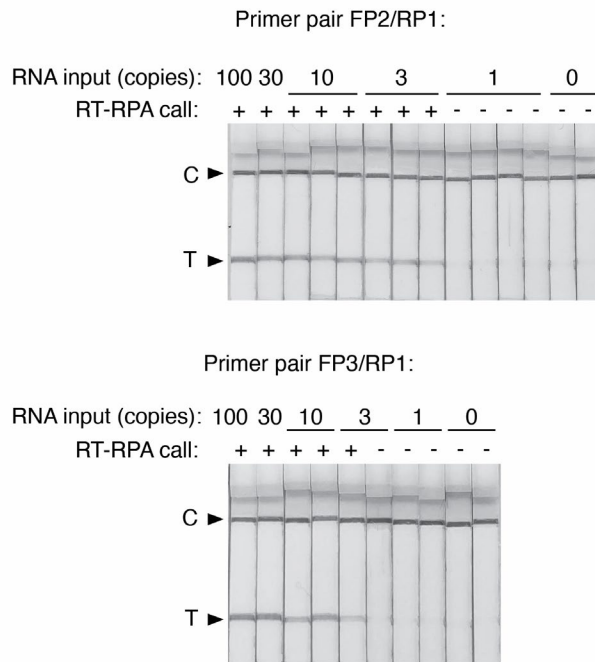
Figure 3-6 Enhanced RT-RPA reactions of SARS-CoV-2 with lateral flow strip readout.
Reactions using primer pairs FP2/FAM-labeled RP1 and FP3/FAM-labeled RP1. All lateral flow strips contain a control (C) and test (T) band.

To allow for increased sample input without sacrificing RPA amplification efficiency, we modified the manufacturer's RPA recipe by using more concentrated reagents (see Materials and Methods section). Altogether, our enhanced recombinase polymerase amplification assay (eRPA) has a detection limit several orders of magnitude better than the manufacturer's RT-RPA assay using ProtoScript II (**Figure 3-3** vs **Figure 3-6, Figure 3-7**).
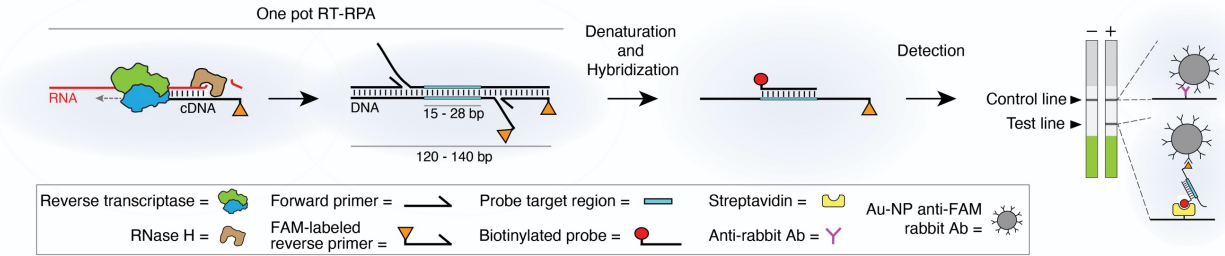
Figure 3-7 Schematic of eRPA.

Viral RNA is first copied to cDNA by RT, then degraded by RNase H. The cDNA product is amplified by RPA using a forward and a FAM-labeled reverse pair of primers specific to the target sequence. The amplified material is then denatured and hybridized to a biotinylated probe. Dual FAM- and biotin-labeled products are detected on lateral flow strips.

eRPA is a sensitive, specific, rapid test for SARS-CoV-2

We designed sensitivity and specificity tests of eRPA assays targeting SARS-CoV-2 N and S genes. The tests were conducted by two independent groups, each of whom randomized the RNA input in a 96-well plate in a checkerboard pattern, then handed the blinded plate to the other group for testing by eRPA (**Figure 3-8**).
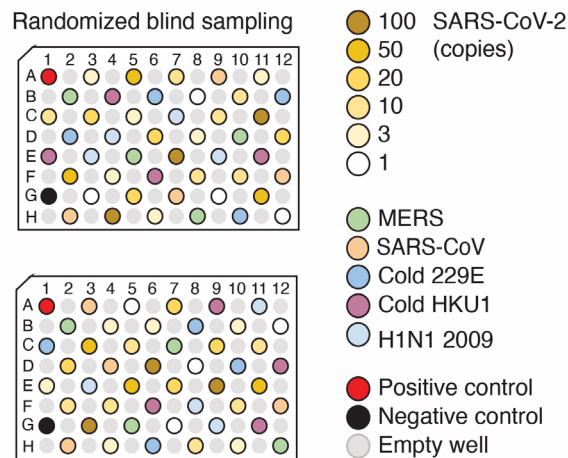
Figure 3-8 Blinded and randomized plate layout used for LOD of eRPA reactions.

For each gene, 52 positive samples were included with a concentration ranging from 100 molecules to 1 molecule of total RNA input (**Figure 3-9**).The titer of the RNA dilutions was confirmed by RT-qPCR (**Figure 3-10, Supp Fig 3-1 Supp Fig 3-2**). Strips were scored at ~20

minutes as this decreases the variability in band intensity that can be observed at low molecule input (**Figure 3-11** and **Supp Fig 3-1**). At or above 10 molecules of RNA input, 31 of 32 N gene samples and 32 of 32 S gene samples were accurately identified as SARS-CoV-2 positive (**Figure 3-9**). Significant detection was achieved even as low as 3 (13 of 24 tests) or 1 (5 of 16 tests) molecules of RNA input. Critically, our assay is also highly specific, showing no cross-reactivity (0 of 80 tests) with 10,000 copies of RNA from other coronaviruses, i.e. MERS, SARS-CoV, CoV-HKU1, or CoV-229E. It also showed no cross-reactivity with the 2009 H1N1 Influenza virus, a respiratory virus with similar initial clinical presentation (**Figure 3-9, Figure 3-11, Supp Fig 3-1**).
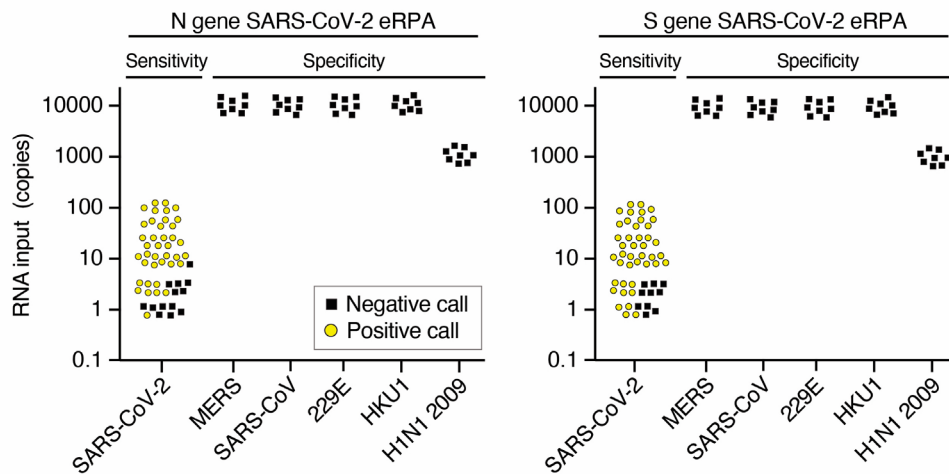
Figure 3-9 eRPA test results for detection of RNA from SARS-CoV-2 or from other viruses.
Synthetic full genome SARS-CoV-2 RNA was amplified by eRPA using primers targeting the N or S gene and reactions were read out by lateral flow strip. The specificity of eRPA was tested against either in vitro transcribed (IVT) RNA of the related viruses MERS and SARS-CoV, or IVT RNA of the common cold coronaviruses HCoV-HKU1 and HCoV-229E, or viral genomic RNA extracted from 2009 H1N1 Influenza. Data points represent positive (yellow circles) or negative (black squares) eRPA tests for each sample tested and are staggered on both axes for visualization.
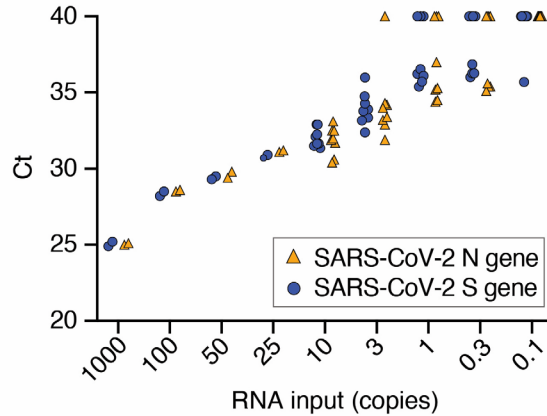
Figure 3-10 RT-qPCR quantification SARS-CoV-2 RNA used as input in the eRPA assay.

The RNA is a synthetic full genome of SARS-CoV-2. Data are Ct values determined using a one-step commercial RT-qPCR assay using primers targeting either the N or S gene of SARS-CoV-2. Data points at Ct=40 represent non-specific or no amplification. N gene (orange triangles) and S gene (blue circles) data are offset on the x-axis for visualization purposes.
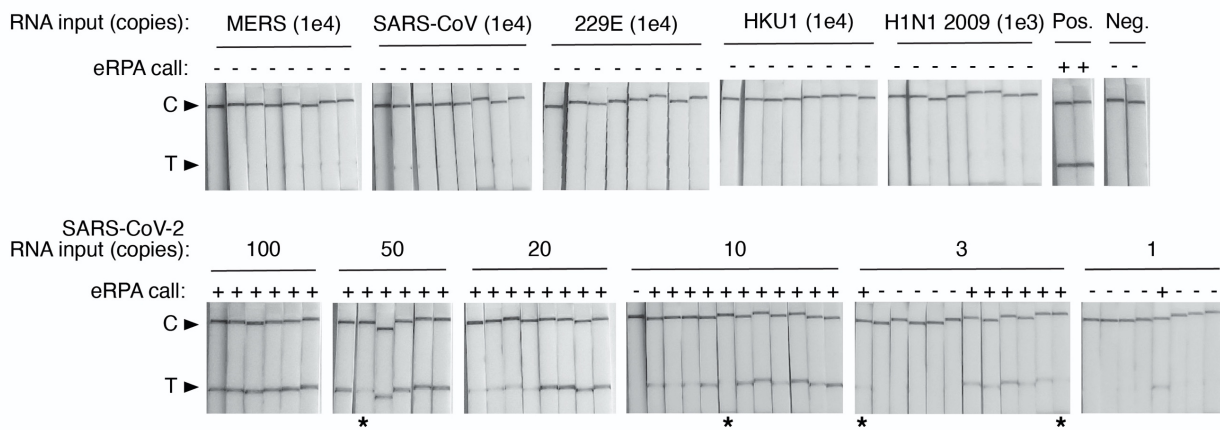
Figure 3-11 Lateral flow strip readouts for all N gene data.

Individual strips (data from Figure 3-9) are labeled with the test call made within 20 mins of detection (positive (+) or negative (-)). The positive (Pos.) eRPA control is 1,000 copies of synthetic full genome SARS-CoV-2 RNA and the negative (Neg.) eRPA control is a water-only input. Images taken for the purpose of display were allowed to dry which reduced the intensity of some weak bands (labeled with asterisks).

For SARS-CoV and MERS, which have the highest target sequence identity with SARS-CoV-2 (91% and 66% respectively), cross-reactivity is dependent on probe choice; we observed cross-reactivity with MERS and SARS-CoV when a longer biotin-probe was used for detection (**Figure 3-12**).
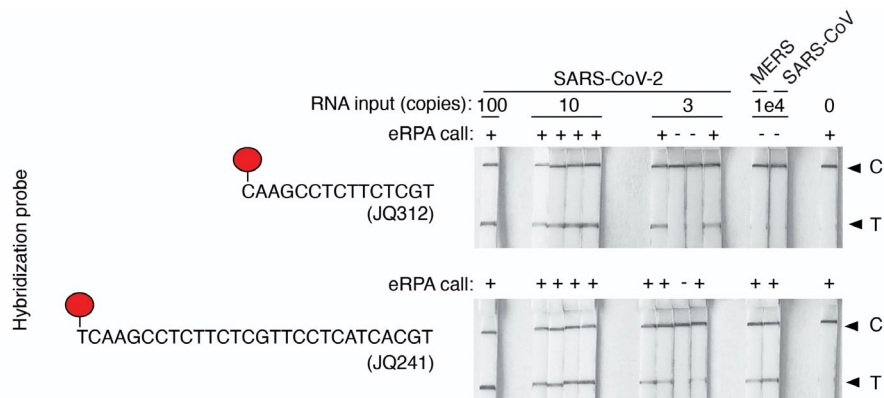
Figure 3-12 Comparison of the specificity and sensitivity of two N-gene probes.
IVT RNA from SARS-CoV-2, MERS, or SARS-CoV was amplified by eRPA. After splitting the reactions in half and hybridizing with a biotinylated probe as shown, each reaction was read out on a lateral flow strip. Individual strips are labeled with the test call made within 20 mins of detection (positive (+) or negative (-)).

We developed an RNA extraction free lysis approach as RNA extraction from clinical samples has become a limiting factor as the global need for SARS-CoV-2 tests has increased. RNA extraction kits are currently hard to obtain, the process of extraction depends on skilled workers, and often involves equipment such as centrifuges. Additionally, extraction free methods would be required for an at-home diagnostic kit. Heat-based lysis has shown promise as a way to rapidly lyse and inactivate viruses for use in diagnostic assays (*23*, *124*). To test whether heat-based sample lysis made viral RNA accessible for eRPA, we initially used packaged reference viral particles, the AccuPlex SARS-CoV-2 verification panel (Seracare). We determined the relationship between temperature and viral lysis by heating for 5 minutes followed by RT-RPA then qPCR for quantification. The replication-deficient virus in the AccuPlex panel is lysed at ~75°C, a temperature that is likely similar to the temperature required to lyse wild-type SARS-CoV-2 (**Figure 3-13**) (*125*).
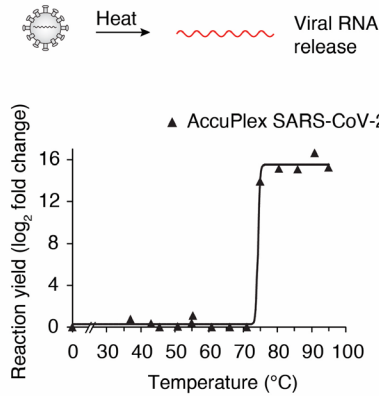
Figure 3-13 Viral particle temperature lysis determination.
AccuPlex packaged SARS-CoV-2 virus was diluted into TCEP buffer and heated for 5 min at the given temperature (see Methods). Released RNA was amplified by eRPA and product formation was quantified by qPCR. Data represent the average of 2 technical replicates.

RNase inhibitors prevent RNA degradation from nasopharyngeal (NP) swabs suspended in viral transport media (VTM), the standard for clinical samples. Our initial experiments using AccuPlex samples or in vitro transcribed (IVT) RNA in VTM yielded poor signal intensities by eRPA. Using RNaseAlert to measure RNase activity, we were surprised to find significant RNase activity in VTM (**Figure 3-14**).
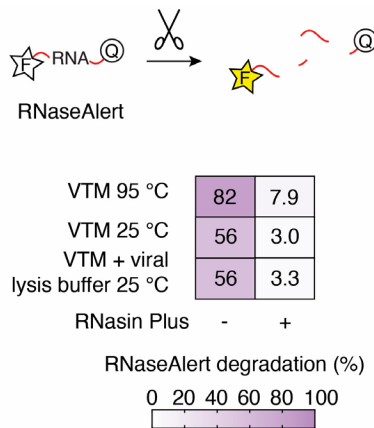
Figure 3-14 Detection of RNase activity of VTM.
RNaseAlert was added to viral transport media (VTM) with or without the addition of RNasin Plus before heating for 5 min at 94°C or added to a 1:1 VTM and viral lysis buffer mix and incubating for 10 min at 25°C. Data represent the average of 4 technical replicates and were determined by normalizing the fluorescence intensity 10 mins after the heating step to a fully degraded control.

In an attempt to address this we tested TCEP, which has been used to inactivate RNases from saliva and urine (*126*). Unfortunately, TCEP and heat treatment of samples with VTM led to gelation, likely due to the presence of gelatin and bovine serum albumin in VTM (**Figure 3-15**).
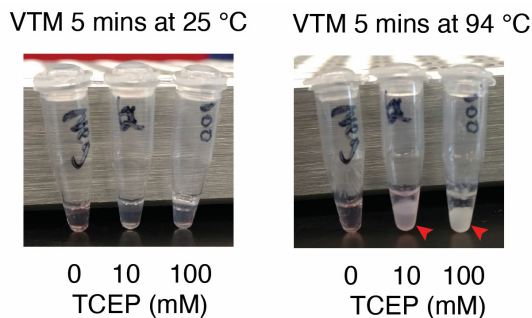
Figure 3-15 Heating VTM in presence of TCEP.
VTM, TCEP and heat leads to the formation of a gelatinous substance (highlighted by red arrowhead).

As an alternative we tested RNasin Plus, a thermostable RNase inhibitor, which significantly protected RNaseAlert from degradation during heat-based lysis in VTM. For future compatibility with PON testing, we also tested a room temperature viral lysis buffer (Intact Genomic FastAmp® Viral and Cell Solution for Covid-19 Testing) and found that RNaseAlert was protected from degradation in the presence of RNasin Plus (**Figure 3-14**).

To confirm that this protocol is effective for patient samples, we tested heat-based lysis of NP-swabs in VTM in the absence and presence of RNasin Plus. We found the addition of RNasin Plus increased RNA yield by ~10-fold and significantly improved the sensitivity of eRPA (**Figure 3 -16**). We also measured the sensitivity of eRPA for AccuPlex viral particles diluted into VTM, PBS, or viral lysis buffer. Sensitivity in these simulated samples, which should more closely reflect what would be achieved from standard samples, was reduced by about 5-fold in comparison to RNA samples in water. Most patients during the initial active phase of infection deliver NP swabs with virus concentrations of $>10^4$ per mL, well within our detection limit (*127–130*).
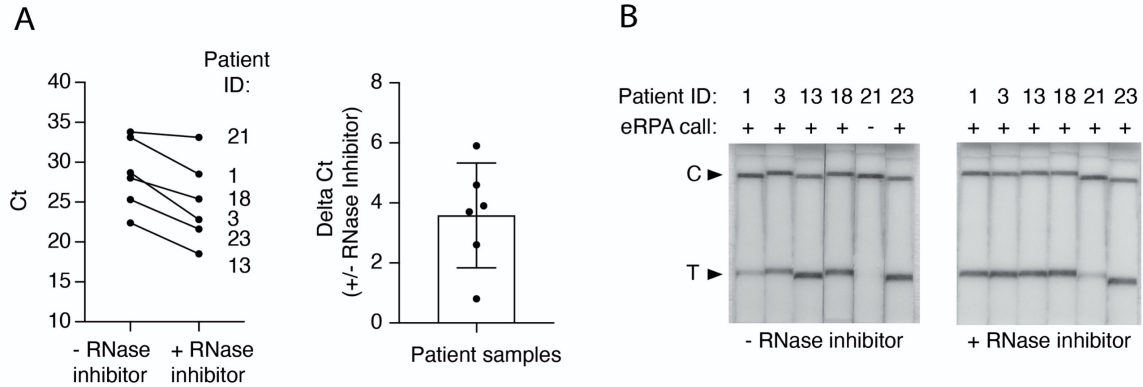
Figure 3 -16 RNase inhibitor increases the RNA titer of patient samples as quantified by RT-qPCR. (A) Unextracted known positive patient samples were heat inactivated for 5 min at 94°C with or without RNasin Plus. Viral RNA was quantified using a commercial one-step RT-qPCR assay. (Left) Ct values for matched samples (n=6 biologically independent samples) with and without RNase inhibitor. (Right) Difference between Ct values in all matched samples (n=6 biologically independent samples) with mean value of 3.6 fold +/- 1.7 SD. Error bars represent +/- 1 standard deviation. (B) Addition of RNase inhibitor to patient samples prior to heat inactivation increases the signal of the eRPA assay. Heat inactivated samples prepared in (A) were tested using eRPA.

Adaptation of eRPA to detect virus in saliva

Given the bottleneck in NP swabs, there has been growing interest in testing saliva instead (*131*). Saliva is a challenging fluid due to the presence of mucins and RNases (*132*, *133*) which can degrade RNA and clog pipettes, leading to a high rate of failed experiments or false negatives. Nevertheless, the viral titer in saliva is sufficient for SARS-CoV-2 detection (*134*). To adapt eRPA to saliva samples, we tested protocols that used TCEP, EDTA, and heat steps (*23*, *124*). The addition of the reducing agent TCEP was critical to decreasing the viscosity of saliva at all temperatures, but the inhibition of RNase activity by TCEP was not complete until the sample was heated above 85°C (**Figure 3-17**).
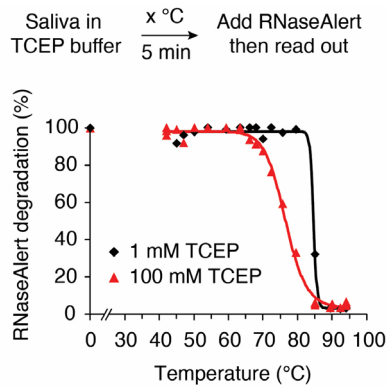
Figure 3-17 Inactivation of RNase activity in saliva by TCEP and heat.

Saliva was first mixed 1:1 with a buffer containing 1 mM (black diamonds) or 100 mM (red triangles) TCEP and heated at the indicated temperature for 5 min. After cooling, RNaseAlert was added and degradation was assessed as in **Figure 3-15**

Because SARS-CoV-2 viral particles lyse at around 75°C (*125*), the period when the sample is being heated from 75°C to 85°C offers a window in which released viral RNA might be degraded during sample preparation. Indeed, RNaseAlert is completely degraded even in the presence of 100 mM TCEP if it is added before the heat inactivation step, but protected if it is added after heat inactivation (**Figure 3-18**).
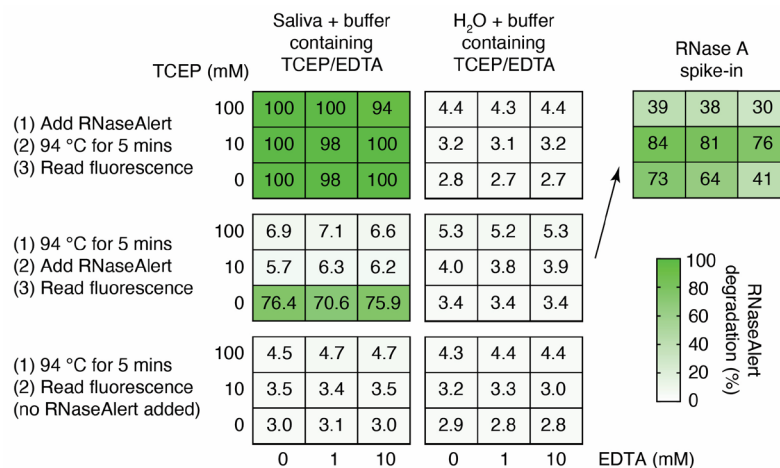
Figure 3-18 TCEP and heat (not EDTA) are required to inactivate the RNase activity in saliva.

Saliva (or water control) was mixed 1:1 with a buffer containing TCEP and EDTA as shown. RNaseAlert was added and the sample was heated as indicated. RNase A was added to a set of water samples post addition of RNaseAlert as control. Data represent the average of 2 technical replicates and was determined from the fluorescence signal 10 mins after the heating step normalized to a fully degraded control.

This distinction is critical as a common method of validating extraction-free saliva sample preparation protocols is to first heat-inactivate the sample and then add viral RNA to determine assay sensitivity (*116*). This method will overestimate assay sensitivity for saliva samples due to the inactivation of salivary RNases. Either murine RNase inhibitor or RNasin Plus helped protect RNA from degradation at low temperatures, with RNasin Plus being more effective at high temperatures (**Figure 3-19**). The combination of RNasin Plus and TCEP protects RNaseAlert from degradation during a heat lysis protocol (**Figure 3-20**).
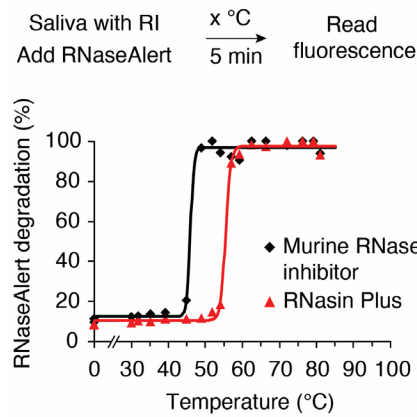
Figure 3-19 RNase inhibitors protect RNA against degradation in saliva at low temperature only.
Saliva was mixed 1:1 with a buffer containing an RNase inhibitor as shown. RNaseAlert was added and the sample heated as indicated. Data represent the fluorescence intensity 10 mins after the heating step normalized to a fully degraded control.
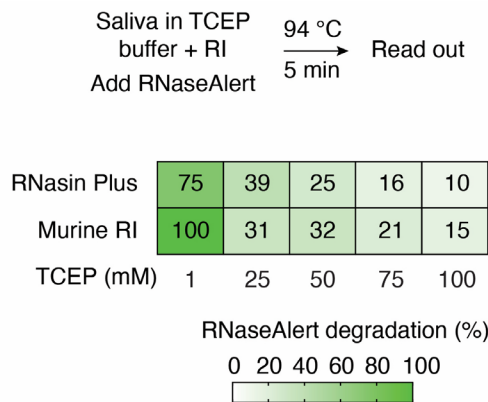
Figure 3-20 The combination of RNase inhibitor and TCEP protects RNA from degradation in saliva. RNaseAlert was added to saliva diluted 1:1 with TCEP buffer containing an RNase inhibitor and treated as shown. RNAseAlert degradation was assessed as in **Figure 3-15**.

Using this protocol (**Figure 3-21A**) we detected SARS-CoV-2 signal in ~70% of samples with 25-100 AccuPlex viral particles in saliva **(Figure 3-21B**), a reduction of 2 to 4-fold compared to the sensitivity of detection in VTM. We saw similar results with IVT SARS-CoV-2 RNA which represents the worst-case scenario for RNA degradation **(Figure 3-21B).** Given that titers of SARS-CoV-2 in saliva are in the range of 10^4 to 10^{10} copies per mL (*134*), this extraction protocol combined with eRPA should be able to identify COVID-19 in a high proportion of infected patients, offering the potential for a high throughput, first pass screening approach that could be important in large-scale testing. We note that we have not yet tested eRPA on actual saliva samples from infected individuals, as these are not readily available.
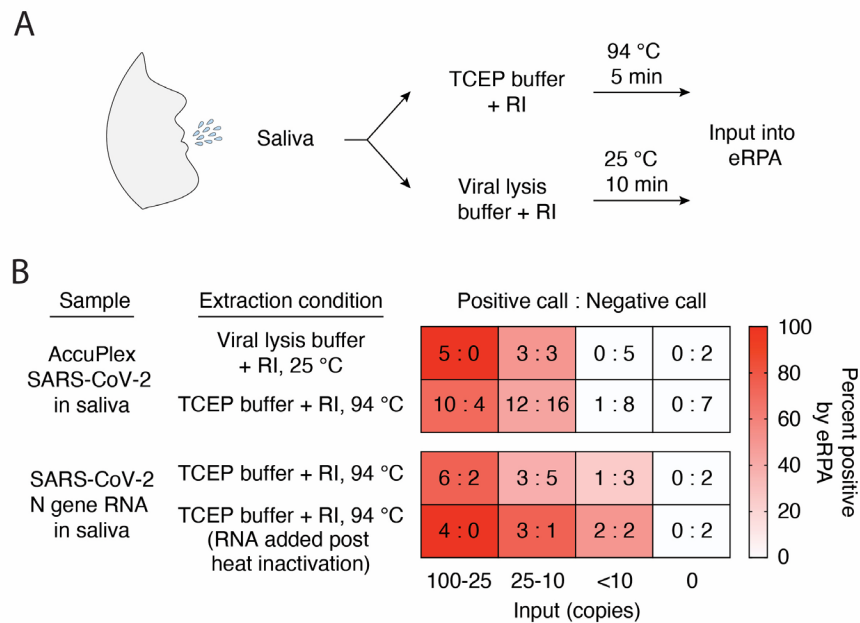
Figure 3-21 Lysis and detection of SARS-CoV-2 N gene from contrived samples.
(A) Schematic of sample processing of patient saliva samples for input into eRPA. (B) Heatmap displaying eRPA test calls for detection of SARS-CoV-2 RNA or AccuPlex packaged virus from saliva treated as displayed in (A). AccuPlex packaged SARS-CoV-2 virus or SARS-CoV-2 N gene IVT RNA were added to saliva and extracted as shown. Values represent the number of positive test calls : number of negative test calls for each condition. Each experiment was repeated three times with similar results.

Comparison of eRPA with RT-qPCR tests on crude clinical samples

To demonstrate that eRPA can detect SARS-CoV-2 in unextracted patient samples, we obtained 30 positive and 21 negative NP swabs from BocaBiolistics (**Table S8**). We processed the samples using our VTM heat lysis method and used this unextracted input, in parallel, in eRPA and in a one-step RT-qPCR assay (**Figure 3-22**).
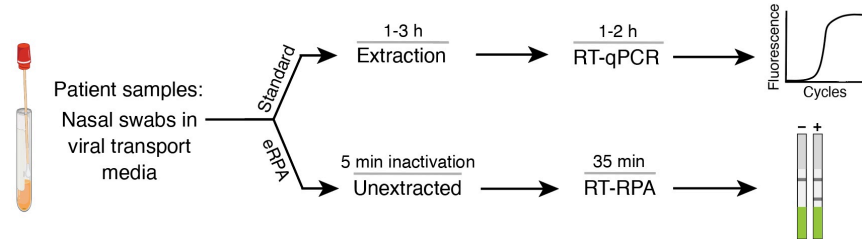
Figure 3-22 Schematic of the workflow for comparing eRPA vs RT-qPCR using patient samples.

We validated our one-step RT-qPCR assay by benchmarking it against the standard CDC N1 RT-qPCR assay (see Methods). All 21 negative samples were negative by eRPA, in duplicate, confirming that the false positive rate for eRPA is very low (**Table S8**). Of the 30 positive samples, 26 had signal by RT-qPCR; 4 may have suffered degradation during transit, see below. For each of the 26 samples that were positive by RT-qPCR, we estimated the number of copies of input RNA into eRPA based on standard curves (**Table S8**). 20 samples had an input of at least 5 molecules of RNA; all 20 of these were positive by eRPA in two repeats (**Supp Fig 3-3**). In 3 samples the input was between 1 and 4 copies; eRPA was positive once, inconclusive once (one positive and one negative of two replicates), and negative once. In 3 samples the input was less than 1 copy, and two of these three samples were inconclusive by RT-qPCR; eRPA was negative twice and inconclusive once (**Figure 3-23**).
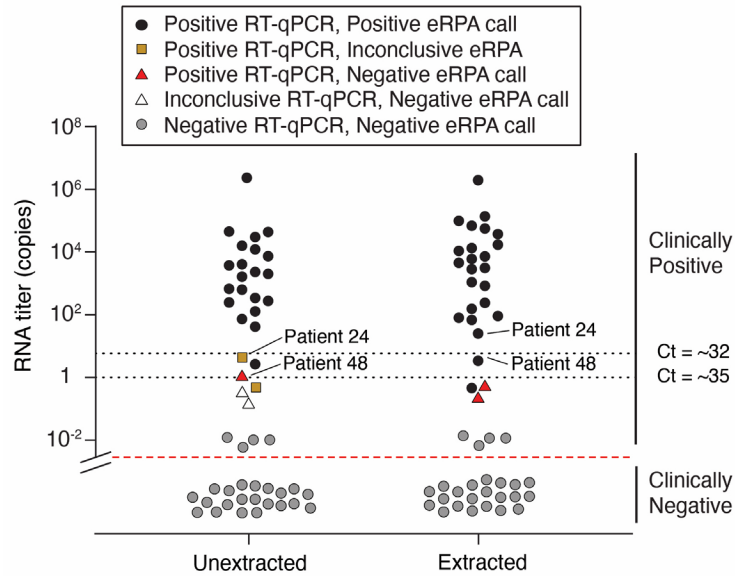
Figure 3-23 Summary of eRPA test results of patient samples and comparison to RT-qPCR.
The y axis represents patient viral titer determined using a commercial one-step RT-qPCR assay from unextracted samples or extracted RNA samples with a standard curve (n=51 biologically independent patient samples).

We repeated the eRPA workflow on the S gene and obtained similar results to the N gene (**Supp Fig 3-4**). In all the patient samples, the RNA copy number of the S gene was on average 4-fold lower than that for the N gene (**Supp Fig 3-5**). This is puzzling, as the detected copy number for both genes are nearly identical from synthetic full genome SARS-CoV-2 RNA, AccuPlex viral particles, and IVT RNA controls (**Supp Fig 3-5**). One possible explanation for this result is that viral transcripts from patient cells present in the NP swab contribute significantly to the total RNA detected. Because the N gene is expressed at up to 10-fold higher levels than the S gene in infected human cells (*135*), if NP swabs collect cells or cellular debris this would bias the observed N gene to S gene copy number ratio. This may be important for other assays as many COVID tests target ORF1ab, which is one of the lower expressed transcripts in human cells.
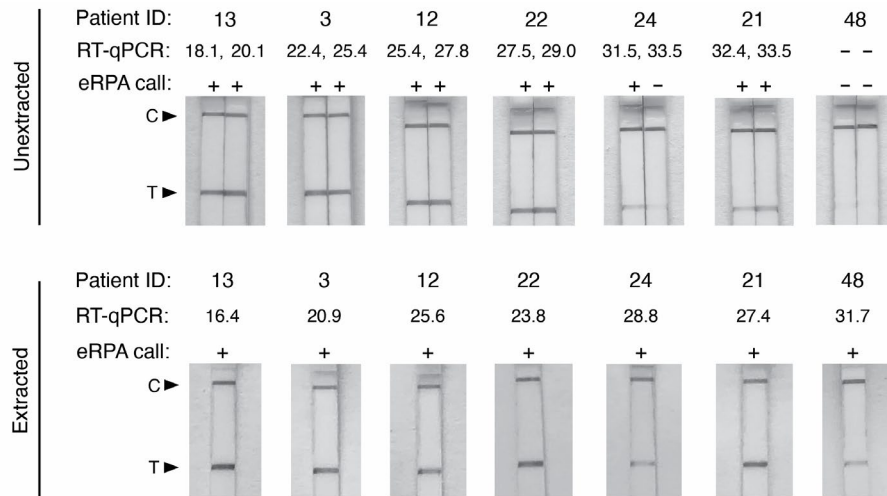
Figure 3-24 Extracted and unextracted lateral flow strip readouts of N gene eRPA tests

Unextracted (top) or extracted (bottom) patient samples (n=7 biologically independent samples) of known infection status. Unextracted patient samples were run in duplicates both by eRPA (calls of positive (+) or negative (-) were made within 20 min of detection) and by one-step RT-qPCR (Ct values shown). RNA was extracted from clinical samples according to standard procedure (see Methods) and was subsequently used as input to eRPA and RT-qPCR.

To validate the results of eRPA and the RT-qPCR results from unextracted samples, we performed RNA extraction on all samples and then repeated RT-qPCR and eRPA (**Figure 3-24**). Overall, RNA extraction increased RNA titer by ~5-fold, matching expectations given that 240 µL of initial sample was concentrated into 50 µL of final volume (**Table S8**). eRPA gives concordant results with RT-qPCR in all extracted samples except those with extremely low titer. Of the 26 extracted samples that were detected as positive by RT-qPCR without extraction, 23 had at least 3 copies of input RNA, and all of these were positively identified by eRPA (**Supp Fig 3-6**). Three samples had <1 copy of input RNA, of which one was identified by eRPA. The four samples with undetectable signal by RT-qPCR before extraction were still negative by both RT-qPCR and eRPA even with extraction. We note that modest changes in sample collection methods would make clinical diagnostics even more sensitive. Currently, NP swabs are typically collected into 3 mL of VTM. Only a small fraction of this volume is used for detection assays. If instead swabs were resuspended in 150-200 µL of liquid, the volume required to cover the head of a swab, the

input would become ~15-20-fold more concentrated without requiring an extraction protocol. This could make the sensitivity of eRPA superior to the current clinical standard of sample extraction followed by RT-qPCR.

Discussion

The eRPA protocol reported here was developed and optimized in just under 3 weeks, with an additional 4 weeks for sample preparation optimization, and patient sample acquisition. In future epidemics and pandemics, this process could be shortened to several days after standardizing sample preparation methods and primer design, and streamlining IRB and COMS approvals. eRPA addresses many of the problems of current SARS-CoV-2 testing methods: it is scalable, compatible with both swabs and saliva samples, can be performed in parallel by minimally trained personnel in low-resource settings (**Supp Fig 3-7**).

A companion manuscript shows that the improvements in RT-RPA we implemented in eRPA also improve other detection approaches such as SHINE (*136*), allowing these assays to become 1-pot, closed-tube, fluorescent readout reactions. eRPA brings us closer to an at-home SARS-CoV-2 test and continued efforts to improve eRPA include: (i) making reactions self-contained and not requiring the addition of reagents before readout, and (ii) making the assay readout quantitative by integration with a lateral flow reader or smartphone app. eRPA is capable of reliably detecting SARS-CoV-2 virus in patient samples that contain as low as 5 viral particles, and is therefore fully adequate to detect infection during the period of peak transmission (*127–130*, *137*, *138*).

Chapter 3 Materials and Methods

<u>RNA template preparation</u>

SARS-CoV-2, SARS-CoV, and MERS N gene containing plasmids were obtained from IDT (2019-nCoV Plasmid Controls). HCoV-229E and HCoV-HKU1 N gene, SARS-CoV-2, SARS-CoV, MERS, HCoV-229E and HCoV-HKU1 S gene were synthesized by Twist Bioscience. All genes were cloned into a T7 promoter expression plasmid. To produce RNA template, in vitro transcription was performed with NxGen® T7 RNA Polymerase (Lucigen #F88904-1) according to the manufacturer's suggested protocol with minor modifications. Final concentrations of the reaction mixture components were 50 units T7 RNA polymerase, 1× reaction buffer, 625 µM NTPs, 10 mM DTT, 500 ng of linearized plasmid template, and RNase-free water to a final volume of 20 µL per reaction. After 10 h at 37 ˚C, 4 units of DNase I (NEB #M0303S) was added and reactions were further incubated for 10 min at 37 ˚C. DNase I was heat inactivated by adding EDTA (5 mM final) and heating at 75 ˚C for 10 min. RNA was purified by RNAClean XP (Beckman Coulter) at 0.6× the volume of the reaction, washed twice with 80% EtOH, then eluted into 20 µL RNase-free water. The size and quality of the RNA product was checked by Bioanalyzer (Agilent) after denaturation at 70 ˚C for 2 min to unfold any RNA structure; all samples were determined to contain the correct RNA product.

To quantify the concentration, each RNA stock was serially diluted in nuclease-free water down to 0.005 molecules/µL and 2 µL was used as input to the RT-qPCR. Poisson distribution based on Ct values of serially diluted samples was used to calculate the stock concentration.

<u>Reverse transcription and RNase H screen</u>

eRPA assay master mixes targeting SARS-CoV-2 N gene were prepared as described below with or without addition of RNase H (NEB) and without reverse transcriptase. The following RT enzymes were added to aliquots of the master mixes: SuperScript III (ThermoFisher), SuperScript IV (ThermoFisher), MMLV (Moloney Murine Leukemia Virus RT, NEB), ProtoScript

II (NEB), Maxima H Minus RT (ThermoFisher). All enzymes were added at 20 U per reaction. N gene IVT RNA diluted with H_2O was used as input to the reactions. Post isothermal amplification, samples were diluted 1:400 in water and products were detected by qPCR using primers JQ289 and JQ223. Specific products were distinguished from primer dimers by analyzing the melting temperature of the qPCR products. The average Ct value of all water control reactions representing primer dimer was used as a baseline to determine the reaction yield (yield = Ct (average water controls) – Ct (specific reaction)).

Primer oligomerization products

Four N gene forward primers (JQ217, CCMS041, CCMS047, and CCMS051) were paired with the reverse primer JQ224. These four primer pairs as well as JQ217 + JQ223 were used in eRPA assays with a water-only sample input. eRPA assays were incubated at 42 ˚C for 10 min. Amplification products were cleaned up using RNA Clean XP (Beckman Coulter) at 2.5× concentration and eluted in 20 μL of nuclease-free water. Purified products were cloned using the Zero Blunt TOPO PCR Cloning Kit (Thermo Fisher Scientific) according to the manufacturer's instructions and Sanger sequenced. The identity of cloned products was determined by first aligning the sequences to the vector sequence using Samtools (v1.9) with an allowed multimapping of k=10000. The sam files were then visualized in IGV (v2.6.2) where the direction, sequence, and copy number of primer oligomers were manually annotated, the direction of the primer indicates a forward or reverse direction with respect to the vector, while the space sequence is the string of nucleotides between two primers. Overlap indicates the primers were overlapping with respect to the vector, "*" indicates there was no space between the primers, and listed nucleotides indicate the sequence between two primers.

SARS-CoV-2 eRPA assay primer sequence alignment

To calculate the percent identity between the SARS-CoV-2 N and S gene primers and the analogous sites in other betacoronaviruses, the RefSeq entries for SARS-CoV-2, SARS-CoV, MERS, HCoV-229E, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 were obtained from NCBI. The sequences were then compared using the EMBL-EBI web tool Clustal Omega to identify indels and mismatches. The subsequences for the forward and reverse primers for both the N gene and the S gene were then located within the SARS-CoV-2 sequence, and the number of mismatches with the antagonist betacoronavirus sequence was tallied. The percent identity was then calculated by dividing the number of matching bases by the length of the primer sequence.

To calculate the number of mismatches between eRPA assay primer and probe sequences and known SARS-CoV-2 variants, the full set of all available SARS-CoV-2 genomes were downloaded from NCBI and were arranged into a single fasta file. This dataset was then converted into a BLAST database using the BLAST+ (v2.6.0) tool and then queried by each of the sequences for the N and S gene eRPA assay. The output from BLAST was then coalesced and filtered to remove any incomplete or partial genomes using R (v.4.0).

Primer screening

Regions of low homology between SARS-CoV-2 and both SARS-CoV and MERS were identified by sequence alignment and were used as target sequences for the biotinylated probe. Unlabeled forward and reverse primers (**Table S7**) were designed to amplify a region of 100-200 nt encompassing the target sequence. Combinations of forward and reverse primers were screened by testing RPA amplification at low RNA input. Reactions were prepared as described below and S gene IVT SARS-CoV-2 RNA was used as input. Post amplification, samples were diluted 1:625 in water and products were detected by qPCR using the same primer pair as for RPA amplification. Specific products were distinguished from primer dimers by analyzing the melting temperature of the qPCR products. Reactions using water only as template were used to

identify primer dimer melting temperatures. All reactions with 10 or 100 copies input but leading only to the formation of primer dimers were labeled as having a reaction yield of zero. The Ct of all reactions leading to specific product formation were converted into an estimated reaction yield by subtracting the raw Ct from the Ct of the lowest specific reaction (for the S gene screen the lowest specific Ct was at 25). Primer pairs with high reaction yields at both 100 and 10 copies input were tested in a secondary screen and the top two primer pairs were subsequently tested by eRPA.

eRPA assay

Isothermal amplification reactions were based on the TwistAmp Basic RPA Kit (TwistDx) with added modifications described below. Each lyophilized pellet was resuspended in a solution of 38 µL rehydration buffer (TwistDx), 1 µL RNase H (5U/µL) (NEB), 0.5 µL SuperScript IV RT (200 U/µL) (ThermoFisher Scientific), and 0.5 µL of forward and reverse primer mix each at 50 µM (N gene, JQ217+JQ235; S gene, CCMS055+CCMS073). This mix was then activated by addition of 1 µL 700 mM magnesium acetate followed by thorough mixing with a pipette. Reactions were prepared by dispensing 8 µL of master mix and 2 µL of input template (RNA, Accuplex virus, or patient samples) per reaction well, mixing the reaction by pipetting, and incubating at 42 ˚C for 25 min. A hybridization mix was prepared by combining 1 µL biotinylated probe at 5 µM (N gene, JQ241 or JQ312; S gene, CCMS069) with 19 µL 10 mM Tris pH 8. 20 µL of hybridization mix was added to each reaction, and samples were heated at 94 ˚C for 3 min followed by a cooling step at room temperature for 3 min. 50 µL of Milenia GenLine Buffer (Milenia Biotec) was added to each reaction, mixed by pipetting, and a lateral flow strip (Milenia HybriDetect) was added. Lateral flow strip signals can be detected and imaged starting 3 min after addition of the strip to the hybridized reaction. Test results were called or imaged within 30 min of strip addition since background bands at the test line can appear over time and low signal test bands can lose intensity as the strip dries.

qPCR and RT-qPCR

SYBR green qPCR reactions were prepared in 10 µL reaction volume using PowerUp SYBR Green PCR Master mix (Thermo Fisher Scientific), 2 µL sample, and 0.4 µM of primers (JQ217 + JQ223 for N gene or CCMS055 + CCMS067 for S gene unless otherwise mentioned). RT-qPCR reactions were prepared in 10 µL reaction volume using the Luna Universal One-Step RT-qPCR kit (NEB), 2 µL sample, and 0.4 µM of primers following the manufacturer's instructions. The CDC one-step RT-qPCR assay used to benchmark our in-house RT-qPCR was performed using the Luna Universal Probe One-step RT-qPCR kit (NEB) and N1 probe/primer mix against SARS-CoV-2 from IDT (2019-nCoV CDC EUA Kit). Reactions were prepared according to the manufacturer's instructions following the CDC protocol. qPCR and RT-qPCR reactions were monitored on either a Bio-Rad C1000 Touch Thermo Cycler (Bio-Rad) or QuantStudio 6 Real Time PCR system (Thermo Fisher Scientific).

Sensitivity and specificity of eRPA with RNA input

Data presented in Figure 2 was generated as a blinded and randomized experiment. Synthetic full genome SARS-CoV-2 RNA (Twist Bioscience) was used as RNA template for eRPA assay on SARS-CoV-2. For the cross-reactivity samples, a single dilution series of RNA input was prepared by mixing at equimolar ratio N and S gene IVT RNA products for each of: SARS-CoV, MERS, HCoV-HKU1, and HCoV-229E. Genomic 2009 H1N1 Influenza (ATCC) was also serially diluted for input to the assay. All dilutions series were made in water and were adjusted for a 2 µL input into the eRPA assay. Two independent groups prepared fully randomized 96-well PCR plates in a checkerboard pattern using those dilutions (**Figure 3-8**). Each group then used the other group's randomized plate as input to eRPA tests targeting either the N gene or the S gene of SARS-CoV-2 performed as described above. All RNA stocks used in these tests were validated by testing dilution series in a one-step RT-qPCR as described above (**Figure 3-10**).

<u>RNaseAlert tests with viral transport media (VTM) and saliva</u>

The RNaseAlert substrate (IDT) was used at 2 µM to assess the RNase activity of saliva and VTM (BD, universal viral transport medium #220220). Fluorescence intensity was determined using an excitation of 485 nm and emission of 528 nm over the course of 10-60 min in a 96-well plate reader (Synergy H1 Plate Reader, BioTek). In general, the degradation of the RNaseAlert substrate was assessed after 10 minutes and fluorescence intensities were averaged over 3 time points and reported normalized to a fully degraded control.

RNasin Plus (Promega) was added to VTM to a final concentration of 1 U/µL and was incubated for 5 min at 25 °C before addition of RNaseAlert. When needed, viral lysis buffer (FastAmp Viral and Cell solution, Intact Genomics) was added 1:1 (v/v) to VTM. TCEP buffer (20 mM Tris pH 8, 10 mM EDTA pH 8, TCEP 1-100 mM) was prepared as a 2× solution and was mixed 1:1 with saliva. RNase inhibitor (RI) was added to 1 U/µL final concentration as shown. For spike-in controls, RNase A (Lucigen) was added to 0.25 µg/µL final concentration. Saliva obtained from two healthy donors was pooled and adjusted to 1 mM TCEP to reduce viscosity. Aliquots of a single pooled sample stored at -20 °C were used for all assays. Saliva samples were obtained from volunteers as approved by the Harvard Medical School Institutional Review Board (IRB 20-0581). Informed written consents were obtained by volunteers.

<u>Virus extraction</u>

The AccuPlex SARS-CoV-2 verification panel (Seracare) containing the N gene, E gene, ORF1a, and RdRp was used as a surrogate to SARS-CoV-2 to optimize the full processing of clinical samples. To determine the temperature lysis of AccuPlex SARS-CoV-2, virus at 1e5 copies/mL was diluted 1:1 in 2× lysis buffer (final: 10 mM Tris HCl pH 8, 5 mM EDTA pH 8, 100 mM TCEP, 1 U/µL RNasin Plus), then incubated for 5 min at a temperature between 55 °C and 95 °C in 5 °C increments. 2 µL of each condition was used as input into eRPA reactions targeting

SARS-CoV-2 N gene as described above. Post amplification, samples were diluted 1:200 in water and products were detected by qPCR using primers JQ289 and JQ223. Specific product formation was distinguished from primer dimer formation by analyzing the melting temperature of the qPCR products and comparison to a water control. The average Ct value of all water control reactions representing primer dimer was used as a baseline to determine the reaction yield (yield = Ct (average water controls) – Ct (specific reaction)).

Detection of AccuPlex SARS-CoV-2 in contrived samples

AccuPlex SARS-CoV-2 was extracted using conditions mimicking patient sample processing. eRPA assays targeting SARS-CoV-2 N gene were performed as above. For extraction in VTM and PBS, AccuPlex SARS-CoV-2 at 100 copies/µL was serially diluted 1:1 (v/v) in either VTM or PBS containing a final concentration of 1 U/µL RNasin Plus. After heating at 94 ˚C for 5 min, samples were kept on ice before being used as input into eRPA. For extraction in viral lysis buffer at 25 ˚C, AccuPlex SARS-CoV-2 at 100 copies/µL was serially diluted 1:1 (v/v) in viral lysis buffer (FastAmp Viral and Cell solution (Intact Genomics)) adjusted with RNasin Plus to 1 U/µL. After 10 min at 25 ˚C, samples were kept on ice before being used as input into eRPA. For extraction of virus in samples containing saliva, 2 vol of AccuPlex SARS-CoV-2 virus at 100 copies/µL was mixed with 1 vol of pooled saliva and 1 vol of 4× TCEP buffer + RI. TCEP buffer + RI was prepared such that final buffer concentrations in the sample were 10 mM Tris HCl pH 8, 5 mM EDTA pH 8, 100 mM TCEP and 1 U/µL RNasin Plus. Lower input samples were prepared by serial dilution with 1:1 (v/v) saliva in 2× TCEP buffer. After heating at 94 ˚C for 5 min, 1/10 vol of 1M H_2O_2 was added and samples were incubated at 25 ˚C for 10 min. Saliva samples were diluted 1:1 with water and kept on ice before being used as input into eRPA. For extraction of virus from saliva with viral lysis buffer, 1 vol of AccuPlex SARS-CoV-2 virus at 100 copies/µL was mixed with 1 vol of pooled saliva and 2 vol of viral lysis buffer adjusted to 2 U/µL RNasin Plus. Lower input samples were prepared by serial dilution with 1:3 (v/v) viral lysis buffer + RI mixed with saliva. For

samples with SARS-CoV-2 RNA, saliva was mixed 1:1 with 2× TCEP buffer + RI. After 5 mins at 25 ˚C, N gene IVT SARS-CoV-2 RNA was spiked into saliva in TCEP buffer and lower input samples were prepared by serial dilution on ice. After heating at 94 ˚C for 5 min, 1/10 vol of 1 M H_2O_2 was added and samples were incubated at 25 ˚C for 10 min. Samples were diluted 1:1 with water and kept on ice before being used as input into eRPA. For RNA added post heat inactivation, a similar protocol was followed using saliva mixed 1:1 with 2× TCEP buffer + RI that was pre-incubated for 5 min at 94 ˚C.

Clinical samples

A cohort of nasal swab patient samples was purchased from BocaBiolistics, FL containing 30 SARS-CoV-2 positive samples and 21 SARS-CoV-2 negative samples. Samples were thawed on ice and 40 µL aliquots were made and subsequently stored at -80 ˚C. At the time of testing, sample aliquots were thawed and RNasin Plus was added to a final concentration of 1 U/µL. The samples were placed on a heat block set to 99 ˚C for 5 min for virus inactivation and lysis. After cooling, samples were spun down and transferred to a 96-well DNA LoBind plate (Eppendorf). 2 µL of the inactivated sample was used as input into eRPA or into RT-qPCR reactions targeting both the N and S gene of SARS-CoV-2. GAPDH was used as a control in RT-qPCR reactions. All patient sample tests included a positive control consisting in 100 copies of synthetic full genome SARS-CoV-2 RNA (Twist Bioscience) and a water only negative control.

Standard RNA extraction from clinical samples

Virions were pelleted by centrifugation at approximately 21,000 ×g for 2 hours at 4 ˚C. The supernatant was removed and 750 µL of TRIzol-LS™ Reagent (ThermoFisher) was added to the pellets and then incubated on ice for 10 min. Following incubation, 200 µL of chloroform (MilliporeSigma) was added, vortexed, and incubated on ice for 2 min. Phases were separated by centrifugation at 21,000 ×g for 15 min at 4 ˚C, and subsequently the aqueous layer was

removed and treated with 1 vol isopropanol (Sigma). GlycoBlue™ Coprecipitant (15 mg/mL) (ThermoFisher) and 100 µL 3M Sodium Acetate (Life Technology) were added to each sample and incubated on dry ice until frozen. RNA was pelleted by centrifugation at 21,000 ×g for 45 min at 4 ˚C. The supernatant was discarded and the RNA pellet was washed with cold 70% ethanol. RNA was eluted in 50 µL of DEPC-treated water (ThermoFisher).

Quantitative SARS-CoV-2 RT-qPCR Assay

Levels of SARS-CoV-2 RNA in extracted samples were detected using the US CDC 2019-nCoV_N1 primers and probe set. Each reaction contained extracted RNA, 1× TaqPath™ 1-Step RT-qPCR Master Mix, CG (ThermoFisher), 500 nM of each the forward and reverse primers, and 125 nM of probe. Viral copy numbers were quantified using N1 qPCR standards to generate a standard curve. The assay was run in triplicate for each sample and two no template control (NTC) wells were included to confirm there was no contamination. Quantification of the Importin-8 (IPO8) housekeeping gene RNA level was performed to determine the quality of sample collection. An internal virion control (RCAS) was spiked into each sample and quantified to determine the efficiency of RNA extraction and qPCR amplification.

In-house RT-qPCR data was converted from Ct values to copies/mL by direct comparison to the CDC RT-qPCR quantitation. In short, the Ct values from the in-house RT-qPCR were plotted against the CDC RT-qPCR Ct values which yielded a linear relationship, $R^2>0.99$, with a slope within error of 1, confirming that the amplification dynamics of both primer sets were similar. The relationship was then re-fit with the slope set to 1 which yielded a line, $R^2>0.99$, with an intercept of between 33 and 34 (95% confidence interval). This fit was then used to directly convert Ct from the in-house qPCR to viral copies/µL.

Chapter 4: A quantitative isothermal diagnostic assay utilizing competitive amplification

Chapter 4 Contributions

The contents of this chapter were from: C.P. Mancuso, Z. Lu, <u>J Qian</u>, S. A. Boswell, M. Springer, A Semi-Quantitative Isothermal Diagnostic Assay Utilizing Competitive Amplification. *Analytical Chemistry. 93, 27, 9541–9548* (2021). The paper was adapted and formatted for this dissertation chapter.

This work is the result of a fruitful collaboration with members of the Springer lab. ZL, JQ, SAB, and MS conceived the study. CPM, ZL, JQ, and SAB performed the key experiments. J.Q. performed the sequencing analysis. CPM made all the figures. More specifically, ZL led proof-of-concept competitive RPA experiments and analyses in **Figure 4-1** and related supplemental figures. JQ led RPA sequence variation experiments and bioinformatics analysis in **Figure 4-2** and related supplemental figures. CPM led in-vitro competitive RPA experiments and analyses in **Figure 4-3** and **Figure 4-4** and related supplemental figures. JQ and CPM co-led competitive RPA experiments on patient samples in **Figure 4-5**. All authors wrote and reviewed the manuscript.

Chapter 4 Acknowledgements

Chapter 4 Abstract

Quantitative diagnostics that are rapid, inexpensive, sensitive, robust, and field deployable are needed to contain the spread of infectious diseases and inform treatment strategies. While current gold-standard techniques are highly sensitive and quantitative, they are slow and require expensive equipment. Conversely, current rapid field-deployable assays available provide essentially binary information about presence of target analyte, not a quantitative measure of concentration. Here we report the development of molecular diagnostic test (qRPA) that utilizes competitive amplification during a recombinase polymerase amplification (RPA) assay to provide quantitative information on a target nucleic acid. We demonstrate that qRPA can quantify DNA, RNA, and viral titers in HIV and COVID-19 patient samples and that it is more robust to environmental perturbations than traditional RPA. These features make qRPA potentially useful for at-home testing to monitor the progress of viral infections or other diseases.

Chapter 4 Introduction

Rapid, inexpensive, and sensitive testing is critical for controlling the spread of diseases with pandemic potential (*1*). For example, as of 2019, 1 in 5 individuals with HIV do not know they are infected (*139*), and in many countries efforts to contain SARS-CoV-2 have been hampered by a lack of adequate testing (*140*). Rapid diagnostic testing with quantitative outputs can lead to more timely treatment and better knowledge of disease progression, leading to improved patient outcomes (*1*). Nucleic acid amplification techniques (NAATs) are the gold standard for identifying infectious diseases (*114*). To enable comprehensive diagnoses, NAATs need to provide specificity, sensitivity, and quantitation. Quantitative Polymerase Chain Reaction (qPCR) is the most commonly used NAAT, but requires expensive and complex instrumentation that cannot realistically be used outside the laboratory. For this reason, researchers have turned to isothermal nucleic acid amplification techniques (INAATs) in an attempt to create a field-deployable diagnostic.

Developing a field-deployable diagnostic is especially challenging if quantitative results are required. In many cases a quality plus/minus result is sufficient, but in other cases viral titer is a key factor in disease treatment. Current field-deployable INAATs are not quantitative. The products of INAATs (such as LAMP (*43*), RPA (*17*), HAD (*50*), and NASBA (*53*)) are generally interpreted in a non-specific manner by monitoring the amount of amplification (using DNA-intercalating or pH-sensitive dyes) or in a sequence-specific manner by detecting amplification products using lateral flow assays (*123*), molecular beacons (*141*), or Cas13-mediated fluorescence (*142*). Because these reactions all proceed to saturation, the endpoint product levels do not depend on the input concentration. Furthermore, time course or "time-to-positive" measurements of these techniques are not robust due to the rapid speed of amplification and the sensitivity of the amplification rate to perturbations (*14*). Finally, due to the lower reaction

temperatures of many INAATs (*114*), non-specific products are produced at higher rates than in qPCR, which affects sensitivity and further limits quantitation.

To overcome these challenges, we developed an assay that is both field-deployable and quantitative by harnessing competitive amplification in a recombinase polymerase amplification (RPA) assay. We selected RPA as the basis of our assay, as it can be performed near ambient temperature (37–42 °C). In RPA, a cocktail of recombinase enzymes, single-stranded binding proteins, and DNA polymerases carry out cyclic strand invasion and polymerization to amplify DNA (*17*). We previously developed an enhanced reverse transcription (eRPA) assay to detect SARS-CoV-2 RNA directly from patient samples with high sensitivity and specificity (*120*). In this new assay, which we term Quantitative Recombinase Polymerase Amplification (qRPA), we introduce reference molecules that are competitively amplified alongside the target (**Figure 4-1A**). Both amplicons are detected separately in a sequence-specific lateral flow assay for quantitative endpoint detection. We demonstrate the efficacy of qRPA for DNA and RNA inputs as well as clinical samples from HIV and COVID-19 patients.

Chapter 4 Results and Discussion

As noted above, endpoint INAATs are not quantitative for two reasons. First, in RPA, amplification continues until reaction components are exhausted, leading to signal saturation. Saturation is reached regardless of the amount of input material and thus input concentration information is largely lost (**Figure 4-1B**). Second, INAATs are prone to non-specific product formation (*114*). The specific amplification is dependent on target level, but the non-specific amplification is dependent on primer levels and non-specific template formation. Thus, when the input level of target is low, non-specific side reactions can out-compete the on-target reaction, consuming significant amounts of the reaction components and reducing the amount of target amplicon produced (**Figure 4-1B**). We reasoned that RPA could be made quantitative by harnessing competition during amplification. Non-specific side reactions cause a decrease in endpoint target amplicon levels, so that the correlation between target input levels and amplicon level is modest at assay completion (**Figure 4-1B**). Rather than relying on stochastic and uncontrolled side reactions, we include a reference molecule in the RPA assay; we refer to this modified RPA reaction as qRPA. In theory, the ratio of the target and reference sequence should stay constant throughout the RPA reaction as long as both the target and reference sequence have the same amplification kinetics (**Figure 4-1C**). The easiest way to achieve this is to have both the target and reference molecules amplified by the same RPA primer set (**Figure 4-1A**).
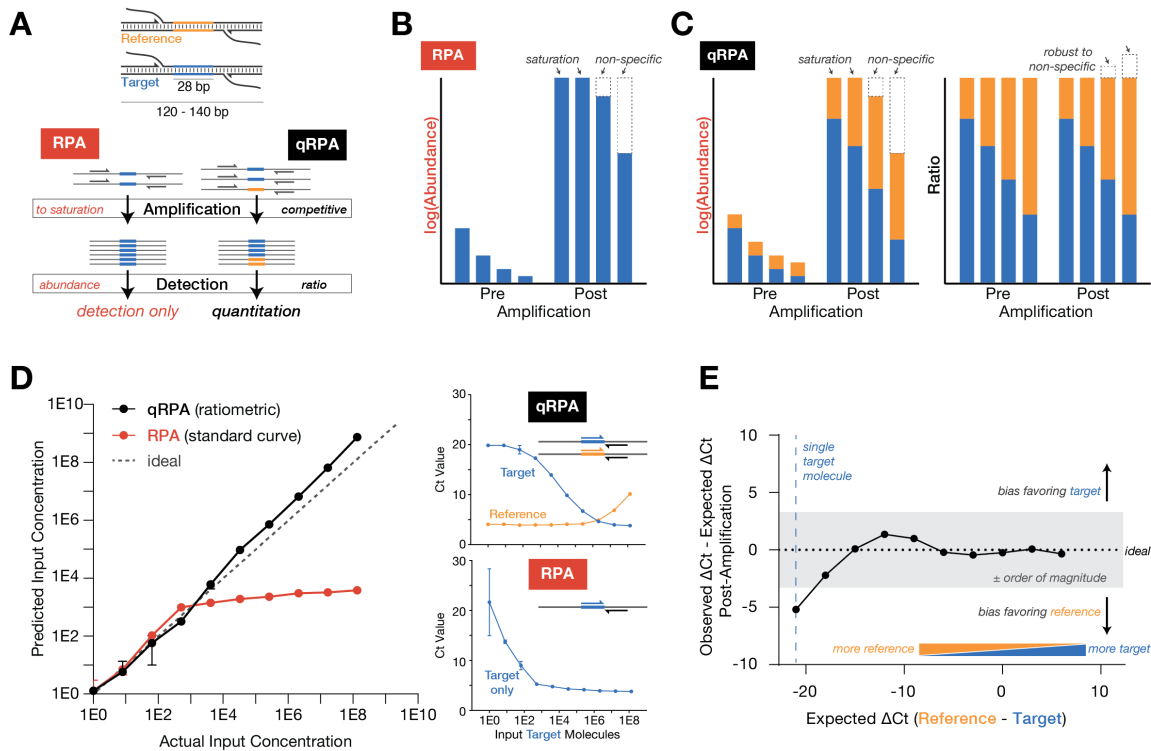
Figure 4-1 Competitive amplification in RPA can be used to maintain target concentration information. (A) In qRPA, addition of a reference molecule at a known concentration can be used to infer target concentration by measuring the ratio of target to reference after amplification. (B) RPA product levels saturate for high input concentrations and are affected by non-specific product formation, depicted by cartoon. (C) In qRPA, though product abundances are still subject to saturation and non-specific products, the ratio is robustly remained and can be used for quantification, depicted by cartoon. (D) Prediction of input concentration by endpoint measurements of RPA and qRPA. qPCR was used to quantify target or reference sequences after amplification by RPA or qRPA. The threshold cycle Ct values (right) were used to calculate a predicted input concentration either by fitting to a standard curve (RPA) or calculating the target to reference ratio (qRPA). (E) Bland-Altman plot of qRPA outputs. The qRPA data was replotted to show the expected and observed ΔCt, which indicates the target/reference ratio. Points are plotted against a horizontal line representing an ideal assay which retains ratiometric information perfectly. In each panel, each point depicts the mean of 3 replicate RPA or qRPA reactions, error bars depict standard deviation.

To compare the performance of RPA and qRPA, we performed both reactions on varying amounts of plasmid DNA containing a synthetic 28-bp barcode target (*122*). For the qRPA reaction, we also included a fixed concentration of a reference plasmid that was identical except in the synthetic barcode sequence. We quantified the products by qPCR using a barcode-specific

primer paired with one universal primer (in a separate qPCR reaction for each barcode) (**Figure 4-1D,** right). In RPA, as the target input reaches single molecule levels there is great variation in the amount of amplification, likely due to the stochastic difference in timing of the initial production of non-specific products (**Figure 4-1D,** right). Amplification appears to be less variable in the qRPA reactions.

We assessed the quantitative capabilities of RPA and qRPA by using endpoint measurements as predictors of the input concentration. We calculated a predicted input concentration for each assay using the cycle threshold values (Ct), which we either fit to a standard curve (RPA) or calculated using the target to reference ratio (qRPA) (**Figure 4-1D,** left). While the RPA assay saturates at high input concentrations, qRPA remains quantitative over a wide range of target input concentrations. In a Bland-Altman plot, if the assay maintains the ratio perfectly, the difference between the observed and expected ΔCt (Cycle threshold of the reference minus the cycle threshold of the target) would be a horizontal line. Indeed, we find that the ratio between target and reference is maintained during amplification for a wide range ($>10^7$ fold difference) of starting target/reference ratios (**Figure 4-1E**). This confirms that input concentration may be inferred from the known amount of reference added and the ratio between target and reference molecules after amplification.

We hypothesized that qRPA should be more robust to environmental perturbation than RPA, since if the amplification kinetics of the target and reference sequences are identical, the ratio of the target and reference amplicons should be maintained regardless of reaction efficiency. To test this prediction, we performed RPA under several environmental perturbations that might be faced in real-world settings, including different temperature, pH, and salt concentration. We quantified the post-amplification products by qPCR as described above. For reactions at pH 2.4, pH 11.4, or with 60mM additional NaCl, the amount of amplification product produced was less predictive of the input level than under standard conditions (**Supp Fig 4-1**); since these effects are mild, we can infer that RPA efficiency is little affected by these conditions. Nevertheless, the

ratiometric product levels are robust to all three variant conditions, so qRPA remains quantitative even when amplification efficiencies are affected by environmental perturbations (**Supp Fig 4-1**). RPA can be performed at different temperatures (*143*), but lower temperatures are known to reduce amplification speed. We found qRPA to be robust to temperature variation and reaction time even when the target Ct levels are reduced by sub-optimal temperatures or sub-saturating reaction lengths which would cause an underestimate if fitting RPA product levels to a standard curve (**Supp Fig 4-2**).

Detection of disease can also be hampered by sequence variation that arises during disease spread. This sequence variation can interfere with amplification reactions, particularly if the variation occurs in the 3' end of the priming region. Since the qRPA method relies on the reaction kinetics for target and reference molecules being equal, quantitation might be affected if the priming efficiencies are affected by mutations. RPA primer efficiencies are known to significantly affect the degree of amplification (*17*, *35*), and polymorphisms in priming regions have been shown to affect the sensitivity of other isothermal amplification diagnostics (*144*). To quantify the potential impact of sequence variation on qRPA, we sequenced a library of RPA primer-region variants before and after amplification with RPA (**Figure 4-2A**). The library was created by degenerate synthesis of a 37-bp RPA primer-region sequence with 91% abundance of the original nucleotide and 3% abundance of the other three nucleotides in the RPA primer binding region. This leads to an average error rate of 3.3 mismatches per template (**Supp Fig 4-3**). In addition, each template had a random 14bp UMI in the region that would be amplified. Thus, by comparing the UMI sequence before and after amplification it is possible to determine which mutations in the forward RPA primer-region affect RPA amplification (**Figure 4-2A**). Amplification bias was largely confined to the 3' end of the RPA primer sequence (**Figure 4-2C**), suggesting that qRPA is robust to the majority of mutations that may occur in the primer-region sequence, even if multiple polymorphisms co-occur (**Figure 4-2B**). This result also has implications for the

optimal design of RPA primers, suggesting that primers with 3' ends aligning to highly-conserved regions in the target may be more robust to inevitable mutational variation.
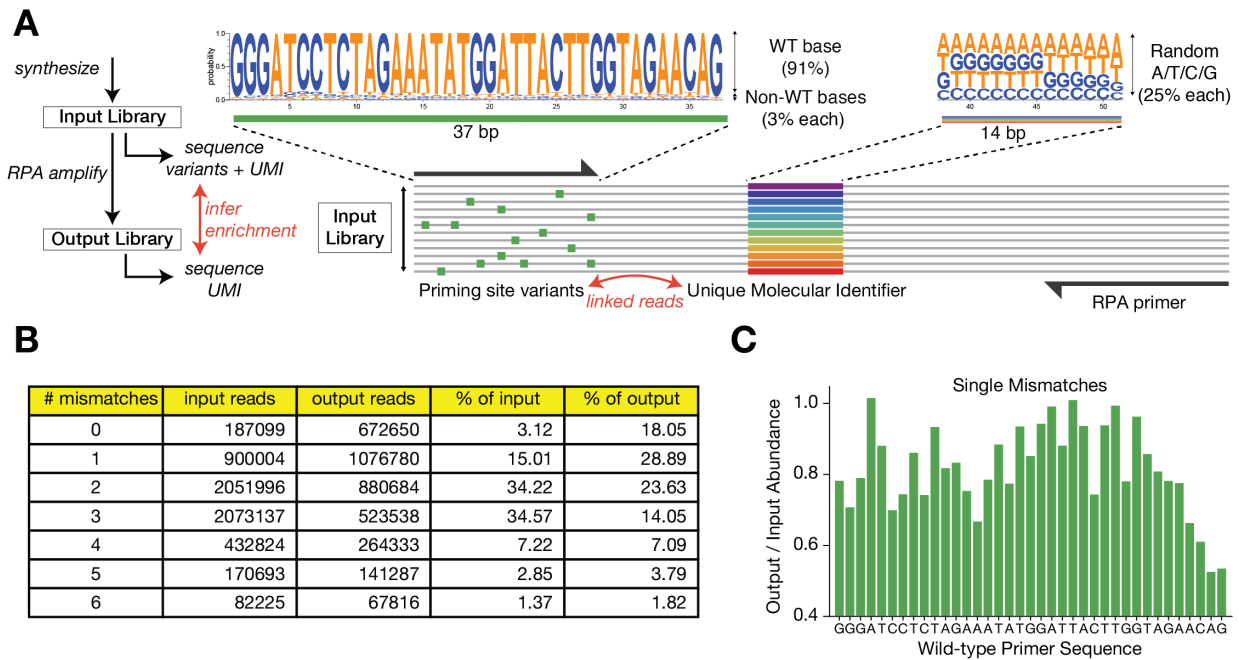
Figure 4-2 qRPA is robust to certain sequence variations.

(**A**) Measurement of amplification bias from sequence variation. A library with random sequence mismatches in the forward priming region each linked to internal Unique Molecular Identifier (UMI) sequences was synthesized. This input library was amplified using RPA, then both input and output libraries were sequenced. (**B**) Frequency table for reads in the input and output libraries, grouped by number of mismatches in priming sequence. (**C**) The abundance of library members with a single mismatch in the priming sequence, plotted by mismatch location, with output library abundance normalized by input library abundance.

qRPA can be combined with a fieldable detection method to measure the ratio of target and reference amplicons and then to infer the input concentration. We followed qRPA with hybridization with two probes, one biotinylated and complementary to the target sequence and one digoxigenin labeled and complementary to the reference sequence. This mixture was heat-denatured, allowed to cool, and then run on a lateral flow strip to quantify the relative abundance of target and reference amplicons in the reaction (**Figure 4-3A**). We chose to use Milenia Biotec HybriDetect 2T lateral flow test strips that contain a streptavidin band, an anti-digoxigenin band, and an anti-Ig band. These complexes are visualized using gold nanoparticle-labeled anti-FAM

90

antibodies that are captured at either test or reference bands on the lateral flow assay strip. To demonstrate this method, we cloned the N gene of SARS-CoV-2 into a plasmid and generated a reference target plasmid in which we replaced the RPA target portion of the sequence with a 28-bp reference barcode. We performed qRPA on a mixture of wild-type and barcoded-reference plasmids to demonstrate quantitation on lateral flow assays. The relative intensity of the reference and target bands for one or more reference concentrations can be used to infer the concentration of a target DNA input with roughly one order of magnitude precision (**Figure 4-3B**).
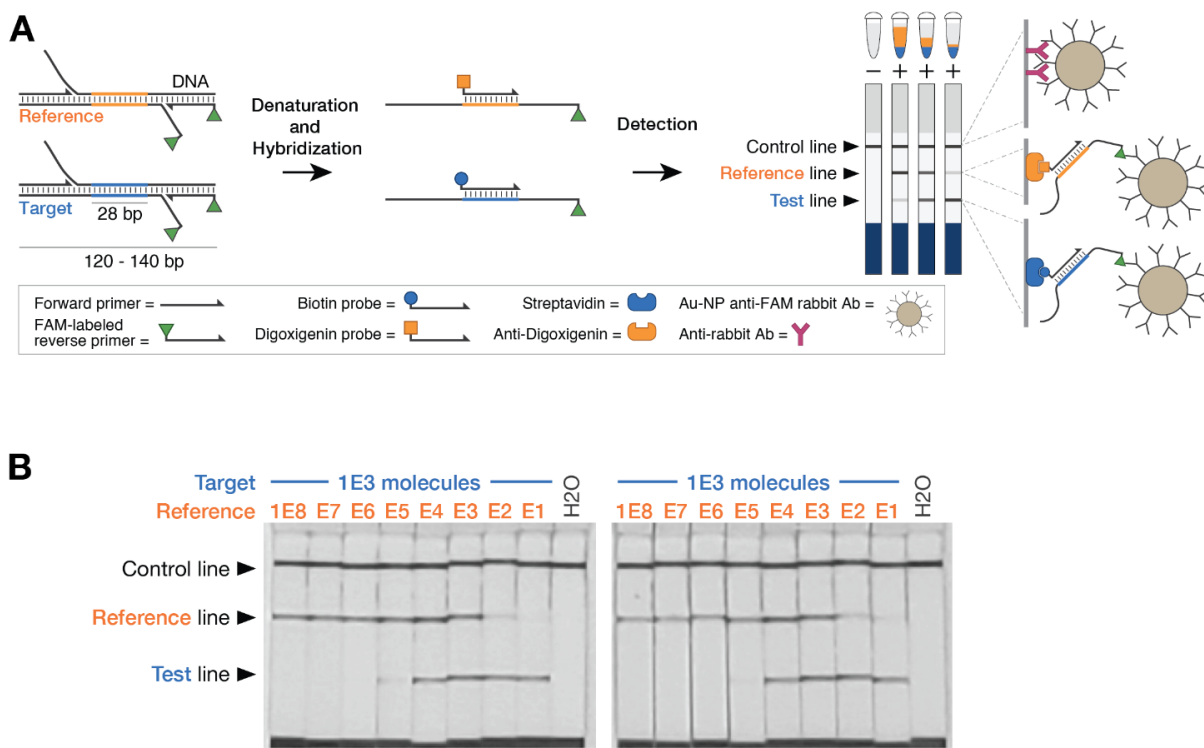
Figure 4-3 qRPA can be combined with a lateral flow assay for fieldable detection.

(A) Target and reference amplicons are detected using different hybridization probes and visualized using gold nanoparticles on lateral flow strips in a multiplexed sandwich assay. The relative intensity of target and reference bands can be used to infer target concentration in the sample. (B) Samples were prepared with 1E3 copies of a plasmid carrying a wild-type SARS-CoV-2 N-gene sequence and a varying number of copies of a reference plasmid in which a 28-bp region of the target was replaced with a synthetic barcode. These samples were amplified using qRPA, hybridized with target and reference probes, and visualized on the photographed lateral flow strips. Bands of equal intensity are observed when the target and reference concentrations are equal.

Many pathogens are composed of RNA genomes, including widespread pathogens such as HIV and SARS-CoV-2. qRPA may be combined with reverse transcription in a one-step RT-qRPA assay for RNA quantitation. We hypothesized that as long as the reverse transcriptase reaction does not exhibit significant bias, then the ratio of RT-RPA products should be retained in an RT-qRPA assay. Building on our previously reported eRPA assay for one-pot RNA amplification (*120*), we tested whether RT-qRPA could be used to quantify RNA input levels. We cloned the integrase gene from HIV into a plasmid and replaced the RPA target portion of the sequence with our 28-bp reference barcode (to be used as the reference sequence). We generated RNA from wild-type and barcoded-reference plasmids using in-vitro transcription (IVT), and then performed qRPA using combinations of these RNA products. Comparing the input ratio of target and reference RNA (pre-amplification) by RT-qPCR and post-amplification by qPCR, we found that RT-qRPA retains the product ratio over several orders of magnitude (**Figure 4-4A**) with comparable performance to TaqMan RT-qPCR (**Supp Fig 4-4**). The amplicons can be detected using lateral flow analysis in a field-deployable manner as we had shown with DNA targets (**Figure 4-4B & Supp Fig 4-5**). We obtained comparable results to the HIV constructs with RNA produced from the wild-type and barcoded-reference SARS-CoV-2 N-gene plasmids described above (**Figure 4-4C, D & Supp Fig 4-5**). Finally, by comparing the results from using several different reference molecule types (e.g. RNA, cDNA, DNA), we determined that for detection of RNA targets, an RNA reference molecule is most appropriate (**Supp Fig 4-6**).
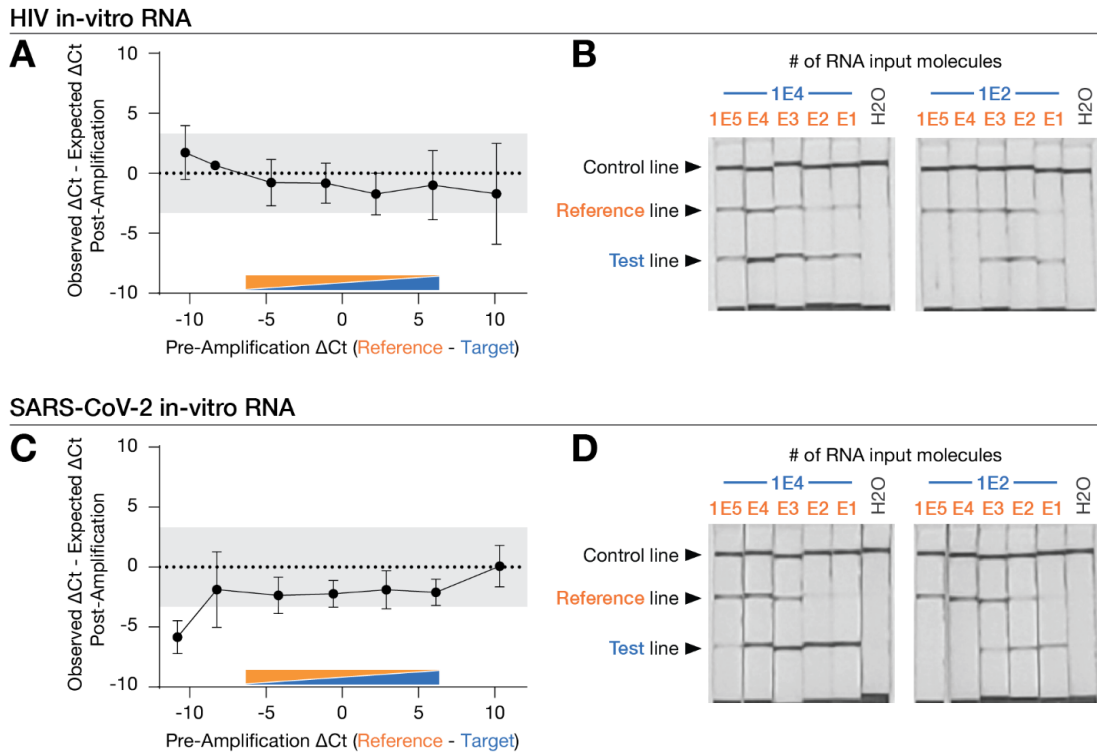
Figure 4-4 One-step RT-qRPA enables quantitation of RNA.

(A) Bland-Altman plot of RT-qRPA outputs for in-vitro HIV RNA samples containing either 1E2 or 1E4 target RNA molecules and a varying amount of reference RNA molecules with a synthetic barcode sequence inserted. The ratio of target to reference is maintained, as measured by ΔCt between target and reference after amplification. (B) The HIV RT-qRPA reaction products were hybridized and loaded on lateral flow strips and photographed. (C) Bland-Altman plot of RT-qRPA outputs for in-vitro SARS-CoV-2 RNA samples performed as above. (D) Photograph of lateral flow assay for RT-qRPA on in-vitro SARS-CoV-2 RNA. In each panel, each point depicts the mean of up to 6 replicate qRPA reactions, error bars depict standard deviation. Reactions in which target or reference were not detected post-amplification were not plotted.

To validate our assay for diagnostic use in quantifying viral load in the context of clinical samples, we developed a qRPA assay for HIV. HIV patients on anti-retroviral therapies may have varying viral loads that vary over time with treatment efficacy (*145*); viral load also correlates with transmissibility (*37*) so inexpensive rapid testing could enable better risk assessment. Viral loads <1500/mL are often considered to be non-transmissible (*37*). We performed qRPA on RNA samples extracted from the blood of individuals that were positive or negative for HIV. Using a trio of qRPA reference standards (0, 10, or 100 copies of reference RNA, corresponding to viral

loads of no, 1500/mL, or 15000/mL in unextracted blood), we classified 13 HIV patient and healthy individuals (**Figure 4-5A & Supp Fig 4-7**). To test the assay over a higher range of viral loads, we created and tested contrived positive samples by spiking our HIV IVT RNA into RNA samples purified from HIV-negative blood samples (**Supp Fig 4-7**). We subsequently asked three different individuals to estimate the titer of blinded samples, and 92% of calls for optimized samples were within an order of magnitude of the true titer (**Table S-9**). Bland-Altman plots show the pre and post qRPA amplification ratios as measured by qPCR were generally conserved (**Supp Fig 4-8**) as expected from our in-vitro validation experiments.
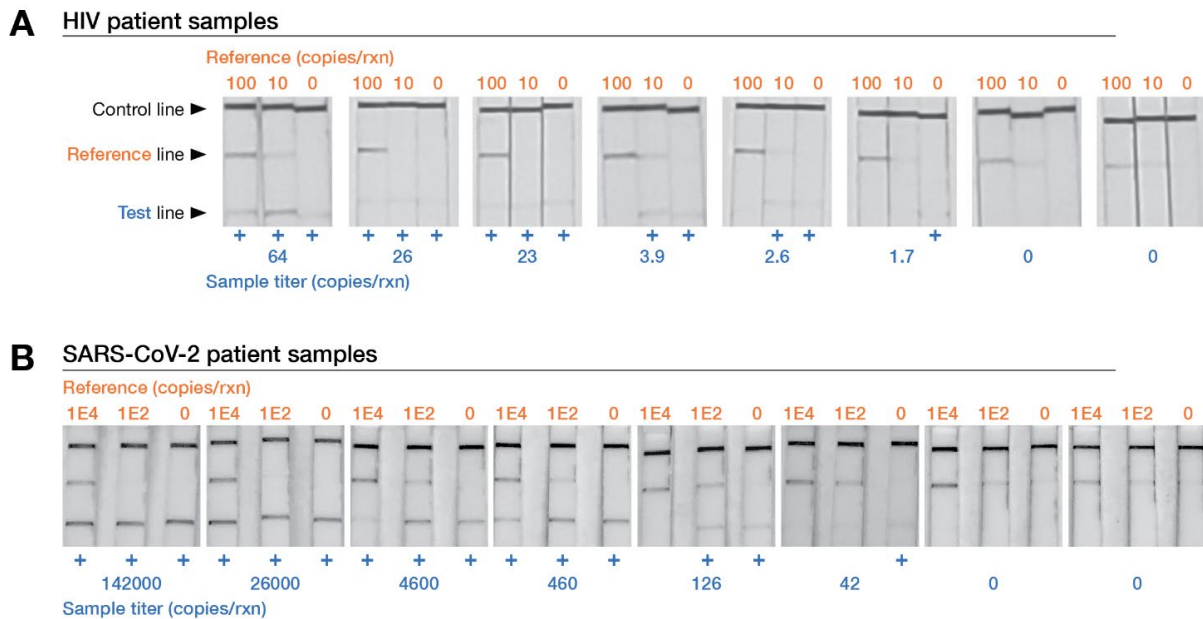
Figure 4-5 RT-qRPA and lateral flow assays can be used to quantify viral titers in patient samples.

(A) Total RNA was extracted from whole blood samples from HIV (+) or (-) patients. Known quantities of in-vitro HIV reference RNA containing a synthetic barcode were added to patient samples and amplified in qRPA. Photograph depicts a selection of lateral flow strips used to sort patient samples by viral titer. (B). Remnant nasopharyngeal swab samples from SARS-CoV-2 (+) or (-) patients were tested without extraction. Known quantities of in-vitro SARS-CoV-2 reference RNA containing a synthetic barcode were added to patient samples and amplified in qRPA. Photograph depicts a selection of lateral flow strips used to sort patient samples by viral titer.

94

Finally, we demonstrated that qRPA can also be used to quantitate SARS-CoV-2 levels in unextracted clinical samples. Blood interferes with many molecular assays and hence some form of extraction is needed, whereas nasal swabs are directly compatible with molecular assays (*120*). We therefore wanted to test the ability of qRPA to quantify SARS-CoV-2 titer directly from nasal swabs. We ran qRPA on remnant nasopharyngeal swab samples suspended in transport medium from individuals that were positive or negative for COVID-19. Using a trio of qRPA reference standards (0, 100, or 10,000 copies of reference RNA), we classified 22 patient samples by viral loads with order of magnitude precision (with one false negative on the sample with lowest viral load) and confirmed 13 negative patient samples (**Figure 4-5B & Supp Fig 4-9**). We subsequently asked three different individuals to estimate the titer of blinded samples, and 95% of these calls were within an order of magnitude of the true titer (**Table S-10**). Post-amplification ratios were biased in favor of the unencapsulated reference over the virus-encapsulated target, which may indicate that RPA is less efficient at releasing and amplifying viral RNA than qPCR (**Supp Fig 4-8**). This bias could be overcome by adjusting the amount of reference loaded, or by developing viral mimics for use as reference species (e.g. Accuplex (*146*)).

Discussion

qRPA uses competitive amplification and multiplexed detection to create an assay that could be used in field-deployable settings that require quantitation. Competitive amplification of the target and reference amplicons overcomes the challenges posed by non-specific amplification for quantitation, and lateral flow detection provides a simple read-out. The method works with both purified DNA and RNA as well as unextracted nasal swab samples.

qRPA overcomes many of the limitations of other isothermal amplification and detection techniques with quantitative outputs. Competitive-LAMP also utilizes competitive amplification to enable order-of-magnitude quantitation of nucleic acids with similar sensitivity and specificity (*38*) but qRPA benefits from simpler assay design (fewer primers) and lower reaction temperatures

(LAMP is performed at 65 °C). In addition, multiplexed detection of target and reference molecules on the lateral flow strip provides more data per reaction, so fewer reactions are needed to determine concentration in qRPA than competitive-LAMP. Real-time RPA utilizing exo probes (and enzymatic hydrolysis) can also be used for time-to-positive dynamic measurements for direct quantitation (*17*) or ratiometric quantitation (*14*). However, this approach requires sensitive real-time fluorescence measurement and internally modified probes that are difficult to manufacture at scale. Finally, although direct sample quantitation without amplification may soon be possible for highly concentrated nucleic acid targets (*147*), qRPA (and other amplification techniques) can achieve higher sensitivities while remaining quantitative.

Further work should broaden the use of qRPA to additional settings, sample types, and detection methods. First, there is a need for INAATs that can be run at ambient temperatures for point-of-need, equipment-free testing. RPA and qRPA can be performed at ambient temperatures but at reduced speed; it may be possible to overcome this by optimizing reaction mixes for ambient temperature reaction rates. Second, some sample types will likely remain challenging for RPA. One notable example, unextracted blood samples, could be made compatible with qRPA by using low-cost nucleic acid extraction methods (*148*). More broadly, validation of ratiometric amplification by post-amplification qPCR will be required in the development stage for each new sample type. Finally, on the detection side, we believe that the competitive amplification methods used in this work combined with multiplexed fluorescence detection using SHERLOCK (*23*, *142*) or molecular beacon (*141*) approaches could overcome challenges in using these detection methods for endpoint samples.

We believe that qRPA is well suited for non-laboratory setting where quantitation is needed for diagnosis. Measuring relative viral load in equipment-limited settings could help combat the AIDS pandemic by tracking the efficacy of anti-retroviral therapies (*145*). qRPA also brings us closer to at-home testing to track the progress of viral infections like SARS-CoV-2 following a positive test, which could help determine transmission risk and ease the load on

96

laboratory testing facilities (*7*). More broadly, higher-throughput, lower-precision testing for infectious diseases outside laboratories enabled by isothermal techniques will have value for contact tracing and epidemiological purposes (*1*) and pandemic control.

Chapter 4 Materials and Methods

RNA template generation

      SARS-CoV-2 N gene containing plasmids were obtained from IDT (2019-nCoV Plasmid Controls). An HIV integrase gene containing plasmids was obtained from Jonathan Li's lab at Brigham and Women's Hospital. Genes were then cloned into a T7 promoter expression plasmid with or without reference sequence insertions using Gibson assembly. To produce RNA template, in vitro transcription was performed with NxGen® T7 RNA Polymerase (Lucigen) according to the manufacturer's suggested protocol with minor modifications. Final concentrations of the reaction mixture components were 50 units T7 RNA polymerase, 1× reaction buffer, 625 µM NTPs, 10mM DTT, 500 ng of linearized plasmid template, and RNase-free water to a final volume of 20 µL per reaction. After 2 h at 37 °C, four units of DNase I (NEB) was added and reactions were further incubated for 10 min at 37 °C. DNase I was heat inactivated by adding EDTA (5mM final) and heating at 75 °C for 10 min. RNA was purified by RNAClean XP (Beckman Coulter) at 0.6× the volume of the reaction, washed twice with 80% EtOH, then eluted into 20 µL RNase-free water. The concentration of each RNA stock was calculated based on a Poisson distribution of RT-qPCR measurements on dilution series of RNA in nuclease-free water down to 0.003 molecules/reaction.

qPCR assays

      Unless otherwise noted, SYBR green qPCR reactions were prepared in 10 µL reaction volume using PowerUp SYBR Green PCR Master mix (Thermo Fisher Scientific), 2 µL sample, and 0.4 µM of primers and the following cycle conditions: (i) UDG activation: 50°C for 2 min, (ii) denaturation: 95 °C for 2 min; (iii) amplification (45 cycles): 95 °C for 15 s, 60 °C for 60 s; (iv) melt curve: 95 °C to 60 °C. RT-qPCR reactions were prepared in 10 µL reaction volume using the Luna Universal One-Step RT-qPCR kit (NEB), 2 µL sample, 0.4 µM of primers and the following cycle

conditions: (i) reverse transcription: 55°C for 10 min, (ii) denaturation: 95°C for 1 min; (iii) amplification (45 cycles): 95 °C for 10 s, 60 °C for 30 s; (iv) melt curve: 95 °C to 60 °C. TaqMan RT-qPCR reactions were prepared in 10 µL reaction volume using the Luna Universal Probe One-Step RT-qPCR kit (NEB), 2 µL sample, 0.4 µM of primers, 0.2 µM probe, and the following cycle conditions: (i) UDG activation: (i) reverse transcription: 55°C for 10 min, (ii) denaturation: 95 °C for 1 min; (iii) amplification (45 cycles): 95 °C for 10 s, 55 °C for 30 s. All qPCR and RT-qPCR reactions were monitored on a QuantStudio 6 Real Time PCR system (Thermo Fisher Scientific) or CFX96 real-time PCR system (Bio-Rad). The primers and probes used in reactions can be found in **Table S-11** and the corresponding amplicons may be found in **Table S-12**.

qRPA and RT-qRPA assays

Unless otherwise noted, qPRA assays were based on the TwistAmp Basic RPA Kit (TwistDx) with added modifications described below. For qRPA, each lyophilized pellet was resuspended in a solution of 39.5 µL rehydration buffer (TwistDx) and 0.5 µL of forward and reverse primer mix each at 50 µM. For RT-qRPA, 1.5 µL of rehydration buffer was replaced by 1 µL RNase H (5 U/µL) (NEB) and 0.5 µL SuperScript IV RT (200 U/µL) (ThermoFisher Scientific). The qRPA or RT-qRPA mix was then activated by addition of 1 µL 700 mM magnesium acetate followed by thorough mixing with a pipette. Reactions were prepared by dispensing 7.5 µL of master mix, 0.5 µL of reference template and 2 µL of sample input per reaction well, mixing the reaction by pipetting, and incubating at 42 °C for 25 or 30 min. Unless otherwise noted, a hybridization mix was prepared by combining 1 µL target-specific biotinylated probe at 5 µM (HIV JQ130; SARS-CoV-2 JQ241), 1 µL reference digoxigenated probe (JQ170) with 18 µL 10 mM Tris pH 8. 20 µL of hybridization mix was added to each reaction, and samples were heated at 94 °C for 3 min followed by a cooling step at room temperature for 3 min. 50 µL of Milenia GenLine 2T Buffer (Milenia Biotec) was added to each reaction, mixed by pipetting, and a lateral flow strip (Milenia HybriDetect 2T) was added. Lateral flow strip signals can be detected and imaged

starting 3 min after addition of the strip to the hybridized reaction. Test results were called or imaged within 45 min of strip addition since background bands at the test and reference lines can appear over time and low signal test bands can lose intensity as the strip dries.

In the pH and salt perturbation qRPA assay, home-made 2X rehydration buffer was prepared to contain 50 mM Tris, 200 mM Potassium Acetate, 11% (w/v) Polyethylene glycol (PEG) 20,000. For pH perturbations, buffer pH was adjusted by hydrochloric acid or sodium hydroxide and measured by pH meter. For salt perturbations, NaCl was added to 120 mM. Each lyophilized pellet (TwistDx) was resuspended by home-made rehydration buffer to form 50 µL of final 1X solution with 25mM Tris, 100 mM Potassium Acetate, 5.5% (w/v) PEG 20,000, 700 mM magnesium acetate and 240 nM primers.

Sequencing analysis of primer mutational library

A mutational template library (JQ75) where each base pair constitution of the forward RPA primer binding site (JQ42) was 91% as wild-type base and 3% of each of 3 remaining bases, was ordered from Integrated DNA Technologies. Additionally, another library (JQ76) with 14-bp randomized UMI was designed 11-bp downstream of the forward RPA primer binding site to associate UMI to template variants (**Figure 4-2**). Template variant and UMI libraries were PCR amplified with sequencing adaptors (JQ77,83-85 & JQ93-96), then indexing primers (**Table S-11**) to build the input library. The input library was used as input template into a 50 ul RPA reaction, where a lyophilized pellet was resuspended in a solution of 29.5 µL rehydration buffer (TwistDx) and 480 nM of forward (JQ42) and reverse primer (JQ92) and 1 ul of input library and supplemented with RNAse-free water until 47.5 ul, then activated by addition of 2.5 µL 280 mM magnesium acetate followed by thorough mixing with a pipette. The post-RPA amplification product was purified with 0.8x AMPure XP beads (Beckman Coulter) and was PCR amplified with sequencing adaptors (JQ97-100 & JQ93-96), then indexing primers (**Table S-11**) to build the output library. The input and output libraries were sent for sequencing (paired-end 150-bp) on

100

NextSeq550 (Illumina). For the bioinformatics analysis, Paired-End ReAd merger (PEAR) was used to merge read 1 and read 2 of both input and output sequenced libraries. Seqtk was then used to filter out reads smaller than 50-bp in length. Then, from the input library sequencing reads, a look-up table of primer variants of 37-bp matched to 14-bp UMIs was built. Individual barcodes that were matched to multiple different primer variants were discarded. Output library 14-bp UMIs were extracted from output sequencing reads and matched to primer variants based on the look-up table.

Clinical HIV samples

A cohort of RNA extract samples were obtained from Jonathan Li's lab at Brigham and Women's Hospital, collected from patients with estimated HIV viral load of 15000 copies/µL or 1500 copies/µL. Replicate 200 µL plasma samples were extracted using a Norgen Biotech Kit and eluted into 55 µL buffer according to manufacturer instructions, then stored at -80 °C. At the time of testing, replicate samples were pooled in groups of 3 and vacuum concentrated to 15-20 µL (such that sample inputs were ~30X more concentrated than plasma). 2 µL of the concentrated sample was used as input into RT-qRPA or RT-qPCR reactions.

Clinical COVID samples

A cohort of nasal swab patient samples was purchased from BocaBiolistics, FL. Samples were thawed on ice and 40 µL aliquots were made and subsequently stored at −80 °C. At the time of testing, sample aliquots were thawed and RNasin Plus was added to a final concentration of 1 U/µL. The samples were placed on a heat block set to 99 °C for 5 min for virus inactivation and lysis. After cooling, samples were spun down and transferred to a 96-well DNA LoBind plate

Supplementary Text

Proper primer design for RPA and qPCR reactions is critical to the development of a qRPA assay. As noted in Materials and Methods, RPA primers in this work were first screened in a modified version of a screening method we previously reported for diluted RPA products (*14*). A collection of forward and reverse primer pairs were used to amplify 100, 10 or 0 template molecules by RPA. The products were quantified using qPCR with non-nested (both primers overlap RPA priming site) or semi-nested (one internal primer and one overlapping primer) primers. Optimal primers were selected based on 1) high amounts of signal for on-target samples with semi-nested primers and 2) low signal from non-nested primers for no template control samples. To quantify post-amplification samples, the use of semi-nested or fully-nested primers is critical to ensure that non-specific amplicons are not detected. Poor primer design for RPA or qPCR primers will cause errors and bias in quantitation during the validation stage of qRPA assay development.

In addition to primer design optimization, lateral flow assay conditions may need to be optimized. The efficiency of hybridization by different probes may lead to quantitation bias, so length and GC content should be matched for target and reference probes whenever possible. Additionally, the amount of probe and sample loaded onto each strip may require optimization. For example, loading 8 µL of reaction product (as done in the in-vitro RNA RT-qRPA assays Fig. 4 and Fig. S4) produced strong bands with high signal, but these saturated signals made it harder to determine relative strength of each band. Loading less RT-qRPA reaction product for the patient samples helped distinguish band strengths more clearly. Additionally, hybridization with 5 µM probe enabled us to cleanly detect SARS-CoV-2 patient samples, but caused faint background signal at the reference band for HIV patient samples (asterisked samples in Fig S6). HIV patient samples also contained much lower input levels of target and reference which produced less product during the reaction time course. HIV samples were best resolved after hybridizing 2.5 µM probe (50% less) with 4 µL sample (100% more). The optimal conditions will

depend on the intended use case. To increase sensitivity at low concentration samples, consider increasing the quantity of reaction product loaded and reducing the concentration of probe to reduce background. Conversely, if high concentration samples are expected and there is a low risk of false positives, consider decreasing the amount of reaction product loaded and increasing the concentration of probe to improve resolution of samples.

For detection of RNA targets, an RNA reference molecule should be used. RNA is less stable than DNA and potentially poses challenges for long-term storage and shipping of RT-qRPA assays mixtures. We therefore we considered the use of alternative reference molecules. We also asked whether variability in the reverse transcription step may cause introduce noise into RT-qRPA measurements relative to qRPA with DNA inputs. We compared the results of our SARS-CoV-2 qRPA assay when starting from RNA, cDNA, or DNA inputs as well as competitively amplifying a wild-type RNA molecule against a DNA reference molecule. We found that replicate RNA vs. RNA reactions had only slightly higher variance than DNA vs. DNA reactions (**Supp Fig 5**). We have shown previously that efficient removal of RNA from RNA-DNA hybrids can improve the efficiency of RPA initiation (*13*). Accordingly, performing reverse-transcription and RNAse H treatment in a separate, prior step to qRPA decreased the variance, while removing RNAse H from the one-step reaction increased variance. Most importantly, competitive amplification of an RNA target molecule against a DNA reference molecule introduced a significant bias in favor of the DNA reference, corresponding to a roughly 30-fold difference in ratio after amplification.

Chapter 5: Conclusion

In this thesis, we developed and applied nucleic acid detection methods to two systems. In Chapter 2, we successfully developed a system that leverages barcoded microbial spores, RPA, and CRISPR-based detection methods to solve the object provenance problem. We have shown evidence that the spores can persist in the environment, that they can be transferred onto various types of objects and rapidly read out with minimal instrumentation. The safety measures built into the design of the engineered spores minimizes spread and horizontal transfer between the inoculation area and the spores themselves. This field-deployable system can be applied in a broad range of application in real-world environments. We have shown proof-of-concept experiments for determining food provenance. This system could also be applied to forensics to help investigators determining objects or person involved in crimes or trafficking. Other applications could be for tracking animal movement for wildlife conservation (*149*) or for determining counterfeiting commercial products in global supply chain (*150*). While this first study suggests that our spore system could work across various environments and different surface, more extensive validation in a wider range of real-world conditions is needed.

Future improvements of our system can be done by engineering spores with better features, including controlled spore growth, active spore destruction, and a timer. Spores that can be tightly controlled for propagation could be used in highly trafficked areas so that spores would not need to be re-inoculated periodically. Controlled growth can be achieved by producing spores from engineering microbial strains that have gene circuits to respond to an input stimulus and have growth as an output (*151*). Additionally, our current spores have passive safety measures built into them that effectively make them unable to germinate and grow. Active safety measures would add another layer of safety. This can be engineered by integrating a toxin/anti-toxin gene circuit into the spores, allowing specific on-demand destruction of spores (*152*). Finally, implementing a timer function (*153*, *154*) to the spores would provide important temporal information about location history and would allow time-resolved trajectory of objects to be determined.

On the detection side of our pipeline, further efforts could be aimed to integrate all the steps into one for ease of use. Currently, NaOH spore lysis, RPA, and CRISPR-based detection steps are separate from each other and thus require manual handling in between each step. Therefore, combining those steps would render field-deployability of this system even more attractive in low-resource settings. The present approach for barcode detection in each sample is relatively low throughput (e.g. 10s to 100s of barcode per sample). To increase the capacity to detect several orders of magnitude more barcodes in many more samples would require a high-throughput detection setup. This could be achieved by using microfluidics, microwells or droplet-based technologies (*155*).

In Chapter 3, with the quickly progressing SARS-CoV-2 pandemic and the experience we gained developing the nucleic acid detection method in chapter 2, we set out to optimize the RPA assay to develop a sample preparation and detection pipeline for SARS-CoV-2 point-of-care diagnostic test. The eRPA assay was developed and optimized in just a few weeks after the genome of SARS-CoV-2 was made available. The assay is rapid, specific, sensitive, and doesn't require expensive instrumentation or highly trained personnel. The eRPA protocol requires minimal preparation of samples and is fully compatible with crude nasal swab samples. The RPA primer screening method we developed can be adapted to any detection targets and is quickly scalable to screen more primer pairs. While eRPA addresses many of limitations of current testing methods and brings us a step closer to an at-home test, there are still some limitations that can be improved in future iterations of this method.

Areas of improvement include 1) making the reaction self-contained in a single tube without needing to opening reaction after sample input, 2) increase sensitivity for saliva samples and using other sample types, 3) lowering optimal reaction temperature, 4) making the assay quantitative. The current eRPA protocol requires opening and closing of post-amplification samples, which could be a source of cross-contamination. Pairing eRPA with other detection methods could address this limitation. CRISPR paired with eRPA allows a fluorogenic readout

(*136*, *156*). Pairing eRPA with fluorescent molecular beacons (*157*) or oligo strand displacement techniques (*117*) could also be explored for ease of use. Additionally, while we have yet to fully integrate saliva samples from patient into eRPA because of the high RNase activity in saliva degrading viral RNA during the lysis step, our initial work on saliva inactivation with RNase inhibitor and TCEP showed a proof-of-concept path forward that could make saliva a universal medium used for future pathogenic diseases. Our saliva inactivation method has been successfully able to lyse viral particles in saliva patient and been integrated into SARS-CoV-2 detection in LAMP assays (*115*). Sample types other than nasal swab and saliva have been used for SARS-Co-V-2 detection, including sputum, bronchoalveolar lavage ocular samples, stool, anal swabs, urine, or blood (*158*) can be explored as input into eRPA.

Future efforts for RPA and all INAATs would be to lower the optimal reaction temperature to ambient or even lower temperature so that these methods can be truly field-deployable. Currently, eRPA requires a heating device at 42C for 30-45 minutes and LAMP requires 65C for 60-120 min because the enzymes used in these methods catalyze reactions optimally at those temperatures. To lower the optimal reaction temperature in INAATs will require the use of new enzymes. The new enzymes or current enzymes would have to be discovered from cold-adapted living organisms or engineered to acquire optimal catalytic activity at low temperatures (*159*), respectively. Enzymes for nucleic acid detection methods have been successfully harnessed from nature, including Taq polymerase which was discovered in Yellowstone from thermostable *T. aquaticus* bacteria and has revolutionized PCR assays (*160*). Bioprospecting of enzymes in psychrophilic or mesophilic organisms have been not been as common since there has not been much a need for enzymes working in those temperatures; however, given the computational resources that we have today to mine genomic data and the potential application for cold-adapted enzymes, the discovery of new polymerases or other enzymes is particularly exciting. Other strategies to develop enzymes with low-temperature optimal catalytic activity is protein engineering (*161*) or directed evolution (*162*) of existing enzymes.

Chapter 4 is a continuation of making improvements to eRPA from Chapter 3. Here, we successfully used the principle of competitive amplification to make the eRPA assay quantitative (qRPA), which would be useful for low-resource environments to monitor the progress of viral infection. We harnessed the non-specific amplification of RPA and used it as a feature for quantitation. By adding a known amount of reference sequence that shares the same primer binding sites as the query sequence to an unknown sample, we were able to quantitatively infer the concentration of the queried sequence quantity by separately detecting each amplicon from the mixed saturated RPA endpoint product. A similar approach has been reported for LAMP (*38*), however qRPA benefits from lower temperature and easier primer designs. Real-time RPA using exo or fgp probes allows dynamic monitoring of RPA amplification; however, the detection method requires expensive instrumentation, and the probes are difficult and expensive to manufacture at scale (*14*). Recent efforts have reported quantitation of samples without amplification, but they remain less sensitive than qRPA (*147*). Future efforts for qRPA include broadening the use to additional settings, sample types, and detection methods.

All together this dissertation spans work across rapid nucleic acid detection methods and their applications in object provenance and viral diagnostics. Field-deployable technologies emerged as the overarching theme across these chapters. This body of work shows only the beginning of the massive potential applications that fieldable nucleic acid detection methods can be used for. Future technologies and discovery of new enzymes can bring us closer to a world where methods are truly field deployable.

Appendix A: Supplementary material related to Chapter 2

Saccharomyces cerevisiae strains			
Strain	Genotype	Alias	Source
BY4743	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0	WT	ATCC
yBC1	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC1	BC1	This work
yBC2	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC2	BC2	This work
yBC3	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC3	BC3	This work
yBC4	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC4	BC4	This work
yBC5	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC5	BC5	This work
yBC6	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC6	BC6	This work
yBC7	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC7	BC7	This work
yBC49	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC49	BC49	This work
yBC50	MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0 hoΔ::BC50	BC50	This work

Bacillus subtilis strains			
Strain	Genotype	Alias	Source
168	Wild-type (trpC2)	WT	BGSC
BDR3897	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72	Δ9	This work
BDR3900	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO::erm	Δ9 ycgO::erm	This work
BDR3916	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 yccR::cat	Δ9 yccR::cat	This work
BDR3917	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 yccR::Pveg-optRBS-mScarlett (Spec)	Δ9 ycvR::Pveg-optRBS-mScarlett (Spec)	This work
BDR4038	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC1 (lox72)	Δ9 ycgO:: V2-BC1	This work
BDR4039	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC2 (lox72)	Δ9 ycgO:: V2-BC2	This work
BDR4040	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC3 (lox72)	Δ9 ycgO:: V2-BC3	This work
BDR4041	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC4 (lox72)	Δ9 ycgO:: V2-BC4	This work
BDR4042	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC5 (lox72)	Δ9 ycgO:: V2-BC5	This work
BDR4043	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC6 (lox72)	Δ9 ycgO:: V2-BC6	This work
BDR4052	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC7 (lox72)	Δ9 ycgO:: V2-BC7	This work
BDR4053	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC8 (lox72)	Δ9 ycgO:: V2-BC8	This work
BDR4054	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC9 (lox72)	Δ9 ycgO:: V2-BC9	This work
BDR4055	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC10 (lox72)	Δ9 ycgO:: V2-BC10	This work
BDR4029	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC11 (lox72)	Δ9 ycgO:: V2-BC11	This work
BDR4030	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC12 (lox72)	Δ9 ycgO:: V2-BC12	This work
BDR4031	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC13 (lox72)	Δ9 ycgO:: V2-BC13	This work
BDR4032	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC14 (lox72)	Δ9 ycgO:: V2-BC14	This work
BDR4033	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC15 (lox72)	Δ9 ycgO:: V2-BC15	This work
BDR4034	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC16 (lox72)	Δ9 ycgO:: V2-BC16	This work
BDR4056	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC17 (lox72)	Δ9 ycgO:: V2-BC17	This work
BDR4057	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC18 (lox72)	Δ9 ycgO:: V2-BC18	This work
BDR4064	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC19 (lox72)	Δ9 ycgO:: V2-BC19	This work
BDR4065	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC20 (lox72)	Δ9 ycgO:: V2-BC20	This work
BDR4066	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC21 (lox72)	Δ9 ycgO:: V2-BC21	This work
BDR4067	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC22 (lox72)	Δ9 ycgO:: V2-BC22	This work
BDR4075	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC23 (lox72)	Δ9 ycgO:: V2-BC23	This work
BDR4068	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC24 (lox72)	Δ9 ycgO:: V2-BC24	This work
BDR4069	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC25 (lox72)	Δ9 ycgO:: V2-BC25	This work
BDR4070	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC26 (lox72)	Δ9 ycgO:: V2-BC26	This work
BDR4071	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC27 (lox72)	Δ9 ycgO:: V2-BC27	This work
BDR4232	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC52 (lox72)	Δ9 ycgO:: V2-BC52	This work
BDR4233	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC53 (lox72)	Δ9 ycgO:: V2-BC53	This work
BDR4234	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC54 (lox72)	Δ9 ycgO:: V2-BC54	This work
BDR4235	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC55 (lox72)	Δ9 ycgO:: V2-BC55	This work
BDR4236	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC56 (lox72)	Δ9 ycgO:: V2-BC56	This work
BDR4237	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC57 (lox72)	Δ9 ycgO:: V2-BC57	This work
BDR4238	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC58 (lox72)	Δ9 ycgO:: V2-BC58	This work
BDR4239	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC59 (lox72)	Δ9 ycgO:: V2-BC59	This work
BDR4240	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC60 (lox72)	Δ9 ycgO:: V2-BC60	This work
BDR4241	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC61 (lox72)	Δ9 ycgO:: V2-BC61	This work
BDR4242	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC62 (lox72)	Δ9 ycgO:: V2-BC62	This work
BDR4243	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC63 (lox72)	Δ9 ycgO:: V2-BC63	This work
BDR4244	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC76 (lox72)	Δ9 ycgO:: V2-BC76	This work
BDR4245	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC77 (lox72)	Δ9 ycgO:: V2-BC77	This work
BDR4246	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC78 (lox72)	Δ9 ycgO:: V2-BC78	This work
BDR4247	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC79 (lox72)	Δ9 ycgO:: V2-BC79	This work
BDR4248	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC81 (lox72)	Δ9 ycgO:: V2-BC81	This work
BDR4249	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC82 (lox72)	Δ9 ycgO:: V2-BC82	This work
BDR4250	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC83 (lox72)	Δ9 ycgO:: V2-BC83	This work
BDR4251	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC84 (lox72)	Δ9 ycgO:: V2-BC84	This work
BDR4252	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC85 (lox72)	Δ9 ycgO:: V2-BC85	This work
BDR4253	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC86 (lox72)	Δ9 ycgO:: V2-BC86	This work
BDR4254	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC87 (lox72)	Δ9 ycgO:: V2-BC87	This work
BDR4255	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC88 (lox72)	Δ9 ycgO:: V2-BC88	This work
BDR4256	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC89 (lox72)	Δ9 ycgO:: V2-BC89	This work
BDR4257	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC90 (lox72)	Δ9 ycgO:: V2-BC90	This work
BDR4258	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC91 (lox72)	Δ9 ycgO:: V2-BC91	This work
BDR4259	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC92 (lox72)	Δ9 ycgO:: V2-BC92	This work
BDR4260	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC93 (lox72)	Δ9 ycgO:: V2-BC93	This work
BDR4261	ΔthrC::lox72 ΔmetA::lox72 ΔtrpC::lox72 ΔpheA::lox72 ΔsleB::lox72 ΔcwlJ::lox72 ΔgerAB::lox72 ΔgerBB::lox72 ΔgerKB::lox72 ycgO:: V2-BC94 (lox72)	Δ9 ycgO:: V2-BC94	This work

Table S1. List of strains used in chapter 2

Bacterial and yeast strains used in this study. List includes wild type strains and mutants generated as part of this study. All unmarked *B. subtilis* mutations are in-frame deletions generated by Cre-mediated recombination and contain a lox72 scar.

Primer Name	Sequence	Target
prCM488	GATAAACACAGGAAACAGCTATGACCATGATTACG	Barcode reverse
prCM515	TCGTCAGTCGTAACTTGGGAACGCACAT	Barcode 1 forward
prCM516	CTTTGGGGTTGGATAAATCTGTCGTGGT	Barcode 2 forward
prCM517	TGAATAAGCGCGGTCCCTAATGTTGGTG	Barcode 3 forward
prCM518	ACCGGTTTATAGTTCGAACAGTCGCAGC	Barcode 4 forward
prCM519	CCCGTGTGGTAACTACGCAAGCCTAAC	Barcode 5 forward
prCM520	GGGCGCGTAACCACTAGCTC	Barcode 6 forward
prCM521	gggctagctAGTGTCCCTTATTCTACTTTGAATTATC	Barcode 7 forward
prCM522	GTTACGGGTCAGGATCATTGCGCAGG	Barcode 8 forward
prCM545	CCGACTAGGTCGGCGATACCATAGCACC	Barcode 24 forward
prCM546	GATGACGCCTCACGTCCGTTGCCCCCTA	Barcode 25 forward
prCM547	TGCCTCAGAATGGGGTTTAACGTTCAGG	Barcode 19 forward
prCM548	ACCGTCCTAGTGCGGATAGTGGGACATG	Barcode 49 forward
prCM549	GTCGATACGGGTCGCTAGATAGGTTGGG	Barcode 50 forward
BCv2-F	GGGATCCTCTAGAAATATGG	universal barcode forward
BCv2-R	TGATAAACACAGGAAACAGC	universal barcode reverse
gRNA_gap_repair_F	ATCAACAACAGAGGACATATGCCC	HO-targeting gRNA
gRNA_gap_repair_R	ATCCTGCACTCATCTACTACCCG	HO-targeting gRNA
BT-1F	CAAACATGCGAAGCGACTTA	detects Bt cry1a
BT-1R	CGCCTCCTTGGATGGTAATA	detects Bt cry1a
JQ 39	gaaatTAATACGACTCACTATAggg	forward primer to anneal to reverse crRNA
JQ 24	GATAAACACAGGAAACAGCTATGACCATGATTACG	RPA reverse primer and realtime PCR reverse primer
JQ 42	GGGATCCTCTAGAAATATGGATTACTTGGTAGAACAG	RPA forward primer
oCB034	cgcGGATCCtctagaaatatggattacttg	megaprimer forward
oCB035	cggCTCGAGataaacacaggaaacagctat	megaprimer reverse
oCB049	gcgAAGCTTacataaggaggaactactatggtttctaaaggt	mScarlett forward
oCB050	GCGGGATCCttacttatacagttcatc	mScarlett reverse
prCM509	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG	16s v3-v4
prCM510	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC	16s v3-v4
Lu_BC90F	TAACTAGCGGCCTATACCCGCTGCACAT	Barcode 90 forward
Lu_BC91F	CAGAGATAGGTCTCGACACTTACGCGGT	Barcode 91 forward
Lu_BC23F	GCAATATTACTAGTCCGTGTCAGCCAGA	Barcode 23 forward

Table S2 List of primers used in chapter 2

Name	Sequence
Group 1 crRNA	TCGTCAGTCGTAACTTGGGAACGCACATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_2	CTTTGGGGTTGGATAAATCTGTCGTGGTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_3	TGAATAAGCGCGGTCCCTAATGTTGGTGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_4	ACCGGTTTATAGTTCGAACAGTCGCAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_5	CCCCGTGTGGTAACTACGCAAGCCTAACGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_6	ACGTAGGGGGGCGCGTAACCACTAGCTCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_7	AGTGTCCCTTATTCTACTTTGAATTATCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_8	GAGTTACGGGTCAGGATCATTGCGCAGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_9	GTGAGTCCGGCCTATCACGTTTGGTAGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_10	TCAGGGGAAACGAGTTAAGCAGAGGCAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_11	CTTGGTCCAATCGTATGCTAAGAGTAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_12	TCGGTTGCACGCGCGTCCTGCACGCTCGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_13	GTTCAAAAGCGGGAGTCCCGGTGAAACCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_14	GAGGGCTTCTACGAAATTGCCTCACCATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_15	GGGATGCCCTGATAACGTAGAGTCTCAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_16	TAGTACGGGCCACGATCTAAGTCGGCGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_17	GTATGTGGGTACGATTGTAGGCTAGAAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_18	AGAACTACCTACGTGGCCACGCAAGCCCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_19	TGCCTCAGAATGGGGTTTAACGTTCAGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_20	ATGTTGGGACGCCAGACTGACACGCAAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_21	CCTACACCTTCACCGAGTGTGGAGCAAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_22	GTTTAACATTCGGGTATCGTCCTTACGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_23	GCAATATTACTAGTCCGTGTCAGCCAGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_24	CCGACTAGGTCGGCGATACCATAGCACCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_25	GATGACGCCTCACGTCCGTTGCCCCCTAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_26	GCTCCGTAGCAATTGATAAACCGCTGAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_27	CTATCGGATTTTTAGATAAATTACGAGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_28	TTTTTAATCGGTAGCGACGATATATGACGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_29	TTGATTCGTGGGGATCCGCGGATAGAAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_30	CTAATATCTGATCCGGTTATAACACCACAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_31	TTGCGGGCCTGGACATTCAATCGCATACGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_32	GCCTAGACCAAAAGTCGCCTTCAGGGCAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_33	ATAGGTAGGTAACGGAACTGACCGATAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_34	CGTACTCCTTACGCCTTACACGGACAATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_35	GATTCTCTGGTCATGAGGGAAGATCGACGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_36	TTTGGACGGGGTCACTACTCTTTACAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_37	TTCATGTATGCTGTCCTACCGCAGTAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_38	TACATAAGGTCGTCCTCCGGTCACGCCCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_39	CAGGCCGCTACTCAAGCGGGGATAGCTTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_40	CGCTTAGCTAGATTGCCCCTGGAAGCGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_41	CGGAGTGTTTGCGAATCCCGAGCATTAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_42	GTCTTTCGAATCGTTCCTTATAACCCGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_43	AGGCTGAATCCTCCGGCGGCTGTTCTAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_44	CAGCCAAACGCTAGCAAAATAGGATATAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_45	CAGACGCTGTCTACCTAGCAGAAGGGCTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_46	AAGGTATAATGTTCAAGCGTGGTGGCTAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_47	ACTGATCTCGGCGTAAATACACCGCAGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_48	TATGTGTATCCTAAGGTACCTTTACGGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_49	ACCGTCCTAGTGCGGATAGTGGGACATGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_50	GTCGATACGGGTCGCTAGATAGGTTGGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_51	ACAGTCAGAGGCGCTTCAATCGTCGCCCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_52	GCGACTCATTTAATTTCAACGGGTAGAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_53	TAGACAGAGGGAGCTATATTGCGCGGCGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_54	GGCATTCAATAGCCAGTAACCTTAGCACGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_55	ACAAAAGTTGGCATTTCGTAGACAAAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_56	TTGCATACTTTGCAGGTTAGGGCCGAGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_57	ACCGGAGTTCAAACGACTTTGGGTCCCTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_58	CACTGGTAGGCTCGCGGGTAAATAGAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_59	CCGAAGCTCAGTGTGATGTGCCTTCCGTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_60	GTACGTATGACTGTGTCCCAGCTATGAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_61	GGACCTTCACGTCAGTCCCGCTGACCTTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_62	GACCAATTATGTTGGCCACGCCACATTAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_63	GTTTTCAAGGGCTCGACACAGCAGAACGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_64	AACCCACTATAGAAGGGCCTTCGAGATTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_65	GGAACCGTCAATTCCGCACTAATAATTTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_66	CATGGCTCCGAAACGCCGTGCAAAGTGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_67	GCTTTCTCATTGTTACTCAGCGCACATTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_68	TCAATCCTGAGGCCGATACTAACAGTCGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_69	ATTCTGGCTATACCGAGCCCGGACTACTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_70	CAACCATGCTTCGGATAGCAATACCACGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_71	CTTGTCTCAAGAATAGTGCCCTATTCGTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_72	CCGTAGGTAGAAACATGTATGGCGCAGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_73	CAGGGCCATGTCGTACTAGTAGCCTGCCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_74	CGAGCCCCCCAAAAAGTCGGGCTACCGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_75	CATATAGGTACATGAAGCCCTACTGTATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_76	CATACCGGATGGTAAACTCCGCGAACGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_77	TGCAATGACCAGGCGATACGAATCGTAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_78	CTCCCCCCCGCAGACACGCACGTGTTGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_79	TTTGTGGTTAGCCTTGTCCCGGAACAGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_80	CCCGGGTTGATGAGGAGTCACTCAACGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_81	ATCCATGTGCGGGGAGAAGTCTGGTGGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_82	CCCCGGTAATTCGATGTTCGCCCCGGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_83	TGGGCTGCTAGTAAGTCGGGAGCCCTTAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_84	CAGGCCGCTAACCTAAACTCGACCTCGAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_85	CAACGCAAGCGCTCTTGGTTTAAGCTGTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_86	GGCACACAACGATCTCCAAATGAACCAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_87	GTAGCGCGACCTCTTTAACTGTAGAAGGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_88	GCTTATCGCATTCCGGGTGAAAACCTAGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_89	AGGGTAATGAAGCCCCGGGATCGAACATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_90	TAACTAGCGGCCTATACCCGCTGCACATGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_91	CAGAGATAGGTCTCGACACTTACGCGGTGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_92	GCCTAACGGCTGCAGCACGATTGCGCAAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_93	GACGCCATTAAGGTGGGTCACCTTTTCAGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
crRNA_Barcode_94	GCATATGGCTAAATTCGCTAGTCGAGCGGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC
Group 2 crRNA	GTCGACCTGCAGGCATGCAAGCGTTTTAGTCCCCTTCGTTTTTGGGGTAGTCTAAATCCCCTATAGTGAGTCGTATTAATTTC

Table S3 List of crRNA used in chapter 2

List of crRNA sequences used for SHERLOCK reactions in this study.

Name	Sequence
Barcode_1	TCGTCAGTCGTAACTTGGGAACGCACAT
Barcode_2	CTTTGGGGTTGGATAAATCTGTCGTGGT
Barcode_3	TGAATAAGCGCGGTCCCTAATGTTGGTG
Barcode_4	ACCGGTTTATAGTTCGAACAGTCGCAGC
Barcode_5	CCCCGTGTGGTAACTACGCAAGCCTAAC
Barcode_6	ACGTAGGGGGGCGCGTAACCACTAGCTC
Barcode_7	AGTGTCCCTTATTCTACTTTGAATTATC
Barcode_8	GAGTTACGGGTCAGGATCATTGCGCAGG
Barcode_9	GTGAGTCCGGCCTATCACGTTTGGTAGG
Barcode_10	TCAGGGGAAACGAGTTAAGCAGAGGCAG
Barcode_11	CTTGGTCCAATCGTATGCTAAGAGTAGC
Barcode_12	TCGGTTGCACGCGCGTCCTGCACGCTCG
Barcode_13	GTTCAAAAGCGGGAGTCCCGGTGAAACC
Barcode_14	GAGGGCTTCTACGAAATTGCCTCACCAT
Barcode_15	GGGATGCCCTGATAACGTAGAGTCTCAG
Barcode_16	TAGTACGGGCCACGATCTAAGTCGGCGG
Barcode_17	GTATGTGGGTACGATTGTAGGCTAGAAA
Barcode_18	AGAACTACCTACGTGGCCACGCAAGCCC
Barcode_19	TGCCTCAGAATGGGGTTTAACGTTCAGG
Barcode_20	ATGTTGGGACGCCAGACTGACACGCAAA
Barcode_21	CCTACACCTTCACCGAGTGTGGAGCAAG
Barcode_22	GTTTAACATTCGGGTATCGTCCTTACGA
Barcode_23	GCAATATTACTAGTCCGTGTCAGCCAGA
Barcode_24	CCGACTAGGTCGGCGATACCATAGCACC
Barcode_25	GATGACGCCTCACGTCCGTTGCCCCCTA
Barcode_26	GCTCCGTAGCAATTGATAAACCGCTGAG
Barcode_27	CTATCGGATTTTTAGATAAATTACGAGG
Barcode_28	TTTTTAATCGGTAGCGACGATATATGAC
Barcode_29	TTGATTCGTGGGGATCCGCGGATAGAAG
Barcode_30	CTAATATCTGATCCGTTATAACACCACA
Barcode_31	TTGCGGGCCTGGACATTCAATCGCATAC
Barcode_32	GCCTAGACCAAAAGTCGCCTTCAGGGCA
Barcode_33	ATAGGTAGGTAACGGAACTGACCGATAA
Barcode_34	CGTACTCCTTACGCCTTACACGGACAAT
Barcode_35	GATTCTCTGGTCATGAGGGAAGATCGAC
Barcode_36	TTTGGACGGGGTCACTACTCTTTACAGC
Barcode_37	TTCATGTATGCTGTCCTACCGCAGTAGC
Barcode_38	TACATAAGGTCGTCCTCCGGTCACGCCC
Barcode_39	CAGGCCGCTACTCAAGCGGGGATAGCTT
Barcode_40	CGCTTAGCTAGATTGCCCCTGGAAGCGC
Barcode_41	CGGAGTGTTTGCGAATCCCGAGCATTAG
Barcode_42	GTCTTTCGAATCGTTCCTTATAACCCGC
Barcode_43	AGGCTGAATCCTCCGGCGGCTGTTCTAA
Barcode_44	CAGCCAAACGCTAGCAAAATAGGATATA
Barcode_45	CAGACGCTGTCTACCTAGCAGAAGGGCT
Barcode_46	AAGGTATAATGTTCAAGCGTGGTGGCTA
Barcode_47	ACTGATCTCGGCGTAAATACACCGCAGA
Barcode_48	TATGTGTATCCTAAGGTACCTTTACGGG
Barcode_49	ACCGTCCTAGTGCGGATAGTGGGACATG
Barcode_50	GTCGATACGGGTCGCTAGATAGGTTGGG
Barcode_51	ACAGTCAGAGGCGCTTCAATCGTCGCCC
Barcode_52	GCGACTCATTTAATTTCAACGGGTAGAA
Barcode_53	TAGACAGAGGGAGCTATATTGCGCGGCG
Barcode_54	GGCATTCAATAGCCAGTAACCTTAGACA
Barcode_55	ACAAAAGTTGGCATTTCGTAGACAAAGC
Barcode_56	TTGCATACTTTGCAGGTTAGGGCCGAGA
Barcode_57	ACCGGAGTTCAAACGACTTTGGGTCCCT
Barcode_58	CACTGGTAGGCTCGCGGGTAAATAGAGC
Barcode_59	CCGAAGCTCAGTGTGATGTGCCTTCCGT
Barcode_60	GTACGTATGACTGTGTCCCAGCTATGAA
Barcode_61	GGACCTTCACGTCAGTCCCGCTGACCTT
Barcode_62	GACCAATTATGTTGGCCACGCCACATTA
Barcode_63	GTTTTCAAGGGCTCGACACAGCAGAACG
Barcode_64	AACCCACTATAGAAGGGCCTTCGAGATT
Barcode_65	GGAACCGTCAATTCCGCACTAATAATTT
Barcode_66	CATGGCTCCGAAACGCCGTGCAAAGTGC
Barcode_67	GCTTTCTCATTGTTACTCAGCGCACATT
Barcode_68	TCAATCCTGAGGCCGATACTAACAGTCG
Barcode_69	ATTCTGGCTATACCGAGCCCGGACTACT
Barcode_70	CAACCATGCTTCGGATAGCAATACCACG
Barcode_71	CTTGTCTCAAGAATAGTGCCCTATTCGT
Barcode_72	CCGTAGGTAGAAACATGTATGGCGCAGA
Barcode_73	CAGGGCCATGTCGTACTAGTAGCCTGCC
Barcode_74	CGAGCCCCCCAAAAAGTCGGGCTACCGC
Barcode_75	CATATAGGTACATGAAGCCCTACTGTAT
Barcode_76	CATACCGGATGGTAAACTCCGCGAACGC
Barcode_77	TGCAATGACCAGGCGATACGAATCGTAG
Barcode_78	CTCCCCCCCGCAGACACGCACGTGTTGC
Barcode_79	TTTGTGGTTAGCCTTGTCCCGGAACAGA
Barcode_80	CCCGGGTTGATGAGGAGTCACTCAACGA
Barcode_81	ATCCATGTGCGGGGAGAAGTCTGGTGGA
Barcode_82	CCCCGGTAATTTCGATGTTCGCCCCGGC
Barcode_83	TGGGCTGCTAGTAAGTCGGGAGCCCTTA
Barcode_84	CAGGCCGCTAACCTAAACTCGACCTCGA
Barcode_85	CAACGCAAGCGCTCTTGGTTTAAGCTGT
Barcode_86	GGCACACAACGATCTCCAAATGAACCAA
Barcode_87	GTAGCGCGACCTCTTTAACTGTAGAAGG
Barcode_88	GCTTATCGCATTCCGGGTGAAAACCTAG
Barcode_89	AGGGTAATGAAGCCCCGGGATCGAACAT
Barcode_90	TAACTAGCGGCCTATACCCGCTGCACAT
Barcode_91	CAGAGATAGGTCTCGACACTTACGCGGT
Barcode_92	GCCTAACGGCTGCAGCACGATTGCGCAA
Barcode_93	GACGCCATTAAGGTGGGTCACCTTTTCA
Barcode_94	GCATATGGCTAAATTCGCTAGTCGAGCG
Universal 2	GTCGACCTGCAGGCATGCAAGC

Table S4 List of barcodes used in chapter 2

List of unique barcode sequences used in this study, see **Figure 2-2** for more detailed barcode design.

Name	Sequence
	forward megamer
F_Barcode_1	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTCGTCAGTCGTAACTTGGGAACGCACATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_2	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTTTGGGGTTGGATAAATCTGTCGTGGTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_3	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTGAATAAGCGCGGTCCCTAATGTTGGTGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_4	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACCGGTTTATAGTTCGAACAGTCGCAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_5	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCCCGTGTGGTAACTACGCAAGCCTAACTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_6	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACGTAGGGGGGCGCGTAACCACTAGCTCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_7	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAGTGTCCCTTATTCTACTTTGAATTATCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_8	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGAGTTACGGGTCAGGATCATTGCGCAGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_9	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTGAGTCCGGCCTATCACGTTTGGTAGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_10	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTCAGGGGAAACGAGTTAAGCAGAGGCAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_11	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTTGGTCCAATCGTATGCTAAGAGTAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_12	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTCGGTTGCACGCGCGTCCTGCACGCTCGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_13	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTTCAAAAGCGGGAGTCCCGGTGAAACCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_14	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGAGGGCTTCTACGAAATTGCCTCACCATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_15	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGGGATGCCCTGATAACGTAGAGTCTCAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_16	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTAGTACGGGCCACGATCTAAGTCGGCGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_17	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTATGTGGGTACGATTGTAGGCTAGAAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_18	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAGAACTACCTACGTGGCCACGCAAGCCCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_19	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTGCCTCAGAATGGGGTTTAACGTTCAGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_20	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTATGTTGGGACGCCAGACTGACACGCAAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_21	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCTACACCTTCACCGAGTGTGGAGCAAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_22	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTTTAACATTCGGGTATCGTCCTTACGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_23	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCAATATTACTAGTCCGTGTCAGCCAGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_24	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCGACTAGGTCGGCGATACCATAGCACCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_25	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGATGACGCCTCACGTCCGTTGCCCCCTATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_26	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCTCCGTAGCAATTGATAAACCGCTGAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_27	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTATCGGATTTTTAGATAAATTACGAGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_28	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTTTTAATCGGTAGCGACGATATATGACTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_29	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTGATTCGTGGGGATCCGCGGATAGAAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_30	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTAATATCTGATCCGTTATAACACCACATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_31	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTGCGGGCCTGGACATTCAATCGCATACTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_32	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCCTAGACCAAAAGTCGCCTTCAGGGCATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_33	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTATAGGTAGGTAACGGAACTGACCGATAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_34	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCGTACTCCTTACGCCTTACACGGACAATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_35	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGATTCTCTGGTCATGAGGGAAGATCGACTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_36	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTTGGACGGGGTCACTACTCTTTACAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_37	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTCATGTATGCTGTCCTACCGCAGTAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_38	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTACATAAGGTCGTCCTCCGGTCACGCCCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_39	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGGCCGCTACTCAAGCGGGGATAGCTTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_40	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCGCTTAGCTAGATTGCCCCTGGAAGCGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_41	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCGGAGTGTTTGCGAATCCCGAGCATTAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_42	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTCTTTCGAATCGTTCCTTATAACCCGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_43	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAGGCTGAATCCTCCGGCGGCTGTTCTAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_44	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGCCAAACGCTAGCAAAATAGGATATATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_45	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGACGCTGTCTACCTAGCAGAAGGGCTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_46	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAAGGTATAATGTTCAAGCGTGGTGGCTATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_47	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACTGATCTCGGCGTAAATACACCGCAGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_48	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTATGTGTATCCTAAGGTACCTTTACGGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_49	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACCGTCCTAGTGCGGATAGTGGGACATGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_50	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTCGATACGGGTCGCTAGATAGGTTGGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_51	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACAGTCAGAGGCGCTTCAATCGTCGCCCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_52	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCGACTCATTTAATTTCAACGGGTAGAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_53	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTAGACAGAGGGAGCTATATTGCGCGGCGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_54	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGGCATTCAATAGCCAGTAACCTTAGACATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_55	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACAAAAGTTGGCATTTCGTAGACAAAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_56	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTGCATACTTTGCAGGTTAGGGCCGAGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_57	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTACCGGAGTTCAAACGACTTTGGGTCCCTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_58	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCACTGGTAGGCTCGCGGGTAAATAGAGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_59	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCGAAGCTCAGTGTGATGTGCCTTCCGTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_60	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTACGTATGACTGTGTCCCAGCTATGAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_61	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGGACCTTCACGTCAGTCCCGCTGACCTTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_62	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGACCAATTATGTTGGCCACGCCACATTATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_63	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTTTTCAAGGGCTCGACACAGCAGAACGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_64	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAACCCACTATAGAAGGGCCTTCGAGATTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_65	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGGAACCGTCAATTCCGCACTAATAATTTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_66	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCATGGCTCCGAAACGCCGTGCAAAGTGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_67	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCTTTCTCATTGTTACTCAGCGCACATTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_68	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTCAATCCTGAGGCCGATACTAACAGTCGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_69	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTATTCTGGCTATACCGAGCCCGGACTACTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_70	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAACCATGCTTCGGATAGCAATACCACGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_71	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTTGTCTCAAGAATAGTGCCCTATTCGTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_72	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCGTAGGTAGAAACATGTATGGCGCAGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_73	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGGGCCATGTCGTACTAGTAGCCTGCCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_74	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCGAGCCCCCCAAAAAGTCGGGCTACCGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_75	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCATATAGGTACATGAAGCCCTACTGTATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_76	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCATACCGGATGGTAAACTCCGCGAACGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_77	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTGCAATGACCAGGCGATACGAATCGTAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_78	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCTCCCCCCCGCAGACACGCACGTGTTGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_79	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTTTGTGGTTAGCCTTGTCCCGGAACAGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_80	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCCGGGTTGATGAGGAGTCACTCAACGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_81	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTATCCATGTGCGGGGAGAAGTCTGGTGGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_82	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCCCCGGTAATTTCGATGTTCGCCCCGGCTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_83	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTGGGCTGCTAGTAAGTCGGGAGCCCTTATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_84	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGGCCGCTAACCTAAACTCGACCTCGATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_85	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAACGCAAGCGCTCTTGGTTTAAGCTGTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_86	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGGCACACAACGATCTCCAAATGAACCAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_87	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGTAGCGCGACCTCTTTAACTGTAGAAGGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_88	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCTTATCGCATTCCGGGTGAAAACCTAGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_89	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTAGGGTAATGAAGCCCCGGGATCGAACATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_90	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTTAACTAGCGGCCTATACCCGCTGCACATTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_91	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTCAGAGATAGGTCTCGACACTTACGCGGTTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_92	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCCTAACGGCTGCAGCACGATTGCGCAATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_93	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGACGCCATTAAGGTGGGTCACCTTTTCATAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG
F_Barcode_94	CTGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGAATTCTAATACGACTCACTATAGGGCTAGCTGCATATGGCTAAATTCGCTAGTCGAGCGTAGATCTCAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCAG

	Reverse megamer
R_Barcode_1	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGATAATTCAAAGTAGAATAAGGGACACTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_2	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGAGCTAGTGGTTACGCGCCCCCCTACGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_3	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGTTAGGCTTGCGTAGTTACCACACGGGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_4	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTGCGACTGTTCGAACTATAAACCGGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_5	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACACCAACATTAGGGACCGCGCTTATTCAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_6	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAACCACGACAGATTTATCCAACCCCAAAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_7	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATGTGCGTTCCCAAGTTACGACTGACGAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_8	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCTGCGCAATGATCCTGACCCGTAACTCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_9	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCTACCAAACGTGATAGGCCGGACTCACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_10	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTGCCTCTGCTTAACTCGTTTCCCCTGAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_11	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTACTCTTAGCATACGATTGGACCAAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_12	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGAGCGTGCAGGACGCGCGTGCAACCGAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_13	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGTTTCACCGGGACTCCCGCTTTTGAACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_14	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATGGTGAGGCAATTTCGTAGAAGCCCTCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_15	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTGAGACTCTACGTTATCAGGGCATCCCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_16	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCGCCGACTTAGATCGTGGCCCGTACTAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_17	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTTCTAGCCTACAATCGTACCCACATACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_18	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGGCTTGCGTGGCCACGTAGGTAGTTCTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_19	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCTGAACGTTAAACCCCATTCTGAGGCAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_20	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTTGCGTGTCAGTCTGGCGTCCCAACATAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_21	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTTGCTCCACACTCGGTGAAGGTGTAGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_22	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCGTAAGGACGATACCCGAATGTTAAACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_23	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCTGGCTGACACGGACTAGTAATATTGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_24	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGTGCTATGGTATCGCCGACCTAGTCGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_25	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATAGGGGGCAACGGACGTGAGGCGTCATCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_26	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTCAGCGGTTTATCAATTGCTACGGAGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_27	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCTCGTAATTTATCTAAAAATCCGATAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_28	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGTCATATATCGTCGCTACCGATTAAAAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_29	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTTCTATCCGCGGATCCCCACGAATCAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_30	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATGTGGTGTTATAACGGATCAGATATTAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_31	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGTATGCGATTGAATGTCCAGGCCCGCAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_32	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATGCCCTGAAGGCGACTTTTGGTCTAGGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_33	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTATCGGTCAGTTCCGTTACCTACCTATAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_34	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATTGTCCGTGTAAGGCGTAAGGAGTACGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_35	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGTCGATCTTCCCTCATGACCAGAGAATCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_36	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTGTAAAGAGTAGTGACCCCGTCCAAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_37	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTACTGCGGTAGGACAGCATACATGAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_38	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGGCGTGACCGGAGGACGACCTTATGTAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_39	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAAGCTATCCCCGCTTGAGTAGCGGCCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_40	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCGCTTCCAGGGGCAATCTAGCTAAGCGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_41	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTAATGCTCGGGATTCGCAAACACTCCGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_42	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCGGGTTATAAGGAACGATTCGAAAGACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_43	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTAGAACAGCCGCCGGAGGATTCAGCCTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_44	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATATATCCTATTTTGCTAGCGTTTGGCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_45	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAGCCCTTCTGCTAGGTAGACAGCGTCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_46	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATAGCCACCACGCTTGAACATTATACCTTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_47	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCTGCGGTGTATTTACGCCGAGATCAGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_48	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCCGTAAAGGTACCTTAGGATACACATAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_49	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACATGTCCCACTATCCGCACTAGGACGGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_50	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCCAACCTATCTAGCGACCCGTATCGACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_51	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGGCGACGATTGAAGCGCCTCTGACTGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_52	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTCTACCCGTTGAAATTAAATGAGTCGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_53	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGCCGCGCAATATAGCTCCCTCTGTCTAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_54	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATGTCTAAGGTTACTGGCTATTGAATGCCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_55	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTTTGTCTACGAAATGCCAACTTTTGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_56	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCTCGGCCCTAACCTGCAAAGTATGCAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_57	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAGGGACCCAAAGTCGTTTGAACTCCGGTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_58	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCTCTATTTACCCGCGAGCCTACCAGTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_59	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAACGGAAGGCACATCACACTGAGCTTCGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_60	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTCATAGCTGGGACACAGTCATACGTACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_61	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAAGGTCAGCGGGACTGACGTGAAGGTCCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_62	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATAATGTGGCGTGGCCAACATAATTGGTCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_63	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGTTCTGCTGTGTCGAGCCCTTGAAAACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_64	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAATCTCGAAGGCCCTTCTATAGTGGGTTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_65	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAAATTATTAGTGCGGAATTGACGGTTCCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_66	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCACTTTGCACGGCGTTTCGGAGCCATGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_67	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAATGTGCGCTGAGTAACAATGAGAAAGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_68	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGACTGTTAGTATCGGCCTCAGGATTGAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_69	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAAGTAGTCCGGGCTCGGTATAGCCAGAATAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_70	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGTGGTATTGCTATCCGAAGCATGGTTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_71	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAACGAATAGGGCACTATTCTTGAGACAAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_72	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCTGCGCCATACATGTTTCTACCTACGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_73	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGGCAGGCTACTAGTACGACATGGCCCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_74	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCGGTAGCCCGACTTTTTGGGGGGCTCGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_75	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATACAGTAGGGCTTCATGTACCTATATGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_76	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCGTTCGCGGAGTTTACCATCCGGTATGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_77	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTACGATTCGTATCGCCTGGTCATTGCAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_78	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCAACACGTGCGTGTCTGCGGGGGGGAGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_79	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCTGTTCCGGGACAAGGCTAACCACAAAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_80	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCGTTGAGTGACTCCTCATCAACCCGGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_81	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCCACCAGACTTCTCCCCGCACATGGATAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_82	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAGCCGGGGCGAACATCGAAATTACCGGGGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_83	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATAAGGGCTCCCGACTTACTAGCAGCCCAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_84	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATCGAGGTCGAGTTTAGGTTAGCGGCCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_85	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAACAGCTTAAACCAAGAGCGCTTGCGTTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_86	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTGGTTCATTTGGAGATCGTTGTGTGCCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_87	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACCTTCTACAGTTAAAGAGGTCGCGCTACAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_88	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACTAGGTTTTCACCCGGAATGCGATAAGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_89	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATGTTCGATCCCGGGGCTTCATTACCCTAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_90	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAATGTGCAGCGGGTATAGGCCGCTAGTTAAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_91	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTAACCGCGTAAGTGTCGAGACCTATCTCTGAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_92	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATTGCGCAATCGTGCTGCAGCCGTTAGGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_93	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTATGAAAAGGTGACCCACCTTAATGGCGTCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG
R_Barcode_94	CTGATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTGAGATCTACGCTCGACTAGCGAATTTAGCCATATGCAGCTAGCCCTATAGTGAGTCGTATTAGAATTCTGTTCTACCAAGTAATCCATATTTCTAGAGGATCCCAG

Table S5 List of barcoded DNA megamers used in chapter 2

List of barcoded DNA megamer sequences used for SHERLOCK reactions in this study.

Table S6: Conditions for incubator scale experiments to test spore persistence

Surface Material	Count	Control Conditions	Perturbation Conditions	Direct Sample Technique	Transfer Sample Technique
Sand	3 control & 3 perturbed	**Temperature:** 25 C Fixed **Wind:** None **Rain:** None	**Temperature:** Weekly oscillation (25-35C) **Wind:** Fans, fixed speed **Rain:** weekly, varying intensity from spraying to watering can	Powersoil	Object Swab, NaOH lysis
Soil	3 control & 3 perturbed	**Temperature:** 25 C Fixed **Wind:** None **Rain:** None	**Temperature:** Weekly oscillation (25-35C) **Wind:** Fans, fixed speed **Rain:** weekly, varying intensity from spraying to watering can	Powersoil	Object Swab, NaOH lysis
Carpet	2 control & 2 perturbed	**Temperature:** 25 C Fixed **Humidity:** 40-50% Controlled **Cleaning:** None	**Temperature:** 25 C Fixed **Humidity:** 40-50% Controlled **Cleaning:** Vacuuming	Surface Swab, NaOH lysis	Object Swab, NaOH lysis
Wood	2 control & 2 perturbed	**Temperature:** 25 C Fixed **Humidity:** 40-50% Controlled **Cleaning:** None	**Temperature:** 25 C Fixed **Humidity:** 40-50% Controlled **Cleaning:** Sweeping	Surface Swab, NaOH lysis	Object Swab, NaOH lysis

Table S6 Description of incubator experiment in chapter 2

Description matrix of conditions used in the incubator experiment, including control conditions, perturbations, and sampling techniques.

114
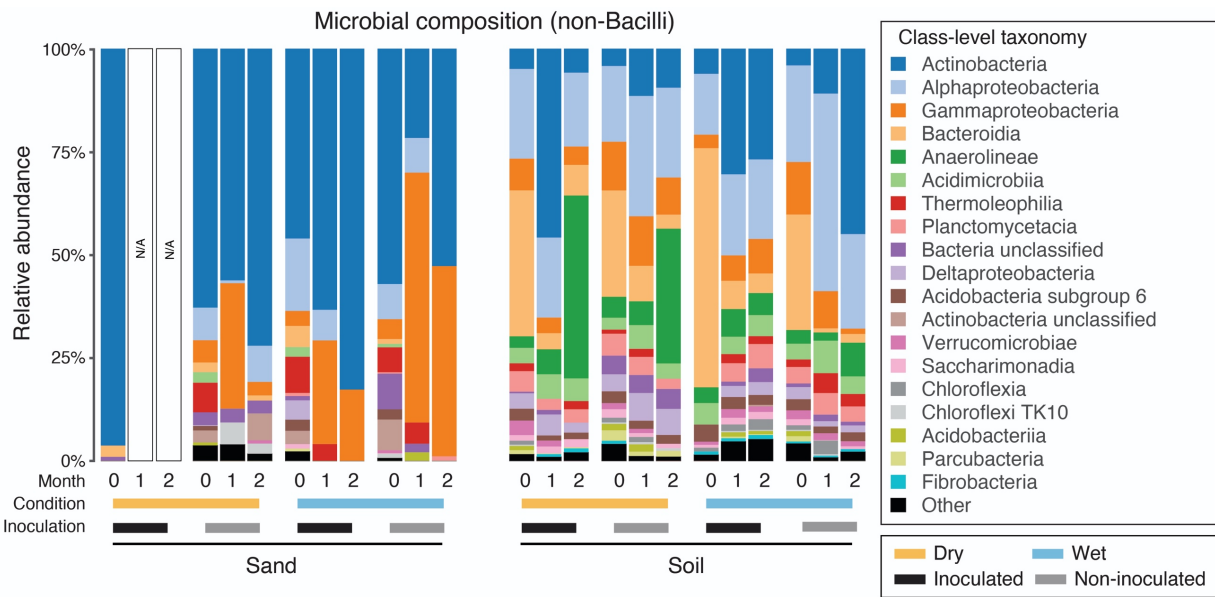
Supp Fig 2-1 Changes in microbial composition of non-inoculated and inoculated sand and soil.

(A) Relative abundance of soil bacterial taxa in samples taken from sand or soil surfaces during a 2-month long experiment. Relative abundance was measured using 16S metagenomics and classified to the class level. The sampling locations differed in surface material, wet/dry status, and inoculation status. Sand samples had extremely low biomass, and most reads from inoculated samples aligned to Bacilli. In all other samples, Bacilli comprised <0.01% of the library.
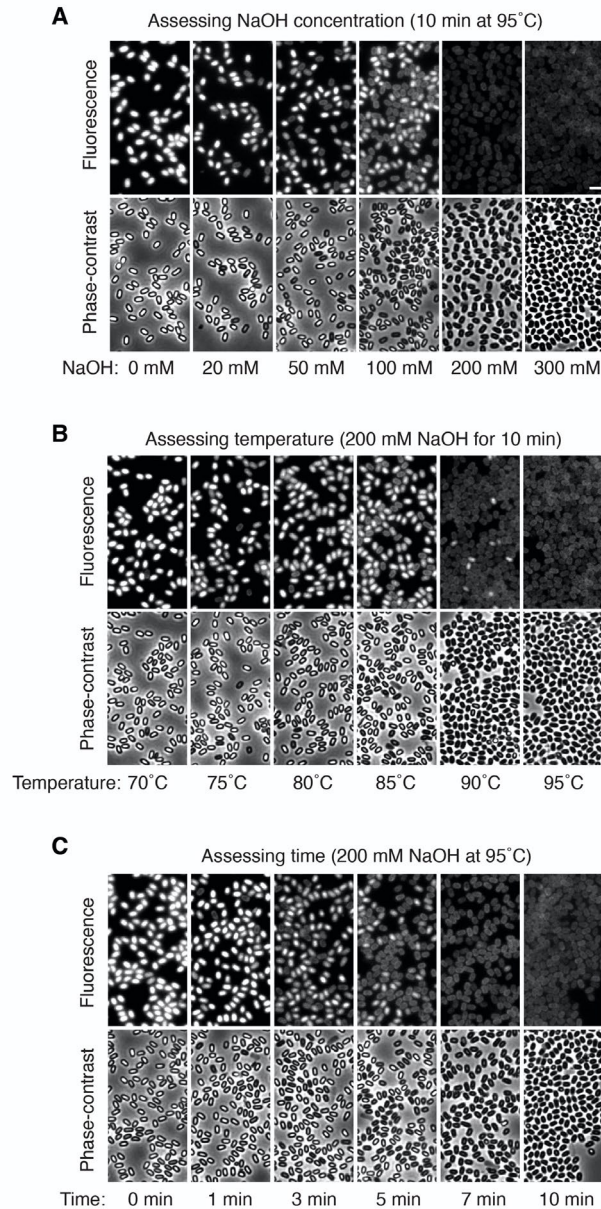
Supp Fig 2-2 Optimization of *B. subtilis* BMS lysis.

(A) To rapidly assess the efficacy of lysis, Δ9 *B. subtilis* BMS harboring a cytoplasmic red fluorescent protein (mScarlet) were analyzed by fluorescence and phase-contrast microscopy after treatment. ~2 x 10^6 Δ9 BMS were resuspended in 50 µL of NaOH at the indicated concentrations and heated for 10 min at 95˚C. (B) ~2 x 10^6 Δ9 BMS were resuspended in 50 µL of 200 mM NaOH and heated at the indicated temperatures for 10 min. (C) ~2 x 10^6 Δ9 BMS were resuspended in 50 µL of 200 mM NaOH and heated at 95˚C for the indicated amount of time. After treatment, BMS were pelleted, washed and resuspended in PBS. An aliquot was then analyzed by fluorescence and phase-contrast microscopy. Loss of fluorescence correlated with the transition from phase-bright to phase-dark BMS. Scale bar indicates 2 µm.

116
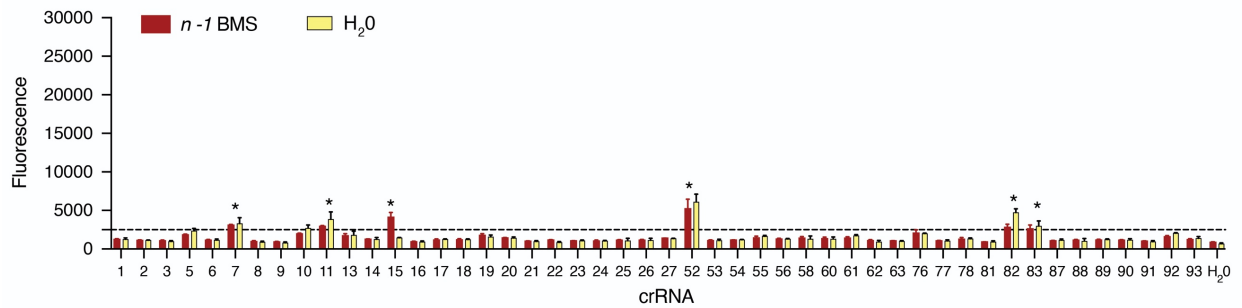
Supp Fig 2-3 *in vivo* BMS and crRNA cross-reactivity assay.

Bars are the SHERLOCK signal from reactions with the *n-1* BMS reactions in red and H_2O RPA reactions in yellow. * denotes crRNA with high background or cross-reactivity (n = 3 technical replicates, bars represent mean + s.e.m.).
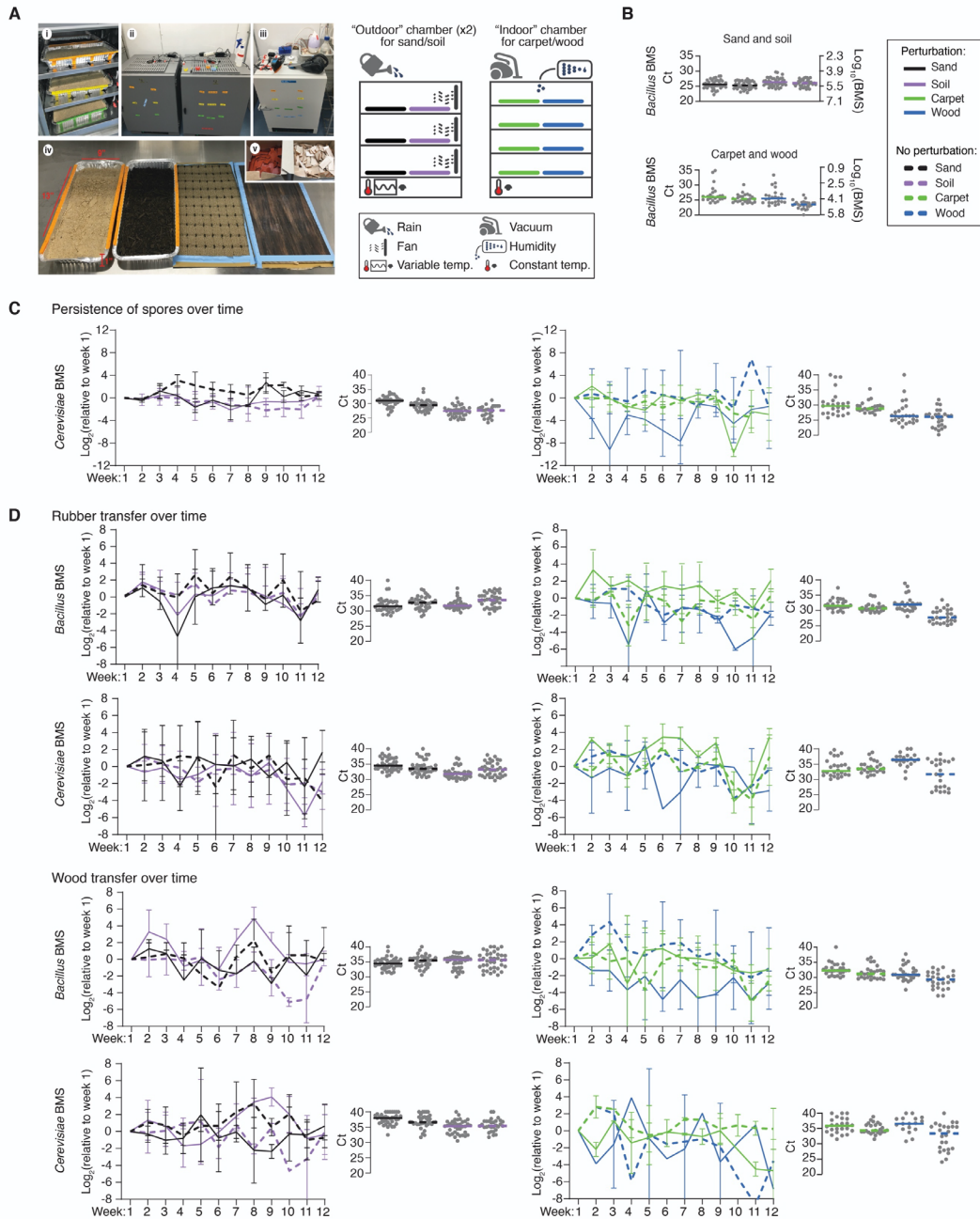
Supp Fig 2-4 Persistence, transferability and maintenance of BMS.

(A) Photos and schematic of lab incubator scale experiments and simulated wind, rain, and vacuuming. (B) Dot plot of real-time qPCR Ct values (left y-axis) and BMS numbers (right y-axis) based on a qPCR standard curve. Each dot represents a different sampling location each week, grouped by trays in the same treatment group for all 12 weeks. (C) BMS persisted on sand, soil, carpet and wood surfaces for at least 3 months. Plots show BMS count number (relative to week 1 values) and qPCR Ct values. (D) BMS were transferable for at least 3 months after inoculation from all 4 test surfaces. Rubber and wood objects placed on inoculated surfaces were used for testing BMS transferability. Plots show relative BMS count numbers and qPCR Ct values.
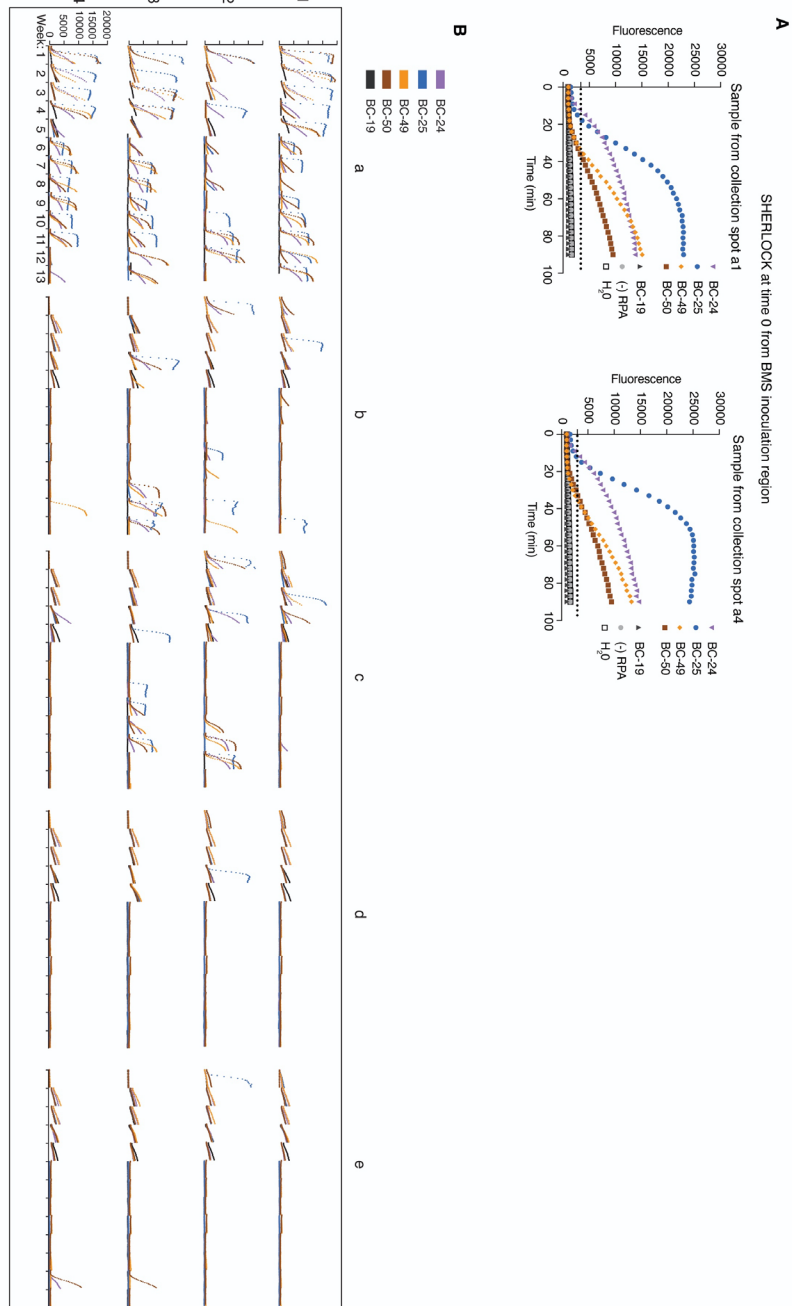
Supp Fig 2-5 Persistence of BMS over time.

(A) SHERLOCK detection of BC-24, 25, 49, 50 BMS from samples taken immediately after BMS inoculation, shown as fluorescence time-courses. BC-19, H_2O, and (-) RPA are negative controls. (B) SHERLOCK detection indicates BMS persistence over 3 months with or without perturbation. SHERLOCK time-course data (y-axis: fluorescence, x-axis: time in min) is shown for each BMS (colors) at each of 20 collection spots (a-e, 1-4 in **Figure 2-10**) for 13 weeks. Signal above the threshold are counted and scored in **Figure 2-11**. The Cas13a batch was changed at week 6, which changed baseline activity, but did not change conclusions of the experiment
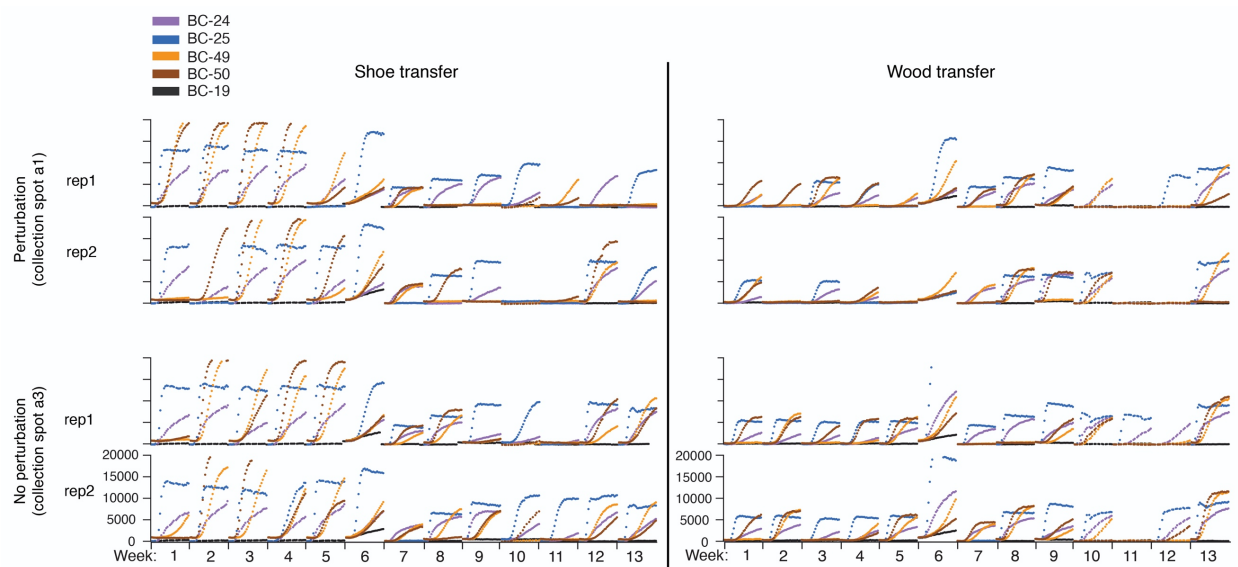
Supp Fig 2-6 Transfer of BMS over time.

SHERLOCK can detect BMS on shoes and wood that come in contact with an inoculated surface. SHERLOCK time-course data (y-axis: fluorescence, x-axis: time in min) for transfer onto shoes or wood is shown for each barcode (colors), for 2 replicates from 2 different locations from the BMS-inoculated region (a1 and a3 in **Figure 2-10**) over 13 weeks. BC-19 is a negative control. The Cas13a batch was changed at week 6, which changed baseline activity, but did not change conclusions of the experiment.
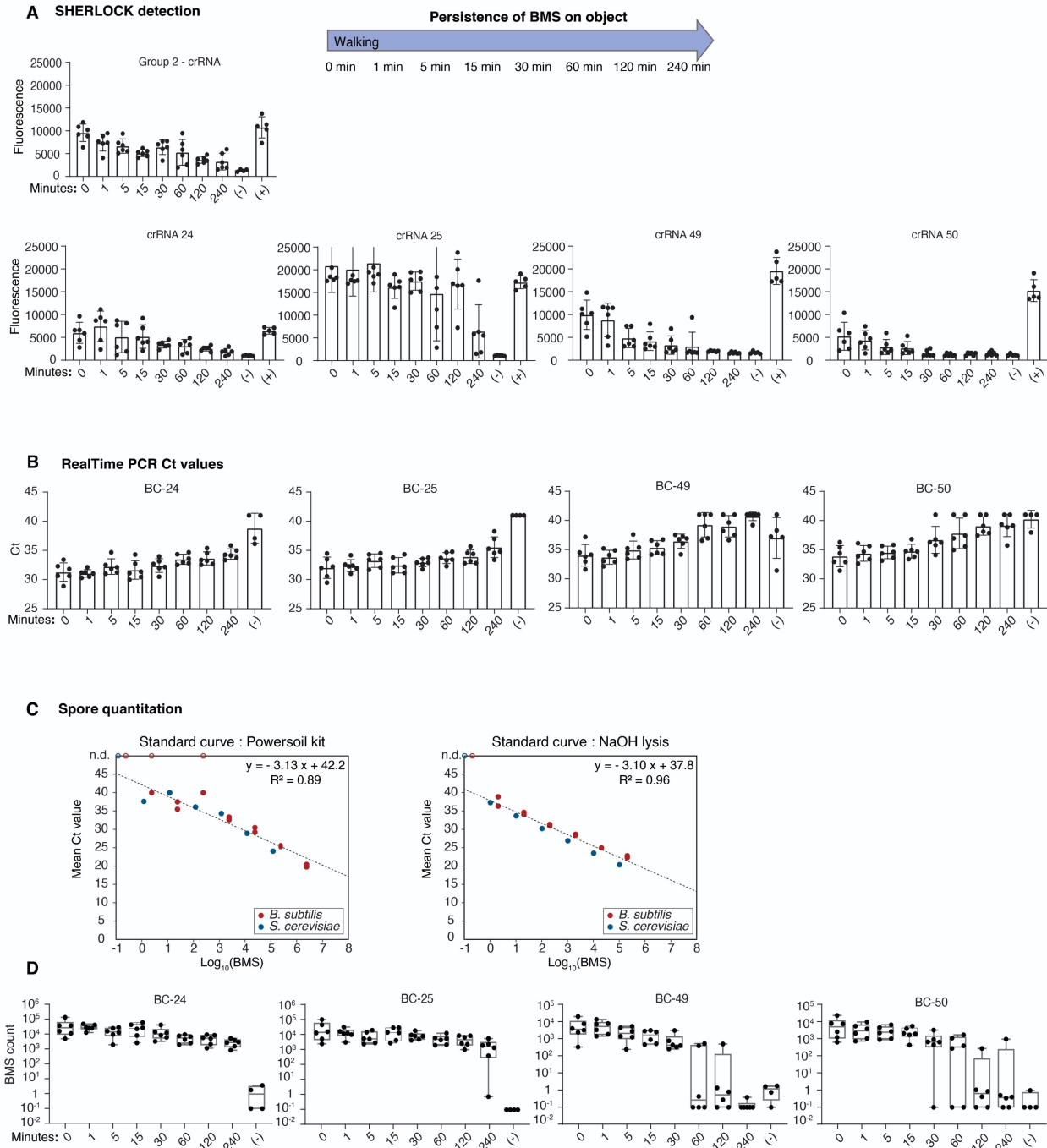
Supp Fig 2-7 Persistence of BMS on object after transfer.

(A) SHERLOCK and (B) qPCR showing BC-24 and 25 *B. subtilis* BMS and BC-49 and 50 *S. cerevisiae* BMS retained on shoes after up to 4 hours of walking on non-inoculated areas. (C) Standard curves constructed from known BMS quantities using either PowerSoil or NaOH lysis methods. (D) Estimating the number of BMS on each shoe from Ct values using the qPCR standard curve.
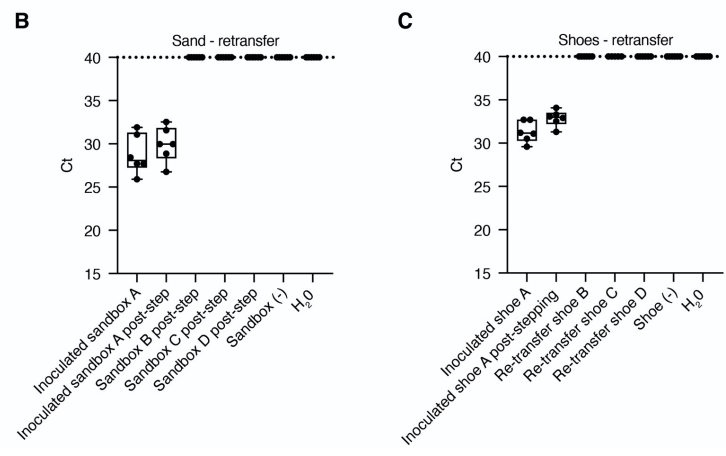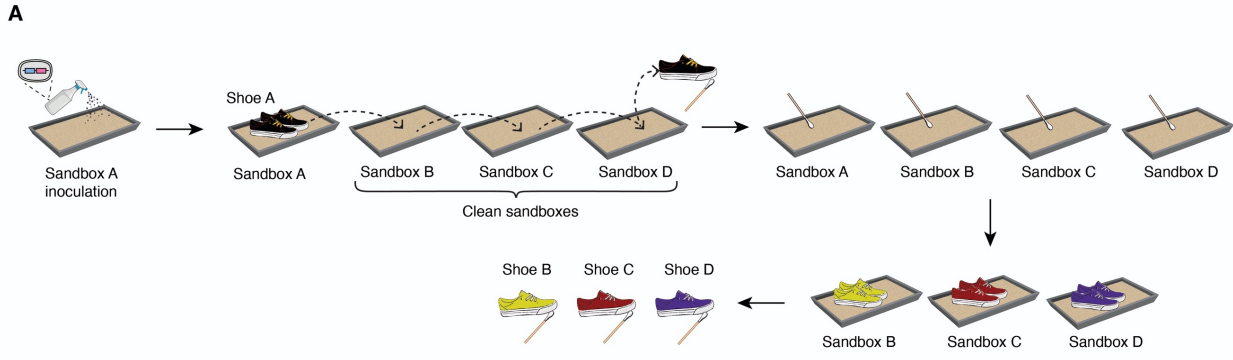
Supp Fig 2-8 Retransfer of BMS to non-inoculated surfaces.

(A) Sandbox A was inoculated with a 4 BMS mixture. Shoe A stepped in inoculated sandbox A and subsequently stepped into 3 clean sandboxes B, C, and D. Sand from sandboxes B, C, and D after shoe A stepping and was sampled and qPCR was performed using BMS-specific primers to quantify BMS. New shoes B, C, and D were stepped in sandboxes B, C, and D, respectively after shoe A had stepped in each sandbox. Shoes B, C, and D were then sampled for BMS. (B) qPCR results of BMS in sand samples from sandboxes A, B, C, and D. (-): non-inoculated sand; H_2O: qPCR water negative control; the horizontal dashed line is the threshold of detection. (C) qPCR results of BMS from swabbed shoes A, B, C, and D. (-): non-inoculated sand; H_2O: qPCR water negative control; the horizontal dashed line is the threshold of detection.
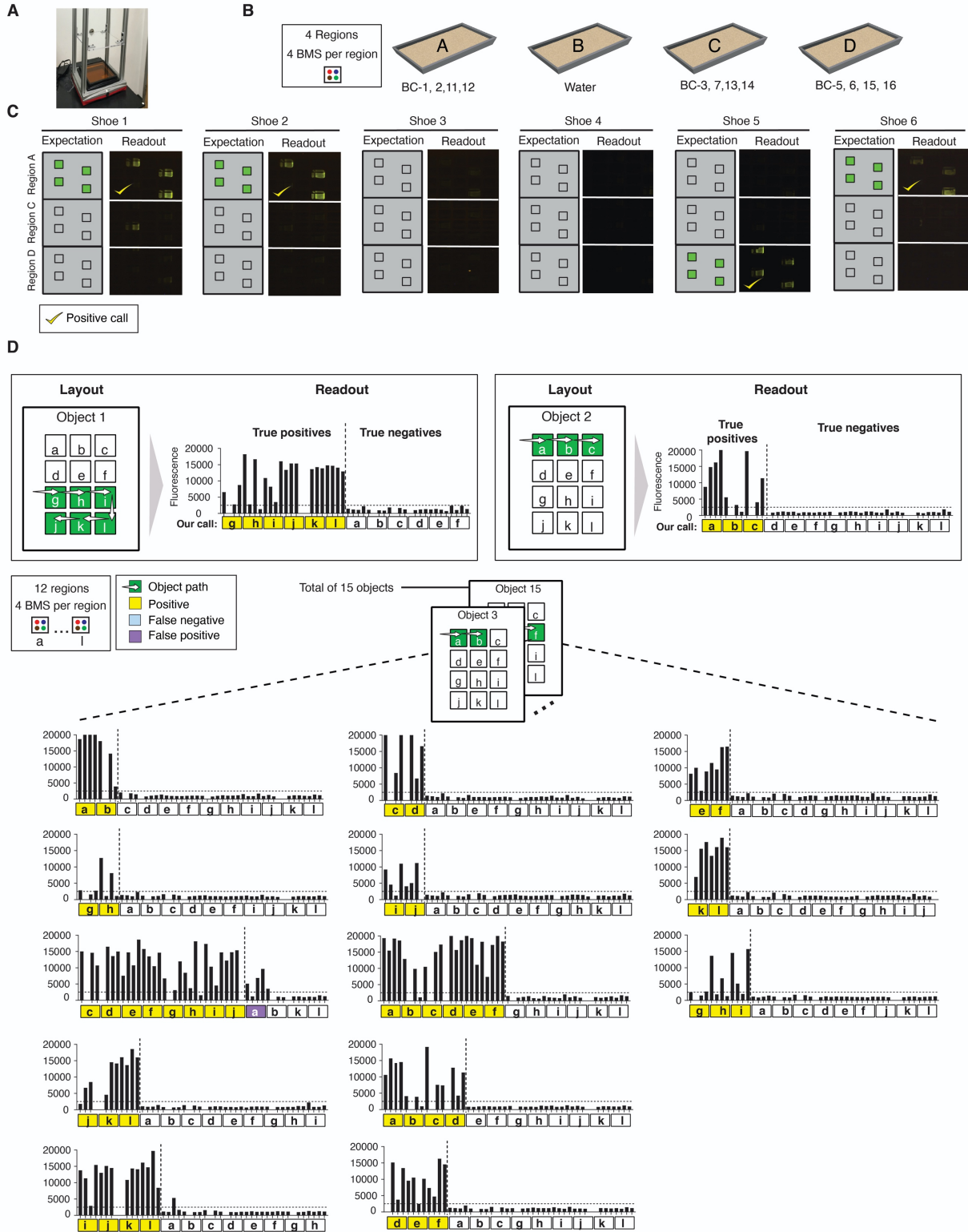
Supp Fig 2-9 Object provenance using 4 BMS per region.

(A) This layout was used to test our field-deployable detection system: a portable light source and an orange acrylic filter were used for imaging SHERLOCK signals. A mobile phone (not shown) was used for photographing the SHERLOCK reaction plates. (B) Four trays were filled with sand, and inoculated by

spraying either 4 unique BMS or H_2O. (C) SHERLOCK reaction plate image for six shoe samples that had stepped in one of the 4 trays. Reaction signal matches expectation shown on left. (D) A subsequent experiment, in which 12 sand regions (squares: a-l) were inoculated with 4 unique BMS each. 15 shoe samples took different paths through the regions and were tested for all possible BMS using SHERLOCK. Endpoint fluorescence values are plotted, with expected positives to the left of the vertical dashed line and expected negatives to the right of the vertical dashed line. The call for each lettered region is denoted by color (yellow: true positive, blue: false negative, purple: false positive), the horizontal dashed line is the threshold for positive calls.
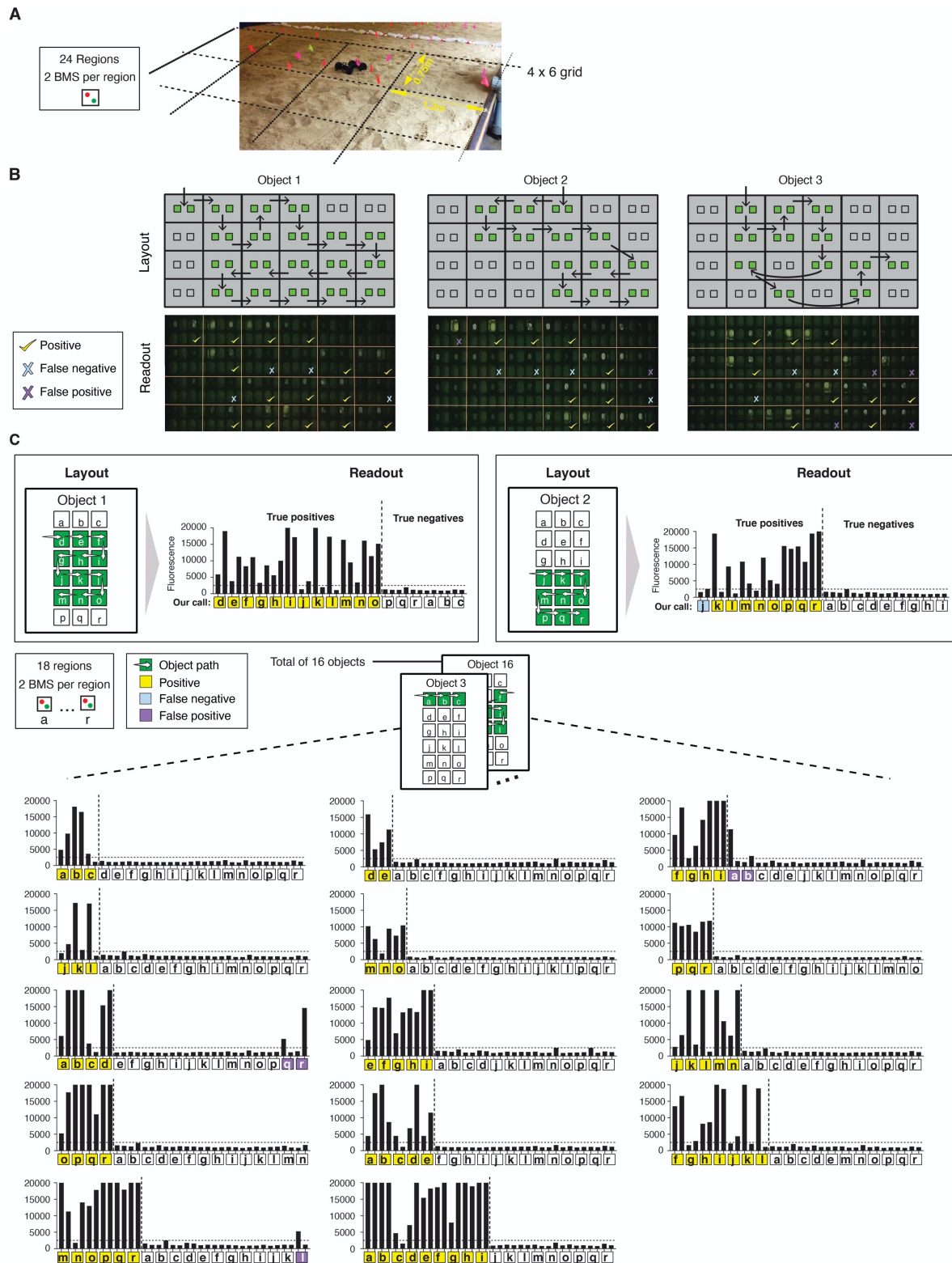
Supp Fig 2-10 Object provenance using 2 BMS per region.

(A) A grid of 24 regions was laid out on a clean area of sand, and each region was inoculated with 2 unique

BMS. (B) Top: path of object overlaid on reaction plate. Bottom: photograph of SHERLOCK reaction plate,

overlaid with correct or incorrect calls. The call for each region is indicated in the corresponding area of the plate (yellow check: true positive, blue cross: false negative, purple cross: false positive). (C) A subsequent experiment, in which 18 trays of sand (squares: a-r) were inoculated with 2 unique barcodes each. 16 shoe samples that took different paths through the regions were tested using SHERLOCK. Endpoint fluorescence values are plotted, with expected positives to the left of the vertical dashed line and expected negatives to the right of the vertical dashed line. The call for each lettered region is denoted by color (yellow: true positive, blue: false negative, purple: false positive), the horizontal dashed line is the threshold for positive calls.

Supp Fig 2-11 Object provenance using 1 BMS per region.

(A) A grid of 20 regions (squares: a-t) were laid out on clean sand, soil, carpet, and wood surfaces, and each region was inoculated with a single unique BMS. 8 remote control car samples, and 24 shoe samples

that took different paths through the regions on different surfaces were tested using SHERLOCK. Endpoint fluorescence values are plotted, with expected positives to the left of the vertical dashed line and expected negatives to the right of the vertical dashed line. The call for each lettered region is denoted by color (yellow: true positive, blue: false negative, purple: false positive), the horizontal dashed line is the threshold for positive calls.

BMS per region	Total number of regions	Objects (*n*)	False positive rate		False negative rate
1 ●	640	32	16.9% (67/397)		19.8% (48/243)
2 ●●	384	20	>= 1 BMS positive per region	4.9% (12/245)	9.4% (13/139)
			2 BMS positive per region	0.4% (1/245)	36.7% (51/139)
4 ●●●●	216	22	>= 1 BMS positive per region	2.6% (4/154)	0% (0/62)
			>= 2 BMS positive per region	0.6% (1/154)	0% (0/62)
			>= 3 BMS positive per region	0.6% (1/154)	12.9% (8/62)
			4 BMS positive per region	0% (0/154)	66.1% (41/62)

Supp Fig 2-12 Statistics for determining provenance with different numbers of BMS.

All the experiments were collated together, and the false positive and false negative rates were calculated using the listed threshold criteria for a positive call. Yellow highlighting indicates the suggested usage

```
                1                          28              1                          28
Plant A    TAGTACGGGCCACGATCTAAGTCGGCGG     Plant J    GTATGTGGGTACGATTGTAGGCTAGAAA
Leaf       TAGTACGGGCCACGATCTAAGTCGGCGG     Leaf       GTATGTGGGTACGATTGTAGGCTAGAAA
Soil       TAGTACGGGCCACGATCTAAGTCGGCGG     Soil       GTATGTGGGTACGATTGTAGGCTAGAAA

Plant B    GGGATGCCCTGATAACGTAGAGTCTCAG     Plant K    TCAGGGGAAACGAGTTAAGCAGAGGCAG
Leaf       GGGATGCCCTGATAACGTAGAGTCTCAG     Leaf       TCAGGGGAAACGAGTTAAGCAGAGGCAG
Soil       GGGATGCCCTGATAACGTAGAGTCTCAG     Soil       TCAGGGGAAACGAGTTAAGCAGAGGCAG

Plant C    GAGGGCTTCTACGAAATTGCCTCACCAT     Plant L    GTGAGTCCGGCCTATCACGTTTGGTAGG
Leaf       GAGGGCTTCTACGAAATTGCCTCACCAT     Leaf       GTGAGTCCGGCCTATCACGTTTGGTAGG
Soil       GAGGGCTTCTACGAAATTGCCTCACCAT     Soil       GTGAGTCCGGCCTATCACGTTTGGTAGG

Plant D    GTTCAAAAGCGGGAGTCCCGGTGAAACC     Plant M    GAGTTACGGGTCAGGATCATTGCGCAGG
Leaf       GTTCAAAAGCGGGAGTCCCGGTGAAACC     Leaf       GAGTTACGGGTCAGGATCATTGCGCAGG
Soil       GTTCAAAAGCGGGAGTCCCGGTGAAACC     Soil       GAGTTACGGGTCAGGATCATTGCGCAGG

Plant E    CTTGGTCCAATCGTATGCTAAGAGTAGC     Plant N    AGTGTCCCTTATTCTACTTTGAATTATC
Leaf       CTTGGTCCAATCGTATGCTAAGAGTAGC     Leaf       AGTGTCCCTTATTCTACTTTGAATTATC
Soil       CTTGGTCCAATCGTATGCTAAGAGTAGC     Soil       AGTGTCCCTTATTCTACTTTGAATTATC

Plant F    ACGTAGGGGGGCGCGTAACCACTAGCTC     Plant O    AGAACTACCTACGTGGCCACGCAAGCCC
Leaf       ACGTAGGGGGGCGCGTAACCACTAGCTC     Leaf       AGAACTACCTACGTGGCCACGCAAGCCC
Soil       ACGTAGGGGGGCGCGTAACCACTAGCTC     Soil       AGAACTACCTACGTGGCCACGCAAGCCC

Plant G    CCCCGTGTGGTAACTACGCAAGCCTAAC     Plant Q    ATGTTGGGACGCCAGACTGACACGCAAA
Leaf       CCCCGTGTGGTAACTACGCAAGCCTAAC     Leaf       ATGTTGGGACGCCAGACTGACACGCAAA
Soil       CCCCGTGTGGTAACTACGCAAGCCTAAC     Soil       ATGTTGGGACGCCAGACTGACACGCAAA

Plant H    TGAATAAGCGCGGTCCCTAATGTTGGTG     Plant R    CCTACACCTTCACCGAGTGTGGAGCAAG
Leaf       TGAATAAGCGCGGTCCCTAATGTTGGTG     Leaf       CCTACACCTTCACCGAGTGTGGAGCAAG
Soil       TGAATAAGCGCGGTCCCTAATGTTGGTG     Soil       CCTACACCTTCACCGAGTGTGGAGCAAG

Plant I    TCGTCAGTCGTAACTTGGGAACGCACAT     Plant S    CTTTGGGGTTGGATAAATCTGTCGTGGT
Leaf       TCGTCAGTCGTAACTTGGGAACGCACAT     Leaf       CTTTGGGGTTGGATAAATCTGTCGTGGT
Soil       TCGTCAGTCGTAACTTGGGAACGCACAT     Soil       CTTTGGGGTTGGATAAATCTGTCGTGGT
```

Supp Fig 2-13 Sanger sequencing of leaf and soil samples to the BMS inoculated on plants.
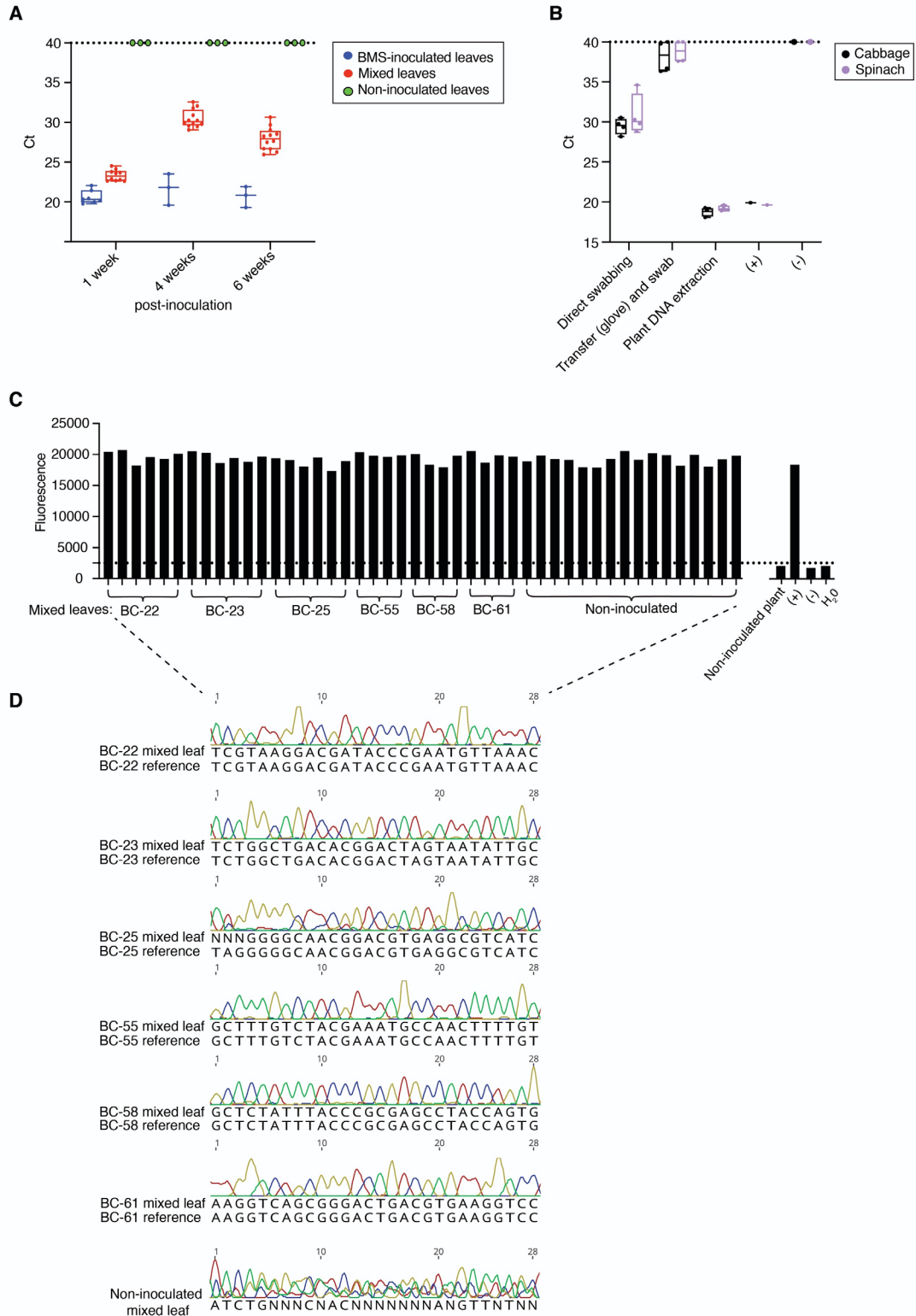
Supp Fig 2-14 BMS remain on plant leaves and can determine leaf provenance.

(A) qPCR measurement of different BMS-inoculated leaves that had been mixed with each other, as shown in **Fig 4C**. Blue: different BMS-inoculated leaves that were mixed; qPCR signal of the specific BMS inoculated on the sampled leaf (leaf numbers, n = 7 for week 1, n = 3 for weeks 4 and 6). Red: different

BMS-inoculated leaves that were mixed. qPCR signal from a different BMS than that of the sampled leaf (leaf numbers, n = 10 for week 1, n = 12 for weeks 4 and 6). Green: qPCR signal of non-mixed and non-inoculated (leaf numbers, n = 3 for week 1, 4 and 6). **(B)** qPCR measurement of BMS from swabbing directly the inoculated plant, swabbing the gloves after the glove touched the inoculated plant, or DNA extractions of different BMS sprayed cabbage and spinach. (-): non-BMS sprayed plant is the negative control. **(C)** SHERLOCK endpoint fluorescence values for BMS-inoculated leaves after mixing. Detection reactions were performed for multiple BMS for each leaf, though each leaf had been inoculated with a single BMS. **(D)** Sanger sequencing alignment identifies the barcode that had been inoculated. Non-inoculated leaves that had been mixed with inoculated leaves has positive SHERLOCK signal, but does not specifically align to any of the barcode references.

Appendix B: Supplementary material related to Chapter 3

Oligo name	Alternate name	Oligo sequence	Description of use
CCMS052	S gene FP1	GTTTTTCTTGTTTTATTGCCACTAGTCTCTA	SARS-CoV-2 S Gene forward primer
CCMS055	S gene FP2	TCTTGTTTTATTGCCACTAGTCTCTAGTCAGT	SARS-CoV-2 S Gene forward primer
CCMS019	S gene FP3	CTCTAGTCAGTGTGTTAATCTTACAACCAGAACT	SARS-CoV-2 S Gene forward primer
CCMS053	S gene FP4	TCAGTGTGTTAATCTTACAACCAGAACTCAAT	SARS-CoV-2 S Gene forward primer
CCMS054	S gene FP5	GTGTTAATCTTACAACCAGAACTCAATTACCCC	SARS-CoV-2 S Gene forward primer
CCMS067	S gene RP1	GAATGTAAAACTGAGGATCTGAAAACTTTG	SARS-CoV-2 S gene reverse primer
CCMS014	S gene RP2	AGAACAAGTCCTGAGTTGAATGTAAAACTGAGG	SARS-CoV-2 S gene reverse primer
CCMS056	S gene RP3	AAGAAAGGTAAGAACAAGTCCTGAGTTGAATGT	SARS-CoV-2 S gene reverse primer
CCMS030	S gene RP4	CCATTGGTCCCAGAGACATGTATAGCATGG	SARS-CoV-2 S gene reverse primer
CCMS008	S gene RP5	CTCTTAGTACCATTGGTCCCAGAGACATGT	SARS-CoV-2 S gene reverse primer
CCMS068	S gene RP6	ATCAAACCTCTTAGTACCATTGGTCCCAGA	SARS-CoV-2 S gene reverse primer
CCMS073		/56-FAM/GAATGTAAAACTGAGGATCTGAAAACTTTG	SARS-CoV-2 S gene reverse primer
JQ217		CAACTTCCTCAAGGAACAACATTGCCAAAA	SARS-CoV-2 N gene forward primer
JQ289		GGCTTCTACGCAGAAGGGAGCAGAGGCGGCAG	SARS-CoV-2 N gene forward primer
JQ214		CAACTGGCAGTAACCAGAATGGAGAACGCA	SARS-CoV-2 N gene forward primer
JQ215		TCGAGGACAAGGCGTTCCAATTAACACCAA	SARS-CoV-2 N gene forward primer
JQ216		CCTAGGAACTGGGCCAGAAGCTGGACTTCC	SARS-CoV-2 N gene forward primer
JQ218		AACTTCTCCTGCTAGAATGGCTGGCAATGG	SARS-CoV-2 N gene forward primer
JQ219		ATGAAACTCAAGCCTTACCGCAGAGACAGA	SARS-CoV-2 N gene forward primer
CCMS041		CAATCCTGCTAACAATGCTGCAATCGTGCTA	SARS-CoV-2 N gene forward primer
CCMS047		CACCAAAAGATCACATTGGCACCCGCAATCC	SARS-CoV-2 N gene forward primer
CCMS051		CTGAGGGAGCCTTGAATACACCAAAAGATCACA	SARS-CoV-2 N gene forward primer
JQ235		/56-FAM/TGGAGTTGAATTTCTTGAACTGTTGCGACT	SARS-CoV-2 N gene reverse primer
JQ223		TGGAGTTGAATTTCTTGAACTGTTGCGACT	SARS-CoV-2 N gene reverse primer
JQ220		CTTCCTTGCCATGTTGAGTGAGAGCGGTGA	SARS-CoV-2 N gene reverse primer
JQ221		CTTGGACTGAGATCTTTCATTTTACCGTCA	SARS-CoV-2 N gene reverse primer
JQ222		GCAGGATTGCGGGTGCCAATGTGATCTTTT	SARS-CoV-2 N gene reverse primer
JQ224		TGGCCTTGTTGTTGTTGGCCTTTACCAGAC	SARS-CoV-2 N gene reverse primer
JQ225		TTGAGTCAGCACTGCTCATGGATTGTTGCA	SARS-CoV-2 N gene reverse primer
XL12		CTCCAGGGAAACTTTGGTGA	MERS N gene RT-qPCR primer
XL13		CCCATAAAAGCACTGGCTGT	MERS N gene RT-qPCR primer
JQ297		GGCGCAAGAATTCAGAACCAGAGATA	HCoV-229E N gene RT-qPCR primer
JQ298		GGATTCCGAGATTGAGATTTGGATTCAC	HCoV-229E N gene RT-qPCR primer
JQ303		ACTACTCAAGAAGCTATCCCTACTAGG	HCoV-HKU1 N gene RT-qPCR primer
JQ304		GAGAACGTGAACCTGGTCGACTATTAG	HCoV-HKU1 N gene RT-qPCR primer
CCMS083		ACGTTGATGTAGGGCCAGAT	MERS S gene RT-qPCR primer
CCMS084		ACCGTCAGCCTTAGAAACATCA	MERS S gene RT-qPCR primer
CCMS086		CCGGTGCACCACTTTTGATG	SARS-CoV S gene RT-qPCR primer
CCMS087		CCCCCTCATAGATGAAGTATGTTGA	SARS-CoV S gene RT-qPCR primer
CCMS077		CTGTTTGCAACGGCTGTGTT	HCoV-229E S gene RT-qPCR primer
CCMS078		ACCACCACTCTCAACAGCAA	HCoV-229E S gene RT-qPCR primer
CCMS080		AAACACCACAGTTCCTCGCA	HCoV-HKU1 S gene RT-qPCR primer
CCMS081		CCCAAACCATAAGAAACATCCACA	HCoV-HKU1 S gene RT-qPCR primer
CC039		TCATGGTCCTACATTGTGGAAA	2009 H1N1 influenza RT-qPCR primer
CC040		GACACTGAGCTCAATTGCT	2009 H1N1 influenza RT-qPCR primer
GAPDH F		GGACTCATGACCACAGTCCA	GAPDH RT-qPCR primer
GAPDH R		GATGTTCTGGAGAGCCCCG	GAPDH RT-qPCR primer
JQ241		/5Biosg/TCAAGCCTCTTCTCGTTCCTCATCACGT	SARS-CoV-2 N gene Hybridization Probe
JQ312		/5Biosg/CAAGCCTCTTCTCGT	SARS-CoV-2 N gene Hybridization Probe
CCMS069		/5Biosg/TGCATACACTAATTCTTTCACACGTGGT	SARS-CoV-2 S gene Hybridization Probe

Table S7 List of all primers used in chapter 3.

Table S8 Data for all patient sample RT-qPCR and eRPA assays performed in chapter 3.
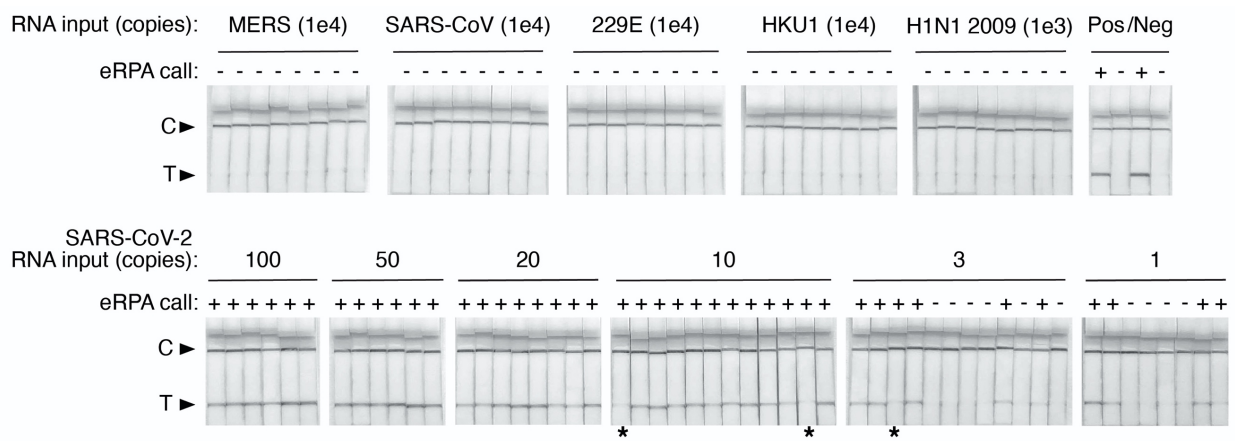
Supp Fig 3-1 Lateral flow strip readouts for all S gene data.

Individual strips (data from Figure 3-9) are labeled with the test call made within 20 mins of detection (positive (+) or negative (-)). The positive (Pos.) eRPA control is 1,000 copies of synthetic full genome SARS-CoV-2 RNA and the negative (Neg.) eRPA control is a water-only input. Images taken for the purpose of display were allowed to dry which reduced the intensity of some weak bands (labeled with asterisks).
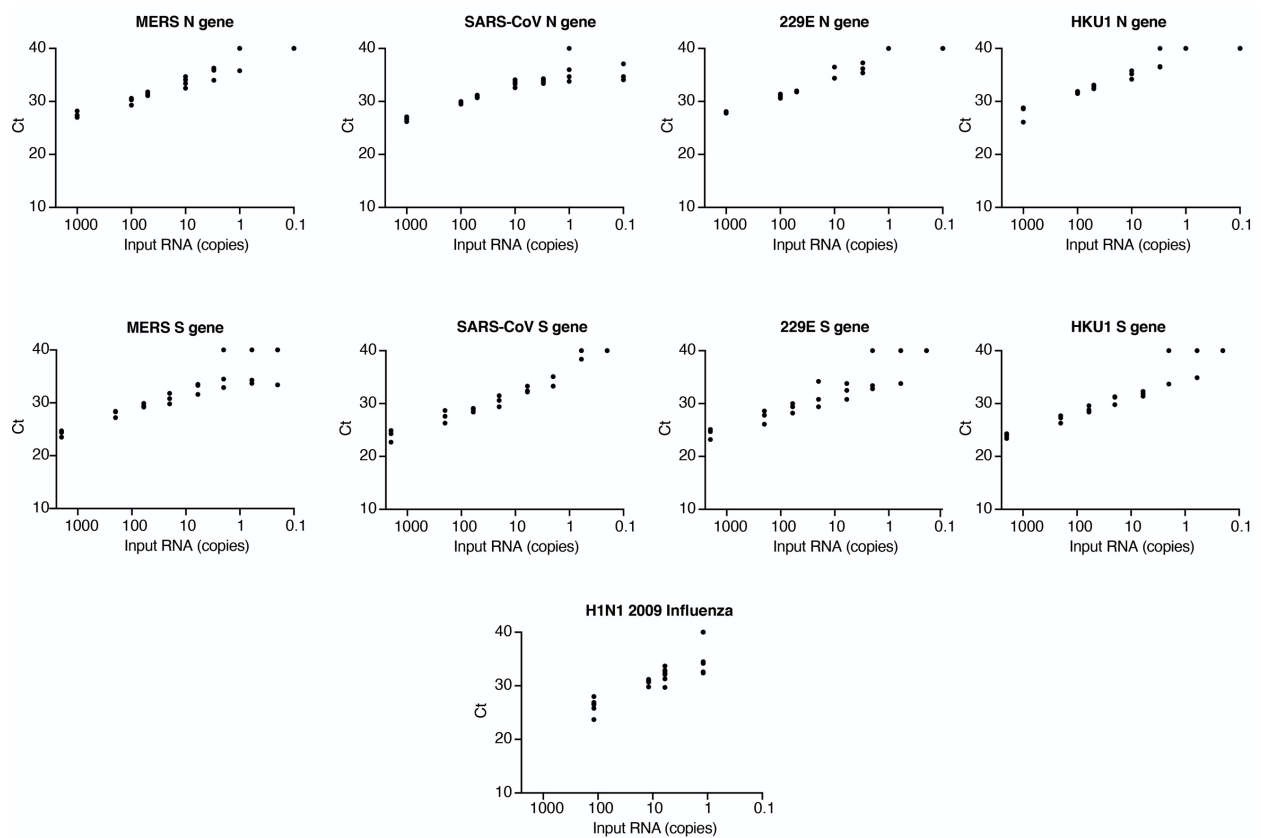
Supp Fig 3-2 RT-qPCR quantification of in vitro transcribed (IVT) RNA
MERS, SARS-CoV, HCoV-229E, and HCoV-HKU1 used as specificity control tests in eRPA; N gene shown in top row and S gene in middle row. Bottom graph RT-qPCR quantification of RNA extracted from 2009 H1N1 Influenza.

Controls: Pos. Neg.
RT-qPCR: 28.7 –
eRPA call: + + – – – –

C ►
T ►

Patient ID: 1 2 3 4 5 6 7 8 9 10 11 12
RT-qPCR: 28.2 , 30.1 – 22.4 , 25.6 25.6 , 27.7 24.1 , 27.3 18.8 , 20.2 22.8 , 25.3 21.5 , 24.3 15.2 , 17.0 24.2 , 26.0 35.3 , 35.1 25.4 , 27.8
eRPA call: + + – – + + + + + + + + + + + + + + + + – + + +

C ►
T ►

Patient ID: 13 14 15 17 18 19 20 21 22 23 24 46 47 48 49 50 51 52
RT-qPCR: 18.1 , 20.1 22.4 , 24.3 18.0 , 20.7 21.5 , 23.6 25.2 , 27.0 26.6 , 28.3 19.5 , 21.9 32.4 , 33.5 27.5 , 29.0 20.7 , 22.6 31.5 , 33.5 – – – – – – 19.9 , 22.2
eRPA call: + – – – – – – – – – – – – + +

C ►
T ►

Patient ID: 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
RT-qPCR: –
eRPA call: –

C ►
T ►

Unextracted

Supp Fig 3-3 Lateral flow strips detection of N-gene from crude clinical samples

Lateral flow strip readouts of all eRPA tests from crude patient samples summarized in **Figure 3-23**. Individual strips are labeled with the eRPA test call made within 20 mins of detection (positive (+) or negative (-)). The positive (Pos.) eRPA control is 100 copies of synthetic full genome SARS-CoV-2 RNA and the negative (Neg.) eRPA control is a water-only input.
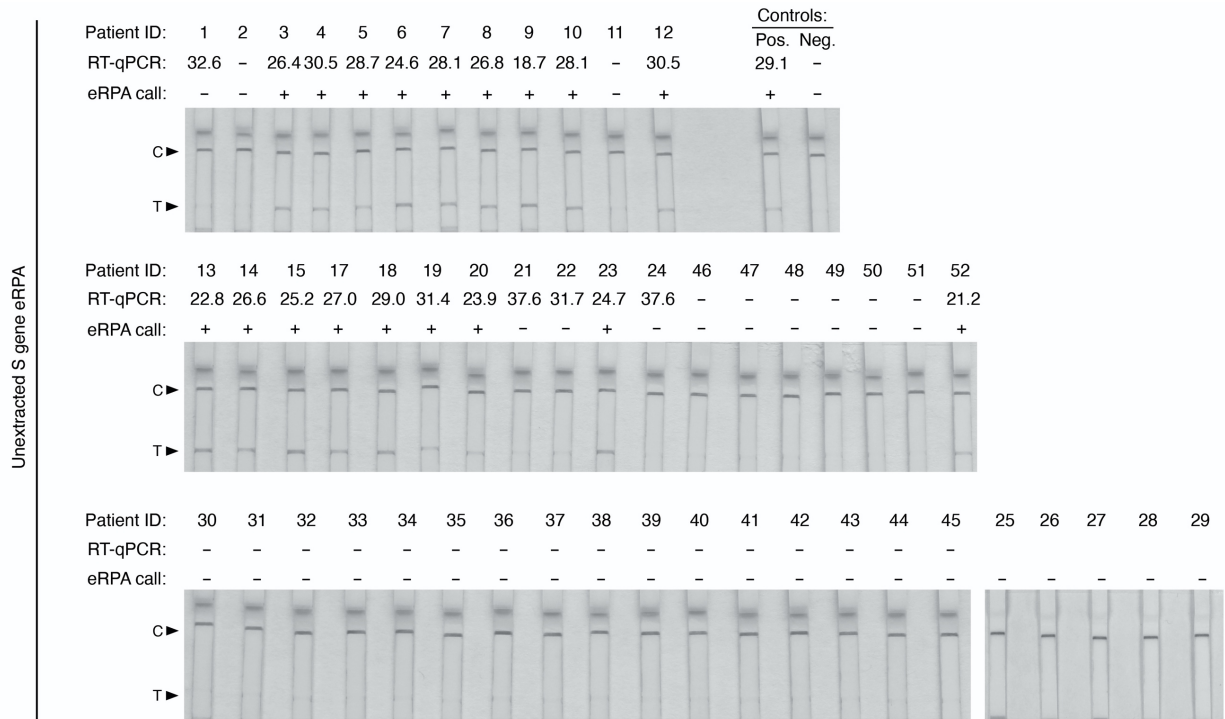
Supp Fig 3-4 Lateral flow strips detection of S-gene from crude clinical samples

Lateral flow strip readouts of S gene eRPA performed on patient samples of known infection status. Individual strips are labeled with the test call made within 20 mins of detection (positive (+) or negative (-)). The positive (Pos.) eRPA control is 100 copies of synthetic full genome SARS-CoV-2 RNA and the negative (Neg.) eRPA control is a water-only input. Negative control samples 25-29 were not screened by RT-qPCR.
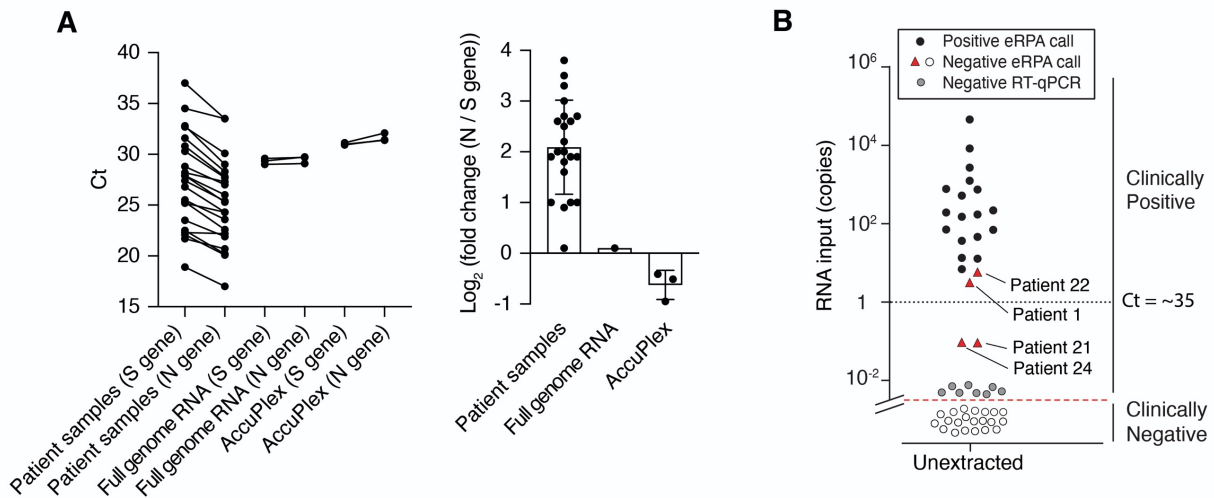
Supp Fig 3-5 Comparison of N and S gene detection from crude patient samples

(A) Comparison of Ct values obtained by RT-qPCR targeting SARS-CoV-2 N and S genes on the same input patient samples. (Left) Matched patient samples (n=22 biologically independent samples) are connected by a solid line. Synthetic full genome SARS-CoV-2 RNA (n=3 biologically independent samples) and AccuPlex packaged SARS-CoV-2 (n=3 biologically independent samples) were used as controls as they both contain an equal amount of N and S gene. (Right) Difference between the Ct values for each patient sample (n=22 biologically independent samples) is plotted with mean value 2.1 fold +/- 0.9 SD. For synthetic full genome SARS-CoV-2 RNA the mean value is 0.1 fold; For AccuPlex packaged SARS-CoV-2 the mean value is -0.6 fold +/- 0.3 SD. (B) Comparison between S gene SARS-CoV-2 eRPA and one-step RT-qPCR performed on the same input samples. The y axis is RNA copies in patient input samples determined by one-step RT-qPCR with standard curve.
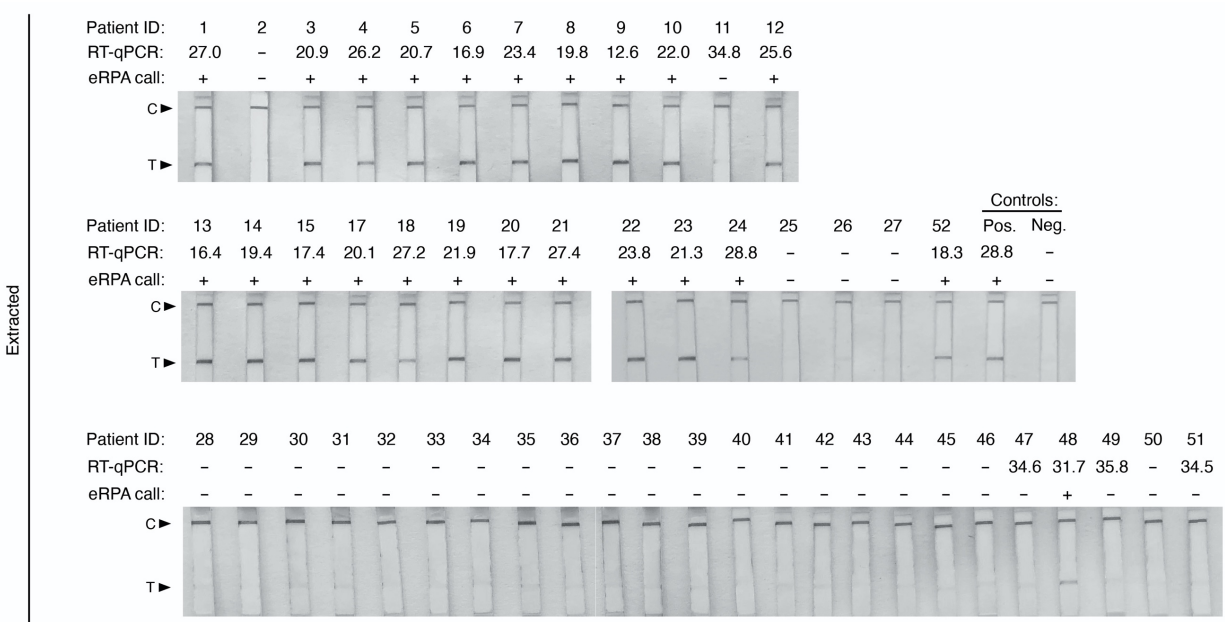
Supp Fig 3-6 Lateral flow strips detection of N-gene from extracted clinical samples

Lateral flow strip readouts of all eRPA tests from extracted patient samples summarized in **Figure 3-23**. Individual strips are labeled with the eRPA test call made within 20 mins of detection (positive (+) or negative (-)). The positive (Pos.) eRPA control is 100 copies of synthetic full genome SARS-CoV-2 RNA and the negative (Neg.) eRPA control is a water-only input.
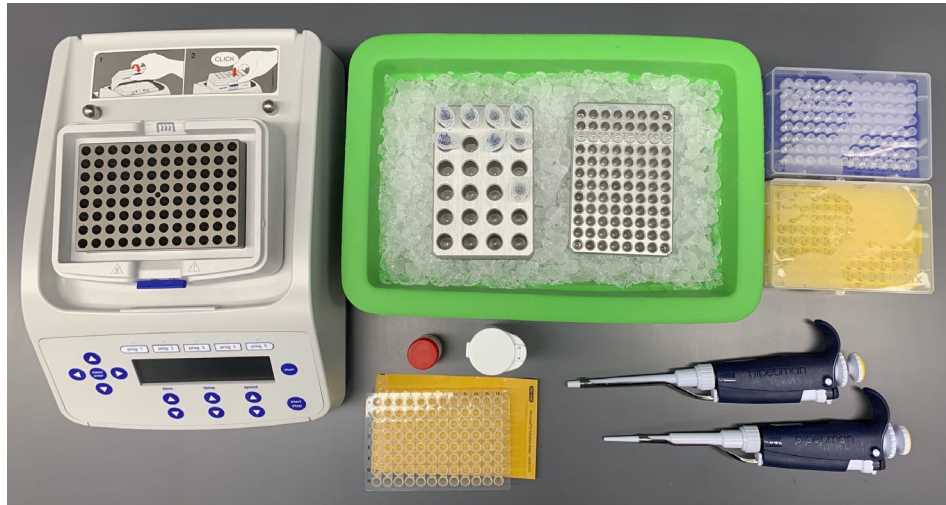
Supp Fig 3-7 Equipment required for eRPA assay

eRPA only requires a limited set of equipment including micropipettes and disposable plastic tips, a heat block capable of reaching 42°C and 94°C, and plastic microtubes or multi-well plates.

Type	Titer	Correct Category	Non-Blinded	Blinded 1	Blinded 2	Blinded 3		HIV Scale	
Spike In	500	3	3	3	3	3		3	>>100 cps/rxn
Spike In	500	3	3	2	3	2		2	~100 cps/rxn
Spike In	250	2	3	3	3	3		1	~10 cps/rxn
Spike In	125	2	2	2	3	2		0	Negative
Spike In	62.5	2	2	2	3	3			
HIV +	76	2*	2	2	2	2			
HIV +	64	2	2	2	2	2			
HIV +	26	1	2	2	1	1			
HIV +	23	1	2	2	1	0			
HIV +	5.9	1*	1	2	1	0			
HIV +	3.9	1	1	2	2	2			
HIV +	2.6	1	1	2	1	1			
HIV +	1.7	1	1	1	1	0			
Negative	0	0*	0	2	1	0			
Negative	0	0*	0	2	1	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			

Table S-9 Classification of HIV lateral flow strips.

Classification calls were made for each trio of strips by three individuals who were blinded to the distribution of titers. A call was also made by one individual unblinded to the distribution of sample titers. Samples marked with * were hybridized with 5 µM probe, causing some background signal in the reference band. All other samples were hybridized using 2.5 µM probe.

Type	Titer	Correct Category	Non-Blinded	Blinded 1	Blinded 2	Blinded 3		COVID Scale	
COVID +	1400000	5	5	5	5	4		5	>>1E4 cps/rxn
COVID +	640000	5	4	4	4	3		4	~1E4 cps/rxn
COVID +	140000	5	5	5	5	4		3	~1E3 cps/rxn
COVID +	100000	5	5	5	5	5		2	~1E2 cps/rxn
COVID +	98000	5	5	5	5	5		1	<<1E2 cps/rxn
COVID +	26000	4	4	5	5	4		0	Negative
COVID +	19000	4	4	4	4	4			
COVID +	11000	4	4	5	4	5			
COVID +	10000	4	4	4	4	4			
COVID +	5000	4	4	4	4	4			
COVID +	4600	3	3	3	3	3			
COVID +	3000	3	3	4	4	3			
COVID +	460	2	3	4	3	3			
COVID +	280	2	3	3	4	3			
COVID +	280	2	2	3	3	3			
COVID +	160	2	2	2	2	2			
COVID +	144	2	2	2	1	1			
COVID +	126	2	2	2	2	2			
COVID +	62	2	2	2	2	2			
COVID +	42	1	1	1	1	1			
COVID +	14	1	1	2	2	2			
COVID +	2.4	1	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	1	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	1	1	0			
Negative	0	0	0	0	0	0			
Negative	0	0	0	0	0	0			

Table S-10 Classification of COVID lateral flow strips**.**

Classification calls were made for each trio of strips by three individuals who were blinded to the distribution of titers. A call was also made by one individual unblinded to the distribution of sample titers.

| Name	Analyte	Oligo Type	Description	Sequence
JQ42	Synthetic barcode	RPA	Barcode RPA for	GGGATCCTCTAGAAATATGGATTACTTGGTAGAACAG
JQ64	Synthetic barcode	RPA	Barcode RPA rev	CCAACGTAAAATTGTGCCTTTGGACTTAAAATGGCGTCG
DP004	Synthetic barcode	qPCR	Barcode 1 qPCR for	CTAGCTCGTCAGTCGTAACTTGGGAACGCACATC
DP019	Synthetic barcode	qPCR	Barcode 5 qPCR for	CTAGCCCCCGTGTGGTAACTACGCAAGCCTAACC
JQ24	Synthetic barcode	qPCR	Barcode qPCR rev	GATAAACACAGGAAACAGCTATGACCATGATTACG
JQ217	SARS-CoV-2	RPA	SARS-CoV-2 RPA for	CAACTTCCTCAAGGAACAACATTGCCAAAA
JQ235	SARS-CoV-2	RPA	SARS-CoV-2 RPA rev	/56-FAM/TGGAGTTGAATTTCTTGAACTGTTGCGACT
CM47	SARS-CoV-2	qPCR	pre-amplification qPCR rev	TGGAGTTGAATTTCTTGAACTG
CM48	SARS-CoV-2	qPCR	Target pre-amplification qPCR for	TCAAGCCTCTTCTCGTTC
CM49	SARS-CoV-2	qPCR	Reference pre-amplification qPCR for	AGTAAGCAGAATCCCGC
CM56	SARS-CoV-2	qPCR	post-amplification qPCR for	GGCTTCTACGCAGAAGGGAG
CM57	SARS-CoV-2	qPCR	Target post-amplification qPCR rev	ACGTGATGAGGAACGAGAAGAGG
CM58	SARS-CoV-2	qPCR	Reference post-amplification qPCR rev	CTCTCTTGACAGCGGGATTCTG
CM46	SARS-CoV-2	TaqMan	TaqMan qPCR for	CAACTTCCTCAAGGAACAAC
JQ383	SARS-CoV-2	TaqMan	TaqMan Target Probe	5'FAM-TCAAGCCTCTTCTCGTTCCTCATCACGT-3'Quencher
JQ382	SARS-CoV-2	TaqMan	TaqMan Reference Probe	5'VIC-AGTAAGCAGAATCCCGCTGTCAAGAGAG-3'Quencher
JQ241	SARS-CoV-2	LFA probe	SARS-CoV-2 biotin probe	/5Biosg/TCAAGCCTCTTCTCGTTCCTCATCACGT
JQ170	SARS-CoV-2 / HIV	LFA probe	Reference digoxigenin probe	/5DigN/AGTAAGCAGAATCCCGCTGTCAAGAGAG
JQ125	HIV	RPA	HIV RPA for	TGAAGGGGCAGTAGTAATACAAGATAATAGT
JQ127	HIV	RPA	HIV RPA rev	/56-FAM/TGTCTACTTGCCACACAATCATCACCTGCCA
CM59	HIV	qPCR	HIV qPCR for	GACATAAAAGTAGTGCCAAGAAGAAAAGCAA
CM60	HIV	qPCR	HIV qPCR rev	GCCATCTGTTTTCCATAATCCCTAATGATC
CM55	HIV	qPCR	Reference qPCR for	GTAAGCAGAATCCCGCTGTCAAG
JQ130	HIV	LFA probe	HIV biotin probe	/5Biosg/ATAAAAGTAGTGCCAAGAAGAAAAGCAA
JQ92	Synthetic barcode	RPA	Barcode RPA rev	ttacgccaagcttgcatgcctgcaggtcgacttaa
JQ75	Synthetic barcode	RPA	Mutational library with 91% bases as wild-type and 3% of each of 3 remaining bases	cacggtcgcagc(N1:03039103)(N1:03039103)(N1:03039103)(N2:91030303)(N3:03030391)(N4:03910303)(N4:03910303)(N3:03030391)(N4:03910303)(N3:03030391)(N2:91030303)(N1:03039103)(N3:03030391)(N3:91030303)(N2:91030303)(N2:91030303)(N3:03030391)(N2:91030303)(N3:03030391)(N1:03039103)(N1:03039103)(N2:91030303)(N3:03030391)(N3:03030391)(N2:91030303)(N4:03910303)(N3:03030391)(N3:03030391)(N1:03039103)(N1:03039103)(N3:03030391)(N2:91030303)(N1:03039103)(N2:91030303)(N2:91030303)(N4:03910303)(N2:91030303)(N1:03039103)gctgcgcgacg
JQ76	Synthetic barcode	RPA	14 bp UMI library	GATAAACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTTAAGGATAATTCAAAGTAGAATAAGGGACACTGCTAGCCC(N:25252525)(N)(N)(N)(N)(N)(N)(N)(N)(N)(N)(N)(N)CGTCGCGCAGC
JQ77	Synthetic barcode	pre-RPA	forward sequencing adaptor	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGcacggtcgcagc
JQ83	Synthetic barcode	pre-RPA	forward sequencing adaptor (staggered 1N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNcacggtcgcagc
JQ84	Synthetic barcode	pre-RPA	forward sequencing adaptor (staggered 2N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNcacggtcgcagc
JQ85	Synthetic barcode	pre-RPA	forward sequencing adaptor (staggered 3N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNcacggtcgcagc
JQ93	Synthetic barcode	pre-RPA	reverse sequencing adaptor	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGttacgccaagcttgc
JQ94	Synthetic barcode	pre-RPA	reverse sequencing adaptor (staggered 1N)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNttacgccaagcttgc
JQ95	Synthetic barcode	pre-RPA	reverse sequencing adaptor (staggered 2N)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNttacgccaagcttgc
JQ96	Synthetic barcode	pre-RPA	reverse sequencing adaptor (staggered 3N)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNttacgccaagcttgc
S504		pre-RPA	Illumina i5 index	S504
N715		pre-RPA	Illumina i7 index	N715
JQ97	Synthetic barcode	post-RPA	forward sequencing adaptor	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGcacggtcgcagcgggatcctctagaaatatgg
JQ98	Synthetic barcode	post-RPA	forward sequencing adaptor (staggered 1N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNCAGcacggtcgcagcgggatcctctagaaatatgg
JQ99	Synthetic barcode	post-RPA	forward sequencing adaptor (staggered 2N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNCAGcacggtcgcagcgggatcctctagaaatatgg
JQ100	Synthetic barcode	post-RPA	forward sequencing adaptor (staggered 3N)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNCAGcacggtcgcagcgggatcctctagaaatatgg
S505		post-RPA	Illumina i5 index	S505
N716		post-RPA	Illumina i7 index	N716

Table S-11 Primers and Probes.

| Name	Analyte	Description	Sequence	Notes
Barcode 5 Target Plasmid	Synthetic Barcode	Synthetic barcode sequence from Qian et al. Science 2020	gggatcctctagaaatatggattacttggtagaacagaattctaatacg actcactatagggctagcCCCCGTGTGGTAACTACGCAAG CCTAACCcttaagtcgacctgcaggcatgcaagcttggcgtaatcat ggtcatagctgtttcctgtgtttatctgcagcgtacgacgccattttaagtcc aaaggcacaattttacgttgg	Barcode 5 capitalized
Barcode 1 Reference Plasmid	Synthetic Barcode	Synthetic barcode sequence from Qian et al. Science 2020	gggatcctctagaaatatggattacttggtagaacagaattctaatacg actcactatagggctagcTCGTCAGTCGTAACTTGGGAAC GCACATcttaagtcgacctgcaggcatgcaagcttggcgtaatcatg gtcatagctgtttcctgtgtttatctgcagcgtacgacgccattttaagtcca aaggcacaattttacgttgg	Barcode 1 capitalized
SARS-CoV-2 Target Plasmid	SARS-CoV-2	N-gene target sequence from Qian et al. Nature Communications 2020	caacttcctcaaggaacaacattgccaaaaggcttctacgcagaagg gagcagaggcggcagTCAAGCCTCTTCTCGTTCCTCATC ACGTagtcgcaacagttcaagaaattcaactcca	Target probe sequence capitalized
SARS-CoV-2 Reference Plasmid	SARS-CoV-2	Reference inserted into N-gene sequence from Qian et al. Nature Communications 2020	caacttcctcaaggaacaacattgccaaaaggcttctacgcagaagg gagcagaggcggcagAGTAAGCAGAATCCCGCTGTCAA GAGAGagtcgcaacagttcaagaaattcaactcca	Reference probe sequence capitalized
HIV Target Plasmid	HIV	Integrase gene target sequence	tgaaggggcagtagtaatacaagataatagtgacATAAAAGTAG TGCCAAGAAGAAAAGCAAagatcattagggattatggaaaa cagatggcaggtgatgattgtgtggcaagtagaca	Target probe sequence capitalized
HIV Reference Plasmid	HIV	Reference inserted into integrase gene target sequence	tgaaggggcagtagtaatacaagataatagtgacAGTAAGCAG AATCCCGCTGTCAAGAGAGagatcattagggattatggaaa acagatggcaggtgatgattgtgtggcaagtagaca	Reference probe sequence capitalized

Table S-12 Amplicons Sequences for each RPA amplicon in the study.
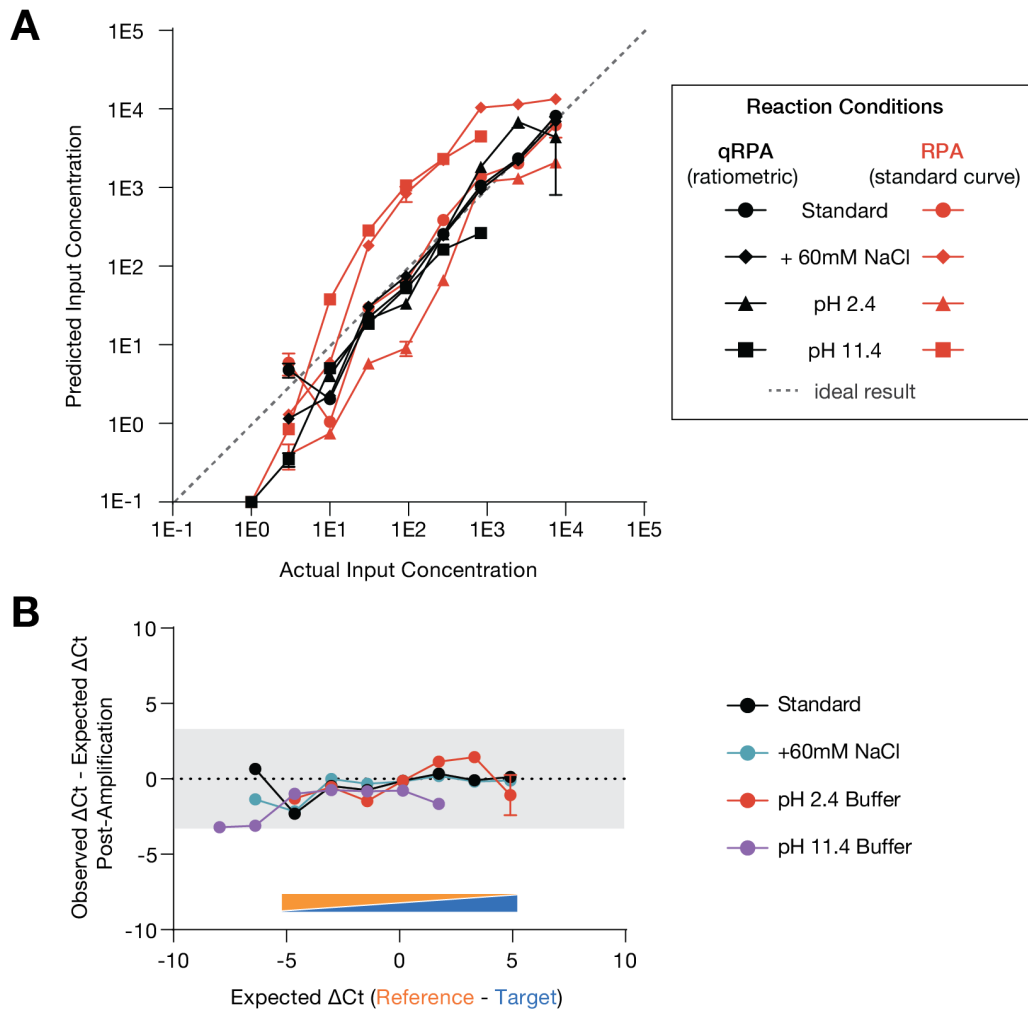
Supp Fig 4-1 qRPA is robust to chemical perturbations.

(A) Prediction of input concentration by endpoint measurements of qRPA performed under various chemical perturbations to the reaction. qPCR was used to quantify target or reference sequences after amplification by qRPA. The threshold cycle Ct values were used to calculate a predicted input concentration either by fitting target product levels to a standard curve (RPA) or calculating the target to reference ratio (qRPA). (B) Bland-Altman plot of qRPA outputs under environmental perturbations. Despite variations in product production, the ratio of target to reference is maintained, as measured by ΔCt between target and reference after amplification. In each panel, each point depicts the mean of 3 replicate RPA or qRPA reactions, error bars depict standard deviation.
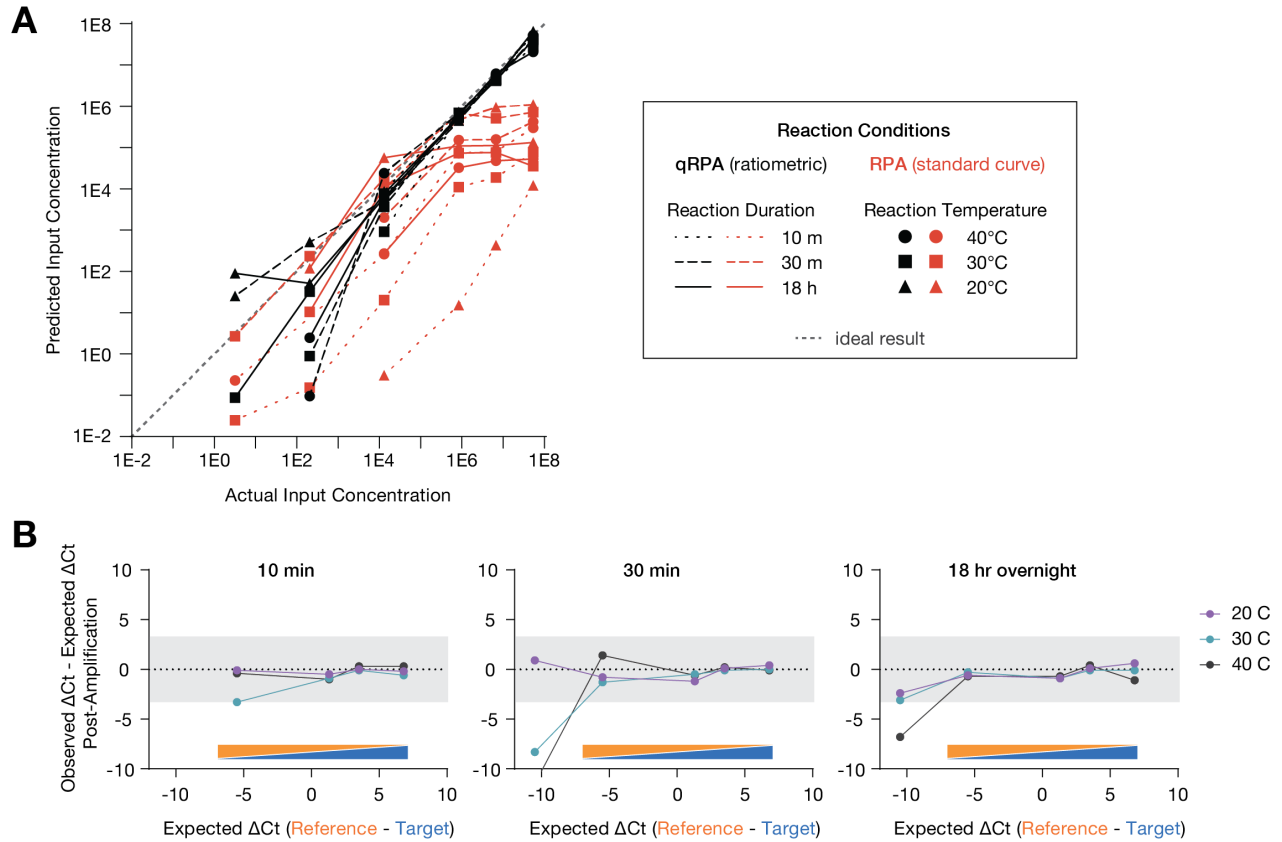
**Supp Fig 4-2** qRPA is robust to time and temperature variation.

(A) Prediction of input concentration by endpoint measurements of RPA and qRPA performed with varying reaction temperatures and durations. Standard qRPA assays are performed at 42 °C for 30 minutes. qPCR was used to quantify target or reference sequences after amplification by qRPA. The threshold cycle Ct values were used to calculate a predicted input concentration either by fitting target product levels to a standard curve (RPA) or calculating the target to reference ratio (qRPA). (B) Bland-Altman plot of qRPA outputs for different reaction temperatures and times. Regardless of amount of product produced, the ratio of target to reference is maintained, as measured by ΔCt between target and reference after amplification. In each panel, each point depicts a single RPA or qRPA reaction.
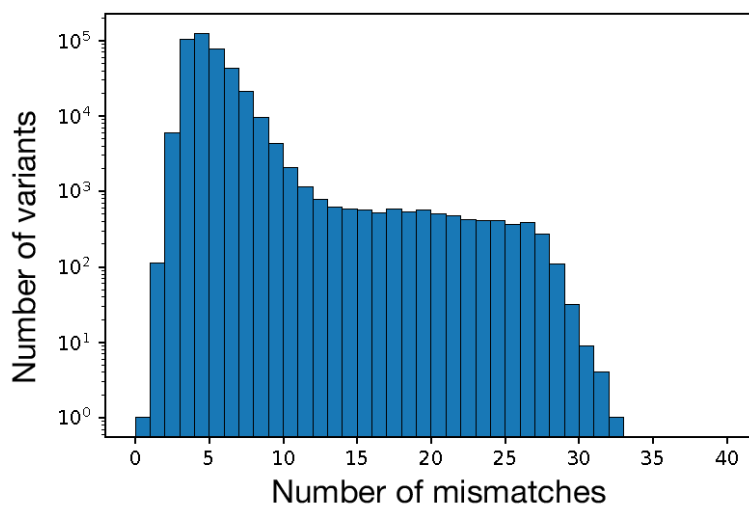
Supp Fig 4-3 Frequency of different numbers of mismatches in input library.

A library was created by degenerate synthesis of a 37-bp target sequence with 91% abundance of the original nucleotide and 3% abundance of the other three nucleotides in the RPA primer binding region. The frequency of sequences with different numbers of mismatches is plotted.
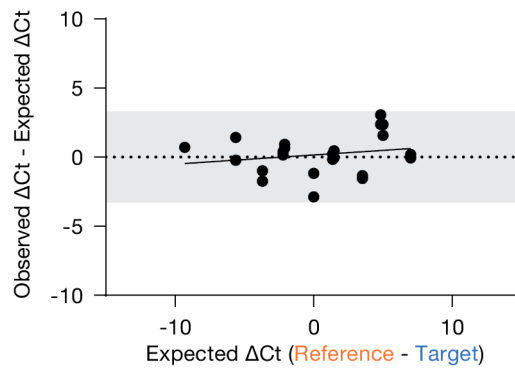
Supp Fig 4-4 Target to reference ratio is maintained in a multiplexed TaqMan assay**.**
Bland-Altman plot of RT-qPCR TaqMan assay on in vitro SARS-CoV-2 RNA. In a multiplexed reaction, wild-type and reference SARS-CoV-2 RNA molecules at different starting ratios were amplified and separately detected with fluorescent hydrolysis probes. The difference between the measured ΔCt in a duplex reaction versus the ΔCt calculated in separate reactions for each sample input was plotted for each ratio. Solid line indicates linear fit.
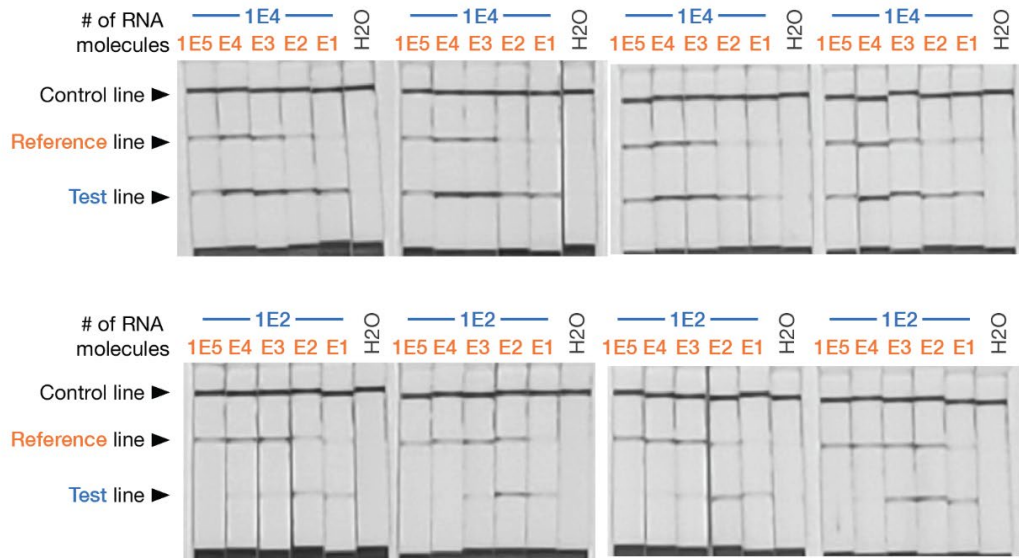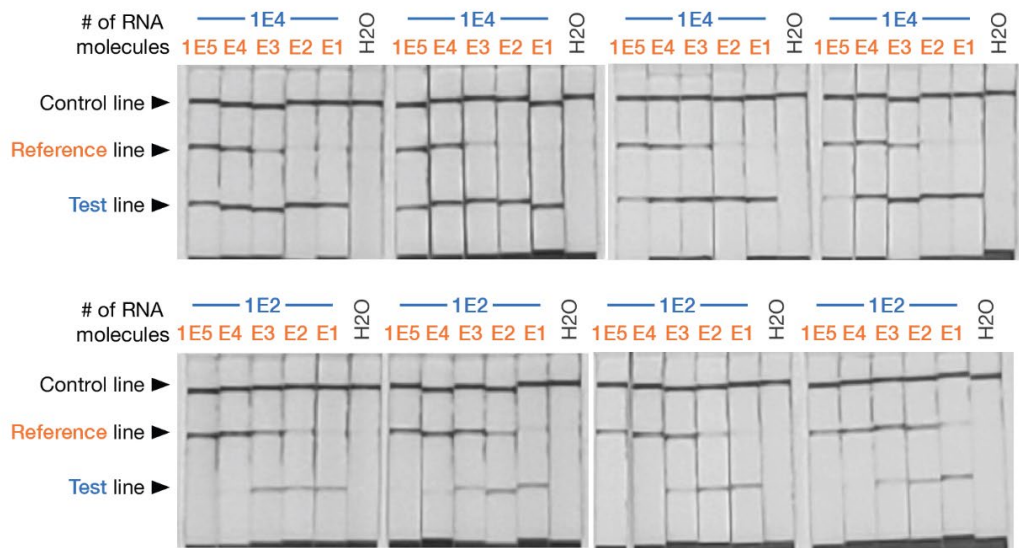
Supp Fig 4-5 Photographs of lateral flow assays run with RT-qRPA for in-vitro RNA samples.
(A) RT-qRPA reactions were performed on in-vitro HIV RNA samples containing either 1E2 or 1E4 target RNA molecules and a varying amount of reference RNA molecules with a synthetic barcode sequence inserted. The HIV RT-qRPA reaction products were hybridized and loaded on lateral flow strips and photographed. (B) Photograph of lateral flow assay for RT-qRPA on in-vitro SARS-CoV-2 RNA performed similarly.

qRPA on 1000 vs 1000 molecules

	RNA vs RNA	RNA vs RNA no RNAse	RNA vs RNA separate RT	RNA vs DNA	DNA vs DNA
Mean	-0.74	-2.10	-0.58	-4.85	-0.66
StDev	0.66	0.78	0.32	0.69	0.55
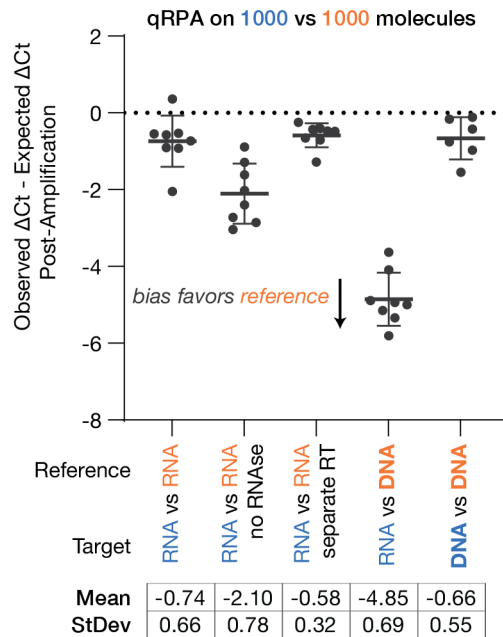
Supp Fig 4-6 RT-qRPA bias and variance depend on choice of reference molecule.

RT-qRPA or qRPA reactions were performed on 1000 target and 1000 reference molecules of varying type. DNA inputs were plasmids encoding SARS-CoV-2 N-gene with or without a synthetic reference barcode inserted. RNA inputs were produced from these plasmids by in-vitro T7 transcription. In the "no RNAse" reaction, RNAseH was excluded from the RT-qRPA mastermix. In the "separate RT" reaction, 1000 molecules each of target and reference DNA were treated with reverse transcriptase for 10 mins at 55 ˚C, then with RNAseH for 10 mins at 37 ˚C, and used as input to a qRPA reaction (no reverse transcriptase). The mean and standard deviation of 8 reactions was calculated to determine bias and noise for each reaction condition.
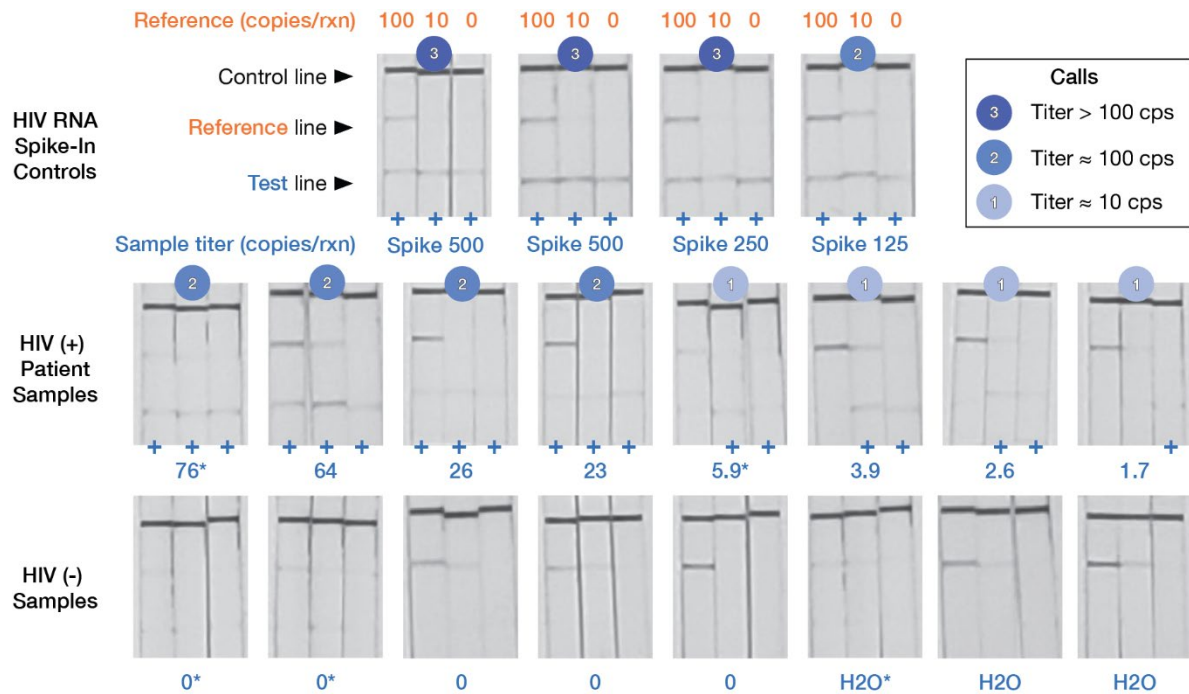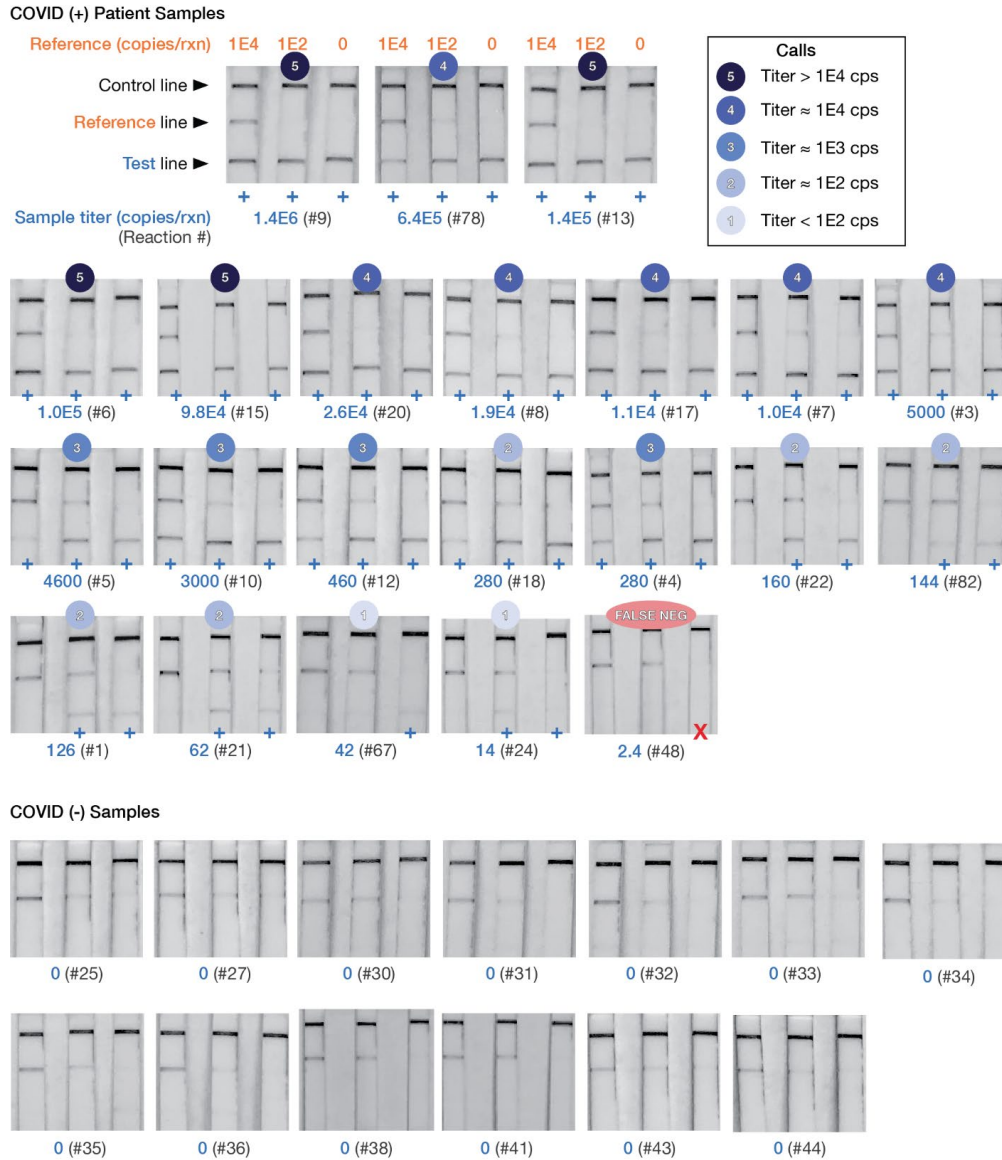
Supp Fig 4-7 Photographs of lateral flow assays run with RT-qRPA for HIV patient samples.

RT-qRPA reactions were performed on RNA extracted from HIV (+) or (-) patient blood samples spiked with 100, 10 or 0 in-vitro HIV RNA molecules containing a synthetic reference barcode. 4 positive control samples were prepared by spiking wild-type in-vitro HIV RNA into (-) patient samples. The RT-qRPA reaction products were hybridized and loaded on lateral flow strips and photographed. Calls were made for each sample based on the relative strength of test and reference lines across all three strips by an individual blinded to the identity of each sample, but unblinded to the distribution of titers. Due to the low concentration of HIV samples, test bands are faint. The titer of each sample was measured by RT-qPCR; the estimated quantity of target molecules in the 2 μL sample input is listed below each set of strips. Samples marked with * were hybridized with 5 μM probe, causing some background signal in the reference band. All other samples were hybridized using 2.5 μM probe
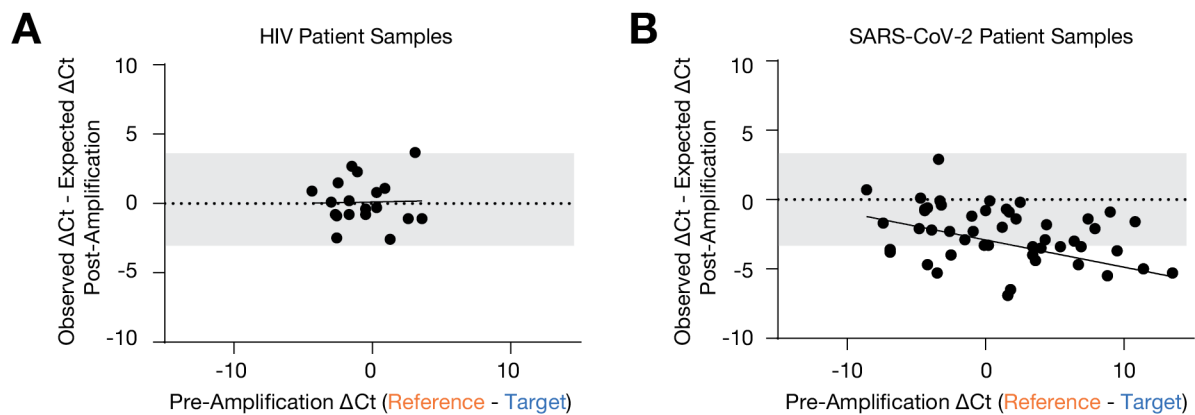
Supp Fig 4-8 Target to reference ration is maintained in RT-qRPA assays for patient samples.

(A) Bland-Altman plot of qRPA outputs for HIV patient samples. The target to reference ratio was measured prior to qRPA using a pair of fully nested qPCR reactions. The difference in ΔCt between target and reference post-amplification vs pre-amplification was plotted for each sample. (B) Bland-Altman plot of qRPA outputs for SARS-CoV-2 patient samples measured as above. In each panel, each point depicts a single qRPA reaction. Dashed line indicates ideal, solid line depicts linear fit.

Supp Fig 4-9 Photographs of lateral flow assays run with RT-qRPA for SARS-CoV-2 patient samples. RT-qRPA reactions were performed on remnant nasopharyngeal swab samples from SARS-CoV-2 (+) or (-) patients were tested without extraction. Samples were spiked with 100, 10 or 0 in-vitro HIV RNA molecules containing a synthetic reference barcode. 4 positive control samples were prepared by spiking wild-type in-vitro HIV RNA into (-) patient samples. The titer of each sample was measured by RT-qPCR; the estimated quantity of target molecules in the 2 μL sample input is listed below each set of strips in blue alongside the sample number in grey. The RT-qRPA reaction products were hybridized and loaded on lateral flow strips and photographed. Calls were made for each sample based on the relative strength of test and reference lines across all three strips by an individual blinded to the identity of each sample, but unblinded to the distribution of titers.

References

1. A. M. Caliendo, D. N. Gilbert, C. C. Ginocchio, K. E. Hanson, L. May, T. C. Quinn, F. C. Tenover, D. Alland, A. J. Blaschke, R. A. Bonomo, K. C. Carroll, M. J. Ferraro, L. R. Hirschhorn, W. P. Joseph, T. Karchmer, A. T. MacIntyre, L. B. Reller, A. F. Jackson, I. D. S. of A. (IDSA), Better Tests, Better Care: Improved Diagnostics for Infectious Diseases. *Clin Infect Dis*. **57**, S139–S170 (2013).

2. M. Euler, Y. Wang, D. Heidenreich, P. Patel, O. Strohmeier, S. Hakenberg, M. Niedrig, F. T. Hufert, M. Weidmann, Development of a Panel of Recombinase Polymerase Amplification Assays for Detection of Biothreat Agents. *J Clin Microbiol*. **51**, 1110–1117 (2013).

3. M. Rojas, I. González, M. A. Pavón, N. Pegels, P. E. Hernández, T. García, R. Martín, Polymerase chain reaction assay for verifying the labeling of meat and commercial meat products from game birds targeting specific sequences from the mitochondrial D-loop region. *Poultry Sci*. **89**, 1021–1032 (2010).

4. S. Uzzau, D. J. Brown, T. Wallis, S. Rubino, G. Leori, S. Bernard, J. Casadesus, D. J. Platt, J. E. Olsen, Host adapted serotypes of Salmonella enterica. *Epidemiology Amp Infect*. **125**, 229–255 (2000).

5. Q. Yang, K. J. Domesle, B. Ge, Loop-Mediated Isothermal Amplification for Salmonella Detection in Food and Feed: Current Applications and Future Directions. *Foodborne Pathog Dis*. **15**, 309–331 (2018).

6. H. C. Metsky, C. B. Matranga, S. Wohl, S. F. Schaffner, C. A. Freije, S. M. Winnicki, K. West, J. Qu, M. L. Baniecki, A. Gladden-Young, A. E. Lin, C. H. Tomkins-Tinch, S. H. Ye, D. J. Park, C. Y. Luo, K. G. Barnes, R. R. Shah, B. Chak, G. Barbosa-Lima, E. Delatorre, Y. R. Vieira, L. M.

Paul, A. L. Tan, C. M. Barcellona, M. C. Porcelli, C. Vasquez, A. C. Cannons, M. R. Cone, K. N. Hogan, E. W. Kopp, J. J. Anzinger, K. F. Garcia, L. A. Parham, R. M. G. Ramírez, M. C. M. Montoya, D. P. Rojas, C. M. Brown, S. Hennigan, B. Sabina, S. Scotland, K. Gangavarapu, N. D. Grubaugh, G. Oliveira, R. Robles-Sikisaka, A. Rambaut, L. Gehrke, S. Smole, M. E. Halloran, L. Villar, S. Mattar, I. Lorenzana, J. Cerbino-Neto, C. Valim, W. Degrave, P. T. Bozza, A. Gnirke, K. G. Andersen, S. Isern, S. F. Michael, F. A. Bozza, T. M. L. Souza, I. Bosch, N. L. Yozwiak, B. L. MacInnis, P. C. Sabeti, Zika virus evolution and spread in the Americas. *Nature*. **546**, 411–415 (2017).

7. M. N. Esbin, O. N. Whitney, S. Chong, A. Maurer, X. Darzacq, R. Tjian, Overcoming the bottleneck to widespread testing: A rapid review of nucleic acid testing approaches for COVID-19 detection. *RNA*. **26** (2020), doi:10.1261/rna.076232.120.

8. T. Ibn-Mohammed, K. B. Mustapha, J. Godsell, Z. Adamu, K. A. Babatunde, D. D. Akintade, A. Acquaye, H. Fujii, M. M. Ndiaye, F. A. Yamoah, S. C. L. Koh, A critical review of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies. *Resour Conservation Recycl*. **164**, 105169 (2021).

9. C. W. Hedberg, J. F. Greenblatt, B. T. Matyas, J. Lemmings, D. J. Sharp, R. T. Skibicki, A. P. Liang, E. D. I. T. S. W. Group, Timeliness of Enteric Disease Surveillance in 6 US States. *Emerg Infect Dis*. **14**, 311–313 (2008).

10. E. Scallan, P. M. Griffin, F. J. Angulo, R. V. Tauxe, R. M. Hoekstra, Foodborne Illness Acquired in the United States—Unspecified Agents. *Emerg Infect Dis*. **17**, 16–22 (2011).

11. T. Briese, G. Palacios, M. Kokoris, O. Jabado, Z. Liu, N. Renwick, V. Kapoor, I. Casas, F. Pozo, R. Limberger, P. Perez-Brena, J. Ju, W. I. Lipkin, Diagnostic System for Rapid and Sensitive Differential Detection of Pathogens. *Emerg Infect Dis*. **11**, 310–313 (2005).

12. G. Palacios, T. Briese, V. Kapoor, O. Jabado, Z. Liu, M. Venter, J. Zhai, N. Renwick, A. Grolla, T. W. Geisbert, C. Drosten, J. Towner, J. Ju, J. Paweska, S. T. Nichol, R. Swanepoel, H. Feldmann, P. B. Jahrling, W. I. Lipkin, MassTag Polymerase Chain Reaction for Differential Diagnosis of Viral Hemorrhagic Fevers. *Emerg Infect Dis*. **12**, 692–695 (2006).

13. H. Li, M. A. McCormac, R. W. Estes, S. E. Sefers, R. K. Dare, J. D. Chappell, D. D. Erdman, P. F. Wright, Y.-W. Tang, Simultaneous Detection and High-Throughput Identification of a Panel of RNA Viruses Causing Respiratory Tract Infections ▿ . *J Clin Microbiol*. **45**, 2105–2109 (2007).

14. Z. A. Crannell, B. Rohrman, R. Richards-Kortum, Quantification of HIV-1 DNA Using Real-Time Recombinase Polymerase Amplification. *Anal Chem.* **86**, 5615–5619 (2014).

15. Y. Zhao, F. Chen, Q. Li, L. Wang, C. Fan, Isothermal Amplification of Nucleic Acids. *Chem Rev*. **115**, 12491–12545 (2015).

16. C. S. Kosack, A.-L. Page, P. R. Klatser, A guide to aid the selection of diagnostic tests. *B World Health Organ*. **95**, 639–645 (2017).

17. O. Piepenburg, C. H. Williams, D. L. Stemple, N. A. Armes, DNA Detection Using Recombination Proteins. *Plos Biol*. **4**, e204 (2006).

18. Z. A. Crannell, B. Rohrman, R. Richards-Kortum, Equipment-Free Incubation of Recombinase Polymerase Amplification Reactions Using Body Heat. *Plos One*. **9**, e112146 (2014).

19. R. K. Daher, G. Stewart, M. Boissinot, M. G. Bergeron, Recombinase Polymerase Amplification for Diagnostic Applications. *Clin Chem*. **62**, 947–958 (2016).

20. W. Gao, H. Huang, P. Zhu, X. Yan, J. Fan, J. Jiang, J. Xu, Recombinase polymerase amplification combined with lateral flow dipstick for equipment-free detection of Salmonella in shellfish. *Bioproc Biosyst Eng*. **41**, 603–611 (2018).

21. J. S. Gootenberg, O. O. Abudayyeh, J. W. Lee, P. Essletzbichler, A. J. Dy, J. Joung, V. Verdine, N. Donghia, N. M. Daringer, C. A. Freije, C. Myhrvold, R. P. Bhattacharyya, J. Livny, A. Regev, E. V. Koonin, D. T. Hung, P. C. Sabeti, J. J. Collins, F. Zhang, Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*. **356**, 438–442 (2017).

22. J. S. Chen, E. Ma, L. B. Harrington, M. D. Costa, X. Tian, J. M. Palefsky, J. A. Doudna, CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science*. **360**, eaar6245 (2018).

23. C. Myhrvold, C. A. Freije, J. S. Gootenberg, O. O. Abudayyeh, H. C. Metsky, A. F. Durbin, M. J. Kellner, A. L. Tan, L. M. Paul, L. A. Parham, K. F. Garcia, K. G. Barnes, B. Chak, A. Mondini, M. L. Nogueira, S. Isern, S. F. Michael, I. Lorenzana, N. L. Yozwiak, B. L. MacInnis, I. Bosch, L. Gehrke, F. Zhang, P. C. Sabeti, Field-deployable viral diagnostics using CRISPR-Cas13. *Science*. **360**, 444–448 (2018).

24. P. Rostron, T. Pennance, F. Bakar, D. Rollinson, S. Knopp, F. Allan, F. Kabole, S. M. Ali, S. M. Ame, B. L. Webster, Development of a recombinase polymerase amplification (RPA) fluorescence assay for the detection of Schistosoma haematobium. *Parasite Vector*. **12**, 514 (2019).

25. Y. D. Wu, M. J. Xu, Q. Q. Wang, C. X. Zhou, M. Wang, X. Q. Zhu, D. H. Zhou, Recombinase polymerase amplification (RPA) combined with lateral flow (LF) strip for detection of Toxoplasma gondii in the environment. *Vet Parasitol*. **243**, 199–203 (2017).

26. G. Choi, J. H. Jung, B. H. Park, S. J. Oh, J. H. Seo, J. S. Choi, D. H. Kim, T. S. Seo, A centrifugal direct recombinase polymerase amplification (direct-RPA) microdevice for multiplex and real-time identification of food poisoning bacteria. *Lab Chip*. **16**, 2309–2316 (2016).

27. I. M. Lobato, C. K. O'Sullivan, Recombinase polymerase amplification: Basics, applications and recent advances. *Trac Trends Anal Chem*. **98**, 19–35 (2018).

28. G. Urtel, M. V. D. Hofstadt, J.-C. Galas, A. Estevez-Torres, rEXPAR: An Isothermal Amplification Scheme That Is Robust to Autocatalytic Parasites. *Biochemistry-us*. **58**, 2675–2681 (2019).

29. S. Das, S. C. Mohapatra, J. T. Hsu, Studies on primer-dimer formation in polymerase chain reaction (PCR). *Biotechnol Tech*. **13**, 643–646 (1999).

30. S. Hoshika, F. Chen, N. A. Leal, S. A. Benner, Artificial Genetic Systems: Self-Avoiding DNA in PCR and Multiplexed PCR. *Angewandte Chemie Int Ed*. **49**, 5554–5557 (2010).

31. N. Sharma, S. Hoshika, D. Hutter, K. M. Bradley, S. A. Benner, Recombinase-Based Isothermal Amplification of Nucleic Acids with Self-Avoiding Molecular Recognition Systems (SAMRS). *Chembiochem*. **15**, 2268–2274 (2014).

32. Z. Yang, C. McLendon, D. Hutter, K. M. Bradley, S. Hoshika, C. B. Frye, S. A. Benner, Helicase-Dependent Isothermal Amplification of DNA and RNA by Using Self-Avoiding Molecular Recognition Systems. *Chembiochem*. **16**, 1365–1370 (2015).

33. Y. Wang, D. Liu, J. Deng, Y. Wang, J. Xu, C. Ye, Loop-mediated isothermal amplification using self-avoiding molecular recognition systems and antarctic thermal sensitive uracil-DNA-glycosylase for detection of nucleic acid with prevention of carryover contamination. *Anal Chim Acta*. **996**, 74–87 (2017).

34. G. A. Denys, Portrait Toxigenic Clostridium difficile assay, an isothermal amplification assay detects toxigenic C. difficile in clinical stool specimens. *Expert Rev Mol Diagn*. **14**, 17–26 (2013).

35. R. K. Daher, G. Stewart, M. Boissinot, D. K. Boudreau, M. G. Bergeron, Influence of sequence mismatches on the specificity of recombinase polymerase amplification technology. *Mol Cell Probe*. **29**, 116–121 (2015).

36. M. Higgins, M. Ravenhall, D. Ward, J. Phelan, A. Ibrahim, M. S. Forrest, T. G. Clark, S. Campino, PrimedRPA: primer design for recombinase polymerase amplification assays. *Bioinformatics*. **35**, 682–684 (2018).

37. G. Marks, L. I. Gardner, C. E. Rose, A. Zinski, R. D. Moore, S. Holman, A. E. Rodriguez, M. Sullivan, T. P. Giordano, Time above 1500 copies. *Aids*. **29**, 947–954 (2015).

38. Y. S. Jiang, A. Stacy, M. Whiteley, A. D. Ellington, S. Bhadra, Amplicon Competition Enables End-Point Quantitation of Nucleic Acids Following Isothermal Amplification. *Chembiochem*. **18**, 1692–1695 (2017).

39. E.-C. Yeh, C.-C. F. Yeh, L. Hu, M.-Y. Huang, L. P. Lee, A Single-Step Digital Nucleic Acid Amplification Platform by Digital Plasma Separation on a Chip. *Biophys J*. **106**, 416a (2014).

40. F. Schuler, F. Schwemmer, M. Trotter, S. Wadle, R. Zengerle, F. von Stetten, N. Paust, Centrifugal step emulsification applied for absolute quantification of nucleic acids by digital droplet RPA. *Lab Chip*. **15**, 2759–2766 (2015).

41. Z. Li, Y. Liu, Q. Wei, Y. Liu, W. Liu, X. Zhang, Y. Yu, Picoliter Well Array Chip-Based Digital Recombinase Polymerase Amplification for Absolute Quantification of Nucleic Acids. *Plos One*. **11**, e0153359 (2016).

42. F. Shen, E. K. Davydova, W. Du, J. E. Kreutz, O. Piepenburg, R. F. Ismagilov, Digital Isothermal Quantification of Nucleic Acids via Simultaneous Chemical Initiation of Recombinase Polymerase Amplification Reactions on SlipChip. *Anal Chem*. **83**, 3533–3540 (2011).

43. T. Notomi, H. Okayama, H. Masubuchi, T. Yonekawa, K. Watanabe, N. Amino, T. Hase, Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res*. **28**, e63–e63 (2000).

44. K. Nagamine, T. Hase, T. Notomi, Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Mol Cell Probe*. **16**, 223–229 (2002).

45. L. Becherer, N. Borst, M. Bakheit, S. Frischmann, R. Zengerle, F. von Stetten, Loop-mediated isothermal amplification (LAMP) – review and classification of methods for sequence-specific detection. *Anal Methods-uk*. **12**, 717–746 (2020).

46. Z. Ali, R. Aman, A. Mahas, G. S. Rao, M. Tehseen, T. Marsic, R. Salunke, A. K. Subudhi, S. M. Hala, S. M. Hamdan, A. Pain, F. S. Alofi, A. Alsomali, A. M. Hashem, A. Khogeer, N. A. M. Almontashiri, M. Abedalthagafi, N. Hassan, M. M. Mahfouz, iSCAN: An RT-LAMP-coupled CRISPR-Cas12 module for rapid, sensitive detection of SARS-CoV-2. *Virus Res*. **288**, 198129 (2020).

47. J. P. Broughton, X. Deng, G. Yu, C. L. Fasching, V. Servellita, J. Singh, X. Miao, J. A. Streithorst, A. Granados, A. Sotomayor-Gonzalez, K. Zorn, A. Gopez, E. Hsu, W. Gu, S. Miller, C.-Y. Pan, H. Guevara, D. A. Wadford, J. S. Chen, C. Y. Chiu, CRISPR–Cas12-based detection of SARS-CoV-2. *Nat Biotechnol*. **38**, 870–874 (2020).

48. R. Wang, C. Qian, Y. Pang, M. Li, Y. Yang, H. Ma, M. Zhao, F. Qian, H. Yu, Z. Liu, T. Ni, Y. Zheng, Y. Wang, opvCRISPR: One-pot visual RT-LAMP-CRISPR platform for SARS-cov-2 detection. *Biosens Bioelectron*. **172**, 112766 (2021).

49. F. V. Shirshikov, Y. A. Pekov, K. A. Miroshnikov, MorphoCatcher: a multiple-alignment based web tool for target selection and designing taxon-specific primers in the loop-mediated isothermal amplification method. *Peerj*. **7**, e6801 (2019).

50. M. Vincent, Y. Xu, H. Kong, Helicase-dependent isothermal DNA amplification. *Embo Rep*. **5**, 795–800 (2004).

51. M. Mahalanabis, J. Do, H. ALMuayad, J. Y. Zhang, C. M. Klapperich, An integrated disposable device for DNA extraction and helicase dependent amplification. *Biomed Microdevices*. **12**, 353–359 (2010).

52. V. Doseeva, T. Forbes, J. Wolff, Y. Khripin, D. O'Neil, T. Rothmann, I. Nazarenko, Multiplex isothermal helicase-dependent amplification assay for detection of Chlamydia trachomatis and Neisseria gonorrhoeae. *Diagn Micr Infec Dis*. **71**, 354–365 (2011).

53. J. Compton, Nucleic acid sequence-based amplification. *Nature*. **350**, 91–92 (1991).

54. P. M. Lizardi, X. Huang, Z. Zhu, P. Bray-Ward, D. C. Thomas, D. C. Ward, Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet*. **19**, 225–232 (1998).

55. P. Moore, Replicating success. *Nature*. **435**, 235–235 (2005).

56. M. Euler, Y. Wang, O. Nentwich, O. Piepenburg, F. T. Hufert, M. Weidmann, Recombinase polymerase amplification assay for rapid detection of Rift Valley fever virus. *J Clin Virol*. **54**, 308–312 (2012).

57. A. A. E. Wahed, A. El-Deeb, M. El-Tholoth, H. A. E. Kader, A. Ahmed, S. Hassan, B. Hoffmann, B. Haas, M. A. Shalaby, F. T. Hufert, M. Weidmann, A Portable Reverse Transcription

Recombinase Polymerase Amplification Assay for Rapid Detection of Foot-and-Mouth Disease Virus. *Plos One*. **8**, e71642 (2013).

58. C. Escadafal, J. T. Paweska, A. Grobbelaar, C. le Roux, M. Bouloy, P. Patel, A. Teichmann, O. Donoso-Mantke, M. Niedrig, International External Quality Assessment of Molecular Detection of Rift Valley Fever Virus. *Plos Neglect Trop D*. **7**, e2244 (2013).

59. A. Aebischer, K. Wernike, B. Hoffmann, M. Beer, Rapid Genome Detection of Schmallenberg Virus and Bovine Viral Diarrhea Virus by Use of Isothermal Amplification Methods and High-Speed Real-Time Reverse Transcriptase PCR. *J Clin Microbiol*. **52**, 1883–1892 (2014).

60. T. A. Mekuria, S. Zhang, K. C. Eastwell, Rapid and sensitive detection of Little cherry virus 2 using isothermal reverse transcription-recombinase polymerase amplification. *J Virol Methods*. **205**, 24–30 (2014).

61. B.-T. Teoh, S.-S. Sam, K.-K. Tan, M. B. Danlami, M.-H. Shu, J. Johari, P.-S. Hooi, D. Brooks, O. Piepenburg, O. Nentwich, A. Wilder-Smith, L. Franco, A. Tenorio, S. AbuBakar, Early Detection of Dengue Virus by Use of Reverse Transcription-Recombinase Polymerase Amplification. *J Clin Microbiol*. **53**, 830–837 (2015).

62. H. M. Amer, A. A. E. Wahed, M. A. Shalaby, F. N. Almajhdi, F. T. Hufert, M. Weidmann, A new approach for diagnosis of bovine coronavirus using a reverse transcription recombinase polymerase amplification assay. *J Virol Methods*. **193**, 337–340 (2013).

63. S. Zhang, M. Ravelonandro, P. Russell, N. McOwen, P. Briard, S. Bohannon, A. Vrient, Rapid diagnostic detection of plum pox virus in Prunus plants by isothermal AmplifyRP® using reverse transcription-recombinase polymerase amplification. *J Virol Methods*. **207**, 114–120 (2014).

64. G. A. Hill-Cawthorne, L. O. Hudson, M. F. A. E. Ghany, O. Piepenburg, M. Nair, A. Dodgson, M. S. Forrest, T. G. Clark, A. Pain, Recombinations in Staphylococcal Cassette Chromosome mec Elements Compromise the Molecular Detection of Methicillin Resistance in Staphylococcus aureus. *Plos One*. **9**, e101419 (2014).

65. M. Euler, Y. Wang, P. Otto, H. Tomaso, R. Escudero, P. Anda, F. T. Hufert, M. Weidmann, Recombinase Polymerase Amplification Assay for Rapid Detection of Francisella tularensis. *J Clin Microbiol*. **50**, 2234–2238 (2012).

66. A. Ahmed, H. van der Linden, R. A. Hartskeerl, Development of a Recombinase Polymerase Amplification Assay for the Detection of Pathogenic Leptospira. *Int J Environ Res Pu*. **11**, 4953–4964 (2014).

67. D. S. Boyle, R. McNerney, H. T. Low, B. T. Leader, A. C. Pérez-Osorio, J. C. Meyer, D. M. O'Sullivan, D. G. Brooks, O. Piepenburg, M. S. Forrest, Rapid Detection of Mycobacterium tuberculosis by Recombinase Polymerase Amplification. *Plos One*. **9**, e103091 (2014).

68. R. K. Daher, G. Stewart, M. Boissinot, M. G. Bergeron, Isothermal Recombinase Polymerase Amplification Assay Applied to the Detection of Group B Streptococci in Vaginal/Anal Samples. *Clin Chem*. **60**, 660–666 (2014).

69. S. Kersting, V. Rausch, F. F. Bier, M. von Nickisch-Rosenegk, Multiplex isothermal solid-phase recombinase polymerase amplification for the specific and fast DNA-based detection of three bacterial pathogens. *Microchim Acta*. **181**, 1715–1723 (2014).

70. K. Krõlov, J. Frolova, O. Tudoran, J. Suhorutsenko, T. Lehto, H. Sibul, I. Mäger, M. Laanpere, I. Tulp, Ü. Langel, Sensitive and Rapid Detection of Chlamydia trachomatis by Recombinase Polymerase Amplification Directly from Urine Samples. *J Mol Diagnostics*. **16**, 127–135 (2014).

71. Z. A. Crannell, A. Castellanos-Gonzalez, A. Irani, B. Rohrman, A. C. White, R. Richards-Kortum, Nucleic Acid Test to Diagnose Cryptosporidiosis: Lab Assessment in Animal and Patient Specimens. *Anal Chem*. **86**, 2565–2571 (2014).

72. S. Kersting, V. Rausch, F. F. Bier, M. von Nickisch-Rosenegk, Rapid detection of Plasmodium falciparum with isothermal recombinase polymerase amplification and lateral flow analysis. *Malaria J*. **13**, 99 (2014).

73. J. F. C. Loo, P. M. Lau, H. P. Ho, S. K. Kong, An aptamer-based bio-barcode assay with isothermal recombinase polymerase amplification for cytochrome-c detection and anti-cancer drug screening. *Talanta*. **115**, 159–165 (2013).

74. Y. Shin, A. P. Perera, K. W. Kim, M. K. Park, Real-time, label-free isothermal solid-phase amplification/detection (ISAD) device for rapid detection of genetic alteration in cancers. *Lab Chip*. **13**, 2106–2114 (2013).

75. D. K. Ghosh, S. B. Kokane, S. Gowda, Development of a reverse transcription recombinase polymerase based isothermal amplification coupled with lateral flow immunochromatographic assay (CTV-RT-RPA-LFICA) for rapid detection of Citrus tristeza virus. *Sci Rep-uk*. **10**, 20593 (2020).

76. S. Lax, D. P. Smith, J. Hampton-Marcell, S. M. Owens, K. M. Handley, N. M. Scott, S. M. Gibbons, P. Larsen, B. D. Shogan, S. Weiss, J. L. Metcalf, L. K. Ursell, Y. Vázquez-Baeza, W. V. Treuren, N. A. Hasan, M. K. Gibson, R. Colwell, G. Dantas, R. Knight, J. A. Gilbert, Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*. **345**, 1048–1052 (2014).

77. S. Lax, J. T. Hampton-Marcell, S. M. Gibbons, G. B. Colares, D. Smith, J. A. Eisen, J. A. Gilbert, Forensic analysis of the microbiome of phones and shoes. *Microbiome*. **3**, 21 (2015).

78. C. Jiang, X. Wang, X. Li, J. Inlora, T. Wang, Q. Liu, M. Snyder, Dynamic Human Environmental Exposome Revealed by Longitudinal Personal Monitoring. *Cell*. **175**, 277-291.e31 (2018).

79. S. Lax, C. R. Nagler, J. A. Gilbert, Our interface with the built environment: immunity and the indoor microbiota. *Trends Immunol*. **36**, 121–123 (2015).

80. J. Fletcher, C. Bender, B. Budowle, W. T. Cobb, S. E. Gold, C. A. Ishimaru, D. Luster, U. Melcher, R. Murch, H. Scherm, R. C. Seem, J. L. Sherwood, B. W. Sobral, S. A. Tolin, Plant Pathogen Forensics: Capabilities, Needs, and Recommendations. *Microbiol Mol Biol R*. **70**, 450–471 (2006).

81. S. R. Tridico, D. C. Murray, J. Addison, K. P. Kirkbride, M. Bunce, Metagenomic analyses of bacteria on human hairs: a qualitative assessment for applications in forensic science. *Investigative Genetics*. **5**, 16 (2014).

82. N. Fierer, C. L. Lauber, N. Zhou, D. McDonald, E. K. Costello, R. Knight, Forensic identification using skin bacterial communities. *Proc National Acad Sci*. **107**, 6477–6481 (2010).

83. C. Huttenhower, D. Gevers, R. Knight, S. Abubucker, J. H. Badger, A. T. Chinwalla, H. H. Creasy, A. M. Earl, M. G. FitzGerald, R. S. Fulton, M. G. Giglio, K. Hallsworth-Pepin, E. A. Lobos, R. Madupu, V. Magrini, J. C. Martin, M. Mitreva, D. M. Muzny, E. J. Sodergren, J. Versalovic, A. M. Wollam, K. C. Worley, J. R. Wortman, S. K. Young, Q. Zeng, K. M. Aagaard, O. O. Abolude, E. Allen-Vercoe, E. J. Alm, L. Alvarado, G. L. Andersen, S. Anderson, E. Appelbaum, H. M. Arachchi, G. Armitage, C. A. Arze, T. Ayvaz, C. C. Baker, L. Begg, T. Belachew, V. Bhonagiri, M. Bihan, M. J. Blaser, T. Bloom, V. Bonazzi, J. P. Brooks, G. A. Buck, C. J. Buhay, D. A. Busam, J.

L. Campbell, S. R. Canon, B. L. Cantarel, P. S. G. Chain, I.-M. A. Chen, L. Chen, S. Chhibba, K. Chu, D. M. Ciulla, J. C. Clemente, S. W. Clifton, S. Conlan, J. Crabtree, M. A. Cutting, N. J. Davidovics, C. C. Davis, T. Z. DeSantis, C. Deal, K. D. Delehaunty, F. E. Dewhirst, E. Deych, Y. Ding, D. J. Dooling, S. P. Dugan, W. M. Dunne, A. S. Durkin, R. C. Edgar, R. L. Erlich, C. N. Farmer, R. M. Farrell, K. Faust, M. Feldgarden, V. M. Felix, S. Fisher, A. A. Fodor, L. J. Forney, L. Foster, V. D. Francesco, J. Friedman, D. C. Friedrich, C. C. Fronick, L. L. Fulton, H. Gao, N. Garcia, G. Giannoukos, C. Giblin, M. Y. Giovanni, J. M. Goldberg, J. Goll, A. Gonzalez, A. Griggs, S. Gujja, S. K. Haake, B. J. Haas, H. A. Hamilton, E. L. Harris, T. A. Hepburn, B. Herter, D. E. Hoffmann, M. E. Holder, C. Howarth, K. H. Huang, S. M. Huse, J. Izard, J. K. Jansson, H. Jiang, C. Jordan, V. Joshi, J. A. Katancik, W. A. Keitel, S. T. Kelley, C. Kells, N. B. King, D. Knights, H. H. Kong, O. Koren, S. Koren, K. C. Kota, C. L. Kovar, N. C. Kyrpides, P. S. L. Rosa, S. L. Lee, K. P. Lemon, N. Lennon, C. M. Lewis, L. Lewis, R. E. Ley, K. Li, K. Liolios, B. Liu, Y. Liu, C.-C. Lo, C. A. Lozupone, R. D. Lunsford, T. Madden, A. A. Mahurkar, P. J. Mannon, E. R. Mardis, V. M. Markowitz, K. Mavromatis, J. M. McCorrison, D. McDonald, J. McEwen, A. L. McGuire, P. McInnes, T. Mehta, K. A. Mihindukulasuriya, J. R. Miller, P. J. Minx, I. Newsham, C. Nusbaum, M. O'Laughlin, J. Orvis, I. Pagani, K. Palaniappan, S. M. Patel, M. Pearson, J. Peterson, M. Podar, C. Pohl, K. S. Pollard, M. Pop, M. E. Priest, L. M. Proctor, X. Qin, J. Raes, J. Ravel, J. G. Reid, M. Rho, R. Rhodes, K. P. Riehle, M. C. Rivera, B. Rodriguez-Mueller, Y.-H. Rogers, M. C. Ross, C. Russ, R. K. Sanka, P. Sankar, J. F. Sathirapongsasuti, J. A. Schloss, P. D. Schloss, T. M. Schmidt, M. Scholz, L. Schriml, A. M. Schubert, N. Segata, J. A. Segre, W. D. Shannon, R. R. Sharp, T. J. Sharpton, N. Shenoy, N. U. Sheth, G. A. Simone, I. Singh, C. S. Smillie, J. D. Sobel, D. D. Sommer, P. Spicer, G. G. Sutton, S. M. Sykes, D. G. Tabbaa, M. Thiagarajan, C. M. Tomlinson, M. Torralba, T. J. Treangen, R. M. Truty, T. A. Vishnivetskaya, J. Walker, L. Wang, Z. Wang, D. V. Ward, W. Warren, M. A. Watson, C. Wellington, K. A. Wetterstrand, J. R. White, K. Wilczek-Boney, Y. Wu, K. M. Wylie, T. Wylie, C. Yandava, L. Ye, Y. Ye, S. Yooseph, B. P. Youmans, L. Zhang, Y. Zhou, Y. Zhu, L. Zoloth, J. D. Zucker, B. W. Birren, R. A. Gibbs, S. K.

167

Highlander, B. A. Methé, K. E. Nelson, J. F. Petrosino, G. M. Weinstock, R. K. Wilson, O. White, Structure, function and diversity of the healthy human microbiome. *Nature*. **486**, 207–214 (2012).

84. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. D. Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. **37**, 852–857 (2019).

85. T. H. Clarke, A. Gomez, H. Singh, K. E. Nelson, L. M. Brinkac, Integrating the microbiome as a resource in the forensics toolkit. *Forensic Sci Int Genetics*. **30**, 141–147 (2017).

86. P. M. (Nel) Wognum, H. Bremmers, J. H. Trienekens, J. G. A. J. van der Vorst, J. M. Bloemhof, Systems for sustainability and transparency of food supply chains – Current status and challenges. *Adv Eng Inform*. **25**, 65–76 (2011).

87. J. Gooch, B. Daniel, V. Abbate, N. Frascione, Taggant materials in forensic science: A review. *Trac Trends Anal Chem*. **83**, 49–54 (2016).

88. P. Buckley, B. Rivers, S. Katoski, M. H. Kim, F. J. Kragl, S. Broomall, M. Krepps, E. W. Skowronski, C. N. Rosenzweig, S. Paikoff, P. Emanuel, H. S. Gibbons, Genetic Barcodes for Improved Environmental Tracking of an Anthrax Simulant. *Appl Environ Microb*. **78**, 8272–8280 (2012).

89. P. A. Emanuel, P. E. Buckley, T. A. Sutton, J. M. Edmonds, A. M. Bailey, B. A. Rivers, M. H. Kim, W. J. Ginley, C. C. Keiser, R. W. Doherty, F. J. Kragl, F. E. Narayanan, S. E. Katoski, S. Paikoff, S. P. Leppert, J. B. Strawbridge, D. R. VanReenen, S. S. Biberos, D. Moore, D. W. Phillips, L. R. Mingioni, O. Melles, D. G. Ondercin, B. Hirsh, K. M. Bieschke, C. L. Harris, K. M. Omberg, V. K. Rastogi, S. V. Cuyk, H. S. Gibbons, Detection and Tracking of a Novel Genetically Tagged Biological Simulant in the Environment. *Appl Environ Microb*. **78**, 8281–8288 (2012).

90. N. Ulrich, K. Nagler, M. Laue, C. S. Cockell, P. Setlow, R. Moeller, Experimental studies addressing the longevity of Bacillus subtilis spores – The first data from a 500-year experiment. *Plos One*. **13**, e0208425 (2018).

91. A. H. Bishop, C. M. O'Sullivan, A. Lane, M. C. B. Ellis, W. J. Sellors, Re-aerosolization of Bacillus thuringiensis spores from concrete and turf. *Lett Appl Microbiol*. **64**, 364–369 (2017).

92. R. I. Adams, A. C. Bateman, H. M. Bik, J. F. Meadow, Microbiota of the indoor environment: a meta-analysis. *Microbiome*. **3**, 49 (2015).

93. G. Sanahuja, R. Banakar, R. M. Twyman, T. Capell, P. Christou, Bacillus thuringiensis: a century of research, development and commercial applications. *Plant Biotechnol J*. **9**, 283–300 (2011).

94. M. Sarwat, M. M. Yamdagni, DNA barcoding, microarrays and next generation sequencing: recent tools for genetic diversity estimation and authentication of medicinal plants. *Crit Rev Biotechnol*. **36**, 191–203 (2014).

95. B.-M. Koo, G. Kritikos, J. D. Farelli, H. Todor, K. Tong, H. Kimsey, I. Wapinski, M. Galardini, A. Cabal, J. M. Peters, A.-B. Hachmann, D. Z. Rudner, K. N. Allen, A. Typas, C. A. Gross, Construction and Analysis of Two Genome-Scale Deletion Libraries for Bacillus subtilis. *Cell Syst*. **4**, 291-305.e7 (2017).

96. A. J. Meeske, L.-T. Sham, H. Kimsey, B.-M. Koo, C. A. Gross, T. G. Bernhardt, D. Z. Rudner, MurJ and a novel lipid II flippase are required for cell wall biogenesis in Bacillus subtilis. *Proc National Acad Sci*. **112**, 6437–6442 (2015).

97. P. Schaeffer, J. Millet, J. P. Aubert, Catabolic repression of bacterial sporulation. *Proc National Acad Sci*. **54**, 704–711 (1965).

98. H. C. Lim, J. W. Sher, F. P. Rodriguez-Rivera, C. Fumeaux, C. R. Bertozzi, T. G. Bernhardt, Identification of new components of the RipC-FtsEX cell separation pathway of Corynebacterineae. *Plos Genet*. **15**, e1008284 (2019).

99. A. A. Horwitz, J. M. Walter, M. G. Schubert, S. H. Kung, K. Hawkins, D. M. Platt, A. D. Hernday, T. Mahatdejkul-Meadows, W. Szeto, S. S. Chandran, J. D. Newman, Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Syst*. **1**, 88–96 (2015).

100. D. M. Gohl, P. Vangay, J. Garbe, A. MacLean, A. Hauge, A. Becker, T. J. Gould, J. B. Clayton, T. J. Johnson, R. Hunter, D. Knights, K. B. Beckman, Systematic improvement of

amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol*. **34**, 942–949 (2016).

101. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. **13**, 581–583 (2016).

102. Q. Chang, Y. Luan, F. Sun, Variance adjusted weighted UniFrac: a powerful beta diversity measure for comparing communities based on phylogeny. *Bmc Bioinformatics*. **12**, 118 (2011).

103. L. A. Klobutcher, K. Ragkousi, P. Setlow, The Bacillus subtilis spore coat provides "eat resistance" during phagocytic predation by the protozoan Tetrahymena thermophila. *P Natl Acad Sci Usa*. **103**, 165–170 (2006).

104. P. S. V. Devi, T. Ravinder, C. Jaidev, Cost-effective production of Bacillus thuringiensis by solid-state fermentation. *J Invertebr Pathol*. **88**, 163–168 (2005).

105. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*. **20**, 533–534 (2020).

106. A. Park, Dr. Anthony Fauci "Not Overly Confident" with U.S. COVID-19 Testing Capabilities. *Time* (2020).

107. J. Temple, Why contact tracing may be a mess in America. *MIT Technology Review* (2020).

108. C. B. F. Vogels, A. E. Watkins, C. A. Harden, D. E. Brackney, J. Shafer, J. Wang, C. Caraballo, C. C. Kalinich, I. M. Ott, J. R. Fauver, E. Kudo, P. Lu, A. Venkataraman, M. Tokuyama, A. J. Moore, M. C. Muenker, A. Casanovas-Massana, J. Fournier, S. Bermejo, M. Campbell, R. Datta, A. Nelson, Y. I. R. Team, C. S. D. Cruz, A. I. Ko, A. Iwasaki, H. M. Krumholz, J. Matheus,

P. Hui, C. Liu, S. F. Farhadian, R. Sikka, A. L. Wyllie, N. D. Grubaugh, SalivaDirect: A simplified and flexible platform to enhance SARS-CoV-2 testing capacity. *medRxiv* (2020), doi:10.1101/2020.08.03.20167791.

109. S. Srivatsan, P. D. Han, K. van Raay, C. R. Wolf, D. J. McCulloch, A. E. Kim, E. Brandstetter, B. Martin, J. Gehring, W. Chen, S. F. S. Investigators, S. Kosuri, E. Q. Konnick, C. M. Lockwood, M. J. Rieder, D. A. Nickerson, H. Y. Chu, J. Shendure, L. M. Starita, Preliminary support for a "dry swab, extraction free" protocol for SARS-CoV-2 testing via RT-qPCR. *Biorxiv* (2020), doi:10.1101/2020.04.22.056283.

110. C. Beltrán-Pavez, C. L. Márquez, G. Muñoz, F. Valiente-Echeverría, A. Gaggero, R. Soto-Rifo, G. P. Barriga, SARS-CoV-2 detection from nasopharyngeal swab samples without RNA extraction. *Biorxiv* (2020), doi:10.1101/2020.03.28.013508.

111. S. K. Wee, S. P. Sivalingam, E. P. H. Yap, Rapid direct nucleic acid amplification test without RNA extraction for SARS-CoV-2 using a portable PCR thermocycler. *Biorxiv* (2020), doi:10.1101/2020.04.17.042366.

112. I. Yelin, N. Aharony, E. S. Tamar, A. Argoetti, E. Messer, D. Berenbaum, E. Shafran, A. Kuzli, N. Gandali, T. Hashimshony, Y. Mandel-Gutfreund, M. Halberthal, Y. Geffen, M. Szwarcwort-Cohen, R. Kishony, Evaluation of COVID-19 RT-qPCR test in multi-sample pools. *medRxiv* (2020), doi:10.1101/2020.03.26.20039438.

113. A. S. Booeshaghi, N. B. Lubock, A. R. Cooper, S. W. Simpkins, J. S. Bloom, J. Gehring, L. Luebbert, S. Kosuri, L. Pachter, Fast and accurate diagnostics from highly multiplexed sequencing assays. *medRxiv* (2020), doi:10.1101/2020.05.13.20100131.

114. G. A. Obande, K. K. B. Singh, Current and Future Perspectives on Isothermal Nucleic Acid Amplification Technologies for Diagnosing Infections. *Infect Drug Resist*. **13**, 455–483 (2020).

115. B. A. Rabe, C. Cepko, SARS-CoV-2 detection using isothermal amplification and a rapid, inexpensive protocol for sample inactivation and purification. *Proc National Acad Sci*. **117**, 24450–24458 (2020).

116. Y. Zhang, N. Odiwuor, J. Xiong, L. Sun, R. O. Nyaruaba, H. Wei, N. A. Tanner, Rapid Molecular Detection of SARS-CoV-2 (COVID-19) Virus RNA Using Colorimetric LAMP. *medRxiv* (2020), doi:10.1101/2020.02.26.20028373.

117. S. Bhadra, T. E. Riedel, S. Lakhotia, N. D. Tran, A. D. Ellington, High-surety isothermal amplification and detection of SARS-CoV-2, including with crude enzymes. *bioRxiv* (2020), doi:10.1101/2020.04.13.039941.

118. J. Joung, A. Ladha, M. Saito, M. Segel, R. Bruneau, M. W. Huang, N.-G. Kim, X. Yu, J. Li, B. D. Walker, A. L. Greninger, K. R. Jerome, J. S. Gootenberg, O. O. Abudayyeh, F. Zhang, Point-of-care testing for COVID-19 using SHERLOCK diagnostics. *medRxiv* (2020), doi:10.1101/2020.05.04.20091231.

119. F. Zhang, O. O. Abudayyeh, J. S. Gootenberg, A protocol for detection of COVID-19 using CRISPR diagnostics (2020) (available at https://static1.squarespace.com/static/5b7c640be2ccd1703a3da4d3/t/5e7773f40df723159b16f262/1584886772519/COVID-19+detection+%28v.20200321%29.pdf).

120. J. Qian, S. A. Boswell, C. Chidley, Z. Lu, M. E. Pettit, B. L. Gaudio, J. M. Fajnzylber, R. T. Ingram, R. H. Ward, J. Z. Li, M. Springer, An enhanced isothermal amplification assay for viral detection. *Nat Commun*. **11**, 5920 (2020).

121. D. Zucha, P. Androvic, M. Kubista, L. Valihrach, Performance Comparison of Reverse Transcriptases for Single-Cell Studies. *Clin Chem*. **66**, 217–228 (2019).

122. J. Qian, Z. Lu, C. P. Mancuso, H.-Y. Jhuang, R. del C. Barajas-Ornelas, S. A. Boswell, F. H. Ramírez-Guadiana, V. Jones, A. Sonti, K. Sedlack, L. Artzi, G. Jung, M. Arammash, M. E. Pettit, M. Melfi, L. Lyon, S. V. Owen, M. Baym, A. S. Khalil, P. A. Silver, D. Z. Rudner, M. Springer, Barcoded microbial system for high-resolution object provenance. *Science*. **368**, 1135–1140 (2020).

123. K. M. Koczula, A. Gallotta, Lateral flow assays. *Essays Biochem*. **60**, 111–120 (2016).

124. G. Ruano, E. M. Pagliaro, T. R. Schwartz, K. Lamy, D. Messina, R. E. Gaensslen, H. C. Lee, Heat-soaked PCR: an efficient method for DNA amplification with applications to forensic analysis. *Biotechniques*. **13**, 266–74 (1992).

125. A. W. H. Chin, J. T. S. Chu, M. R. A. Perera, K. P. Y. Hui, H.-L. Yen, M. C. W. Chan, M. Peiris, L. L. M. Poon, Stability of SARS-CoV-2 in different environmental conditions. *medRxiv* (2020), doi:10.1101/2020.03.15.20036673.

126. B. L. Pasloske, W. Wu, Method and reagents for inactivating ribonucleases RNase A, RNase I and RNase T1. *Google Patents* (2004).

127. X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, G. M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*. **26**, 672–675 (2020).

128. S. Zheng, J. Fan, F. Yu, B. Feng, B. Lou, Q. Zou, G. Xie, S. Lin, R. Wang, X. Yang, W. Chen, Q. Wang, D. Zhang, Y. Liu, R. Gong, Z. Ma, S. Lu, Y. Xiao, Y. Gu, J. Zhang, H. Yao, K. Xu, X.

Lu, G. Wei, J. Zhou, Q. Fang, H. Cai, Y. Qiu, J. Sheng, Y. Chen, T. Liang, Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. *BMJ*. **369**, m1443 (2020).

129. Y. Pan, D. Zhang, P. Yang, L. L. M. Poon, Q. Wang, Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect Dis*. **20**, 411–412 (2020).

130. Y. M. Bar-On, A. Flamholz, R. Phillips, R. Milo, SARS-CoV-2 (COVID-19) by the numbers. *Elife*. **9**, e57309 (2020).

131. E. Pasomsub, S. P. Watcharananan, K. Boonyawat, P. Janchompoo, G. Wongtabtim, W. Suksuwan, S. Sungkanuparph, A. Phuphuakrat, Saliva sample as a non-invasive specimen for the diagnosis of coronavirus disease 2019: a cross-sectional study. *Clin Microbiol Infec*. **27**, 281–285 (2020).

132. H. R. Hasan, Y. A. H. Al-Issa, A study of Rnase enzymes in saliva samples from women with breast tumors. *Med. J. Islamic World Acad. Sci.*, 21–26 (2011).

133. A. B. Acquier, A. K. D. C. Pita, L. Busch, G. A. Sánchez, Comparison of salivary levels of mucin and amylase and their relation with clinical parameters obtained from patients with aggressive and chronic periodontal disease. *J Appl Oral Sci*. **23**, 288–294 (2015).

134. A. L. Wyllie, J. Fournier, A. Casanovas-Massana, M. Campbell, M. Tokuyama, P. Vijayakumar, B. Geng, M. C. Muenker, A. J. Moore, C. B. F. Vogels, M. E. Petrone, I. M. Ott, P. Lu, A. Venkataraman, A. Lu-Culligan, J. Klein, R. Earnest, M. Simonov, R. Datta, R. Handoko, N. Naushad, L. R. Sewanan, J. Valdez, E. B. White, S. Lapidus, C. C. Kalinich, X. Jiang, D. J. Kim, E. Kudo, M. Linehan, T. Mao, M. Moriyama, J. E. Oh, A. Park, J. Silva, E. Song, T. Takahashi, M. Taura, O.-E. Weizman, P. Wong, Y. Yang, S. Bermejo, C. Odio, S. B. Omer, C. S. D. Cruz, S.

Farhadian, R. A. Martinello, A. Iwasaki, N. D. Grubaugh, A. I. Ko, Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *medRxiv* (2020), doi:10.1101/2020.04.16.20067835.

135. J.-H. Seol, H.-J. Kim, Y.-J. Yang, S.-T. Kim, H.-D. Youn, J.-W. Han, H.-W. Lee, E.-J. Cho, Different roles of histone H3 lysine 4 methylation in chromatin maintenance. *Biochem Bioph Res Co*. **349**, 463–470 (2006).

136. J. Arizti-Sanz, C. A. Freije, A. C. Stanton, B. A. Petros, C. K. Boehm, S. Siddiqui, B. M. Shaw, G. Adams, T.-S. F. Kosoko-Thoroddsen, M. E. Kemball, J. N. Uwanibe, F. V. Ajogbasile, P. E. Eromon, R. Gross, L. Wronka, K. Caviness, L. E. Hensley, N. H. Bergman, B. L. MacInnis, C. T. Happi, J. E. Lemieux, P. C. Sabeti, C. Myhrvold, Streamlined inactivation, amplification, and Cas13-based detection of SARS-CoV-2. *Nat Commun*. **11**, 5921 (2020).

137. H.-Y. Cheng, S.-W. Jian, D.-P. Liu, T.-C. Ng, W.-T. Huang, H.-H. Lin, T. C.-19 O. I. Team, Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods Before and After Symptom Onset. *Jama Intern Med*. **180**, 1156–1163 (2020).

138. R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, C. Wendtner, Virological assessment of hospitalized patients with COVID-2019. *Nature*. **581**, 465–469 (2020).

139. Joint United Nations Programme on HIV/AIDS. *Communities at the Centre* (2019).

140. M. P. Cheng, J. Papenburg, M. Desjardins, S. Kanjilal, C. Quach, M. Libman, S. Dittrich, C. P. Yansouni, Diagnostic Testing for Severe Acute Respiratory Syndrome–Related Coronavirus 2: A Narrative Review. *Ann Intern Med*. **172**, 726–734 (2020).

141. S. Tyagi, D. P. Bratu, F. R. Kramer, Multicolor molecular beacons for allele discrimination. *Nat Biotechnol*. **16**, 49–53 (1998).

142. J. S. Gootenberg, O. O. Abudayyeh, M. J. Kellner, J. Joung, J. J. Collins, F. Zhang, Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science*. **360**, eaaq0179 (2018).

143. L. Lillis, D. Lehman, M. C. Singhal, J. Cantera, J. Singleton, P. Labarre, A. Toyama, O. Piepenburg, M. Parker, R. Wood, J. Overbaugh, D. S. Boyle, Non-Instrumented Incubation of a Recombinase Polymerase Amplification Assay for the Rapid and Sensitive Detection of Proviral HIV-1 DNA. *Plos One*. **9**, e108189 (2014).

144. N. L. Washington, S. White, K. M. S. Barrett, E. T. Cirulli, A. Bolze, J. T. Lu, S gene dropout patterns in SARS-CoV-2 tests suggest spread of the H69del/V70del mutation in the US. *medRxiv* (2020), doi:10.1101/2020.12.24.20248814.

145. T. Peter, D. Ellenberger, A. A. Kim, D. Boeras, T. Messele, T. Roberts, W. Stevens, I. Jani, A. Abimiku, N. Ford, Z. Katz, J. N. Nkengasong, Early antiretroviral therapy initiation: access and equity of viral load testing for HIV treatment monitoring. *Lancet Infect Dis*. **17**, e26–e29 (2017).

146. C. Plachot, C. Huang, R. Vemula, J. Wu, H. J. Lee, B. Anekella, Use of recombinant virus technology to produce non-infectious, whole process controls for emerging viruses such as Ebola, Chikungunya, Dengue-2, Norovirus GII, MERS-CoV and Zika. *J Clin Virol*. **82**, S49 (2016).

147. P. Fozouni, S. Son, M. D. de L. Derby, G. J. Knott, C. N. Gray, M. V. D'Ambrosio, C. Zhao, N. A. Switz, G. R. Kumar, S. I. Stephens, D. Boehm, C.-L. Tsou, J. Shu, A. Bhuiya, M. Armstrong, A. R. Harris, P.-Y. Chen, J. M. Osterloh, A. Meyer-Franke, B. Joehnk, K. Walcott, A. Sil, C. Langelier, K. S. Pollard, E. D. Crawford, A. S. Puschnik, M. Phelps, A. Kistler, J. L. DeRisi, J. A. Doudna, D. A. Fletcher, M. Ott, Amplification-free detection of SARS-CoV-2 with CRISPR-Cas13a and mobile phone microscopy. *Cell*. **184**, 323-333.e9 (2021).

148. N. Kolluri, N. Albarran, A. Fan, A. Olson, M. Sagar, A. Young, J. Gomez-Marquez, C. M. Klapperich, SNAPflex: a paper-and-plastic device for instrument-free RNA and DNA extraction from whole blood. *Lab Chip*. **20**, 3386–3398 (2020).

149. M. Staats, A. J. Arulandhu, B. Gravendeel, A. Holst-Jensen, I. Scholtens, T. Peelen, T. W. Prins, E. Kok, Advances in DNA metabarcoding for food and wildlife forensic species identification. *Anal Bioanal Chem*. **408**, 4615–4630 (2016).

150. M. Ibrahim, T.-C. Liang, K. Scott, K. Chakrabarty, R. Karri, Molecular Barcoding as a Defense Against Benchtop Biochemical Attacks on DNA Fingerprinting and Information Forensics. *Ieee T Inf Foren Sec*. **15**, 3595–3609 (2020).

151. S. G. Kim, M. H. Noh, H. G. Lim, S. Jang, S. Jang, M. A. G. Koffas, G. Y. Jung, Molecular parts and genetic circuits for metabolic engineering of microorganisms. *Fems Microbiol Lett*. **365** (2018), doi:10.1093/femsle/fny187.

152. F. Stirling, L. Bitzan, S. O'Keefe, E. Redfield, J. W. K. Oliver, J. Way, P. A. Silver, Rational Design of Evolutionarily Stable Microbial Kill Switches. *Mol Cell*. **68**, 686-697.e3 (2017).

153. D. Pinto, S. Vecchione, H. Wu, M. Mauri, T. Mascher, G. Fritz, Engineering orthogonal synthetic timer circuits based on extracytoplasmic function σ factors. *Nucleic Acids Res*. **46**, gky614- (2018).

154. D. J. Burgess, CRISPR as a molecular clock. *Nat Rev Genet*. **22**, 200–200 (2021).

155. C. M. Ackerman, C. Myhrvold, S. G. Thakku, C. A. Freije, H. C. Metsky, D. K. Yang, S. H. Ye, C. K. Boehm, T.-S. F. Kosoko-Thoroddsen, J. Kehe, T. G. Nguyen, A. Carter, A. Kulesa, J. R. Barnes, V. G. Dugan, D. T. Hung, P. C. Blainey, P. C. Sabeti, Massively multiplexed nucleic acid detection with Cas13. *Nature*. **582**, 277–282 (2020).

156. M. Patchsung, K. Jantarug, A. Pattama, K. Aphicho, S. Suraritdechachai, P. Meesawat, K. Sappakhaw, N. Leelahakorn, T. Ruenkam, T. Wongsatit, N. Athipanyasilp, B. Eiamthong, B. Lakkanasirorat, T. Phoodokmai, N. Niljianskul, D. Pakotiprapha, S. Chanarat, A. Homchan, R. Tinikul, P. Kamutira, K. Phiwkaow, S. Soithongcharoen, C. Kantiwiriyawanitch, V. Pongsupasa, D. Trisrivirat, J. Jaroensuk, T. Wongnate, S. Maenpuen, P. Chaiyen, S. Kamnerdnakta, J. Swangsri, S. Chuthapisith, Y. Sirivatanauksorn, C. Chaimayo, R. Sutthent, W. Kantakamalakul, J. Joung, A. Ladha, X. Jin, J. S. Gootenberg, O. O. Abudayyeh, F. Zhang, N. Horthongkham, C. Uttamapinant, Clinical validation of a Cas13-based assay for the detection of SARS-CoV-2 RNA. *Nat Biomed Eng*. **4**, 1140–1149 (2020).

157. Q. Wu, C. Suo, T. Brown, T. Wang, S. A. Teichmann, A. R. Bassett, *Biorxiv*, in press, doi:10.1101/2020.06.01.127019.

158. R. M. Martinez, Clinical Samples for SARS-CoV-2 Detection: Review of the Early Literature. *Clin Microbiol Newsl*. **42**, 121–127 (2020).

159. M. Santiago, C. A. Ramírez-Sarmiento, R. A. Zamora, L. P. Parra, Discovery, Molecular Mechanisms, and Industrial Applications of Cold-Active Enzymes. *Front Microbiol*. **7**, 1408 (2016).

160. A. Chien, D. B. Edgar, J. M. Trela, Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J Bacteriol*. **127**, 1550–1557 (1976).

161. T. A. Coulther, H. R. Stern, P. J. Beuning, Engineering Polymerases for New Functions. *Trends Biotechnol*. **37**, 1091–1103 (2019).

162. J. N. Milligan, R. Shroff, D. J. Garry, A. D. Ellington, Evolution of a Thermophilic Strand-Displacing Polymerase Using High-Temperature Isothermal Compartmentalized Self-Replication. *Biochemistry-us*. **57**, 4607–4619 (2018).