# CRISPR Biosensors for Resource-limited Nucleic Acid Detection

by

Deborah Anne Najjar

B.Eng, The Cooper Union (2016)
S.M, Massachusetts Institute of Technology (2018)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022



Author .....
Program in Media Arts and Sciences,
School of Architecture and Planning,
August 19, 2022

Certified by .....
Joseph A. Paradiso
Alexander W. Dreyfoos (1954) Professor
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by .....
Tod Machover
Academic Head, Program for Media Arts and Sciences

# CRISPR Biosensors for Resource-limited Nucleic Acid Detection

by

Deborah Anne Najjar

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on August 19, 2022, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences

## Abstract

To achieve a healthier relationship with our environment and each other, we must continue to improve our ability to rapidly and sensitively monitor pathogens both inside and outside the body. Unfortunately, biological sensing has lagged behind electronic and chemical sensors both in cost and accessibility, primarily due to the need for specialized equipment, sterile work spaces, and sensitive reagents to operate biosensors. The COVID-19 pandemic has only emphasized the need for more sensitive, rapid, and decentralized biosensing solutions that can provide in-the-moment data for personal and public health-related decision making. Recent advances in CRISPR-based biosensors has allowed for a new class of diagnostics with sequence-specific nucleic acid detection capabilities that can provide a rapid response without the need for traditional laboratory infrastructure.

The research presented in this dissertation aims to further characterize and expand the applicability of CRISPR-based biosensor systems for resource-limited contexts by non-specialist users. Contributions include a minimally instrumented implementation of a CRISPR-based SHERLOCK assay for rapid and decentralized point-of-care detection of SARS-CoV-2 RNA and variants, a microfluidic platform for SARS-CoV-2 antibody and CRISPR-based RNA detection through multiplexed electrochemical sensing, and a field-deployable magnetic bead-based waterborne pathogen concentration and CRISPR-based detection system for environmental monitoring. This work also examines how local and indigenous knowledge figure into the conceptualization, collection, and utilization of environmental data within local monitoring programs and considers how novel biosensing tools could generate data at the appropriate resolution for community monitoring needs.

Thesis Supervisor: Joseph A. Paradiso
Title: Alexander W. Dreyfoos (1954) Professor
Program in Media Arts and Sciences

**CRISPR Biosensors for Resource-limited Nucleic Acid Detection**

Deborah Anne Najjar

This dissertation has been reviewed and approved by the following committee members:

Thesis Supervisor .....

Professor Joseph A. Paradiso
Alexander W. Dreyfoos (1954) Professor
Program in Media Arts and Sciences
Massachusetts Institute of Technology

**CRISPR Biosensors for Resource-limited Nucleic Acid Detection**

Deborah Anne Najjar

This dissertation has been reviewed and approved by the following committee members:

Member, Thesis Committee .....

Professor James J. Collins
Termeer Professor of Bioengineering
Department of Biological Engineering /Institute for Medical Engineering and Sciences
Massachusetts Institute of Technology

**CRISPR Biosensors for Resource-limited Nucleic Acid Detection**

Deborah Anne Najjar

This dissertation has been reviewed and approved by the following committee members:

Member, Thesis Committee .....

Professor Jennifer Gabrys
Chair in Media, Culture and Environment
Department of Sociology
University of Cambridge

# Acknowledgments

Oh! What a time, to be here at the acknowledgements page of my dissertation. I am, I must admit, a bit surprised and not a little bit proud to have reached this moment. My PhD path had a great deal of twists and turns, wherein I changed my research and my advisors and then entered into a global pandemic that changed my research even more. But through it all, I have felt like the luckiest person, to be able to keep busy and do research that allowed me to grow a great deal as a researcher and as a person. It is impossible to have been able to achieve even small bits of this alone.

So first and foremost – to Prof. Joe Paradiso and Prof. Jim Collins – who took me in with arms open and very few questions asked, and who trusted that providing me with leeway would result in some very interesting and publish-able results! You both have been so supportive and I could not have asked for a better combination of sounding boards. To Prof. Jennifer Gabrys – who I cold emailed after I reread your papers and realised I should just take a chance and reach out! Your work is so deeply inspiring to me and you have provided so many resources and recommendations that have deepened my inquiry into the worlds of sociology and participatory ecological monitoring.

To my lab mates!
On the ResEnv side – Brian, Caroline, DD, Fangzheng, Julianna, Irmandy, Patrick, Ali, David, Sam, Ariel, Mark – you are all so endlessly knowledgeable, so kind, so fun, and quick to help. It has been such a treat to get to work in the pen and listen in on all of the projects that are constantly cooking.

On the Collin Lab side – Helena and Rose, you are two of my most favorite people to collaborate with. Thank you for teaching me so much and getting excited about all of the little 3D printed gadgets. To Nikki, Nina, Xiao, Angelo, Nico, Makoto, Kehan, Woojung, Nayoung, Peter, and Evan, coming to the Wyss every day (even and especially during COVID times) was such a true joy. You are all wells of endless insights and have provided an equal amount of support and comic relief throughout this whole experience.

To Mrs. Greenstein, my kindergarten and then eight grade math teacher, who changed my whole life.

To the communities that I work with – thank you from the bottom of my heart for your generosity of time and spirit. Being on the Mystic and at Mahia were not only very welcome reprieves from the daily grind of wet lab life, but hugely informative and highly motivating experiences that I will keep close to me far beyond these pages.

To the MAS offices, who made things happen over the course of my time – Linda, Monica, Mahy, Rebecca, Sarra, and to the Directors office, Comms team, and Finance office who make sure that things keep running, thank you thank you for all the

(kosher) food and time and support and laughs. To Candido, Cornell, and Kevin, who keep the building (and coffee machines) running no matter what. To Will and Joana and the whole NecSys team who have helped me with computer nightmares on many an occasion, always with an abundance of competence and grace. To Hannah, who always has the biggest smile on her face and makes sure that things are tidy and sanitary.

To all of my MAS friends! Oh boy. Too many to list, which I think is a very great sign. Thank you thank you thank you. To my StudCom friends (especially Shannon, Pedro, Carmelo, Bianca, and Sunanda), and to all those from the Student Advocacy group who came together to support each other when it felt like there were no other options. To my cohort, full of such inspiring, interesting, and curious people who I learn so much from – it is such a pleasure to know you all. To the 99F crew, who welcomed me in when I feared the the parties were dying and let me spend more time on coordinating drinks and DJs than research (for a brief period of time). To Open Oceans – Katy, Jenni, and Alexis in particular, you have all been a life raft for me during the rough seas I experienced during this degree. To our Muddy Monday crew, one of the more unfortunate COVID casualties. To Maggie and Mehak, who are extremely smart and also very funny and hot, it's been a pleasure. To Daniella, who I hope will be my eyes and ears on MAS things when I'm gone. And to my dearest: Adam and Jake and Agnes and Sara and Susan! My love and admiration and gratitude is so big and bright. You are pieces of me that I will carry forever, and I am all the better for it.

To all of my non-MAS friends who have nonetheless been there for me throughout this journey, even as I was physically/mentally/emotionally unavailable for significant chunks of time– Noam, Jordan, Haley, Eytan, Sivan, Sarah, Lilah, Shira, Dalia, Ezra, Noah, Yonina, Ariana, Josh, Josh, Wade, Daphne, Olive, Mia, Emily, Adira, Mairav, Ahuva, Kayla, Chris, Jon, Eli, Nick, Jacob, and so many more – thank you for sticking with me anyway! I will try and make it up to all of you starting after this document is handed in.

To my dear friend (cringe) Avery (MSN) – it feels impossible to consider this document (within and) without you. Thank you for being a listening ear, enabler, proofreader, musician, culture snob, foodie, MMA fighter, movie critic, and gremlin. I think we've moved away from our co-dependancy during the almost unbearable Early Years and into something more mature and healthy (watching movies in our beds alone or else playing Yahtzee and Rummikub in a public park). You have pushed me in almost all the ways and it may have not made me the best version of myself, but it certainly made me the most authentic. I hope that when you release The Files one day that people turn to this acknowledgement and feel pity for me that I ever trusted you as a friend. Writing this feels so dumb and gross and corny – but that's showbiz, baby. And to Eva Paige, what can I say but Spaciba! You are not only a genius, fashion icon, and chess (grand)master, but also a truly harasho person. Queen Stays Queen. And to Barb and Rob and Bowie (and sure, Ryan too) for letting me into the

family and allowing me to cook Thanksgiving that one time, which was really insane in retrospect.

To the Brattle Theater, BSO, Assembly Row, Felipes, The Kosher Wok, Vittorio's (RIP), MIT's ceramics studio (Darrell!), The Muddy, Tatte, Dunkin, Clover, JP Licks, the MBTA, Amtrak, – I simply could not have done it without you all. My lifeblood!

To my family!

My grandparents who would have (I hope!) been proud to see me – Chaim, Avigayil, Philip, and Gloria. I can only try and bring about as much love and good to the world that you all did in your lifetimes. To my grandfather Philip in particular, who completed coursework for a PhD in teaching but never finished it because he had a family to support, I always told you that I would be dedicating my PhD to you. You laughed, but here I am doing it! I wish you all could have been here to celebrate with me.

To Renanint, who reminded me that no one can be happy for me, I have to be happy for myself.

To my aunts, uncles, cousins, and extended family! You are the best parts of all of the important things of life. I wish that I had lived closer these past six years so I could attend more parties and get-togethers. I hope to continue celebrating the moments in life with you all forever!

My mom and dad, Ellen and Abe, who have given me lots of leeway to stray as far from home as the Greater Boston area (!) to spend lots of time getting this degree. Thank you for your support and for letting me be a little freak, it means the world to me. You have both given me the tools throughout my childhood to think critically, engage deeply, and to refuse to accept that the world as it is is how it should be. I am grateful for all the time in natural history museums, aquariums, and museums and sorry that it led me on a career path that has, so far, not produced much financially. Maybe one day! To Rebecca, Josh, Shira, Deena, Yair, Noam, and Aleh – you are all my greatest cheerleaders and greatest critics and I work very hard to try and make you all proud every day. I hope it worked. I love you all and want nothing more than to spend time with you (as long as it's not crammed in a car) loudly talking over each other.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Research Questions

With a title that indicates this dissertation will focus on biological sensing, sensing that uses biological molecules to detect a specific chemical or biological signature, I would first like to bring up the biological sensors that we use every day, the main ways that we interface with the world, our on-body sensors. Our bodies are an elaborate and intricate amalgamation of sensors that allow us to perceive the world around us (Fig. 1-1), and while these sensors are helpful for allowing us to navigate and regulate, they are calibrated within a very specific regime of frequencies – leaving most of the world invisible because it is imperceptible. This is the very reason why we design and build so many sensors that extend and augment our sensing abilities. From scopes and spectrometers to imagers and integrated circuits, we have been able to see from the galactic down to the subatomic and uncover the hidden molecular makeup of our surroundings. Sensors allow us to develop a fuller, richer, more complex, and more nuanced picture of our world and the systems within it.

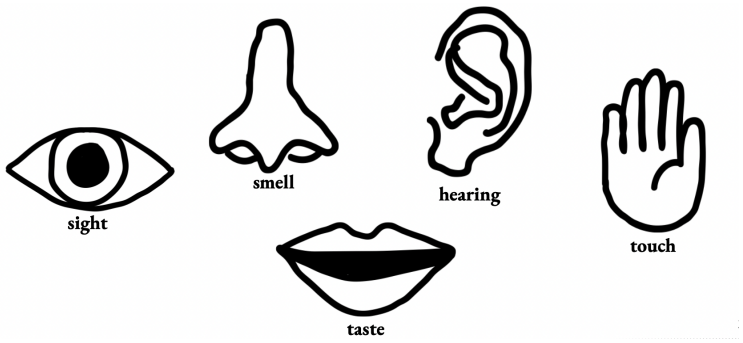



Figure 1-1: **The five senses.**

In recent years, there has been a move to open source and democratize these sensors so that anyone with time, interest, and some basic tools can monitor their surroundings. However, while it is fairly simple to design a system that will provide data on abiotic factors such as temperature, pressure, chemicals, radiation, and particulate

matter within air and water samples, there are currently few sensors widely accessible that provide a comparable level of sensitivity for monitoring biotic factors.

This is not due to the lack of highly sensitive tools available, which the field of biology has been developing at an astounding pace in recent decades, but more a disconnect in the pipeline between laboratory and lay person access to these technologies [27]. There is no better scenario to illustrate this issue than the current global pandemic. COVID-19 testing has been highly centralized, creating a massive strain on facilities and causing significant lags in the delivery of results to those who submit samples. These tests are often necessary for immediate decision making on how to safely engage in their environment, which means any delay increases local an global risks.

Why is it that there are so few tools available for people to sensitively and consistently sense things like microbes and pathogens outside of a laboratory? Why is it that there needs to be large amounts of money, time, effort, and even risk needed in order to determine these data? Similarly, testing for environmental pathogens such as waterborne pathogens from run-off and contamination is currently difficult for a non-laboratory facility to quickly or routinely perform. With climate change causing erratic weather events, be it storms or draughts, to be more prevalent [194], it is necessary for those who interact with and rely upon specific water sources to quickly and sensitively monitor aquatic pathogens. Why is it that it is currently so difficult to do? Therefore, the questions at the center of this research are:

1. Can we design highly sensitive biological sensing systems for improved pathogen detection and monitoring related to human and environmental health that are affordable, easy to use, and functional in resource-limited contexts?
2. Can we understand how these biosensors and the data they produce figure into existing monitoring frameworks by engaging with local and indigenous monitoring programs?

## 1.2 Biosensor Systems

There are countless examples of how sensors have improved our understanding of how the world operates. One example that has inspired this research is Tidmarsh, which is a project in the Responsive Environments group that monitors a retired cranberry bog in Massachusetts as it rewilds [145]. The project is a deep exploration in how a mesh of sensor nodes together with distributed audio and video feeds can provide insights into how a local ecology fluctuate over time. Another example is the AirKit project, a distributed air quality sensing system that collects data on particulate matter, temperature and humidity [140]. The researchers provided in-depth instructions for how to build the air sensor, as well as tools for data analysis and mapping so that the collected data can be shared online between citizens in real time. These projects are successful because they are powered by electrochemical sensors, which have become increasingly more affordable to fabricate and connect to bluetooth or

wifi networks for data transfer and data logging. Why is it that biosensing is not yet able to be miniaturized and automated on the same scale?

### 1.2.1 Current methods of biological sensing

One possibility is that biological sensing is a still a young field, with nucleic acids only first visualized in the 1950's [16]. The most common method (which is also the oldest, far predating the discovery of nucleic acids) for biosensing is directly culturing microbes. While this method can allow for quantitation, which means the ability to know the relative amount of microbes present in the sample, and doesn't require sequence specific knowledge of the microbe to try and isolate it, it typically take 48-72 hours to get results (this is only if the microbe can be grown in the media) and additionally requires someone to bring a sample to a laboratory and pay a premium [44]. Another common method of biosensing is polymerase chain reaction, or PCR, which was discovered in 1985 and was the first method available to amplify nucleic acids [154]. PCR is one of the most robust and well-developed methods in molecular biology, having been modified and updated over the years to even allow for workflows that provide real-time quantitation of nucleic acids [93]. PCR reactions are considered to be highly specific and sensitive to their target while being simple to design as long as the sequence of interest in known. However, there is a similarly long turnaround time for samples processed if done in a centralized laboratory facility (which is what most people who do not have access to a laboratory would need to do) and the additional cost of extracting and running the microbial material can be high.

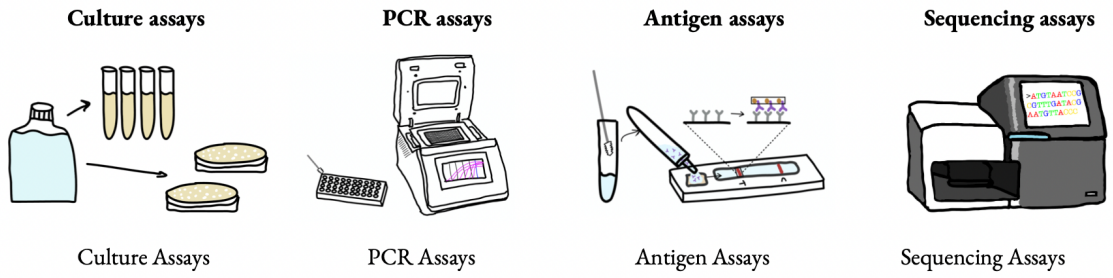



Figure 1-2: **Current methods in biosensing.** (L-R) Culturing assay, PCR assay, lateral flow assay, and sequencing assay

A third biosensing method is through antigen testing, which can have plate or lateral flow assay formats to allow for detection of biomolecules or analytes of interest in a laboratory or at home on the order of minutes. These assays typically sense biomolecules rather than nucleic acids, which makes these tests less specific than nucleic acid-based approaches [116]. Additionally, these assays often taking months or more to manufacture due to the laboratory development process, which requires the

isolation and affinity of specific antigens and/or antibodies to assure that the test is accurate. Finally, high-throughput sequencing has become a leading biosensing method in recent years, which has the ability to gather genetic information on not only one species of interest, but as many as are present within the sample [196]. However, the high-throughput nature of sequencing also means that running a reaction often produces gigabytes of data that requires complicated computational assembly and analysis to interpret. Additionally, while sequencing a specific DNA region from a purified samples has gone down in cost, sequencing complex samples or performing metagenomic analysis is still quite expensive. Descriptions of these instruments can be seen in (Fig. 1-2).

For any biosensing system to function in a resource-limited (which is to say, outside of a laboratory setting) context, a sensing system would need to be able to have a raw sample as an input and give some perceivable output. Only the antigen test is capable of doing this from the techniques mentioned above, but the sensitivity and specificity leave much to be desired. Are there methods that are outside of the these paradigms that we can adapt to allow us to sensitively monitor nucleic acid material in an out-of-lab setting? Below will highlight some of the advances within raw sample preparation, nucleic acid amplification, detection of the nucleic acid material, and low-cost readers which can help make next-generation biosensor systems a reality.

### 1.2.2 Sample preparation

One of the more challenging steps in making biosensors viable outside of a laboratory context is sample preparation. Many diagnostics require a variety of sample types such as blood, urine, stool, saliva, soil, or water. These sample types often contain inhibitors which can prevent biosensors from properly amplifying the desired nucleic acids or introduce a false positive signal into the assay. Sample also typically require a lysis step to release the nucleic acids and make them accessible for sensing [85]. Typically, samples are processed in a laboratory setting using nucleic acid extraction and concentration protocols that require expensive and specialized equipment, costly kits, and toxic chemicals to produce a product that is appropriate for a biosensor. Recently, there have been a variety of novel sample preparation protocols that implement combinations of heat, chemicals such as detergents, chelators, chaotropic salts, and enzymatic stabilizers, and mechanical disruption to both release nucleic acids and inactivate inhibitors and nucleases. Some of these protocols have the potential to improve low-resource sample preparation for out-of-lab biosensing applications.

After inactivation and lysis of the sample, the nucleic acids in the sample must be concentrated and purified. The most common methods for concentration has been paper- and filter-based sample preparations which use paper, glass, and thermoplastic filters [95, 206, 110, 244]. Another makes use of functionalized magnetic beads that can attract either whole microbes or lysed nucleic acid material [22, 2, 31, 5, 68, 37]. Paper- and magnetic-based methods of purification often follow extraction and pro-

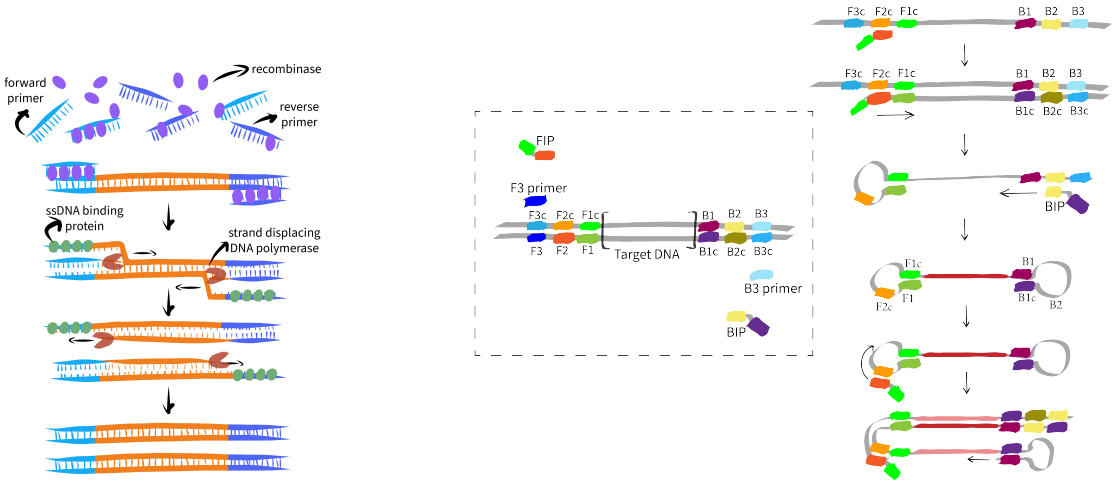
ceed a wash step, which further purifies the sample. In the case of many low-resource methods, there are trade-offs regarding overall purity of the final sample and nucleic acid loss due to the extraction method. The amplification method that follows purification will typically dictate the level of purity required for the nucleic acid material.

### 1.2.3 Nucleic acid amplification

Biosensors often need an amplification step (the same way a circuit may need an op-amp) to ensure a robust output signal even when sample concentrations are low. PCR is not ideal for resource-limited settings, because it requires a thermocycler and cold chain-dependent enzymes. Avoiding the use of a thermocycler greatly increases the usability of a method, which is why there have been a variety of isothermal amplification methods, including loop-mediated amplification(LAMP) [162], nucleic acid sequence-based amplification (NASBA) [48], rolling circle amplification (RCA) [4], strand displacement amplification (SDA) [227], and recombinase polymerase amplification (RPA) [203], that researchers have developed to to simplify the process. LAMP typically operates at 63-65℃; NASBA typically operates at 41℃; RCA typically operates at 25-37℃, SDA typically operates at 37℃ and RPA typically operates at 37-41℃.

For many low-resource biosensor applications, RPA or LAMP tend to be the preferred method of amplification (Fig. 1-3a). For RPA, the amplification process takes around 30 minutes and uses a single set of primers together with single-stranded DNA (ssDNA) binding proteins and a strand-displacing DNA polymerase to exponentially amplify the DNA target of interest. RPA has been shown to work well for a wide range of organisms and sample types [135] and can amplify RNA if a reverse transcriptase is added to the reaction.

On the other hand, LAMP requires a higher operating temperature of 63-65℃, which makes it difficult to perform without a heating element. It is a rapid amplification method that can work on both DNA and RNA templates and uses 2-3 pairs of primers together with strand-displacing activity and polymerases to rapidly generate large concatemers of DNA, with signal saturation occurring within 15-30 minutes after the start of the reaction. There are direct colorimetric output methods for both RPA and LAMP, with RPA implementing a labeling process in the amplification that allows for a lateral flow assay-based readout and LAMP implementing a pH-based color change that relies on the production of phosphate ions as a byproduct of the amplification process. LAMP has the advantage of being widely available through multiple providers, whereas RPA is patented and produce by a single company and therefore subject to issues with pricing, quality control, and production.




(a) (b)

Figure 1-3: **Nucleic acid amplification methods.** (a) RPA amplification scheme (Adapted from [135]) (b)LAMP amplification scheme. (Adapted from [216])

### 1.2.4 Nucleic acid detection

In recent years, there have been many new ways of detecting nucleic acids in addition to direct detection of amplified products, which can have issues with sensitivity and specificity with respect to the nucleic acid target of interest. Methods such as Förster resonance energy transfer (FRET)-based nucleic acid sensing [174], surface-enhanced raman scattering (SERS) spectroscopy-based nucleic acid sensing [73], nucleic acid aptamer-based sensors and [117], and toehold switch-based sensors [84] to name a few, have been developed in recent years to try and improve sensitivity. While all of these methods are promising, the method that this dissertation will focus on is that of CRISPR-based biosensing [127], which has expanded rapidly in the past few years as CRISPR-Cas systems have been further characterized.

CRISPR-Cas, an immune-like response employed by some microbes to fight viruses, is a family of enzymes that can be implemented with only two molecular components: a Cas enzyme, which acts as a pair of molecular scissors, and a guide RNA, which acts as a set of instructions that directs the Cas enzyme to the target region of DNA or RNA to cut. Depending on the choice of nuclease and amplification method, CRISPR-based biosensors can be operated in resource-restricted areas on raw samples and can provide quantitative data regarding the target of interest.

When choosing a CRISPR nuclease, there are many characteristics to consider, including the size of the CRISPR enzyme, guide requirements, cutting activity, and temperature for optimal activity. Type V CRISPR enzyme Cas12a (Fig. 1-4) and Type VI CRISPR enzyme Cas13a in particular have helped to have ushered in a new class of diagnostic biosensors due to their ability to activate collateral cleavage of

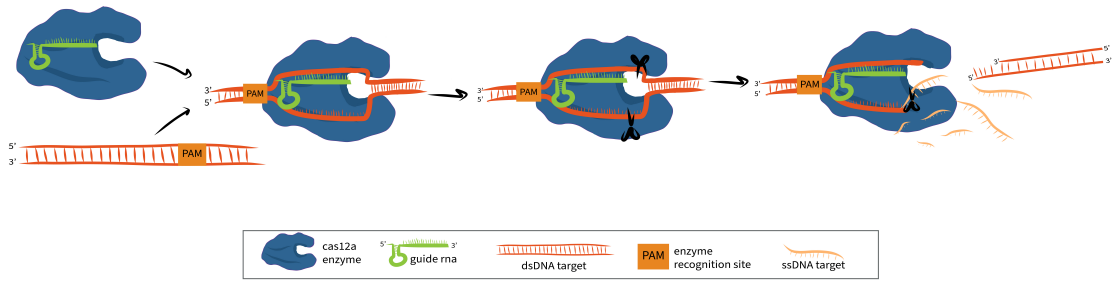



Figure 1-4: **Cas12a collateral cleavage.** A diagram showing the Cas12a-guide RNA complex localizing to a target strand of DNA and cleaving, which activates the single-stranded DNA collateral cleavage activity of the enzyme.

single-stranded nucleic acid (ssDNA in the case of Cas12a and ssRNA in the case of Cas13a) upon binding to and cutting their target sequence. This feature was first reported by [81, 80] and [35] as a novel and sensitive method of nucleic acid detection called SHERLOCK (Specific High-sensitivity Enzymatic Reporter unLOCKing). The biosensor works by adding short ssDNA or ssRNA with a blocked fluorescence molecule into a CRISPR reaction. If the Cas enzyme finds and cuts its target, it will cut and unblock the single-stranded nucleic acid, releasing a fluorescence signal (Fig. 1-5) [112].

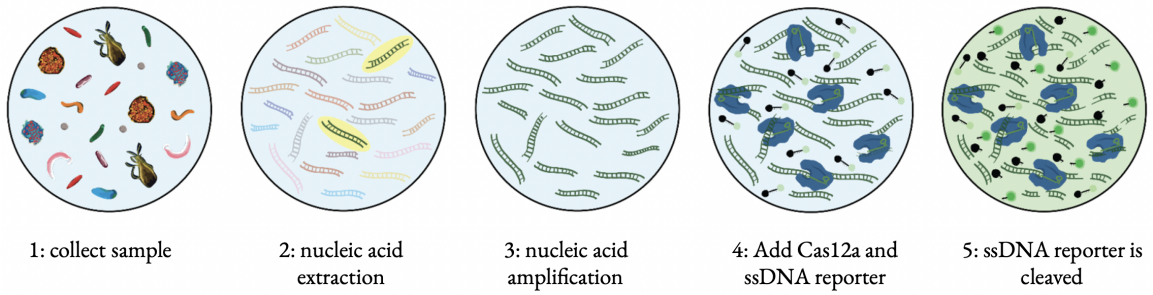
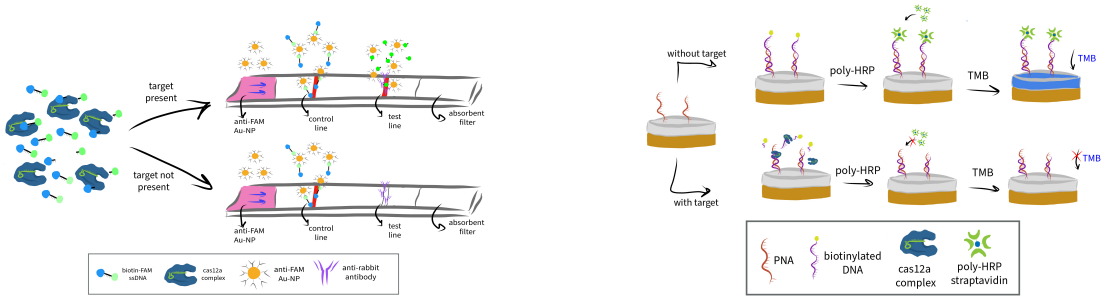



Figure 1-5: **SHERLOCK workflow.**An example of a SHERLOCK biosensor workflow, where a sample with microbial species is collected, the nucleic acids are extracted and amplified, the CRISPR mixture with complexed Cas12a and a quenched ssDNA fluorescent probe, which is unquenched upon the Cas12a finding and cleaving its target.

SHERLOCK reactions have allowed for highly sequence-specific nucleic acid-targeting biosensors, able to detect even single base changes, that provide a response on the order of minutes rather than hours to days typically needed to sequence or culture samples [105]. Additionally, these sensors have detected targets present at the attomolar ($10^{-}18$) level and have successfully been freeze-dried, which makes dilute field sample testing a possibility [155, 106, 121, 105].




(a) (b)

Figure 1-6: **Nucleic acid detection methods.** (a) CRISPR-based lateral flow read-out using a FAM-biotin reporter (b) CRISPR-based electrochemical read-out using a biotin-based HRP-TMB reporter

### 1.2.5 Sample read-out

In order to read a CRISPR biosensor output, it is necessary to have a detection mechanism. For a biotin-based lateral flow reaction, no additional hardware is needed to real the gold nanoparticle-based colorimetric output that a user can see (Fig. 1-6a). For a biotin-based electrochemical assay, a sensitive charged surface is required in order to monitor the output (Fig. 1-6b) [187]. For fluorescence-based reactions, a fluorometer or transilluminator is needed in order to visualize the reaction output (Fig. Fig. 1-7).

For many scientists, making these biological assays and sensors available in an out-of-lab context can be daunting, because validating and publishing their work typically requires meticulously controlled experiments that are extensively validated on expensive, complex, and sensitive instruments. There is a messiness involved in performing biological science outside of a wet lab that is difficult to frame in a traditional (publishable) experimental context. Furthermore, much of the work of getting biosensors to users involves translational work, such as product testing for functionality, viability, and life cycle and design testing for legibility and overall user experience, which all tend to be outside the purview of a traditional molecular biologist. Examples of biosensors that make it out to the field and few and far between [104, 108, 27]. In order to approach this issue, bioengineers can step into the world of participatory research, which can aid in the to development of new devices and field-tested frameworks that can move biosensing outside the lab.

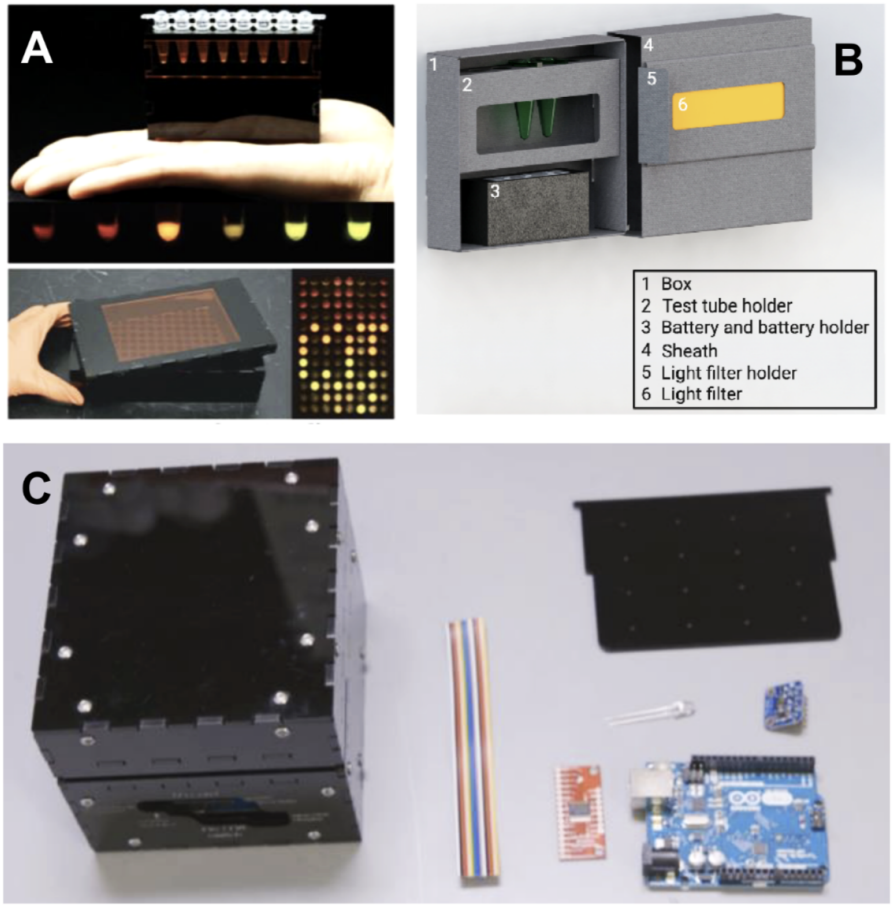



Figure 1-7: **Low-cost fluorometer.** (a) the BioBits™Bright reader and incubator was used in classroom settings for gene circuit education [201] (b) a paper-based transilluminator with optional heater for point-of-care diagnostics (c) and a portable electronic reader for toehold-based reactions.

## 1.3 Biosensing and Local Data

Pathogens spreading from point sources – be it a human, a contaminated piece of chicken, a well, a nonhuman animal, or a sink handle – are a few of the countless microbial interactions that humans encounter a daily basis. While much has been written about the benefits of low-cost and point-of-care diagnostics in improving the world's diagnosis and treatment of illness, particularly for resource-limited settings [220, 236, 94], I believe that these diagnostics for disease detection are a specific case within the larger category of environmental sensing. As discussed in the above section, we have only recently begun developed the type of technology capable of making visible these relationships. Focusing on the ways in which we can improve our understanding of and interactions with our shared environment can bring about mutual understanding that impacts the ways in which we engage at every level.

### 1.3.1 Evidencing the Environmental

Biological sensing is a choice to participate in multi-species encounters that have the ability to shift perspectives and ways of relating to, and being in, the world. It also allows for humans to more deeply interrogate the roles they play within larger ecosystems. Microbes are often sentinels that alert humans of changes within an ecosystem [76]. These microorganisms are critical to long-term environmental health and currently at risk due to the impacts of climate change on the environment [32]. Building sensors systems that allow humans to understand the changing landscape of microorganisms is a way of not only building stronger relationships of understanding between humans and microbes, but also to evidencing the harm that is and will continue to impact these vulnerable communities and consequently, our larger ecosystems.

Looking through the broad lens of environmental monitoring opens up methodologies from the sciences and environmental studies, but also political ecology, participatory design, and citizen science. Combining the scientific with the social will be necessary to develop an understanding of how microbial monitoring practices could operate at both the personal and communal scale [198]. An environmental approach shifts the focus from humans towards a more expansive set of ecologies involved in our planetary well-being. In order to understand these relationships, we must design infrastructure that uncovers their existence. But like most relationships, ecological relationships are complicated. Generating sensor data, while at the heart of this dissertation, is only the first step, because it only one stream of data. Biosensors can, and should, be used together with other qualitative and quantitative data collection methods for a multi-modal approach to sensing that produces data sets with the same richness as their source material.

### 1.3.2 Biological Datafication

Therefore, future efforts to sustainably manage our planet will rely in part on our ability to implement monitoring programs that provide insights on how communities are being impacted by pandemics, climate change, industrial activity, urbanization, and other environmental and public health concerns. Environmental monitoring is always in-the-moment, but insights are often only gleaned when looking back on long-term data that is collected. Therefore, keeping methods and tools consistent means that the data collected is consistent with the rest of the data set, providing more to look back on and compare to – which is to say, time and consistency can make data more useful. Even when a monitoring programs moves to a new site, keeping methods and tools consistent means that the data from that site can be compared with other data sets previously collected (Fig. 1-8. What's more, each new data set collected provides more legitimacy for the previous data collected. This consistency of data can be a double-edged sword though, ultimately working against one of the benefits of monitoring, which is that it can be dynamic, allowing for iterations as it works to properly resolve its subject. In the case of academic researchers, they have grants,

areas of expertise, standard methodologies and specialized equipment that they have experience with (or else simply not enough time or funding to get new ones) which may keep them using a method long past when it is useful. In all cases (academic, local, governmental), there is the above question of the data set, and trying to collect data that serves the data set rather than the site [26].

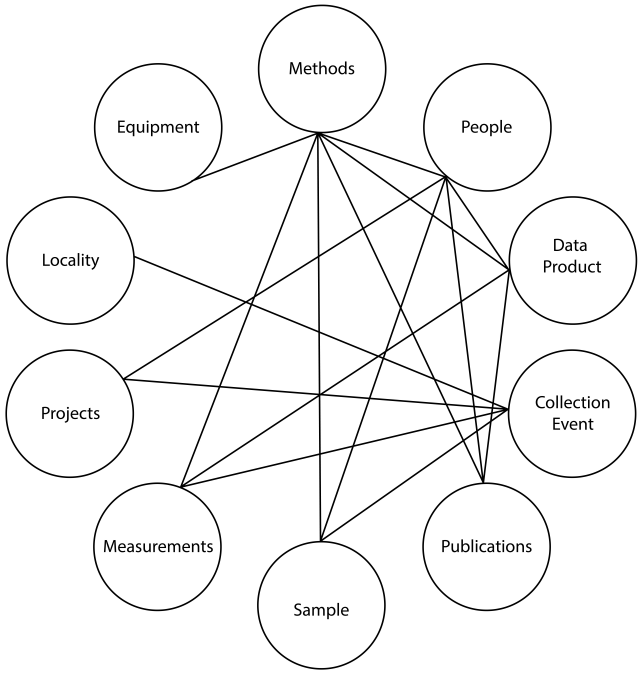



Figure 1-8: **A simplified description of data types and their relationships in environmental monitoring.** Circle represent data types and lines represent the relationships between the data types. (Figure re-made from: [211])

### 1.3.3 Participatory science and citizen engagement

Which is all to say, the world of research and monitoring is not always generous enough to provide the time and space for a flexible approach to developing a robust and responsive environmental monitoring project, even if that is the original intent. However, it is possible that incorporating citizen science and participatory monitoring can act as a frictional force (hopefully a good, productive, generative friction) against these pressures. Citizen science, where volunteers provide time, expertise, and momentum behind civic researcher efforts, has been studied by researchers who have an interest in how they form, how citizens participate, and the quality of data the programs generate [97, 103, 40, 158, 3]. Citizen science and participatory monitoring increases the stakeholders, the perspectives, the goals, the lived experience, and the expertise in a monitoring program, which has the ability to steer the data

and collection methods away from entrenched positions. A participatory approach to environmental monitoring can inform an iterative research design, because there are other people with their own thoughts and opinions and insights and interests who are involved. These extra voices necessarily slow down the process, but can also allow for a fuller discussion from a broader range of voices and opinions [82].

Engaging the public in efforts in citizen science work not only allows for large-scale and distributed research that would otherwise be impossible to accomplish, but the very act of engaging with the public allows for types of "misbehavior" and "overspill" from the intended protocol that can generate new and often unintended discoveries [151]. It also allows for the collection of data that may not be as highly controlled or "clean" as what can be collected with specialized equipment or highly trained personnel, but certainly can be "just good enough" to accomplish the monitoring task at hand and set the groundwork for future analysis or inquiry [70]. Developing multi-modal and participatory environmental sensing has the unique possibility of simultaneously producing new social relations among participants as well as new understandings and data on environmental change, allowing for both epistemological and ontological transformations [199].

## 1.4 Dissertation Overview

This dissertation is meant to provide a space to share my research on improving biosensing systems that have one foot in the wet lab and the other in the field, as well as ethnographic musings on how data and knowledge figure in citizen science and participatory monitoring programs undertaken by communities and how that can shape future programs and technologies within the space.

In **Chapter 2**, I discuss *miSHERLOCK*, a decentralized testing and disease monitoring platform that was designed during the COVID-19 pandemic to address the need for nucleic acid-based diagnostics for at-home use. The sample-to-answer system combines saliva sample inactivation and concentration, CRISPR-based combined amplification and detection reaction targeting SARS-CoV-2 and variants thereof, a customized reaction housing for incubation and transillumination, and a phone application to read reaction results and allow for data transmission to local authorities.

In **Chapter 3**, I discuss how the *miSHERLOCK* system can be adapted for another low-resource setting, field-based waterborne pathogen monitoring. This section also discusses two modifications to the reaction housing design, one to allow for concentration of nucleic acid material from large dilute environmental samples and another to accommodate zero-gravity reaction conditions.

In **Chapter 4**, I discuss *MELODe*, a lab-on-a-chip platform that allows for automated detection of multiple biomolecules of interest with an electrochemical-based output signal. The system uses a peristaltic pump-based microfluidic chip with areas for sample inactivation and concentration, reservoirs for the amplification and CRISPR-based detection reagents, a central reaction chamber for nucleic acid capture, amplification, and CRISPR-based detection, and an electrochemical chip for a highly sensitive current-based sensor output.

In **Chapter 5**, I discuss the political ecology of environmental sensing and data collection, how participatory monitoring methods can allow for multi-modal data streams, and an ethnography of two volunteer water monitoring programs which highlights how novel sensing technologies can figure into existing monitoring frameworks.

In **Chapter 6**, I discuss the practical knowledge gained over the course of this PhD and the and future directors for these biosensors.

# Chapter 2

# Minimally instrumented SHERLOCK for CRISPR-based point-of-care diagnosis

This chapter has been adapted from:

de Puig, H.*, Lee, R. A*., Najjar, D.*, Tan, X.*, Soenksen, L. R., Angenent-Mari, N. M., ... & Collins, J. J. (2021). Minimally instrumented SHERLOCK (miSHERLOCK) for CRISPR-based point-of-care diagnosis of SARS-CoV-2 and emerging variants. Science Advances, 7(32), eabh2944.

*Equal contribution, listed alphabetically

In this Chapter, I describe research that I contributed to in designing a minimally instrumented version of SHERLOCK that can be built and run in an at-home setting for distributed biosensing. This project was started in May 2020, during the COVID-19 pandemic. Therefore, we demonstrated the utility of this platform by designing a one-pot SHERLOCK assay that uses a saliva sample as an input and has the ability to differentiate between different variants of the SARS-CoV-2 genome. My contributions to this work include designing and building the miSHERLOCK device, collaborating on the experimental design, testing, and optimization of the SHERLOCK assays, troubleshooting the overall sample-to-answer workflow, co-writing the manuscript, and designing all figures for the manuscript. I would like to thank all of the other authors on the paper for their contributions to this work, it would have been impossible to accomplish all that we did in such a short span of time without coordination and teamwork.

## 2.1 Introduction

Significant progress has been made in the use of CRISPR (clustered regularly interspaced short palindromic repeats)/Cas components of adaptive microbial immunity in molecular diagnostics [205, 46, 163]. Several Cas effectors have been used as highly specific nucleic acid sensors that cause detectable collateral cleavage of engineered nucleic acid probes after target binding. SHERLOCK (specific high-sensitivity enzymatic reporter unlocking) [81, 80] and DETECTR (DNA endonuclease-targeted CRISPR trans reporter) [35] use Cas13a or Cas12a to create ultra-sensitive molecular diagnostics for a variety of targets, including infectious diseases such as Zika virus, Cytomegalovirus, BK virus, and Plasmodium species [121, 106, 127], with simplified readouts including lateral flow assays [80, 106, 127, 104].

Multiple CRISPR/Cas diagnostics have been created to target SARS-CoV-2 using varied viral purification, amplification, and detection methods [6, 28, 180, 11]. However, the vast majority of workflows still require multiple liquid-handling steps and laboratory equipment such as pipettes, centrifuges, and heating blocks, as well as the technical skills for their use. Although there are several home diagnostic tests approved for use by the US Food and Drug Administration (FDA), the vast majority of the tests involve self-collection followed by mailing to a central laboratory or are based on rapid antigen tests, which have been shown to be less accurate than nucleic acid-based testing with the potential for relatively high false negative and false positive results [66]. There is only one FDA-approved at-home nucleic acid-based test for SARS-CoV-2, which costs $50.00 USD and requires a physician's prescription [238]. These diagnostic methods have contributed to ongoing efforts to make SARS-CoV-2 testing as widely available as possible. However, despite these advances, all of the tests described above only allow for the general detection of SARS-CoV-2 and not of specific strains.

The relationship between specific strains and mutations and changes in virulence and viral behavior are actively under study [209]. New variants may affect transmissibility, treatment efficacy, and the degree of immunity that is generated by both natural infection and immunization [91]. Of particular concern are variants B.1.1.7 (originally discovered in the United Kingdom), B.1.351 (originally discovered in South Africa), and P.1 / B.1.1.28.1 (originally discovered in Brazil / Japan) [229, 212](20–22). The N501Y spike mutation is common to all three of these variants and causes a 4- to 10-fold increased affinity to the human ACE2 receptor for SARS-CoV-2 [134], which is hypothesized to contribute to the observed increase in transmissibility of B.1.1.7. The B.1.351 and P.1 variants have additional receptor binding domain (RBD) mutations, such as E484K, that show significantly reduced neutralization by antibodies generated by current vaccines and by prior natural infection presumably from non-variant SARS-CoV-2 strains [212, 232, 83, 219]. These variants also have additional mutations in the spike N-terminal domain and appear to be resistant to several therapeutic monoclonal antibodies targeting that region [212]. As new studies advance, it is clear that the variants and their associated mutations will have a significant

impact on public health and the efficacy of SARS-CoV-2 control measures such as social restrictions, vaccinations, and therapies. While variant identification through specialized epidemiological sequencing centers is useful, the lack of global access to this resource and delay in result availability has hampered the tracking of and response to the spread of new SARS-CoV-2 variants. There is an urgent need for POC diagnostics for SARS-CoV-2 variants. Of particular benefit would be a system that is easy to use, simple to setup, and smartphone-integrated to enable distributed, non-centralized data collection, rapid adoption, and scaled-up deployment in response to outbreaks [122].

Here we describe the development of a low-cost, self-contained, POC diagnostic called miSHERLOCK (Minimally Instrumented SHERLOCK) that is capable of concurrent universal detection of SARS-CoV-2 as well as specific detection of the B.1.1.7, B.1.351, or P.1 variants. The miSHERLOCK platform integrates an optimized one-pot SHERLOCK reaction with an RNA paper-capture method compatible with in situ nucleic acid amplification and Cas detection. miSHERLOCK combines instrument-free, built-in sample preparation from saliva, room-temperature stable reagents, battery-powered incubation, and simple visual and mobile phone-enabled output interpretation with a limit of detection that matches US Centers for Disease Control and Prevention (CDC) RT-qPCR assays for SARS-CoV-2 of 1,000 copies/mL (cp/mL) (Fig. 2-1).

Although saliva is not a commonly used clinical sample, several studies demonstrate comparable performance between saliva and nasopharyngeal samples for the detection of SARS-CoV-2 [234]. Additionally, in paired collection samples in hospitalized patients, salivary SARS-CoV-2 viral load has been shown to be marginally higher than nasopharyngeal swabs and positive for a greater number of days [99]. There are also FDA-approved home-based saliva collection kits for mail-in SARS-CoV-2 diagnosis [13]. Saliva offers the significant advantage of easy, instrument-free, non-invasive self-collection, which avoids dependence on limiting equipment such as swabs and transport media, and decreases infectious risk to medical personnel and use of personal protective equipment during collection [165]. However, saliva samples typically require several processing steps prior to use. We describe a novel combined filtration and concentration step from untreated saliva that is directly processed on our platform without separate processing steps and significantly enhances our assay sensitivity.

Our platform only requires two simple user steps, is readily adjustable for additional variants or pathogen targets, and does not require transfer of amplicons, which significantly reduces the risk of cross-contamination by lay users. We anticipate the usage of miSHERLOCK for general SARS-CoV-2 detection as well as the specific detection of N501Y and E484K mutations with the goal of locally tracking variant strains and assessing the need for variant-specific booster vaccines, such as those targeting E484K due to its effects on the efficacy of current vaccines. To showcase the flexibility of miSHERLOCK, we also demonstrate high performance detection of the Y144del mu-

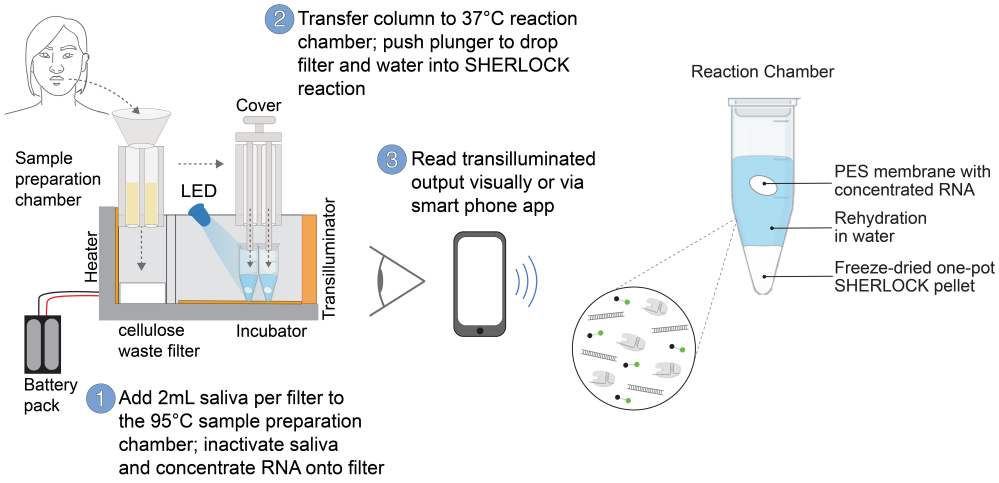



Figure 2-1: **Schematic of miSHERLOCK**. MiSHERLOCK integrates instrument-free viral RNA extraction and concentration from unprocessed saliva, one-pot SHERLOCK reactions that detect SARS-CoV-2 and variants, fluorescent output, and accessory mobile phone app for automated result interpretation. Step 1: The user turns on the device and introduces 4 mL of saliva into the sample preparation chamber (2 mL per filter) and adds 40 µL of 1 M DTT and 500 mM EGTA lysis buffer. Saliva flows by gravity and capillary action through a PES membrane, which accumulates and concentrates viral RNA. Step 2: The user transfers the flow columns into the reaction chamber and depresses the plunger cover to release the PES membrane and sealed stored water into freeze-dried, one-pot SHERLOCK reaction pellets. Step 3: The user returns after 55 min and visualizes the assay directly or using a smartphone app that quantifies fluorescent output and automates result interpretation. The app may also be used for distributed remote result reporting.

tation.

## 2.2 Results

### 2.2.1 Bioinformatic analysis and selection of SARS-CoV-2 target regions

A key aspect of coronavirus replication is nested transcription, which produces high levels of sub-genomic RNA from the 3' end of the SARS-CoV-2 viral genome during active infections [225] including the nucleoprotein (N) gene (Fig. 2-2) [65]. To identify potential targets for our assays, we performed bioinformatic analysis of conserved regions with minimal secondary structure near the 3' end of the SARS-CoV-2 genome. We identified a region of the N gene that was highly conserved among SARS-CoV-2 sequences and that did not show significant homology to other coronaviruses (Fig. 2-2). We designed SHERLOCK assays for this target.

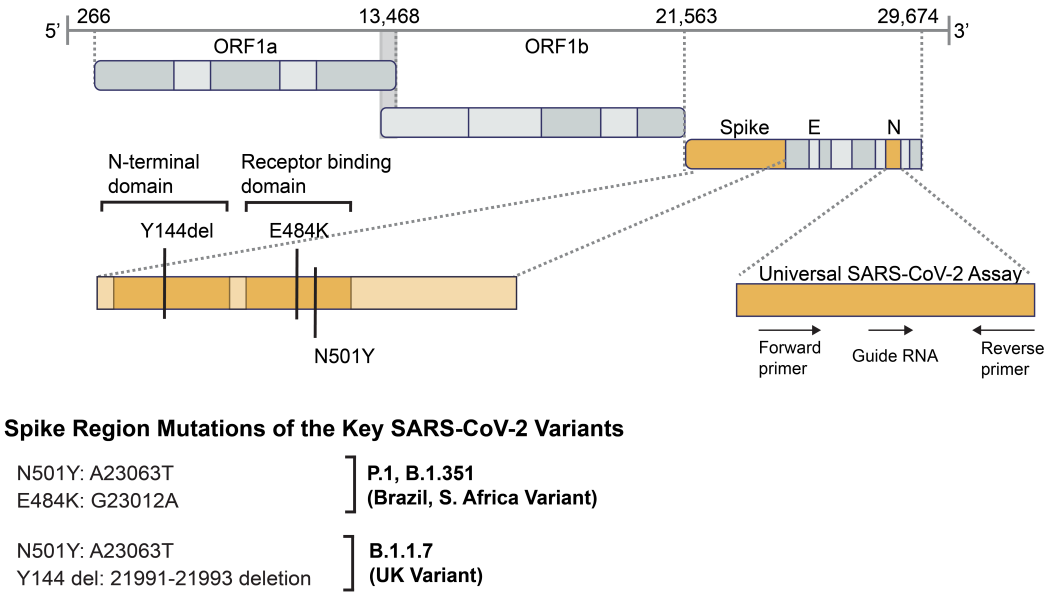



Figure 2-2: **SARS-CoV-2 genomic map indicating regions that are targeted in this study.** The N gene target is used for a universal SARS-CoV-2 assay. SARS-CoV-2 variants are detected by targeting key mutations in the N-terminal and RBD regions of the SARS-CoV-2 spike protein, including N501Y, Y144del, and E484K.

SHERLOCK consists of two components: isothermal nucleic acid amplification and Cas-mediated detection. We systematically evaluated recombinase polymerase amplification (RPA) primer sets and gRNAs to determine the most sensitive combinations using commercially obtained full-length synthetic SARS-CoV-2 genomic RNA standards (Twist Biosciences MT106054.1). The best performing N gene gRNA from a set of 30 tested showed exact target matches in 90.7% of all full-length SARS-CoV-2 sequences deposited at the US National Center for Biotechnology Information (NCBI) ( 43k genomes) (Table A.1). Mismatches were mostly due to a 3' C>T single nucleotide polymorphism (SNP) in 7.4% of sequences, which is not expected to affect Cas12a targeting since Cas12a gRNA function has been shown to be mediated primarily through 5' interactions [204]. The best performing N gene RPA primers from a set of 100 pairs tested showed exact target matches in 97.4% (forward primer) and 97.0% (reverse primer) of NCBI SARS-CoV-2 genomes.

To ensure the universality of the N gene gRNA and RPA primers for SARS-CoV-2 variants, full-length high-quality sequences (>29,000 nucleotides (nt), <1% Ns, and <0.05% unique amino-acid mutations not seen in other genomes) were obtained from GISAID [57] for the B.1.1.7 (50,001 genomes), B.1.351 (577 genomes), and P.1 (78 genomes) variants and aligned as separate groups. The N gene gRNA exactly matched 99.7%, 100%, and 100% of B.1.1.7, B.1.351, and P.1 genomes, respectively. The N gene forward RPA primer matched 99.8%, 99.5% and 100% of B.1.1.7, B.1.351, and

P.1 genomes, respectively. The N gene reverse RPA primer exactly matched only 0.06% of B.1.1.7 genomes due to a C>T SNP at the +3 position from the 5' end in 99.94% of B.1.1.7 genomes. This primer SNP is not expected to have any significant effect on RPA amplification efficiency based on the length of RPA primers and prior studies on RPA primer design, which demonstrate a tolerance of 1-3 nucleotide mismatches [133]. The N gene reverse RPA primer exactly matched 100% of B.1.351 and P.1 genomes.

The N gene SHERLOCK assay limit of detection (LOD) using a dilution series of heat-inactivated SARS-CoV-2 RNA [American Type Culture Collection (ATCC) VR-1986HK] spiked into water was 20,000 copies/mL with a SHERLOCK reaction time of 55 min (Fig. 2-3). This LOD is comparable to high-performance SARS-CoV-2 RT-qPCR assays [172], with a faster time to result. Additionally, our assay did not show any significant cross-reactivity against Human coronavirus OC43 or Human coronavirus 229E genomic RNA spiked in water (Fig. 2-3).

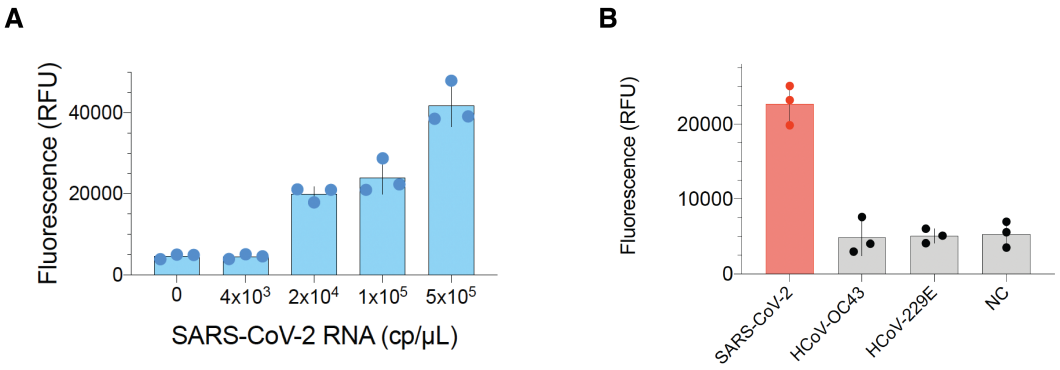



Figure 2-3: **Sensitivity and specificity of SARS-CoV-2 assay.** A. Testing of the universal SARS-CoV-2 N gene assay using full-length synthetic RNA spiked in water. The limit of detection was 20,000 cp/mL. B. Universal SARS-CoV-2 N gene assay showed no significant cross-reactivity with common endemic Human coronaviruses OC43 and 229E. Template is RNA spiked into water. NC indicates no template negative control.

### 2.2.2 SHERLOCK assays that specifically identify SARS-CoV-2 variants

To identify the B.1.1.7, B.1.351, and P.1 SARS-CoV-2 variants, we designed SHERLOCK assays that targeted a panel of key spike protein mutations that are currently representative of these variants: N501Y, Y144del, and E484K (Fig. 2-2). For each mutation, several gRNA sequences were designed and tested with up to 110 primer pairs in order to obtain the lowest LOD (Fig. 2-4).

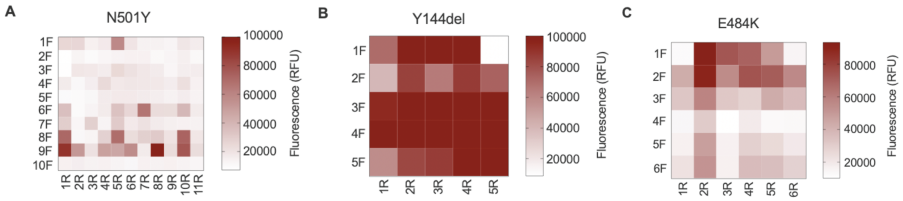



Figure 2-4: **Heatmaps of RPA primer screens to detect SARS-CoV-2 variants.** For each mutation, up to 110 different primer sets (forward denoted by "F", reverse denoted by "R") were screened in pairs to identify those that yielded the highest fluorescence signal in SHERLOCK assays. (A-C) Heatmaps of primer pairs for variant mutations N501Y, Y144del, and E484K.

N501Y is a mutation in the spike RBD resulting from an A23063U SNP that is shared by the B.1.1.7, B.1.351, and P.1 SARS-CoV-2 variants. The N501Y gRNA exactly matched 99.8% of B.1.1.7 genomes, 98.8% of B.1.351 genomes, and 100% of P.1 genomes (Table A.1). The N501Y RPA reverse primer exactly matched 99.6%, 95.3% and 100% of B.1.1.7, B.1.351, and P.1 genomes, respectively. The N501Y RPA forward primer exactly matched 98.7%, 0.17%, and 0% of the B.1.1.7, B.1.351, and P.1 genomes, respectively. The lack of exact matches to the B.1.351 and P.1 genomes is due to a G>A SNP at the genomic position corresponding to the +7 position from the 5' end of the forward RPA primer in 99.3% of B.1.351 and 100% of P.1 genomes. This primer SNP is not expected to have any significant effect on RPA amplification efficiency [133]. Our assay targeting N501Y effectively discriminated mutant versus wildtype virus with an LOD of 100,000 copies/mL using full-length synthetic B.1.1.7 variant SARS-CoV-2 RNA spiked into water (Fig. 2-5A-B).

The Y144del spike mutation is a 3 nt deletion characteristic of B.1.1.7 SARS-CoV-2 variants that is not present in B.1.351 and P.1 variants. Its presence together with N501Y strongly suggests a B.1.1.7 variant. The Y144del gRNA showed exact matches to 98.0% of B.1.1.7 genomes and 0% of B.1.351 and P.1 genomes. The forward and reverse RPA primers exactly matched 99.7% and 99.8% of B.1.1.7 genomes, respectively (Table A.1). Our assay clearly distinguished between wildtype and Y144del target RNA to an LOD of 10,000 copies/mL using full-length synthetic B.1.1.7 variant RNA diluted in water (Fig. 2-5C-D).

The E484K mutation is a critical spike RBD mutation present in the B.1.351, B.1.525, and P.1 SARS-CoV-2 variants, which has drawn significant attention. This mutation has been identified as a potential major contributor to reduced efficacy of current vaccinations and immunity resulting from natural non-variant SARS-CoV-2 infections with demonstrably lower viral neutralizing potency from convalescent and post-vaccinated patient sera and significantly reduced susceptibility to several therapeutic monoclonal antibodies against SARS-CoV-2 [212, 232, 134]. The G23012A SNP that causes E484K also creates a new TTTN protospacer adjacent motif (PAM) site in the antisense strand that is needed for maximal Cas12a function. This is expected to al-

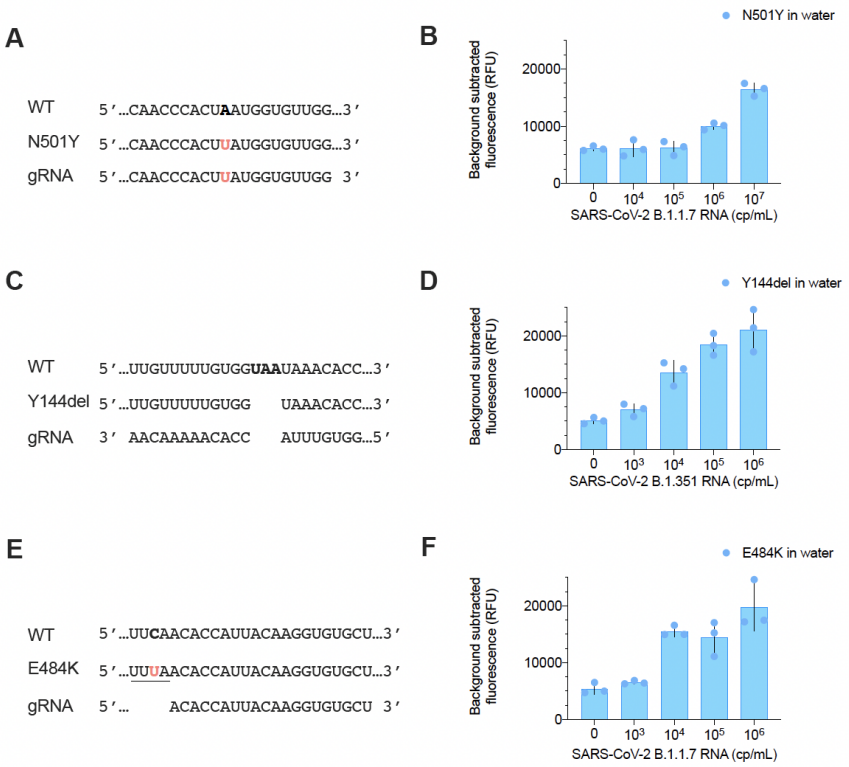



Figure 2-5: **Testing of N501Y, Y144del, and E484K mutation-specific SARS-CoV-2 SHERLOCK assays in water.** Mutation-specific gRNAs for N501Y (A-B), Y144del (C-D), and E484K (E-F) show high SHERLOCK activity when tested against full-length variant RNA matching mutant target RNA. Error bars represent the standard deviation of triplicate experiments.

low differentiation between E484K mutant RNA and wildtype viral RNA. The E484K gRNA and forward and reverse RPA primers all exactly matched 100% of B.1.351 and P.1 genomes (Table A.1). The E484K gRNA and RPA primers also exactly matched to nearly 100% of B.1.1.7 genomes, but because B.1.1.7 lacks the G23012A SNP that causes E484K and therefore the TTTN PAM site, our E484K assay will have significantly reduced activation by B.1.1.7 genomes. We confirmed clear differentiation between E484K mutant and wildtype viral full-length RNA spiked into water at an LOD of 10,000 copies/mL after a reaction time of 55 minutes (Fig. 2-5E-F).

### 2.2.3 SHERLOCK assay optimization and instrument-free viral RNA capture and concentration

For POC testing targeted toward non-specialist users, it is critical to minimize the number of user steps in order to reduce the likelihood of user error and contamination. We applied several engineering solutions to simplify and enhance sample preparation, one-pot SHERLOCK reactions, signal readout, and result interpretation.

SHERLOCK reaction conditions were extensively optimized by varying buffers [35, 106, 11], reverse transcriptases, and reporter concentrations to obtain the lowest LOD (Fig. 2-6A-C). In agreement with a prior report (15), we found that the addition of RNase H to the SHERLOCK reaction improved reaction kinetics and increased overall fluorescent output (Fig. 2-6 D), likely due to enhanced reverse transcriptase efficiency via degradation of inhibitory RNA:DNA hybrid intermediates.

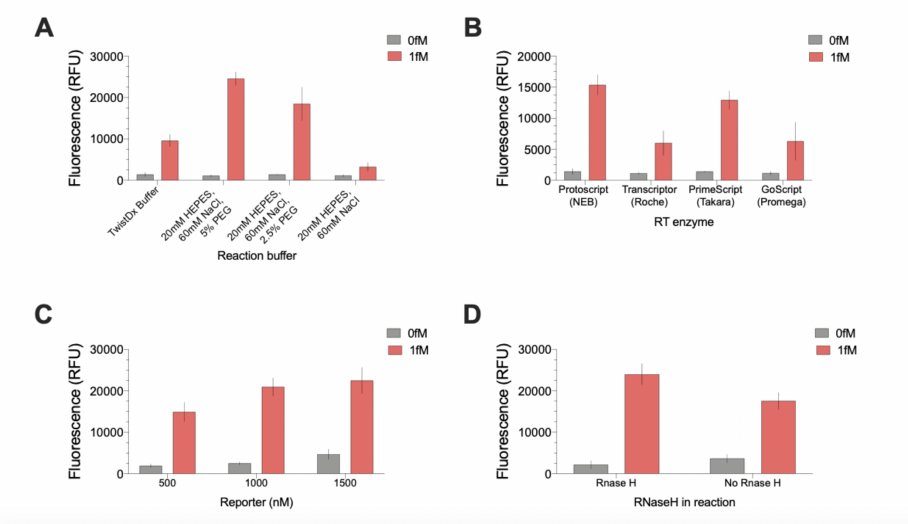



Figure 2-6: **Optimization of multiple reaction conditions of SHERLOCK assays in solution.** We optimized different reaction conditions in one-pot SHERLOCK reactions that combined RPA with Cas12a amplicon detection. (A) Buffer with HEPES and PEG had a higher signal-to-noise ratio as compared to standard RPA buffer supplied by commercial vendor TwistDx. (B) RT enzymes from different manufacturers were tested in one-pot SHERLOCK reactions. Protoscript from NEB was used in miSHERLOCK as it led to higher signal-to-noise ratios. (C) Quenched fluorescent reporter concentrations for SHERLOCK reactions were optimized. (D) RNase H (0.05U/μL) improved the signal-to-noise ratio in SHERLOCK assays. Error bars represent standard deviations of triplicate technical replicates.

We next adapted SHERLOCK assays for use with saliva, which has been identified as an alternative to nasopharyngeal and nasal swabs for SARS-CoV-2 diagnosis. Saliva has several advantages, including being readily available, easy to self-collect, and not requiring swabs or other collection equipment aside from a simple container, which enables mass collection [234, 21]. Unprocessed saliva cannot be used directly in a SHERLOCK assay without pre-treatment due to salivary nucleases that hydrolyze quenched fluorescent reporters and lead to high false positive signals (Fig. 2-7). Unprocessed saliva is also viscous and typically requires several sample preparation steps including centrifugation and a series of manual manipulations in order to release genomic material from viral particles and purify them from inhibitors of nucleic acid amplification and detection reactions [21]. POC nucleic acid tests generally require

prior sample preparation via commercial kits [6, 234, 21], which are not suitable for applications in resource-limited settings or for at-home use by non-specialist users.

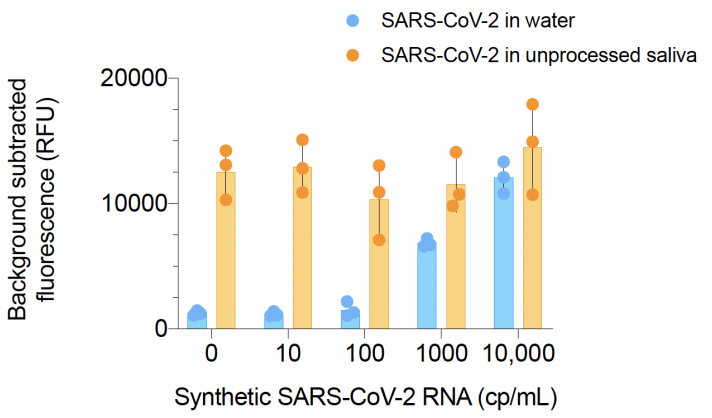



Figure 2-7: **Unprocessed human saliva non-specifically activates SHERLOCK assays.** SARS-CoV-2 RNA spiked in unprocessed human saliva (orange) and in water (blue) demonstrate that saliva sample pre-treatment is required to avoid false positive signals in SHERLOCK assays.

To avoid nucleic acid purification kits that are costly, labor-intensive, and require specialized equipment and user training, we developed a novel technique to inactivate nucleases in unprocessed saliva, lyse viral particles, and concentrate resultant nucleic acids onto a porous membrane that can be directly added to SHERLOCK detection reactions. We started by testing a variety of buffers and heating conditions (Fig. 2-8) to inactivate nucleases and release nucleic acids from viral particles. We found the addition of 10 mM dithiothreitol (DTT) and 5 mM ethylene glycol-bis(2-aminoethylether)- N,N,N',N'-tetraacetic acid (EGTA) followed by heating to 95℃ for 3 minutes effectively eliminated the false positive signal associated with salivary nucleases without inhibiting the performance of downstream SHERLOCK-based target nucleic acid detection.

Nucleic acid capture and concentration onto porous membranes compatible with in situ amplification has been described previously as a sample preparation method for nucleic acid tests [130, 185, 191]. We engineered a column that collects 2 mL of user saliva (Fig. 2-1) and flows the saliva to a 4 mm polyethersulfone (PES) membrane (Millipore) through gravity and capillary action from an absorbent cellulose filter applied under the membrane. The overall efficiency of nucleic acid capture is determined by the transfer rate of RNA from bulk solution to the capture matrix and the subsequent and separate binding rate of RNA to the capture matrix. We found that the highest capture efficiency is achieved by transport rates of at least 1 min/mL (Fig. 2-9). The RNA flow rate can be finely tuned by changing the diameter of the PES membrane (Fig. 2-9). Slowing the flow rate to 1.5 min/mL did not significantly increase RNA recovery, but given the sharp drop-off in RNA recovery

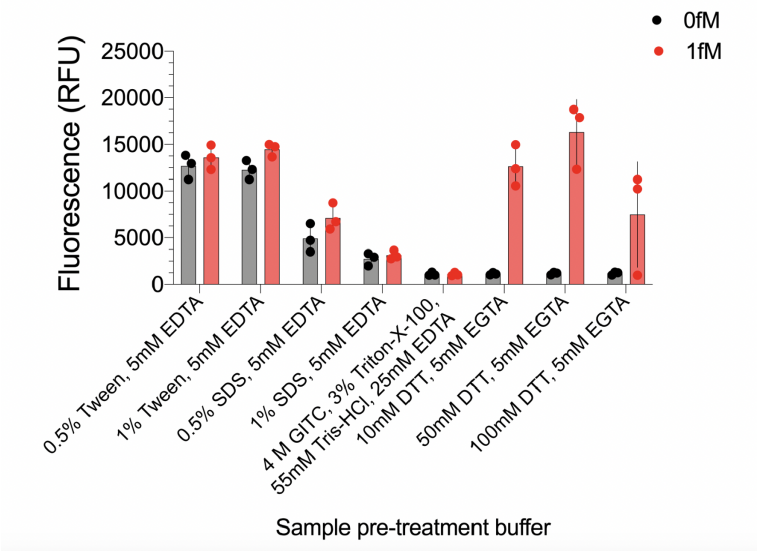



Figure 2-8: **Chemical and heat pre-treatment inactivates salivary nucleases.** Healthy human saliva alone and saliva spiked with 1fM full-length synthetic SARS-CoV-2 RNA were pre-treated with a range of buffers and 95℃ heating for 5 min. Buffers that contained EGTA and DTT showed the highest signal-tonoise ratios in subsequent SHERLOCK reactions. Error bars represent standard deviations of triplicate technical replicates.

at flow rates faster than 1 min/mL, we aimed for a total flow time of 3-6 minutes to maximize RNA capture from our 2 mL sample volume, which was achieved by flowing through a 4 mm-diameter PES membrane (Fig. 2-9). The PES membrane contains 0.22 μm pores and is functionalized with a hydrophilic surface treatment that serves as a porous matrix to capture and concentrate nucleic acids including SARS-CoV-2 RNA. The simplicity of this design allows instrument-free, intuitive liquid-handling while simultaneously achieving significant specimen concentration to enable 2-20 fold improvement in overall signal.

### 2.2.4 Construction of the miSHERLOCK integrated POC diagnostic device

We next endeavored to combine our instrument-free method of heat- and chemical-based inactivation of salivary nucleases and viral lysis and our instrument-free method of nucleic acid recovery and concentration from saliva into a low-cost, easy-to-use, integrated diagnostic device. To this end, we created miSHERLOCK, which incorporates our sample preparation methodology with SHERLOCK reactions and enables direct visual readout. In miSHERLOCK, on-device sample preparation and RNA concentration onto a PES membrane are followed by on-device physical transfer of the RNA-containing capture membrane into one-pot SHERLOCK reactions.

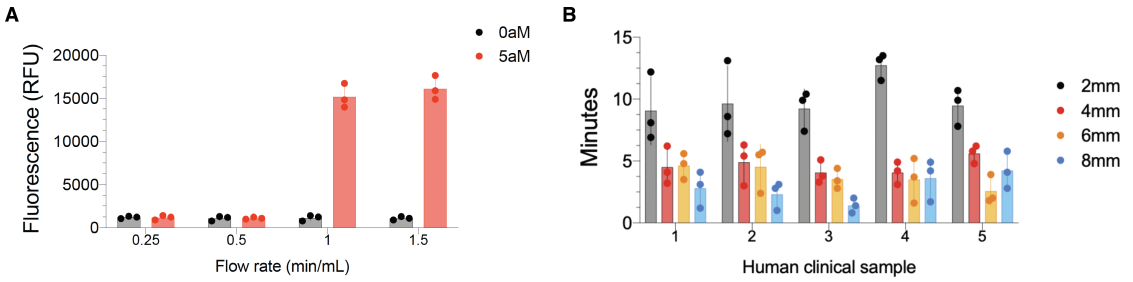



Figure 2-9: **PES filter flow rate characterization.** A. Flow rate through PES membranes affects RNA capture efficiency. We measured the capture efficiency of 5 aM synthetic, full-length SARS-CoV-2 RNA in water filtered through a PES membrane at different flow rates. PES membranes that contained concentrated RNA were then added to one-pot SHERLOCK reactions. Flow rates slower than 1 min/mL were ideal to concentrate SARS-CoV-2 RNA. Error bars represent the standard deviation of triplicate technical replicates. B. Human saliva flow rate is faster with larger aperture diameters of the PES membrane. The flow rate in miSHERLOCK's sample preparation chamber was optimized by flowing 2.5 mL of human saliva through different aperture diameters (mm) of impermeable tape covering the PES membrane at the bottom of the flow column. Smaller apertures led to slower flow rates. Error bars represent the standard deviation of triplicate technical replicates.

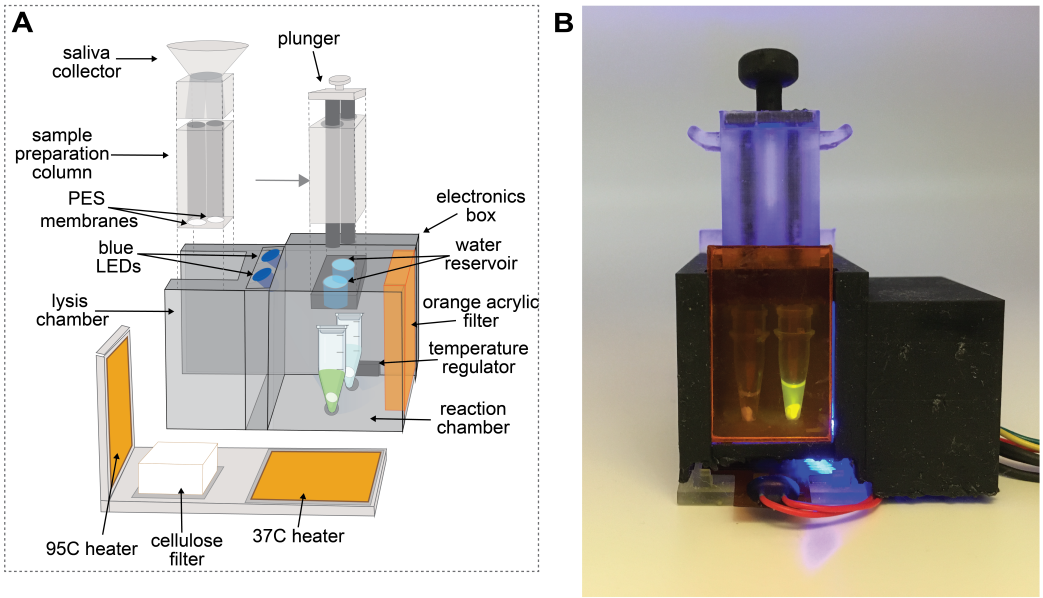



Figure 2-10: **miSHERLOCK device schematic.** A. Exploded schematic of an integrated duplexed miSHERLOCK device with the two modules shown. B. Photograph of miSHERLOCK device after reaction with positive and negative saliva samples. Human saliva flow rate is faster with larger aperture diameters of the PES membrane.

We designed the device with two zones: a high-heat 95°C lysis area that contains an absorbent cellulose filter that wicks saliva for filtration, and a low-heat 37°C reaction area that regulates SHERLOCK reaction temperature. The low-heat area contains LEDs and an orange acrylic optical filter for transillumination and fluorescent readout (Fig. 2-10, A-2). A duplexed device (for two SHERLOCK reactions) is demonstrated and validated here (Fig. A-3, 2-11), but the platform was designed to be scalable and modular, and we also constructed triplex (three reactions) and quadruplex (four reactions) miSHERLOCK versions (Fig. A-1). The heater and temperature regulator units are detachable and reusable. Our diagnostic device is estimated to cost $15 (Table A.3), but reuse of electronics/heaters lowers costs to $6 per duplexed assay, which is mostly comprised of the cost of commercially obtained enzymatic components of the amplification and detection reactions (Tables A.3, A.4).

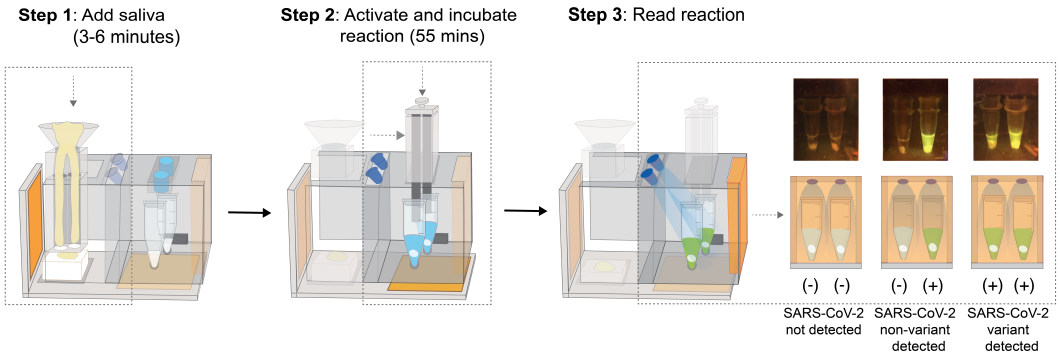



Figure 2-11: **Workflow timing and representative examples of MiSHERLOCK reaction.**

In the miSHERLOCK diagnostic workflow, the user introduces 2 mL of saliva into the collector, which contains preloaded lysis reagents. The user activates the heater on the device, and after 3-6 minutes, viral particles have been lysed, salivary nucleases have been inactivated, and the saliva has been wicked into the filter, leaving concentrated purified RNA on the PES membrane. The user then removes the collector and transfers the sample preparation column to the reaction chamber. The user pushes a plunger into the column, which punctures a water reservoir to rehydrate and activate the SHERLOCK reaction as well as deposits the PES membrane inside the reaction chamber. The user returns in 55 minutes to observe the visual fluorescence readout through the transilluminator (Fig. A-1), 2-12, 2-11).

While SHERLOCK results are easily visually assessed by most users, we created a companion mobile phone application (app) to help provide automated quantitation and simplified interpretation of SHERLOCK results. The app uses the embedded camera in a smartphone in combination with a color segmentation algorithm to detect and quantify observed fluorescence at the end of the incubation period as compared to a fluorescence standard placed on the same reader (Fig. 2-12). Our app quantifies the number of pixels corresponding to the selected fluorescence color to provide a

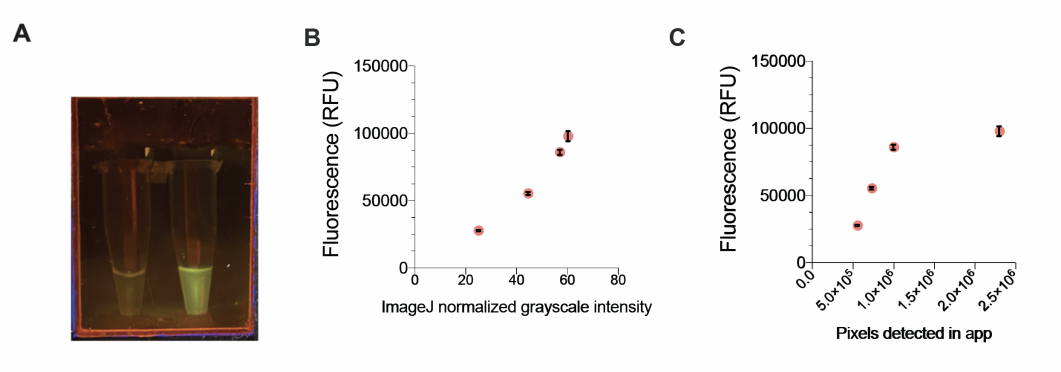



Figure 2-12: **Fluorescence readouts in the miSHERLOCK app are comparable to those measured in standard laboratory plate readers.** A. Representative photo of visual fluorescence signal of a single pair of negative and positive reactions in the miSHERLOCK device. A dilution series of reactions was performed, and fluorescence values were quantified utilizing a BioTEK NEO HTS plate reader and used as a comparison to B. quantitation using our custom-built phone app and C. ImageJ quantitation of the fluorescent area of the photo. We observed that both methods can be used to interpret the results from miSHERLOCK devices semi-quantitatively. Error bars represent the standard deviation of triplicate measurements.

simple qualitative metric of "positive" or "negative" result (Fig. A-3). Test results can also be sent to an online database for real-time distributed disease reporting and strain tracking as required.

### 2.2.5 Evaluation of miSHERLOCK performance and validation with clinical samples

To assess the analytical sensitivity of our assays, we performed LOD studies to determine the lowest concentration at which greater than or equal to 95% of all (true positive) replicates test positive. For initial tests of miSHERLOCK performance, we used heat-inactivated SARS-CoV-2 (ATCC VR-1986HK) spiked into commercially obtained healthy human saliva to measure the LOD of our N gene universal SARS-CoV-2 assay in the integrated device. The miSHERLOCK device functioned well across a range of concentrations and showed a LOD of 1,240 copies/milliliter (cp/mL) [95% confidence interval (CI): 730 – 10,000] on the device (Fig. 2-13, A-4).

To ensure that saliva samples contain enough genetic material for testing, we designed a SHERLOCK assay that targets the human RNaseP gene and validated its performance using clinical samples (Fig. A-5). This assay serves as a positive control for RNA extraction and reagent stability in the miSHERLOCK platform. We tested full-length synthetic RNA representative of the SARS-CoV-2 variants B.1.1.7 (Twist B.1.1.7_601443), P.1 (Twist EPI_ISL_792683), and B.1.351 (Twist EPI_ISL_678597) containing N501Y, Y144del, and E484K mutations, and spiked

A

| | | | | | | LOD (cp/ml) | 95% CI (cp/ml) |
|---|---|---|---|---|---|---|---|
| **Universal SARS-CoV-2 assay** | RNA (cp/ml) | 3200 | 1000 | 320 | 100 | 1200 | 730–10,000 |
| | Replicates (+/total) | 10/10 | 9/10 | 3/10 | 0/5 | | |
| **N501Y** | RNA (cp/ml) | 32,000 | 10,000 | 3200 | 1000 | 49000 | 21,000–89,000 |
| | Replicates (+/total) | 10/10 | 3/9 | 4/10 | 0/10 | | |
| **E484K** | RNA (cp/ml) | 3200 | 1000 | 320 | 100 | 1200 | 660–19,000 |
| | Replicates (+/total) | 10/10 | 9/10 | 5/10 | 0/5 | | |
| **Y144Del** | RNA (cp/ml) | 3200 | 1000 | 320 | 100 | 1100 | 590–15,000 |
| | Replicates (+/total) | 10/10 | 9/10 | 6/10 | 0/5 | | |

B

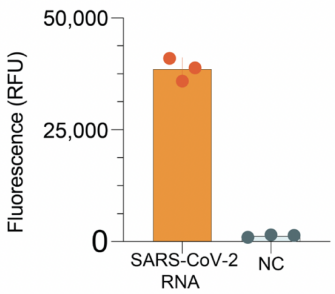



Figure 2-13: **SARS-CoV-2 Assay Limit of Detection.** A. Table summarizing performance near the LOD of four SARSCoV- 2 assays in the miSHERLOCK device. B. Universal SARS-CoV-2 assay using 100,000 cp/mL spiked in saliva compared to healthy saliva negative control (NC).

them into commercially obtained healthy human saliva.

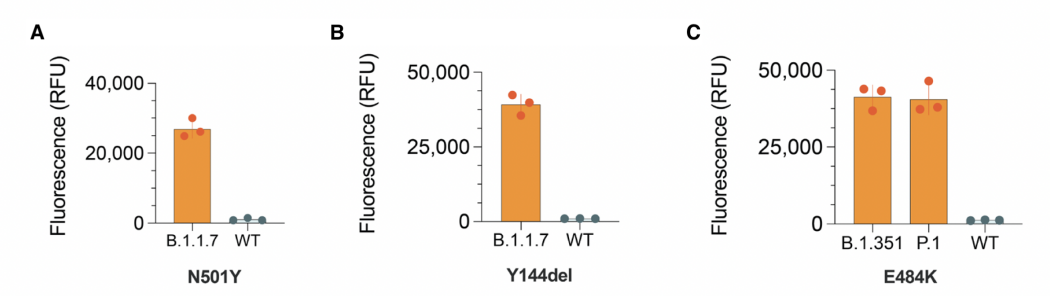



Figure 2-14: **Characterization of SARS-CoV-2 mutant assays.** A-C.Sequences of the wild type (WT) and N501Y, Y144del, and E484K mutant SARS-CoV-2 genomic regions and gRNAs. Mutation-specific gRNAs show high SHERLOCK activity when tested against full-length viral RNA containing each of the indicated mutations (orange), but only minimal SHERLOCK activation when challenged with wild-type full-length SARS-CoV-2 RNA (gray). For (A) to (C), WT control reactions were tested with WT full-length SARS-CoV-2 RNA (2,000,000 cp/mL). Error bars represent the SD of triplicate experiments.

The miSHERLOCK platform worked well across a range of concentrations and also lowered the LOD 2-20 fold, as compared to one-pot SHERLOCK assays performed with equivalent concentrations of synthetic full-length SARS-CoV-2 variant RNA spiked in water (Fig. 2-5). Using the miSHERLOCK device, we determined the LODs with observed positive rates $\geq 95\%$ for the N501Y, Y144del, and E484K assays were

49,000 cp/mL (95% CI: 21,000-81,000), 1,100 cp/mL (95% CI: 590 – 15,000), and 1,200 cp/mL (95% CI: 660 – 19,000) (Fig. 2-13, 2-14, A-4). Given the successful integration of sample handling, salivary nuclease inactivation, viral RNA extraction, purification, and concentration, and one-pot SHERLOCK in the miSHERLOCK device with improved overall signal output, we obtained clinical saliva samples from 27 RT-qPCR positive COVID-19 patients (Boca Biolistics) to test the performance of miSHERLOCK with clinical samples. We also tested 21 healthy human saliva control samples (BioIVT). The miSHERLOCK device demonstrated 96% sensitivity and 95% specificity (Fig. 2-15A) for the detection of SARS-CoV-2 in clinical saliva samples across a range of viral loads as confirmed by concurrent RT-qPCR (cycle threshold (CT) range 14-38) (Fig. 2-15B).

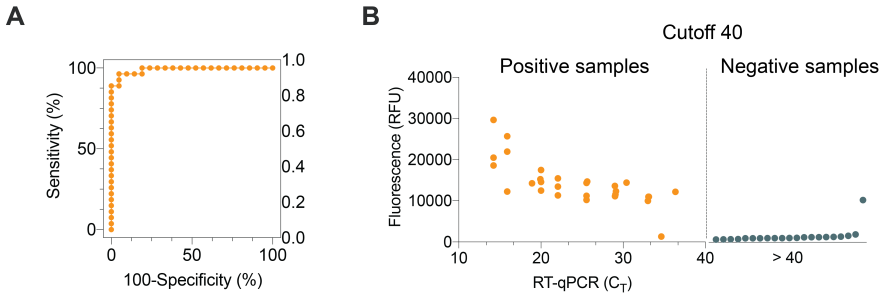



Figure 2-15: **[MiSHERLOCK performance on patient samples.** A. Receiver operating characteristic curve analysis of the patient sample data collected for the universal SARS-CoV-2 assay using results from 27 RT-qPCR confirmed positive and 21 negative human saliva samples. B. Clinical COVID-19 saliva sample RT-qPCR cycle threshold (CT) plotted against fluorescent readout on miSHERLOCK demonstrates dose-dependent semiquantitative results. RT-qPCR–positive saliva samples (CT range: 14 to 38) are plotted in orange, and RT-qPCR–negative samples (CT $>$ 40) are sorted by fluorescence and plotted in gray.

## 2.3 Discussion

We describe the design of miSHERLOCK as a low-cost, portable, self-contained, and integrated diagnostic capable of highly sensitive universal detection of SARS-CoV-2 that equals CDC RT-qPCR performance guidelines, as well as being the only POC diagnostic capable of specific detection of SARS-CoV-2 variants. Several innovative features of our design address critical limitations of current diagnostics in the areas of assay sensitivity, ability to detect viral genomic mutations, simplicity of use, and prevention of laboratory amplicon contamination. One key feature of our design is the incorporation of a unique instrument-free method of RNA isolation from saliva that does not require laboratory equipment, yet achieves specimen filtration and concentration of sample RNA and increases assay sensitivity 2-20 fold. This was accomplished with a simple low-cost membrane by engineering the flow-rate mediated

by gravity and capillary action. Other CRISPR-based diagnostics for SARS-CoV-2 have been described, but used commercial RNA extraction kits for sample preparation [180, 50], or use simplified lysis that still require several pipetting steps in order to perform the detection reactions [6, 160].

Another innovative feature is the ability of the miSHERLOCK device to accept modular target assay components that can be easily exchanged and scaled for multiplexing as needed. Consistency of signal output interpretation is enhanced via an automated mobile phone app, which also allows distributed tracking and reporting. We demonstrate highly sensitive universal detection of SARS-CoV-2 as well as three high-performance variant diagnostic modules. Notably, these can be easily and rapidly adapted for future variants or pathogens and deployed in accordance with local conditions and diagnostic goals. The reusable heater and temperature regulator electronics minimize the cost, waste, and environmental footprint of miSHERLOCK. The device can be printed using off-the-shelf 3D printers with commonly available biodegradable polylactic acid to further reduce plastic waste. Cost analysis indicates the miSHERLOCK device has a total cost of $15, but reusing the electronics and heaters would reduce the cost to $11 per duplexed assay, mostly due to commercially obtained enzyme reagent costs that may be significantly reduced with large scale purchasing and manufacturing (Table A.4). Lastly, amplicon contamination in nucleic acid testing is a pervasive problem that has affected several COVID-19 clinical and research laboratories [184]. By eliminating the need to handle and transfer post-amplification reactions, we significantly reduce the risk of cross contamination, which is especially important for non-specialist users.

Limitations of our study include the small set of clinical COVID-19 saliva samples tested due to the fact that saliva is not routinely collected in most biorepositories and was difficult to obtain within the context of a proof-of-concept exploratory study. Similarly, we were unable to test clinical samples of SARS-CoV-2 variants due to lack of availability. However, miSHERLOCK showed highly sensitive and specific detection with commercially sourced full-length variant RNA spiked into control human saliva and showed near-perfect concordance with RT-qPCR when detecting SARS-CoV-2 from clinical samples of unprocessed saliva. Our E484K variant testing relied on the difference in signal obtained in the presence of a mutated Cas12a PAM site. As-Cas12a and LbaCas12a have both been shown to exhibit reduced but still present cis nucleic acid cleavage despite a TTTV to TTCV mutation, with reductions of 80% and 60% in cleavage efficiency in human cells, respectively [36]. However, the effect of PAM mutations on collateral cleavage is unclear. Collateral cleavage in the absence of a TTTG PAM site has been reported for LbaCas12a [160], but we observed a significant difference in collateral cleavage signal between a mutated PAM site and a canonical PAM site. It is conceivable that a PAM site mutated at a single SNP (i.e., TTTA to TTCA) has a larger inhibitory effect than the complete absence of a PAM site. It is possible that eventually the genetic linkage between different mutations amongst variants would disappear as genetic drift continues. It is also possible that background mutations surrounding clinically important mutation sites may eventu-

ally make non-sequencing-based nucleic acid diagnosis of these mutations challenging. However, the number of genomes that show even a few background SNPs adjacent to the clinically relevant mutations mentioned above is extraordinarily small, and we anticipate that CRISPR-based SARS-CoV-2 mutation characterization will continue to have meaningful utility.

Given the rapid time to result of one hour, we believe that miSHERLOCK POC testing for SARS-CoV-2 variant strains will be highly useful for the control and management of the COVID-19 pandemic. For example, one duplex configuration could include modules for universal SARS-CoV-2 identification as well as identification of the N501Y mutation, which would detect the B.1.1.7, B.1.351, and P.1 variants. This may trigger decisions about increased social distancing or lockdowns in response to the increased infectivity associated with N501Y variants. Another possibility is to detect the E484K mutation in order to guide the distribution of potential vaccine boosters targeted to this variant due to the observed reduction in vaccine efficacy associated with the B.1.351 and P.1 strains [229, 212]. As ongoing clinical studies progress, mutation-specific diagnostics may also guide specific protocols in treatment and hospital infection control. For example, we anticipate that our diagnostic may be most useful at the point of care in low-resource settings. While commercially prepared therapeutic cocktails of monoclonal antibodies are unlikely to be widely available in such environments due to cost, convalescent sera is expected to be more readily available and has been shown to be effective in reducing the progression to severe COVID-19 [128]. However, current studies indicate that B.1.1.7 variants remain generally susceptible to convalescent sera while E484K-containing variants are highly resistant to neutralization by convalescent sera [212]. The use of our variant-focused diagnostic may therefore optimize the usage of this treatment or guide infection control policies. Although targeted to specific known mutations, miSHERLOCK may also be used as a method to triage clinical samples for further analysis with full genome sequencing for detailed epidemiological monitoring.

As new SARS-CoV-2 variants continue to evolve, the ability to identify variants and rapidly adapt diagnostics to track them will be critical to the successful treatment and containment of the ongoing COVID-19 pandemic. The streamlined workflow and flexible modular design of miSHERLOCK represent important, timely advances in the translation of CRISPR-based assays to field-applicable POC tests, particularly for low-resource settings.

## 2.4 Materials and Methods

### 2.4.1 Bioinformatic analysis of SARS-CoV-2 genomes and gRNA and RPA primer design

For universal SARS-CoV-2 detection, 43,305 full-length sequences were downloaded from NCBI and aligned using MAFFT [109]. For B.1.1.7 SARS-CoV-2 variants, 50,001 full-length high-quality sequences (>29,000 nt, <1% Ns, <0.05% unique amino acid mutations) genomes were downloaded from GISAID [57] and aligned using MAFFT. For B.1.351 SARS-CoV-2 variants, 577 full-length, high-quality sequences were downloaded from GISAID and aligned using MAFFT. For P.1 SARS-CoV-2 variants, 78 full-length, high-quality sequences were downloaded from GISAID and aligned using MAFFT.

Cas12a gRNAs consist of two parts: the handle region (UAAUUUCUACUAAG-UGUAGAU) that the Cas protein recognizes and binds, and a user-defined spacer region added to the 3' end of the handle that determines the specificity to the target. Spacer regions were selected following established guidelines [74]. For RPA amplification, we designed 10 to 21 forward and reverse RPA primers for each variant target. RPA primers for the universal SARS-CoV-2 assay [167] were selected after testing a range of RPA primers, including some obtained from the literature. Primers were 25-40nt and total amplicon size was 100-200 bp.

Sequences were analyzed using Biopython [39] and JalView [230]. Exact binding percentages of gRNAs and RPA primers for each assay to SARS-CoV-2 and variant genomes are shown in A.1. All gRNA and RPA primer sequences are listed in Table A.2).

### 2.4.2 Clinical samples and ethics statement

De-identified clinical samples from the Dominican Republic were obtained from Boca Biolistics under their ethical approvals. RT-qPCR was performed by Boca Biolistics using the Perkin Elmer New Coronavirus Nucleic Acid Detection kit. The Institutional Review Board at the Wyss Institute and Harvard University as well as the Harvard Committee on Microbiological Safety approved the use of the clinical samples in this study.

### 2.4.3 Simulated clinical samples

Simulated SARS-CoV-2 (wildtype) samples were prepared by diluting commercially purchased heat-inactivated SARS-CoV-2 (ATCC VR-1986HK) quantified by qPCR into water or commercially purchased human saliva (BioIVT). Specificity targets of purified genomic RNA for Human Coronavirus OC43 and Human Coronavirus-229E

were purchased from the American Type Culture Collection (ATCC) and diluted in water.

Simulated variant SARS-CoV-2 samples were prepared by spiking full-length commercially purchased variant strains for B.1.1.7 (Twist Biosciences B.1.1.7_601443), P.1 (Twist Biosciences EPI_ISL_792683), and B.1.351 (Twist Biosciences EPI_ISL_678597) in water or human saliva (BioIVT) followed by serial dilutions. Synthetic RNA of mutant target regions were also generated for initial assay characterization. To produce mutant RNA target sequences, synthetic DNA with an upstream T7 promoter sequence (5' GAAATTAATACGACTCACTATAGGG 3') was purchased from Integrated DNA Technologies (IDT) and in vitro transcribed to generate 150-500 base-pair RNA targets for different mutant regions using the HiScribe T7 High Yield RNA Synthesis kit from New England Biolabs (NEB). Reactions were incubated for 16 hours at 37℃, treated with DNase I (NEB), and purified using the RNA Clean & Concentrator-25 kit (ZymoResearch). RNA was quantified (ng/μL) on a Nanodrop 2000 (Thermo Fisher Scientific). The concentration of target RNA was calculated by RT-qPCR, using a standard curve with quantified gene block DNA. Table A.2 lists the synthetic targets and qPCR primers.

### 2.4.4 SHERLOCK RPA primer and gRNA screening

RPA primers were ordered from IDT and the design strategy for the nucleoprotein and variant spike gene regions is described in the Results. To synthesize gRNA, DNA sequences with an upstream T7 promoter sequence (IDT) were transcribed with the HiScribe T7 High Yield RNA Synthesis kit (NEB). Reactions were incubated for 16 hours at 37℃, treated with DNase I (NEB), and purified using the RNA Clean & Concentrator-25 kit (ZymoResearch). We performed gRNA screens in 10 μL volumes using 100 nM Cas12a (NEB), 200 nM gRNA, 1x NEB 2.1 buffer (NEB), 1 μM ssDNA fluorescent quenched reporter (56-FAM/TTATT/3IABkFQ, IDT) and 100 pM spiked target DNA standard diluted in water. Reactions were incubated at 37℃ for 30 minutes and fluorescence kinetics were measured using a BioTek NEO HTS plate reader (BioTek Instruments) with readings every 2 minutes (excitation: 485 nm; emission 528 nm). For the SARS-CoV-2 wildtype assay, gRNAs with the highest fluorescent signal to be tested against RPA primer combinations were selected. For variant assays, we additionally tested gRNAs against 100 pM concentrations of a wildtype dsDNA gene block template of the target region to ensure that Cas12a could effectively discriminate between the two targets. Best performing gRNAs with the highest fluorescent signal and discriminating ability were screened against multiple RPA primer sets (Fig. 2-4). RPA screens were performed as per manufacturer's instructions using 7.5 μL reaction volumes from the RPA Liquid Basic kit (TwistDx) with addition of 10 U/μL of Protoscript reverse transcriptase (NEB). 10 μL of RPA primer screen reaction was added to a 1.25 μL Cas reaction with the same reaction conditions as described for the gRNA screen. Selected RPA primers and gRNAs are shown in Table A.2.

### 2.4.5 Sample preparation and concentration

To demonstrate paper-based RNA capture and concentration, we used an Aladdin single-syringe infusion pump (World Precision Instruments) to test different rates of low-concentration RNA flow through PES membranes (Millipore, cat. GPWP04700). We used an Integra surgical biopsy punch (Fisher Scientific) to create a 3 mm PES membrane disc that was compressed into the tip of a 21-gauge hypodermic needle (Becton Dickinson) and secured onto the tip of a 5 mL syringe (Becton Dickinson). The syringe was loaded onto an Aladdin infusion pump and 2 mL volumes of SARS-CoV-2 RNA (Twist Biosciences MT106054.1) at a concentration of 500 copies/mL in water were flowed at rates of 0.25 min/mL, 0.5 min/mL, 1 min/mL, and 1.5 min/mL through the syringe.

To optimize the flow column parameters, a 1 cm x 1 cm square of PCR sealing tape (ThermoScientific) was cut and fitted to the bottom of the column. Different flow-through apertures were punched into the squares using Integra surgical biopsy punches at 1 mm, 3.5 mm, 5 mm, and 7 mm diameters, and the flow columns were loaded with PES membranes at 2 mm, 4 mm, 6 mm, and 8 mm diameters, respectively. One mL, 2 mL, and 2.5 mL of SARS-CoV-2 RNA (Twist Biosciences) at a 1000 copies/mL concentration in saliva were then flowed through the column. To test the sample pre-treatment reagents to lyse virions and to inactivate the nucleases, we tried multiple detergents with and without the addition of 5mM EGTA including 0.5% Tween-20 (Sigma Aldrich), 1% Tween, 0.5% Sodium dodecyl sulfate (SDS) (Sigma Aldrich), 1% SDS. We tested a lysis buffer comprised of 4 M guanidinium thiocyanate (GITC, Sigma Aldrich), 55 mM Tris-HCl (Sigma Aldrich), 25 mM EDTA (Ethylenediaminetetraacetic acid, Sigma Aldrich), and 3% (v/v) Triton X-100 (Sigma Aldrich). We tested the reducing agent, dithiothreitol (DTT, Thermo Fisher Scientific), at 10 mM, 50 mM, and 100 mM.

### 2.4.6 One-pot lyophilized SHERLOCK assay

CRISPR-based sensor reactions were prepared using 200 nM EnGen Lba Cas12a (New England Biolabs), 400 nM gRNA, 1x NEB buffer 2.1, 430 nM each of the RPA primers (IDT), 5 U/μL Protoscript reverse transcriptase (NEB), 0.05 U/μL RNase H (Ambion), 20 mM HEPES pH 6.8 (Thomas Scientific), 60 mM NaCl (Sigma Aldrich), 5% PEG (Sigma Aldrich), 1 μM fluorophore-quenched ssDNA fluorescent reporter (56-FAM/TTATT/3IABkFQ) (IDT), 14 mM magnesium acetate (TwistDx), and 1 TwistAmp Basic RPA pellet (TwistDx). Prepared sensor reactions excluding the magnesium acetate were deposited into 0.2 mL PCR tubes that were snap-frozen. Magnesium acetate was then added to the pellet and the PCR tube was snap-frozen again prior to lyophilization (Labconco) for 4-6 hours. In-device activation of sensors was achieved by rehydration with 50 μL water and deposition of the PES membrane

with captured RNA into the reaction. SHERLOCK reactions were activated by RNA triggers diluted in water alone. Reactions proceeded for 60-120 minutes at 37℃ and fluorescent readout was measured either continuously in a Biotek NEO HTS plate reader (BioTek Instruments) or at the beginning and end of a run when reactions were performed in the miSHERLOCK device. Measurements from the miSHERLOCK device were performed via extraction of 3 µL aliquots from the reaction tube and quantitation in the plate reader, although naked eye visual fluorescent readout was also observable in the device.

### 2.4.7 Construction of point-of-care diagnostic device

We designed the miSHERLOCK platform using Autodesk's Fusion 360 3D CAD software. The housing and components were printed using a Formlabs Form 3 printer (FormLabs). Black resin was chosen to print the housing to minimize reflectance when reading the fluorescence assays. A 2 mm orange acrylic sheet (McMaster-Carr) was laser cut (Universal Laser Systems VLS2.30) to 2.75 cm x 2.25 cm, 2.75 cm x 3.2 cm, or 2.75 cm x 4.0 cm for the duplex, triplex, or quadruplex transilluminator filter, respectively. Double-sided tape (Scotch) was used to tightly line the water reservoir with aluminum foil (Reynolds) and 50 µL of nuclease-free water was loaded for each run. Twenty sheets of Whatman gel blotting paper GB003 (Sigma Aldrich) were loaded into the sample preparation zone for absorption of filtered saliva. Electronic components for the polyimide heaters (Alibaba), temperature controller (DigiKey), and LED lights (Adafruit) were soldered, with heat shrink applied to all wires. Product numbers are listed in Table A.3. The set point for the temperature controller circuit was programmed to 37℃ by selecting a 120 k Ω resistor (DigiKey) and confirmed using a Dallas DS18B20 digital temperature sensor. The LEDs were soldered in series with a 220 Ω for the duplex, a 100 Ω resistor for the triplex, with no resistor needed for the quadruplex (DigiKey) to allow for a current of 25 mA when attached to the 12 V battery source. The temperature sensor circuit and LEDs were mounted to the housing and connected to the battery pack. Fig. A-6 illustrates the miSHERLOCK circuit diagrams, Fig. A-7 shows the electronics placement in the device, and Fig. A-1 shows triplexed and quadruplexed versions of the miSHERLOCK platform.

### 2.4.8 Point-of-care diagnostic device mobile application

In this study, we built a mobile app using Xcode, C++, Objective-c, OpenCV 3.1 and Swift for iOS ((Fig. A-3). The mobile application architecture consists of a camera interface that assists in continuous capturing of fluorescence images as produced by the testing device, which are segmented based on the image colors selected by the user when the user clicks over the screen showing the fluorescent regions in a standard sample. The OpenCV libraries were primarily used for image processing, which included pixel-level color detection, filtering, binarization, and masking. From a usability perspective, the software is presented as an iOS native app icon. Upon

loading the app, the user can select the desired color for detection from the standard assay, to then proceed towards measuring fluorescence on user-collected test tubes (Movie S1). The analysis events can be screenshot and saved on the smartphone to report assay results for epidemiological purposes.

### 2.4.9 On-device miSHERLOCK reactions

We performed experiments on the miSHERLOCK platform to validate the sample preparation and one-pot lyophilized SHERLOCK reaction with patient samples. First, the 95℃ heater was attached to a 24V battery source (two 12V batteries) and the 37℃ heater, temperature regulator, and LEDs were attached to a 12V battery source. The LEDs were inserted into their slots above the reaction chamber (Fig. A-7B) and the temperature regulator was inserted into the electronics box (Fig. A-7C).

The water reservoir was covered with a piece of foil held in place with double-sided tape and filled with 50µL of water for reaction rehydration, the lyophilized reactions were placed within the reaction chamber, the transilluminator filter was slotted into place, the cellulose absorbent filter was placed at the bottom of the lysis chamber, and the PES filters were attached to the bottom of the sample preparation column (Fig. 2-10A).

The sample preparation column was placed within the lysis chamber and topped with the saliva collector. For a duplexed (two-target) reaction 4 mL of saliva, 40µL of a 1 M DTT and 500mM EGTA solution was added (final concentration 10mM DTT and 5mM EGTA) and then deposited into the saliva collector whereupon saliva was separated by gravity into separate sample preparation and lysis chambers. Within the lysis chamber, the saliva was inactivated by the 95℃ heater and the RNA was captured on the filter at the bottom of the column (Fig. 2-11, step 1). Following the concentration, the saliva collector was removed and the sample preparation column was moved above the water reservoir. The plunger was used to deposit the filter and the water within the reservoir into the lyophilized reactions in the reaction chamber.

The plunger additionally acted as a cover for the reactions and prevented evaporation during incubation (Fig. 2-11, step 2). The SHERLOCK reactions incubated for 55 to 120 minutes and were periodically monitored visually by observing the fluorescence through the transilluminator (Fig. 2-11, step 3). Results were typically visible within 55 minutes of incubation and were further confirmed through the mobile application.

### 2.4.10 Data analysis

Fluorescence values are reported as absolute for all experiments used for LOD calculation. Background-subtracted fluorescence, in which the fluorescence value at the initial time point (0 minutes) is subtracted from the end time point (usually 60 minutes)

is reported for initial screening experiments. Due to different baseline initial fluorescence units between runs, it was more effective to compare background-subtracted fluorescence than raw fluorescence. The relationship between the proportion of replicates testing positive and the corresponding sensitivity was examined using Probit regression to estimate 95% LOD and 95% CI of each target. Image analysis of the fluorescence intensities by ImageJ (Fig. 2-12): images of the tubes inside a miSHERLOCK device were captured with a mobile phone camera. We measured grayscale signal intensities of the areas inside the tubes and subtracted the values from background noise to obtain normalized grayscale intensities. All data were plotted and statistical tests were performed using GraphPad Prism 8. Figures were created using Biorender or Adobe Illustrator 2020.

**Author contributions:** H.d.P., R.A.L., D.N., and X.T. designed and performed experiments, analyzed the data, and wrote the manuscript. L.R.S., N.M.A.-M., N.M.D., N.E.W., C.F.N., P.Q.N., A.O., A.S.M., T.C.F., G.L., H.S., and J.N. performed experiments and edited the manuscript. J.J.C. directed overall research and edited the manuscript.

# Chapter 3

# Minimally Instrumented SHERLOCK for Waterborne Pathogen Detection

In this Chapter, I describe research that I contributed to on adapting our miSHERLOCK system for low-resource environmental sample types. Helena de Puig led the data collected on the *E. coli* sensors and magnetic bead-based concentration method. I would like to thank Helena for all her support – the waterborne pathogen biosensors were the first project I worked on when I joined the Collins lab in Fall 2019. My contributions to this work include the experimental design, testing, and optimization of all SHERLOCK assays detailed, troubleshooting the overall sample-to-answer workflow, and designing and building the miSHERLOCK devices. The work described in this chapter work was funded in part through grants from Natick Soldier Systems Center and NASA TRISH Seed grant through the Media Lab's Space Exploration Initiative. I would like to thank them both for the support in this research endeavor.

## 3.1 Introduction

To sustainably manage our planet, we will need to develop and implement environmental monitoring programs that provide deeper insights on how local ecosystems are changing due to the impacts of climate change, industrial activity, urbanization, and globalization. One of greatest determinants of ecosystem health is the presence of microbes in food and drinking water, with pathogens posing a greater threat as pollution and storm events interact with our food and water [215]. Biosensors can provide critical information on ecosystem threats through monitoring chemical, physical, and microbial population changes, as well as gathering longitudinal data needed to advocate for increased environmental protections. Therefore, increasing the availability of biological sensors for environmental monitoring in general, and community-steered microbial monitoring in specific, can increase the number of people who can track ecological changes.

Though there has been an increase in low-cost and open-source electronic and chemical

sensors that hobbyists, concerned citizens, and grassroots organizations have access to for gathering longitudinal data on air and water contaminants, [181], biological sensing has lagged behind. This is primarily due to the high costs associated with laboratory-grade testing, need for specialized equipment, and requirements around miniaturization [139, 27]. The ability to use biology to better sense in the biological dimension is key to generating a higher-resolution portrait of our microbial ecosystems, helping us to monitor and react more quickly to changes that are occurring.

An additional burden for microbial monitoring is that many environmental pathogens are in a state referred to as "viable but not culturable" (VBNC), which means that the transcriptome of the organism is in a semi-dormant state [126, 178]. Because these VBNC microbes are functioning on an altered metabolism, using a system that senses through enrichment (culturing) or through proteins (antibody or antigen-based test), would prove to be difficult or impossible. Therefore, we need nucleic acid-based biosensors that can be quick and reliable. Similar to resource-limited field contexts, the International Space Station has expressed interest in finding new methods for microbial analysis, with the current state of the art being performing PCR experiments or culturing microbes on-board, or else sending samples to be processed in labs on Earth [170, 33, 195, 34]. NASA has proposed integrating low-cost and disposable methods into the microbial monitoring for ISS, but has been slow to act on it.

One promising technology recently discovered has been the application of CRISPR-Cas sensor systems towards a programmable enzymatic approach to nucleic acid detection. CRISPR-Cas – an immune-like response employed by bacteria to defend against viral invasion – requires only two components: a Cas enzyme, which acts as a pair of molecular scissors, and a guide RNA, which acts as a set of instructions that directs the Cas enzyme to the target region of DNA or RNA to cut [81]. Synthetic biologists have adopted the system to more easily edit, activate, and repress genes across a range of organisms [52, 118, 143, 105].

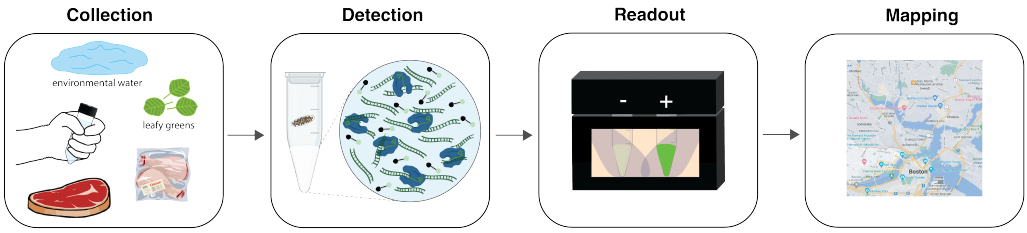



Figure 3-1: **Low-resource CRISPR biosensor workflow.** CRISPR biosensors run with a variety of sample inputs, use lyophilized Cas-based enzymes for nucleic acid detection, can be read using low-cost hardware, and results can be mapped for spatially-localized data aggregation.

CRISPR biosensors can provide localized, low-cost, robust, easy-to-interpret data

on the order of minutes to hours, which makes them ideal for out-of-laboratory use (Fig. 3-1) [127, 105]. Recently, we published a project called miSHERLOCK [47], which demonstrated that SHERLOCK biosensors can be run as a one-pot combined amplification and detection reaction in low-cost and easily fabricated housing that performs all of the necessary steps to go from raw sample to signal. Applying these CRISPR biosensor systems for monitoring purposes can hopefully be one in a broad and diverse landscape of monitoring methods and modalities that can help communities continue to identify baselines and changes in their local environments and foster a continued culture of stewardship and sustainability.

## 3.2 Results

### 3.2.1 SHERLOCK assay optimization for pathogen detection

Common pathogens associated with waterborne and foodborne illness are of great concern to human and environmental health. One of the unique benefits of designing a SHERLOCK biosensors is that it senses nucleic acids, which allows the sensor to discriminate between subspecies and serovars of microbes that have both pathogenic and non-pathogenic states. We focused on designing a system that not only targeted pathogen species, but programmed the sensors to target specific genes directly implicated in the pathogenicity of these microbes. The targets chosen for our set of exemplary SHERLOCK sensors were: *Cryptosporidium parvum* ([87]), *Giardia lamblia* ([224]), *Salmonella enterica* ([9]), and *Escherichia coli* ([120].

RPA was chosen as the amplification method for these biosensors [203] to keep the assays compatible with the miSHERLOCK platform and retain the potential for a one-pot reaction. We chose SHERLOCK genomic target sites based both on the implication in pathogenicity as well as for the availability of published RPA primers available for those sites. For *Cryptosporidium parvum*, we chose the *18s rRNA* gene, which is a unique genomic site used for *C.parvum* identification and had the additional benefit of published and validated RPA primers [42]. For *Giardia lamblia*, we chose the *beta giardin* gene, which is species-specific gene integral to the structural integrity of the parasite [100]. For *Salmonella enterica*, we chose the *InvA* gene, which is implicated in the pathogenicity of salmonella serovars and also had the benefit of published and validated RPA primers [131]. For *Escherichia coli*, we chose the *stx-1* and *stx-2*, which are two related toxins that are present in enterohemorrhagic *E. coli* [120].

After identifying the candidate RPA primers, we designed Cas12a gRNA guides that were within the primers' amplified products. The candidate guides were tested at a range of high concentrations (pM) with short synthetic DNA that contained the target. While all of the guides performed well, we chose to move forward with Salmonella guide 2, Cryptosporidium guide 2, and Giardia guide 1 (Fig. 3-2. It should be noted that there have been published reports of assays that employ multiple guides for the

same target to increase signal or time to answer, which can be explored in these assays if desired [67].

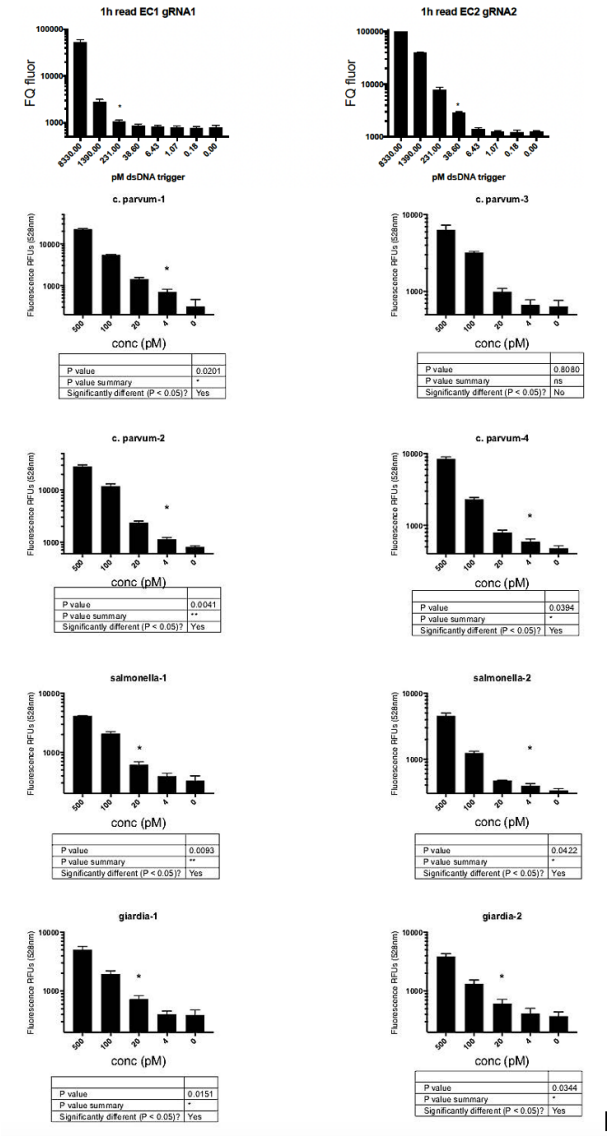



Figure 3-2: **Waterborne pathogen guide RNA assays.** We tested the guides on high concentrations of synthetic DNA target. While all of the guides made good choices, we chose to move forward with Salmonella guide 2, Cryptosporidium guide 2, and Giardia guide 1.

Following guide selection, we combined the RPA amplification with the CRISPR detection into a two-pot SHERLOCK assay. We used serially dilutions of short synthetic DNA targets to test the limit of detection for the sensors. The assays showed attomolar sensitivity and clear signal within one hour of the CRISPR assay activation (Fig. 3-3, 3-4), indicating that these sensors can detect clinically relevant levels of pathogens in samples. The finalized primer and guide sequences can be seen in

Table 3.1.

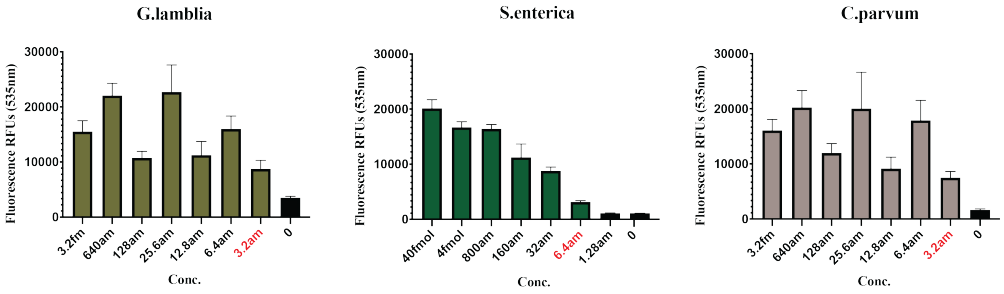



Figure 3-3: **SHERLOCK limit of detection for waterborne pathogens.** The limit of detection for *G.lamblia*, *S.enterica*, and *C.parvum* were all seen to be at single attomolar sensitivity.

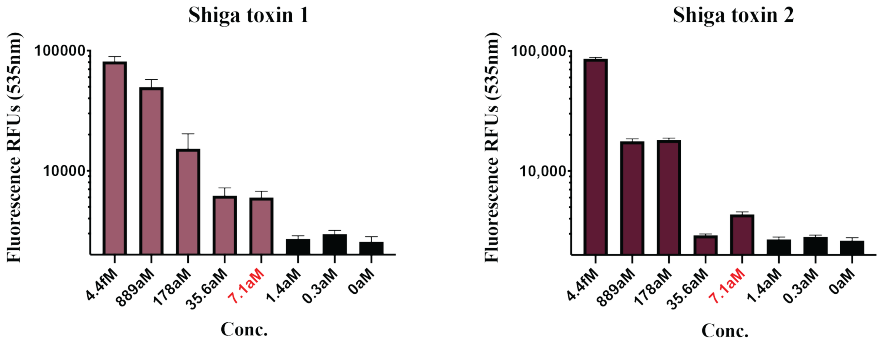



Figure 3-4: **SHERLOCK limit of detection for shiga toxins.** The limit of detection for *shiga-1* and *shiga-2* were also seen to be at single attomolar sensitivity.

Following the success of the SHERLOCK assay on the synthetic targets, we began testing individual assays on genomic DNA or cell extracts of our target organisms. Below, we will focus on the *S.enterica* and *E.coli* sensors, which were chosen first to be further characterized and were used to validate the pre-concentration methods used in this research.

### 3.2.2 SHERLOCK assays on genomic DNA targets

Shiga toxin sensors were tested against shiga toxin producing *E. coli* (STEC) strain O157:H7, which produces both *stx-1* and *stx-2*. We used the enteropathogenic (EPEC) strain O127:H6, which does not produce shiga toxin, as a negative control. The *E.coli* were cultured and diluted into neutral pH water solution (containing 0.2g/l of sea salts, 0.5µM Tannic acid and 2mM sodium bicarbonate at pH 7) and inactivated

| Sequences for waterborne pathogen SHERLOCK assays. | |
|---|---|
| Name | Sequence |
| ssDNA-FQ | /56-FAM/ TTATT /3IABkFQ/ |
| T7 promoter | GCGCTAATACGACTCACTATAGG |
| Se-RPA-F | GGCGATAGCCTGGCGGTGGGTTTTGTTGTCTT |
| Se-RPA-R | ACTTCATCGCACCGTCAAAGGAACCGTAAA |
| Se-gRNA | GGCGAAGCGTACTGGAAAGGATCTACACTTAGTAGAAATTACC |
| Gl-RPA-F | CTGGAGCTGCGAGTGCTGCACAGAGATTCC |
| Gl-RPA-R | TGAGCGTACGGGTGGAGGTGAACATAGACATCCTT |
| Gl-gRNA | TGGGCTCAAATTTTTGCGCGATCTACACTTAGTAGAAATTACC |
| Cp-RPA-F | GTGGCAATGACGGGTAACGGGGAATTAGGG |
| Cp-RPA-R | AATTGATACTTGTAAAGGGGTTTATACTTAACTC |
| Cp-gRNA | ATACAGGGAGGTAGTGACAAATCTACACTTAGTAGAAATTACC |
| Stx1-RPA-F | TTTTTCACATGTTACCTTTCCAGGTACAACAGCGGTTACA |
| Stx1-RPA-R | AACCGTAACATCGCTCTTGCCACAGACTGCGTCAGTGAGG |
| Stx1-gRNA | GGGUAAUUUCUACUAAGUGUAGAUUCUGCAUCCCCGUACGACUG |
| Stx2-RPA-F | ATATATCAGTGCCCGGTGTGACAACGGTTTCCATGACAACGG |
| Stx2-RPA-R | AACTGCTCTGGATGCATCTCTGGTCATTGTATTACCACTG |
| shiga2-gRNA | GGGUAAUUUCUACUAAGUGUAGAUAUCAUAUCUGGCGUUAAUGG |

Table 3.1: **Sequences for waterborne pathogen SHERLOCK assays.**

with heat before experimentation. Our experiments show a limit of detection lower than (103) cfu/ml (Fig. 3-5).

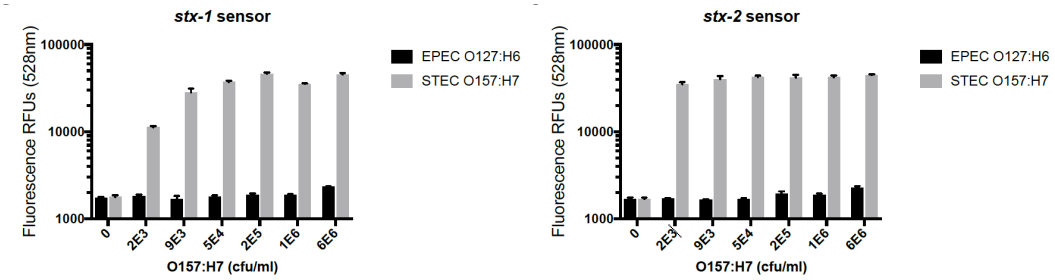
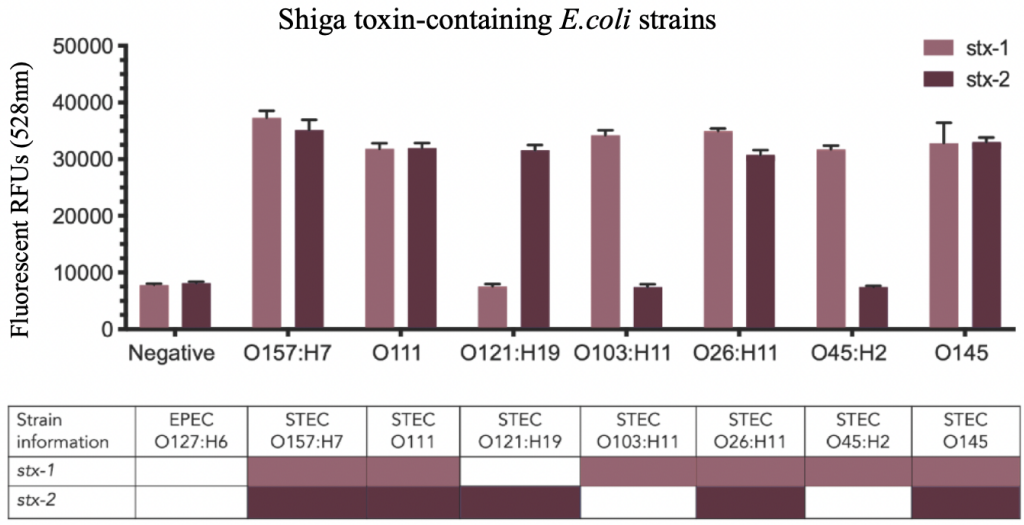



Figure 3-5: **Shiga toxin sensor sensitivity.** SHERLOCK reactions against (L) *stx-1* and (R) *stx-2*, on E. coli strains. O157:H7 (produces both shiga toxins) and EPEC strain O127:H6 (does not produce shiga toxin).

We then tested our Shiga toxin sensors against STEC *E.coli* strain O157:H7, as well

as other shiga-toxin producing *E.coli* strains (Fig. 3-6). We validated the performance of our sensors against a panel of non-O157 STEC serogroups. Our panel (from ATCC: MP-9 and MP-10) provided a robust combination of shiga toxin gene expression from varied *E.coli* sources. Genomic extracts from the *E. coli* serogroups, as well as O157:H7 as a positive control and a non-STEC negative control *E. coli*, were tested with our SHERLOCK sensors. We observed that both *stx-1* and *stx-2* sensors accurately identified the correct *E. coli* shiga toxins within the strains, showing that the SHERLOCK sensors provide shiga-dependent activation. Furthermore, both sensors accurately detected *stx-1* and *stx-2* genes in a variety of *E. coli* serogroups, indicating that the genomic detection regions we chosen for both *stx-1* and *stx-2* sensors were highly conserved among serogroups.




| Strain information | EPEC O127:H6 | STEC O157:H7 | STEC O111 | STEC O121:H19 | STEC O103:H11 | STEC O26:H11 | STEC O45:H2 | STEC O145 |
|---|---|---|---|---|---|---|---|---|
| *stx-1* | | | | | | | | |
| *stx-2* | | | | | | | | |

Figure 3-6: **Shiga toxin panel.** Sensors *stx-1* and *stx-2* tested against different strains of shiga toxin producing *E. coli*. The bar graph shows the fluorescence readings for sensors *stx-1* (light purple) and *stx-2* (dark purple). The table below shows the expected results according to *E. coli* strain information. Results indicate excellent correlation between the shiga sensors and expected toxin presence.

The *S. enterica* SHERLOCK sensor was tested with quantitative genomic DNA for *S. enterica*, serovar Typhimurium (ATCC 700720DQ). We followed the same workflow as previously described for the synthetic DNA targets, but spiked varying concentrations of S. enterica genomic DNA, with the estimation that one genomic copy of *S. enterica* represents one colony forming unit (CFU) of *S. enterica*, given that the *InvA* gene is a single-copy gene [64]. The results of these experiments showed that we were able to detect a clearly distinguished signal from as few as 50 CFU/ml of genomic DNA before implementing any concentration protocols, which indicated that our *S.enterica* sensors did not lose any sensitivity when transitioning to a more complex sample type (Fig. 3-7).

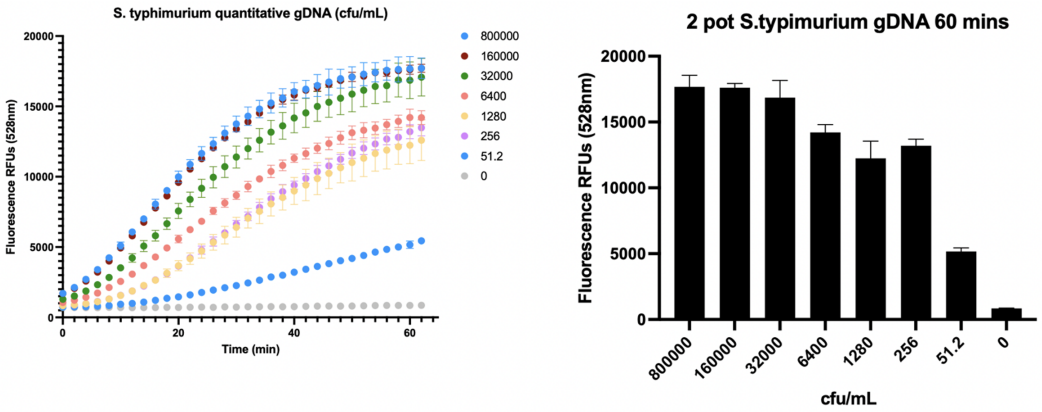



Figure 3-7: ***S. enterica* genomic DNA sensitivity** *S. enterica* genomic DNA in a two-pot SHERLOCK reaction showed a sensitivity of 50 CFU/ml. The endpoint was taken at 60 minutes.

### 3.2.3 Pre-concentration of pathogen target

Following the success of the SHERLOCK sensors, we focused on developing a novel pre-concentration method to allow for high volume sample concentration, which is particularly important for dilute samples. Though there are many methods for sample concentration, the two most common laboratory techniques used are functionalized magnetic bead-based concentration and filter membrane-based spin column concentration, both touched upon in Section 1.2.2. We experimented with both methods to determine which would be most successful for our research goals. The first method that we optimized was magnetic bead-based concentration. Magnetic beads, which often range from 10−7-10−5m in diameter, are a convenient method to capture, purify, and concentrate biomolecules or organisms of interest. Magnetic beads have a magnetic core, which allow for simple concentration in a sample through the the presence of a magnet, with a variety of surface coatings and conjugates that allow the beads to more easily and specifically attract their intended target. In our case, we looked for a system that would allow us to attract pathogenic species, which led us to use magnetic beads functionalized with FcMBL(the Fc region of human IgG1 linked to and engineered mannose binding lectin (MBL) protein), which was a capture method initially developed for broader range nonspecific pathogen concentration from septic patients [31].

First, we generated a calibration curve that could allow us to correlate *E. coli* concentrations (CFU/ml) determined by plated agar cultures and fluorescence of the labeled bacteria. We then repeated the experiments to validate that *E. coli* O157:H7 could also be bound to FcMBL-coated magnetic beads, by culturing *E. coli* O157:H7 and introduced it at varying known concentrations into aqueous buffer solution. Mannose was used as a positive control to verify FcMBL-magnetic bead binding, as well as a calibrant representing MBL binding units (Fig. 3-8A). We were able to quantify concentrations of *E. coli* on the order of 103 CFU/ml of *E. coli* O157:H7 (Fig. 3-8B)

through using FcMBL-coated magnetic beads. Our results correlated well with prior data obtained calibrating capture of nonpathogenic *E. coli* results.

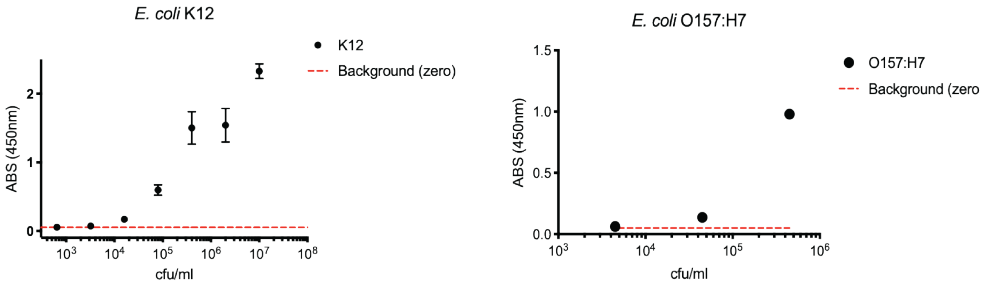



Figure 3-8: **Magnetic concentration of microbial target.** a) Mannose calibration shows that less than 1ng/µl mannose can be detected by using magnetic beads coated with FcMBL. b) calibration on increasing concentrations of *E. coli* O157:H7 indicate that 103 *E. coli* O157:H7 CFU/ml can be detected by using FcMBL-coated magnetic beads.

Following this, we compared two different brands of functionalized magnetic beads, Dynabeads and TurboBeads, to test for improved concentration efficiency. Both beads were conjugated with streptavidin before conjugating them with biotinylated MBL (the Fc region of the IgG1 was removed for the second round of our experiments after concentration was seen to work well using only the MBL). We performed a titration for each bead type to determine the ideal quantity in a 50ml reaction, and determined that 100µg of Dynabeads and 20µg of TurboBeads provided the best results. We then determined the limit of detection for the bead using a range of concentrations from $10^5$-$10^{-1}$ CFU/ml of short synthetic DNA for the *stx-1* and *stx-2* targets. The TurboBeads had a lower limit of detection than the Dynabeads, with an ability to concentrate down to $10^{-1}$ CFU/ml of target (Fig. 3-9). In previous SHERLOCK experiments, we determined that up to 5µg of beads could be added to a SHERLOCK reaction without significantly impacting reaction kinetics. These experiments should be repeated with the TurboBeads to determine whether higher concentrations of beads can be added to the reaction. It should be noted that the background signal for the experiments in Fig. 3-9 were higher than usual, which could be due to the presence of nucleases that were not completely inactivated in the sample lysis. Hopefully this background can be be minimized by chemical additives to assist in proper nuclease denaturation.

Next, we aimed was to evaluate filter-based capture methods for the concentration of pathogens and nucleic acids to improve assay sensitivity. Filter membranes are a method for simple and selective capture and concentration of biological material [207] and offer a wide range of properties including pore size, mechanical strength, surface

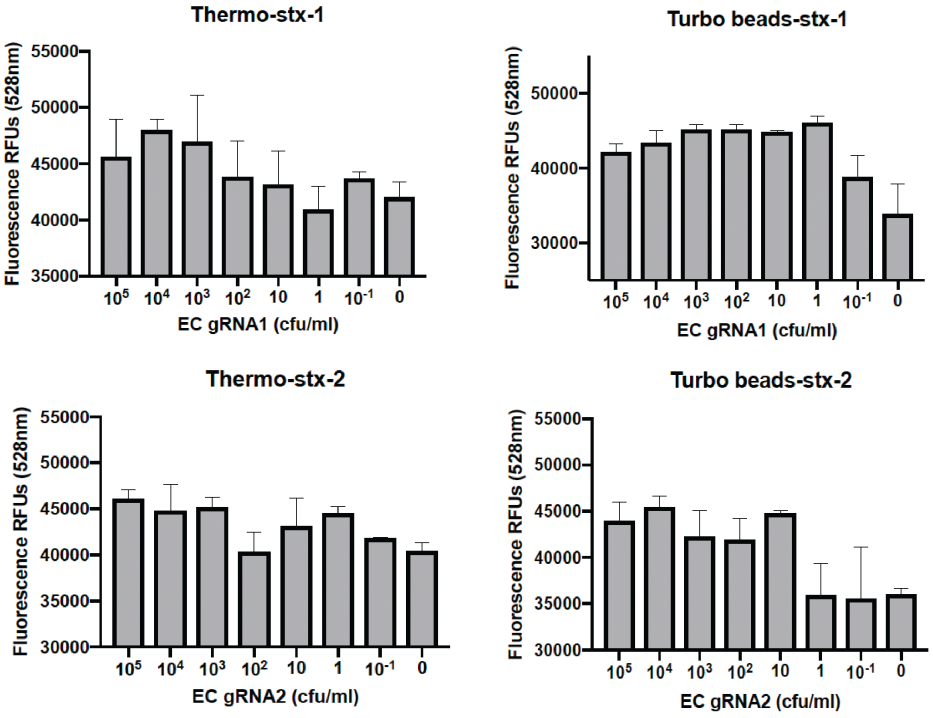



Figure 3-9: **Comparison of magnetic beads for MBL-mediated target concentration.** Dynabead-streptavidin conjugates (L) and TurboBead-streptavidin conjugates (R) coated with biotinylated MBL were used to concentrate a titration of bacteria from 50ml to 100µl for *stx-1* (top) and *stx-2* (bottom). TurboBeads were seen to be more efficient at concentrating the target species. All experiments were performed in triplicate.

charge, wettability, and functionalization which can be tuned for a variety of uses. Recent advances in filter-based nucleic acid concentration emphasize their benefits in a resource-limited settings [168, 207], with the added benefit that filters are easily disposable and can be incinerated, which is useful when outside of a laboratory. Based on the literature, as well as previous research in nucleic acid capture published by our group, [130, 121, 47], we chose a 0.22 µM polyethersulfone (PES) membrane. The combination of mechanical strength, wettability and hydrophilic surface allow for PES membranes to have a higher flow rate of liquid sample with high binding affinity for nucleic acids while maintaining low affinity to proteins.

We tested the PES filter nucleic acid capture efficiency using the *S. enterica* genomic DNA as the target. We placed three 4mm diameter PES filters in 50ml of DPBS solution spiked with known quantities of quantitative genomic DNA solution to shake and incubate at 37℃ for 30 minutes. Following this, we removed the filters and assayed the supernatant and filters separately in a two pot SHERLOCK reaction

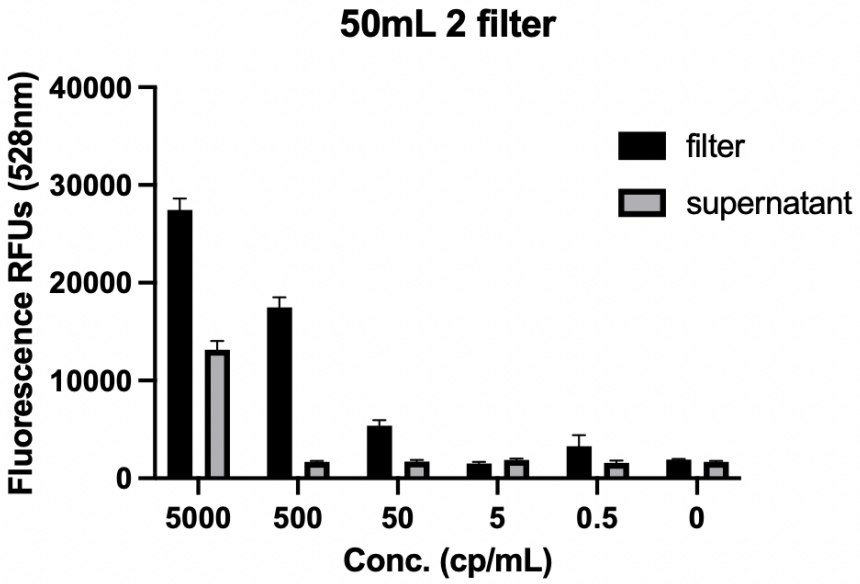



Figure 3-10: **Filter concentration of microbial target.** Varying concentrations of genomic DNA were spiked into 50ml of PBS. Results indicate that the PES filter was able to bind nucleic acids and increase sensitivity of the assay by two orders of magnitude.

(Fig. 3-10). While we were able to pick up signal at 5000 CFU/ml from both the filter and supernantant samples, only the filtered samples were able to generate signal for concentrations as low as 50 CFU/ml. This indicates that the filter membrane did successfully bind and concentrate nucleic acids to its surface and allows for improved signal without increasing background signal.

While these concentration results are promising, further experimentation can be performed to understand how the SHERLOCK reaction can be modified to accommodate higher volumes of magnetic beads into a sample, as well as whether adding the filter together with magnetic beads can offset the impacts of higher magnetic bead concentrations within the reactions.

### 3.2.4 Construction of miSHERLOCK integrated device for field-based biosensing

After we confirmed the sensitivity and efficacy of both the magnetic bead and filter-based concentration methods, we designed a modified miSHERLOCK device [47] for environmental water samples that combined both methods to increase the potential capture capability of the system. The device has three components: a cap that screws on to a 50ml conical tube and contains a slider where a 4-5mm PES filter

can be fitted, a base containing a magnet and blotting paper, and a reaction chamber for two PCR tubes that has a resistive heater and two royal blue LEDs (Fig. 3-11).

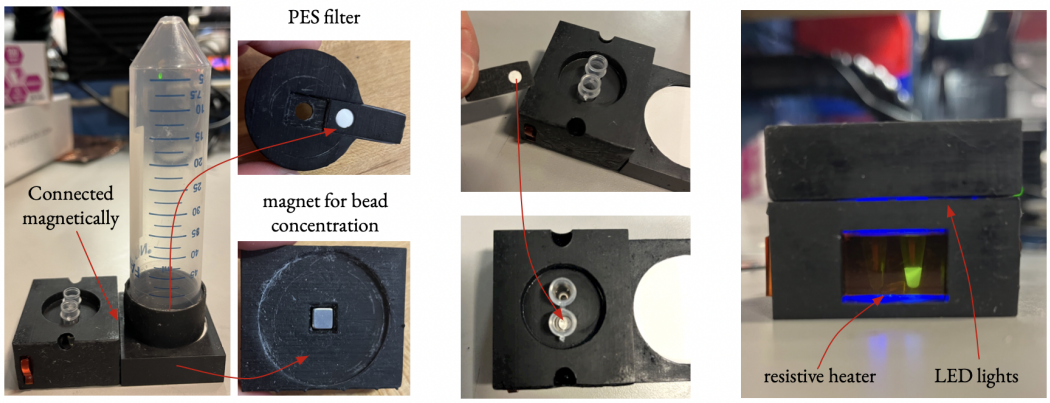



Figure 3-11: **miSHERLOCK for waterborne pathogen detection**. Images of the fabricated modified miSHERLOCK device for waterborne pathogen samples, showing (L-R) concentration of pathogen through magnetic beads and PES filter, transfer of the PES filter with magnetic beads into the lyophilized SHERLOCK reaction, and reaction chamber with incubation and LED-based transillumination to visualize the assay results.

The workflow is as follows: The sample is collected in the 50ml conical tube and 5 µg of MBL-conjugated beads are added and the tube and incubated in a shaking incubator for 30 minutes. Following this, the custom cap with PES filter membrane is screwed to the tube and placed onto the blotting paper/magnet, where the magnetic beads are concentrated onto the PES filter. The slider is then removed from the cap and the filter is dropped into the lyophilized SHERLOCK reaction in the reaction chamber. 50µ of water are added to each tube and then the tubes are sealed. The resistive heater is turned on and the magnetic concentration acts as a cover for the reaction chamber to keep temperature consistent during incubation. One hour later, the blue LED lights together with the orange acrylic transilluminator are used to determine whether the pathogen was present in the sample through visual detection or image processing software or app. A simplified illustration of this workflow can be seen in Fig. 3-12.

### 3.2.5 Construction and evaluation of miSHERLOCK integrated device for biosensing in microgravity

We built a second modified miSHERLOCK reader to test SHERLOCK biosensors in a simulated microgravity, another low-resource environment in need of quick and simple biological sensing. The device contains a sample deposition chamber which moves on a track, which allows for a nucleic acid concentration step to be incorporated prior to the reaction if a user requires it. The deposition chamber also has

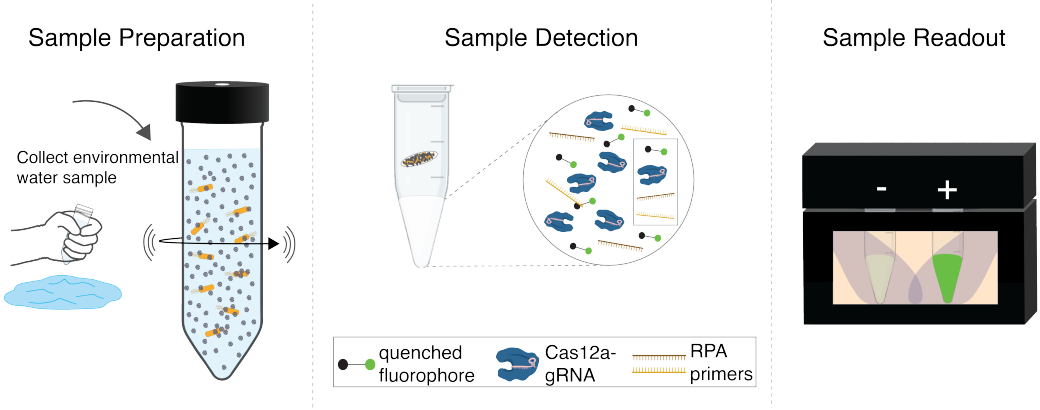
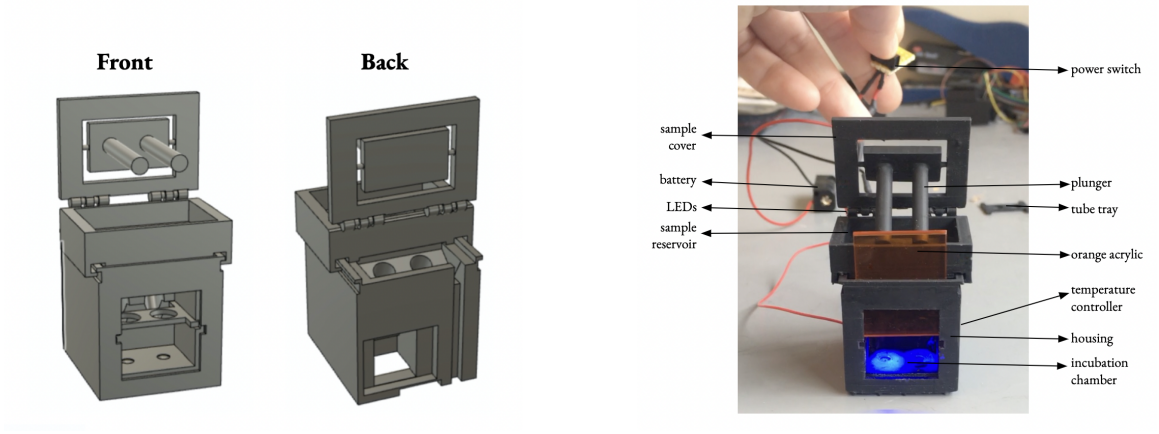



Figure 3-12: **miSHERLOCK for waterborne pathogen summary.** This animation shows a summary of the methods used for magnetic bead-based sample preparation, CRISPR-based sample detection, and sample readout.




(a) (b)

Figure 3-13: **Zero-G miSHERLOCK device.** The miSHERLOCK device designed for use in microgravity. All of the components are connected, preventing the device and assay from disassembling during the run.

two plungers that swicel, which force the filtered sample into the lyophilized assay (similar to the original miSHERLOCK design) and additionally prevent the reaction from escaping the tube while the assay runs in microgravity. A polyimide heater sits beneath the housing and is controlled by a trip-point temperature sensor, providing a 37℃ chamber for reaction incubation. Two blue LEDs angled beneath the sample reservoir together with an orange acrylic act as a transilluminator are used to determine whether the pathogen was present in the sample through visual detection or image processing software or app. Images of the CAD model and an annotated image of the fabricated device can be seen in Figs. 3-13a, 3-13b.

(a) (b)

Figure 3-14: The microgravity miSHERLOCK during flight testing on a Zero-G flight in May 2021.

### 3.2.6 Flight testing the miSHERLOCK in microgravity

We had the opportunity to flight-test our microgravity miSHERLOCK device through MIT Media Lab's Space Exploration Initiative to confirm that it would function properly in low gravity settings. The experiment was coordinated and approved with the Zero-G Corporation to assure appropriate containment and safety measures were taken. Two readers were assembled, one containing two lyophilized reactions with the *S. enterica* sensor and the other containing two lyophilized reactions with the *E. coli* sensors. Power for the heater and LEDs were provided by a 9V battery. Pipettes pre-loaded with either deionized water or deionized water spiked with target were assembled before take-off and dispensed into the reaction chamber at the start of the flight. Below are images taken from a GoPro inside of our experimental set-up, containing the zero-G miSHERLOCK devices (back left). (Fig. 3-14a, 3-14b ).

The device functioned well, with the 9V efficiently powering the LEDs and heater on the device. The plungers stayed inside the tubes and the sample preparation lid stayed closed during the duration of the flight through the 20 cycles of simulated zero gravity, indicated that the designed miSHERLOCK device was properly connected and can be used without concern for separation in a low gravity environment.

## 3.3 Discussion

After the success of our miSHERLOCK device for SARS-CoV-2 detection from raw saliva samples, we wanted to extend the utility of the underlying device platform to other sample types. Our goal was to design of tools for environmental monitoring and baselining that could be used in resource-limited settings where there is no routine access to laboratory equipment or wall wart power sources. We designed ultrasensitive SHERLOCK assays that can be run in custom miSHERLOCK devices that allow for

the detection of the key waterborne and foodborne pathogens *G. lamblia*, *S. enterica*, *C. parvum*, and *E. coli*. These biosensor systems were simple to fabricate and can be cheap when produced at scale.

We made modifications to the original miSHERLOCK device design to allow for use in microgravity or with dilute environmental samples. While the device is designed to minimize user actions, there is still more work to be done before the modified miSHERLOCK devices can be be field-ready outside the laboratory. Further research should be done to determine whether there are alternatives to a shaking incubator for the bead and filter membrane-based concentration in the waterborne pathogen miSHERLOCK workflow. It is possible that a low-power coil-based mixer using AC power and a polyimide heater can mimic the environment. A variety of sample preparation methods can also be designed and tested for the miSHERLOCK for microgravity samples, which were not tested on the zero-G flight. Further testing in microgravity of the overall zero-G workflow should be done, particularly on whether plunging the rehydration solution from a reservoir is sufficient for activation of the reaction similar to the method used in the flight of pipetting. It is possible that a paper-based system, which allow for flow by way of wicking and capillary action, can be incorporated into the miSHERLOCK system to allow for efficient rehydration of reactions while avoiding mixing. Additionally, the SHERLOCK biosensors are currently optimized for a two-pot reaction and would benefit from additional testing to determine the formulation that is best suited for the sample type.

The design of the original miSHERLOCK reader and the modified designs were all influenced by ethnographic and field work performed with two water monitoring communities (see Chapter 5). However, because the device fabrication took place during the COVID-19 pandemic, it was not possible to test these devices with these communities. In the future, it would be useful to design multiple versions of the device and perform user testing to assure the device is as intuitive and easy to use as possible. Instructions to fabricate the devices can also be shared with communities that have access to soldering and 3D printing facilities to allow them to test the devices on their own and provide feedback.

## 3.4 Materials and Methods

### 3.4.1 SHERLOCK RPA primer and gRNA screening

RPA primers were ordered from IDT based off of the literature research and guide design described in the results. To synthesize gRNA, DNA sequences with an upstream T7 promoter sequence (Integrated DNA technologies (IDT)) were transcribed with the HiScribe T7 High Yield RNA Synthesis kit (NEB). Reactions were incubated for 16 hours at 37℃, treated with DNase I (NEB), and purified using the RNA Clean & Concentrator-25 kit (ZymoResearch). We performed gRNA screens in 10 µL vol-

umes using 100 nM Cas12a (NEB), 200 nM gRNA, 1x NEB 2.1 buffer (NEB), 1 μM ssDNA fluorescent quenched reporter (56-FAM/TTATT/3IABkFQ, IDT) and 0-100 pM spiked target DNA standard diluted in water. Reactions were incubated at 37℃ for 30 minutes and fluorescence kinetics were measured using a BioTek NEO HTS plate reader (BioTek Instruments) with readings every 2 minutes (excitation: 485 nm; emission 528 nm). For the assay, gRNAs with the highest fluorescent signal and lowest background signal to be tested against RPA primer combinations were selected. RPA screens were performed as per manufacturer's instructions using 7.5 μL reaction volumes from the RPA Basic kit (TwistDx). 10 μL of RPA primer screen reaction was added to a 1.25 μL Cas reaction with the same reaction conditions as described for the gRNA screen. Selected RPA primers and gRNAs are shown in Table 3.1.

### 3.4.2 Magnetic concentration of microbes in simulated water.

Varying concentrations of *E. coli* species were added to 50ml of aqueous buffer solution (0.2g/l of sea salts, 0.5μM Tannic acid and 2mM sodium bicarbonate at pH 7) and mixed them with 5 μg of magnetic beads coated with FcMBL fusion. A KingFisher Flex (Thermo) was used to collect and wash the bacteria-coated magnetic beads. Captured *E.Coli* were detected with MBL-conjugated to horse radish peroxidase (MBL-HRP). *E.Coli* detection was visualized by incubation with tetramethylbenzidine (1-step Ultra TMB-ELISA). TMB development was terminated by adding 1M sulphuric acid, and the absorbance was determined at 450nm wavelength.

### 3.4.3 Filter concentration of microbes in simulated water.

Varying concentrations of *S. enterica* genomic DNA (ATCC 700720DQ) were added to 50ml of DPBS buffer (Thermo Fisher 14190144) and incubated with 1-3 PES filters (Millipore, cat. GPWP04700) that were punched using 4 mm Integra surgical biopsy punches for 30 minutes at 37℃in a shaking incubator. Following this, the filters were removed using sterile tweezers and added to an RPA reaction, which was was performed as per the manufacturer's instructions using 25 μL reaction volumes from the RPA Basic kit (TwistDx). 10 μL of RPA primer screen reaction was added to a 1.25 μL Cas reaction with the same reaction conditions as described for the gRNA screen.

### 3.4.4 Construction of waterborne pathogen miSHERLOCK device

The miSHERLOCK platform was designed using Autodesk's Fusion 360 3D CAD software. The housing and components were printed using a Formlabs Form 3 printer (FormLabs). Black resin was chosen to print the housing to minimize reflectance when reading the fluorescence assays. An orange acrylic sheet (McMaster-Carr) was laser cut to fit within the slot. Double-sided tape (Scotch) was used to line the water

reservoir with aluminum foil (Reynolds) and 50 µL of nuclease-free water was loaded. A neodymium magnet (0.25" x 0.25" x 0.25") (McMaster-Carr 5848K91) was added to the sample preparation area, followed by 2-3 sheets of Whatman gel blotting paper GB003 (Sigma Aldrich) for absorption of water. A PES filter between 4-5mm was loaded into the top of the 50ml conical tube attachment. A high-power resistor (Digikey 355018RJT) was used for the heating element. The LEDs (Adafruit 301) were soldered in series. The heater and LEDs were mounted to the housing and connected to a 3.7V 2500mAh LiPo battery (Adafruit 328).

### 3.4.5 Construction of zero-G miSHERLOCK device

The miSHERLOCK platform was designed using Autodesk's Fusion 360 3D CAD software. The housing and components were printed using a Formlabs Form 3 printer (FormLabs FLM-PLA- 175-BLK). Black resin was chosen to print the housing to minimize reflectance when reading the fluorescence assays. A 2 mm orange acrylic sheet (McMaster-Carr) was laser cut (Universal Laser Systems VLS2.30) to 2.75 cm x 2.25 cm for the transilluminator filter. Double-sided tape (Scotch) was used to tightly line the water reservoir with aluminum foil (Reynolds) and 50 µL of nuclease-free water was loaded. Electronic components for the polyimide heaters (Alibaba), temperature controller (DigiKey TC622VAT), and LED lights (Adafruit 301) were soldered, with heat shrink applied to all wires. The set point for the temperature controller circuit was programmed to 37℃ by selecting a 120 k Ω resistor (DigiKey S120KCACT) and confirmed using a Dallas DS18B20 digital temperature sensor. The LEDs were soldered in series with a 200 Ω resistor (DigiKey S200CACT) to allow for a current of  20 mA when attached to the 9V battery source. The temperature sensor circuit and LEDs were mounted to the housing and connected to the battery pack.

### 3.4.6 Data analysis

Fluorescence values are reported as absolute for all experiments used for LOD calculation. Background-subtracted fluorescence, in which the fluorescence value at the initial time point (0 minutes) is subtracted from the end time point (usually 60 minutes) is reported for initial screening experiments. All data were plotted and statistical tests were performed using GraphPad Prism 8. Figures were created using Adobe Illustrator 2020.

# Chapter 4

# Multiplexed electrochemical sensor for point-of-care simultaneous RNA and antibody detection

This chapter has been adapted from:

**Najjar, D.***, Rainbow, J.*, Timilsina, S. S.*, Jolly, P.*, de Puig*, H., Yafia, M., ... & Ingber, D. E. (2022). Lab-on-a-chip multiplexed electrochemical sensor enables simultaneous detection of SARS-CoV-2 RNA and host antibodies. *Nature Biomedical Engineering.*

*Equal contribution

In this Chapter, I describe research that I contributed to in designing a lab-on-a-chip microfluidic platform that incorporates multiplexed electrochemical sensing of nucleic acids and antibodies from a raw saliva sample. This project began around 2018 with a collaboration between the Collins and Ingber lab. I joined on in Spring 2021 to help design, optimize, and validate the LAMP primers and CRISPR guides for the CRISPR-based RNA detection, as well as perform the assays for all of the saliva patient samples for RNA detection in the initial set of experiments. In Fall 2021, I worked closely with the team to integrate the saliva-based RNA and antibody detection into a lab-on-a-chip microfluidic system, which was critical for automating the extraction, amplification, and detection of the RNA in particular. I should note that "chip" in this context is that of a microfluidic chip, which tend to be on the order of centimeters in size with channels that are on the order of microns in width. I would like to thank Mohamed for all the time that he took to constantly change the chip design each time I ran an experiment and had slight modifications. I worked with Mohamed and Sanjay to run all of the clinical samples on the LOC microfluidic system, re-wrote the manuscript, and re-designed and figures. I am grateful to all those who were involved in this (very long) process and provided so much guidance and insight. It was a wonderful team effort.

## 4.1 Introduction

The COVID-19 pandemic has highlighted the need for cost-effective diagnostics for SARS-CoV-2 RNA as well as for detection of antibodies generated by the host in response to infection. This type of multi functional detection platform will be particularly useful for diagnosis of both acute and convalescent infections, as well as for assessing patient immunization status following vaccination. The clinical timeline of SARS-CoV-2 infection consists of an acute phase, when viral RNA is detectable in clinical samples, such as saliva or nasopharyngeal swabs, followed by a convalescent phase when serology biomarkers, such as IgG antibodies, are present in saliva and serum [98]. Therefore, simultaneous analysis of these different biomarkers in clinical samples as the disease progresses could provide more accurate results for disease monitoring and management.

Molecular (nucleic acid) diagnostics that detect the presence of viral RNA are key to detecting the virus during the first 5 days of infection, with a viral load peak around day 4 (Fig. 4-1) [233, 166, 243]. Following the first few days of infection, the host produces IgM, IgA, and IgG antibodies in a process known as seroconversion. These antibodies often become stable after the first 6 days of seroconversion and their titers can remain stable over months[86, 72].

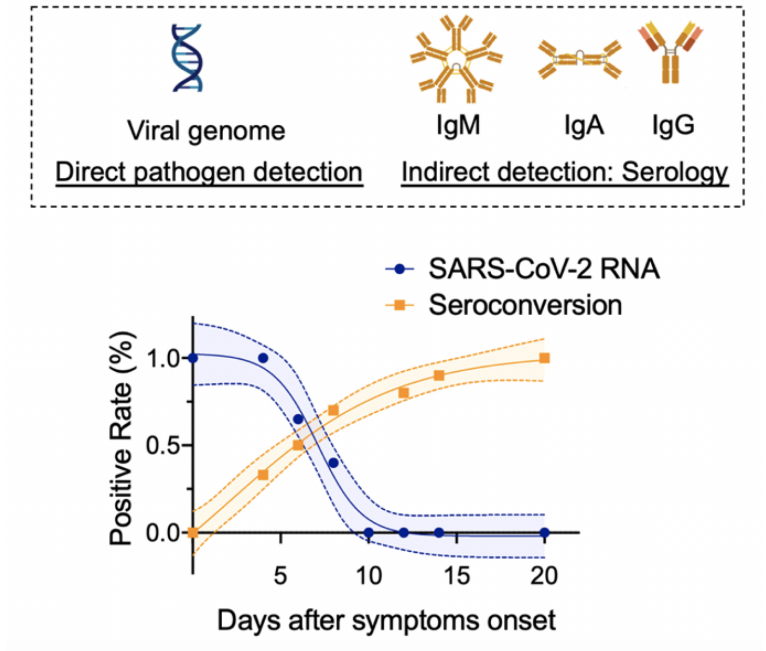



Figure 4-1: **Direct pathogen detection approaches detect the virus, its genome or viral antigens.** These assays are useful during the acute phase of disease, typically 0-7 days after symptoms. Serology assays measure the presence of host antibodies, such as IgG, IgM or IgA, that the patient generates during adaptive immune response to the disease. Data replotted from [98, 180]

The presence of different antibody types varies during infection [222] and correlates with disease severity [51, 159]. In particular, extensive cohort studies in hospitalized patients show that IgG antibodies against different viral proteins (e.g., nucleocapsid or spike proteins) correlate with disease severity and outcomes. For example, antibodies against the spike (S) protein are more specific [63] and correlate with virus neutralization [51], while antibodies against the nucleocapsid (N) protein have longer clearance rates [142] and might appear earlier during infection [150]. Therefore, serological assays that detect the host's antibodies developed after an infection can widen the testing window for SARS-CoV-2 beyond the molecular diagnostic time frame and provide insights into the patient's progression.

Such assays also may be used to determine whether titers are maintained over time and thus the effectiveness of vaccination responses. For example, SARS-CoV-2 vaccine efficacy trials have highlighted a direct correlation between the titer of antibodies targeting the Receptor Binding Domain (RBD) found in the S1 subunit of the S protein, the neutralizing antibody titer, and vaccine efficacy [78]. Thus, the development of serological assays targeted to individual SARS-CoV-2 viral antigens could have important implications for predicting the efficacy of vaccines and estimating the need for boosters.

Molecular diagnostics, on the other hand, commonly involve methods such as quantitative Polymerase Chain Reaction (qPCR), which require rigorous sample preparation and temperature control, cold storage of reagents, expensive instrumentation (requiring routine maintenance), and trained personnel to run the tests. While there are home diagnostic tests approved for use by the U.S. Food and Drug Administration (FDA), many of these tests either involve self-collection and mailing to a central laboratory or are based on rapid antigen tests, which have shown to be less accurate than nucleic acid tests, such as qPCR [141].

During the last few years, powerful diagnostic techniques that capitalize on CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) technology and associated programmable endonucleases have gained significant interest, due in part to their high specificity, programmability, and capacity to work at physiological conditions [35, 81, 80, 121, 106]. CRISPR-based diagnostics capitalize on endonucleases, such as Cas12a, which has a specific cleavage activity towards double-stranded DNA (dsDNA) fragments matching its guide RNA (gRNA) sequence. Once the Cas12a-gRNA complex binds to its dsDNA target, it activates and subsequently engages in indiscriminate collateral hydrolysis of nearby single-stranded DNA (ssDNA) [74, 58].

Electrochemical (EC) methods of CRISPR-based nucleic acid detection typically combine nucleic acid probes conjugated on an electrode with CRISPR Cas effectors and have detection limits in the picomolar-femtomolar ($10^{-12}$-$10^{-15}$) range [29, 88, 235, 125, 240]. Unfortunately, those detection limits are inadequate for diagnosing SARS-CoV-2 in clinical samples, which require ultrasensitive RNA detection at the attomolar ($10^{-18}$) scale. To overcome this limitation, amplification techniques such as

loop-mediated isothermal amplification (LAMP) can be performed prior to Cas12a detection, which improves the sensitivity of fluorescent CRISPR-based assays by orders of magnitude [35, 81, 80, 106, 58].

The combination of serological and nucleic acid diagnostics can improve the overall accuracy of SARS-CoV-2 diagnosis [132] and provide qualitative data on the patient's disease severity and state of progression. However, workflows for SARS-CoV-2 diagnosis that integrate serological and nucleic acid tests are currently limited as these assays require laboratory equipment such as pipettes, centrifuges, and heating blocks, as well as specialized technical skills. Despite the advances to make SARS-CoV-2 tests as widely available as possible, a lab-on-a-chip (LOC) platform for point-of-care (POC) use that can detect both SARS-CoV-2 RNA and serological markers is not yet available.

EC biosensors offer a particularly promising solution to achieve ultra-sensitive, selective, multiplexed, quantitative, and cost-effective LOC detection of both nucleic acids and proteins; they also offer the potential to interface with electronic medical records, integrated cloud systems, and telemedicine. Despite their potential, EC diagnostic platforms have only been used to detect either nucleic acid or proteins individually [29, 217] and have been limited to platforms that require multiple liquid handling steps and specialized equipment. Here, we describe a low-cost, 3-D printed, self-contained, LOC diagnostic platform that is capable of concurrent detection of both SARS-CoV-2 nucleic acids and host antibodies directed against the virus from unprocessed saliva samples. This device integrates microfluidics that enable automated liquid handling for sample preparation and a simple and sensitive read-out for both viral RNA as well as host antibodies. The simplicity of this device makes it user-friendly and should enable its use for POC testing within hospitals and in COVID-19 testing clinics.

## 4.2 Results

### 4.2.1 Design of the microfluidic electrochemical lab-on-a-chip platform

In considering the design of an LOC diagnostic device, we focused on accuracy, ease of use, and a capacity to integrate with digital health platforms. Towards this goal, we fabricated a microfluidic chip capable of processing untreated saliva to detect both viral RNA and host antibodies on the same EC sensor chip (Fig. 4-2).

We chose saliva over the more common nasopharyngeal swab specimens on our device due to saliva's ease of self-collection and increased clinical sensitivity due to a high viral load [234, 165]. Because unprocessed saliva is viscous and contains nucleases that could inhibit a downstream reaction, viral RNA is typically purified from saliva using costly and complicated nucleic acid purification kits. However, to eliminate the

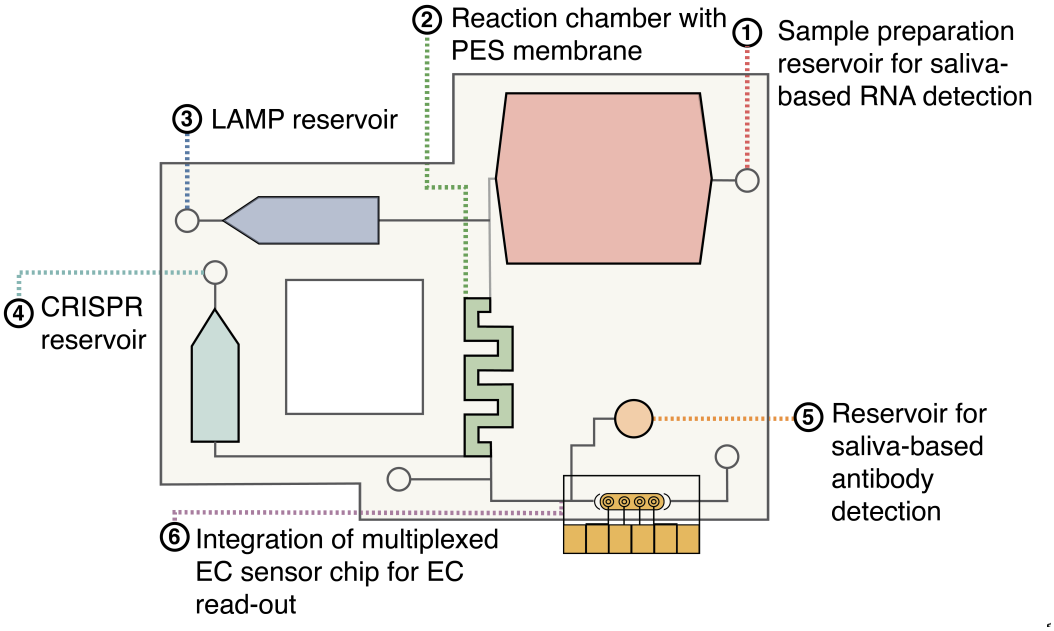



Figure 4-2: **Overview of microfluidic chip for LOC EC-based multiplexed detection of SARS-CoV-2 biomarkers.** Overview of the microfluidic chip designed for LOC sample-to-answer saliva detection of SARS-CoV-2 RNA and antibodies.

need for purification prior to reaction, we experimented with various sample preparation workflows that would allow for viral lysis and nuclease inactivation (Fig. 4-3). Among different buffers tested, we found that incubating the saliva sample with a Proteinase K [119] solution at 55℃ followed by a 95℃ inactivation effectively lysed the sample and eliminated false positive signals without inhibiting the performance of the downstream LAMP and CRISPR-based detection steps.

Following the sample preparation step, we needed a method to capture the nucleic acids from the saliva sample. We found that a polyethersulfone (PES) membrane was able to concentrate the RNA without inhibiting the downstream reactions [185, 130, 191, 47]. Overall, the device incorporates a 500µL sample preparation chamber which is used for the saliva-based RNA detection, a 50µL reaction chamber that contains the PES membrane, a 25µL LAMP reservoir, a 22µL CRISPR reservoir, and a 20µL reservoir for direct saliva-based antibody detection. The compact design allowed for the use of two high-power resistors as heating elements (Fig. 4-4, B-1), one beneath the sample preparation reservoir and the other beneath the reaction chamber. The LAMP and CRISPR reservoirs were isolated from the sample preparation reservoir and reaction chamber to assure that there was no carryover heating from the heated elements during the sample preparation and concentration.

To optimize chip functionality, we tested a variety of reservoir and channel dimensions and placements to ensure uniform heating and precise fluid flow (Fig. B-2). The chip was designed with different channel dimensions to establish resistors to the fluid flow

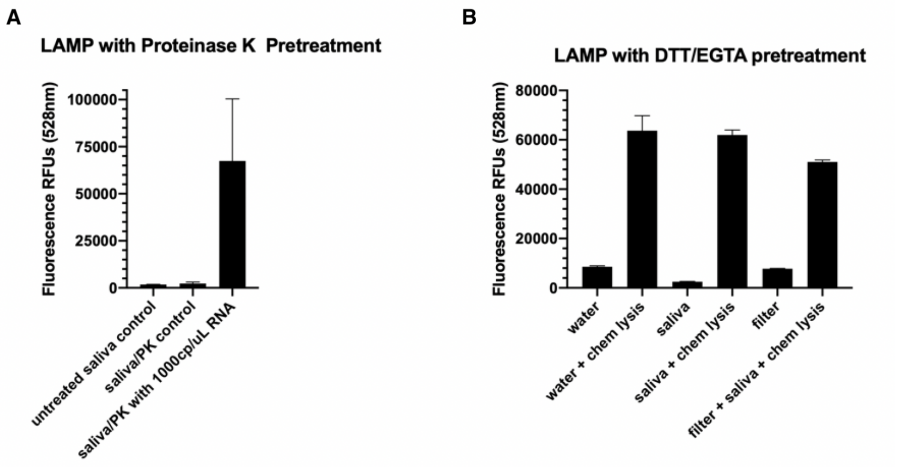



Figure 4-3: **Lysis testing for saliva sample preparation** A. Proteinase K incubated with saliva at a 1:10 concentration for 15 mins at 55℃ followed by a 95℃ inactivation for 5 min, was found to be an effective strategy for lysing the salvia sample and inactivating nucleases. B. A mixture of 100mM DTT and 500mM EGTA caused a false negative signal when run with the LAMP reaction.

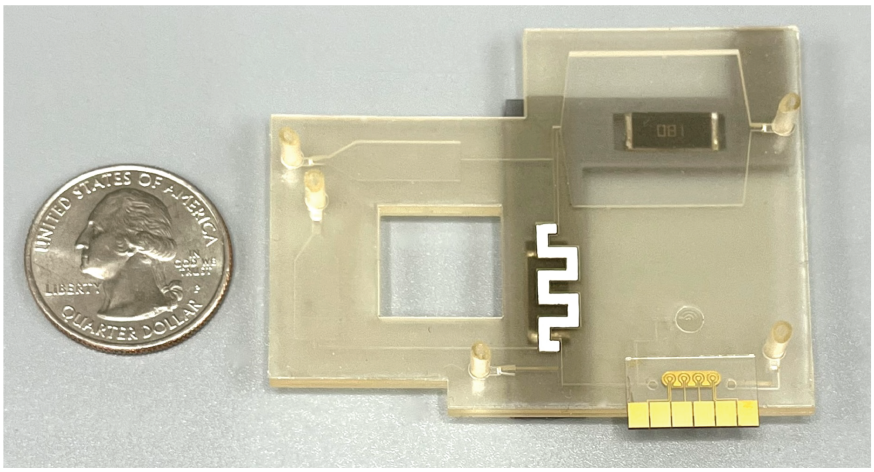



Figure 4-4: **Photograph of the microfluidic system.** Photograph of the microfluidic system, consisting of the EC sensor chip, LOC microfluidic chip, and power resistor heaters.A quarter dollar is provided for scale.

while the fluids were controlled by a small, fingertip-sized DC peristaltic pump with a maximum flow rate of 200 µl/min controlled by an Arduino microcontroller. The channels leading to the reservoirs were designed with higher resistance to avoid any backflow or cross-contamination with the reservoirs (400 x 150 µm) (Fig. 4-5), while the ends of the channels leading to the reservoirs were designed with a stepped stop valve to stop any flow from the reaction chamber to the reservoirs.

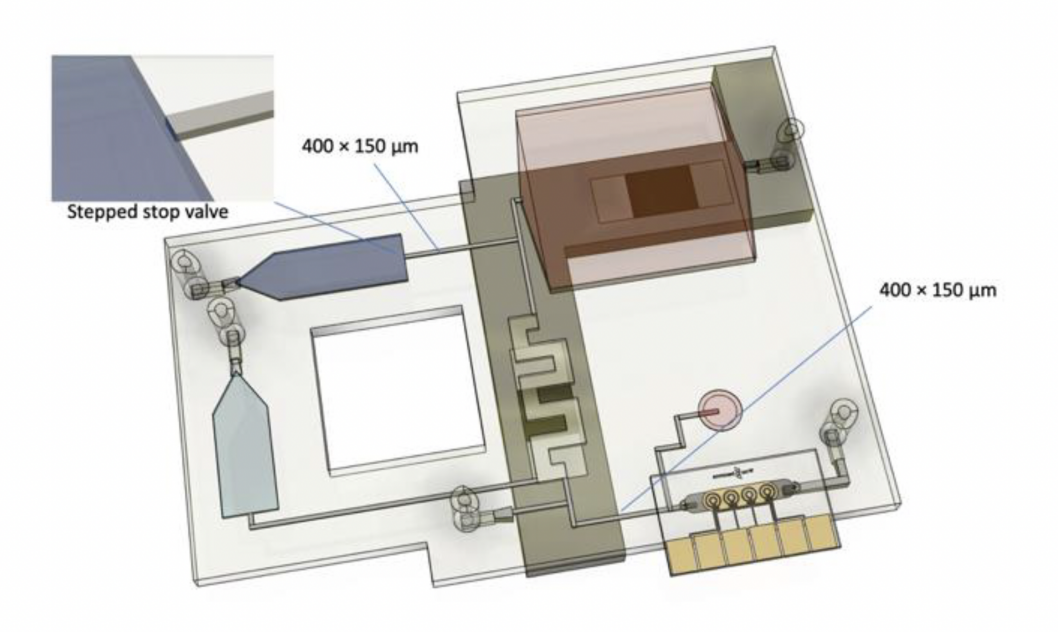



Figure 4-5: **Dimensions of the microfluidic chip.** Dimensions of the microfluidic chip. The width of the channels on the microfluidic device were designed to have precise control over the fluid movements within the chip.

We tested a variety of reaction chamber geometries to ensure that the PES membrane surface area was adequate for nucleic acid capture and that the chamber had optimal fluid retention over the course of the reaction, settling on a serpentine shape (Fig. B-2) and designed higher resistance on the flow path to the EC sensor chip to avoid unwanted flow before the desired time point.

The workflow of the sample-to-answer microfluidic detection device is as follows, with all pumping steps referring to the use of the Arduino-controlled small peristaltic pump and all heating steps referring to the high-power resistors beneath the microfluidic chip (Fig. 4-2): The user manually inputs saliva onto the antibody detection reservoir (Fig. 4-2, 5) and the RNA sample preparation reservoir (Fig. 4-2, 1). In the RNA sample preparation chamber, the saliva combines with a Proteinase K solution and is heated to 55℃ for 15 mins followed by 95℃ for 5 mins to allow for virus lysis and nuclease inactivation [119]. The saliva sample is then pumped over the PES membrane within the reaction chamber, where the RNA binds to the membrane. The reaction chamber is heated to 95℃ for an additional 3-5 min to ensure denaturation of potential reaction inhibitors. The LAMP solution is then pumped from the reservoir into the reaction chamber and incubated at 65℃ for 30 min, followed by the CRISPR mixture and incubated at 37℃ for an additional 30 min, after which it is pumped over the EC sensor chip to incubate. Saliva for antibody detection is then directly pumped over the EC sensor chip and incubated. The chip is washed with Phosphate Buffered Saline with Tween® 20 (PBST) followed by the addition of Polystreptavidin-Horseradish Peroxidase (HRP) and a precipitating form of tetram-

ethylbenzidine (TMB), after which the chip is read using a potentiostat. A more in-depth visual of the microfluidic chip and its workflow is provided in Fig. B-3).

### 4.2.2 Integrated CRISPR-based molecular diagnostic assay

The CRISPR-Cas RNA detection experiments reported here capitalize on Cas12a from a Lachnospiraceae bacterium, which has specific cleavage activity towards ds-DNA fragments matching its guide RNA (gRNA) sequence. Upon target binding, activated Cas12a-gRNA engages in collateral cleavage of nearby single-stranded DNA (ssDNA) [35] that can be read optically as an increase in fluorescence due to the hydrolysis of a fluorophore-quencher labeled ssDNA reporter. As a result, the signal obtained from the fluorescent assay will be high in the presence of its SARS-CoV-2 target and low when it is not detected. LAMP primers [175, 138, 186, 156, 237] and Cas12a-gRNAs were evaluated from a range of conserved regions in the SARS-CoV-2 genome to determine the most sensitive combinations using commercially available, synthetic full-length SARS-CoV-2 genomic RNA. The ORF1a assay, which targets a highly conserved region in the SARS-CoV-2 viral genome, had a limit of detection (LOD) of 2.3 viral RNA copies/μL with a reaction time of 50 min (Table B.1, Fig. 4-6).

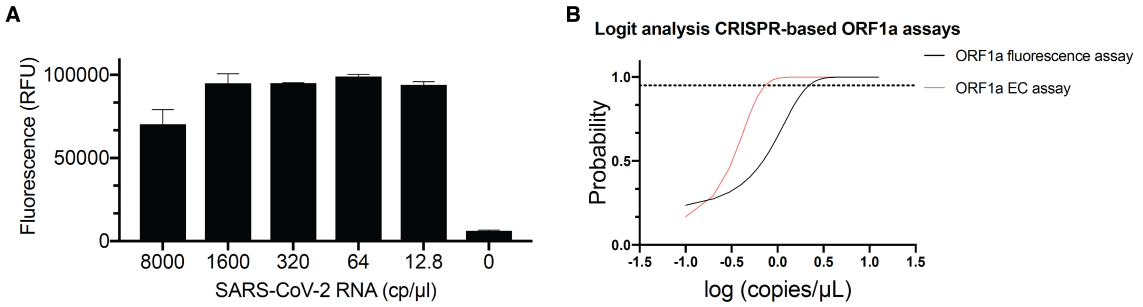



Figure 4-6: **Limit of detection of the CRISPR-based assay with a fluorescence output.** A. Serially diluted full-length SARS-CoV-2 RNA was spiked in water and amplified by RT-LAMP. Results show that dilutions down to 12.8 cp/μL of viral RNA had a clear positive signal (Student's t test p value $<0.001$) in our fluorescence-based assays. Error bars represent the standard deviation of triplicate experiments, biological replicates. B. Sensitivity of the viral RNA electrochemical assay was higher than the fluorescence output by comparing logit regression curves and their fit characteristics. Serially diluted full-length SARS-CoV-2 RNA was spiked in water and amplified by RT-LAMP. Results show that dilutions down to 0.8 cp/μL of viral RNA had 95% probability of leading to a clear positive signal in our electrochemical platform (red)as compared to the fluorescence-based platform of 2.3cp/μl (black). Table B.1 contains the raw data used to fit the logit function. Each concentration was probed with five independent biological replicates.

This LOD is comparable to high-performance SARS-CoV-2 reverse transcription

quantitative polymerase chain reaction (RT-qPCR) assays [62], with half the time to result. To confirm that the assay was able to detect SARS-CoV-2 active virus, the ORF1a assay primers and guide RNA were further validated using 11 SARS-CoV-2 RT-qPCR negative patient saliva samples and 19 SARS-CoV-2 RT-qPCR positive saliva samples with a range of cycle threshold (CT) values (Fig. 4-7, Table B.2).

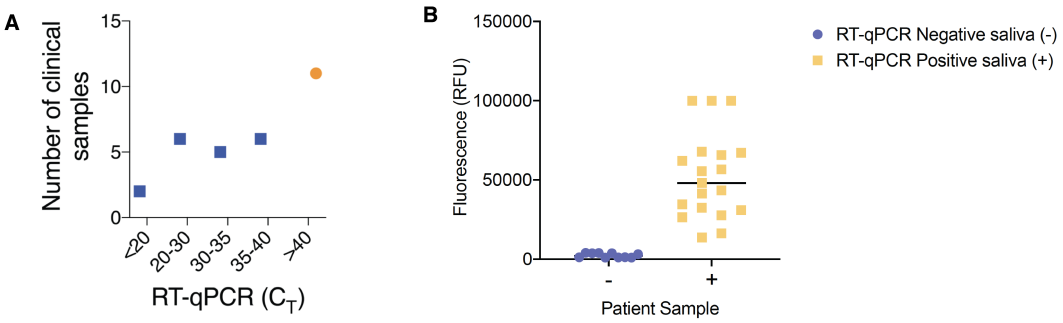



Figure 4-7: **Clinical patient saliva sample characterization.** A. Clinical SARS-CoV-2 saliva samples with a wide range of RT-qPCR cycle thresholds (CT) plotted against the number of samples tested in our device. (blue squares at RT-qPCR positive and orange circle is RT-qPCR negative). B. CRISPR-based fluorescence assays can diagnose SARS-CoV-2 positive and negative clinical samples. SARS-CoV-2 RT-qPCR positive saliva (+, yellow square) shows a high fluorescence signal whereas negative samples (-, purple circle) show a low fluorescence signal. Unpaired Student's t test $p<0.0001$ for differences in fluorescence between SARS-CoV-2 positive and negative samples.

To integrate the CRISPR-based molecular assay onto the EC sensor platform, we designed a biotinylated ssDNA reporter probe (RP) that partially hybridized to peptide nucleic acid (PNA) capture probes immobilized on the surface of the antifouling composite-coated gold electrodes [245] (Fig. 4-8). Functionalized EC biosensors were incubated with samples containing the LAMP/Cas12a mix which includes the biotinylated ssDNA RP. In the presence of SARS-CoV-2 target RNA, Cas12a collaterally cleaved the biotinylated ssDNA reporter, leading to a reduction of binding of poly-HRP-streptavidin and thus, a reduction in the precipitation of TMB deposited locally on the surface of the electrode [245].

Reduced precipitation of TMB was recorded as peak current, which was measured using cyclic voltammetry (CV) by sweeping the voltage between -0.5 and 0.5 V (Fig. 4-9A). As a result, the signal obtained from the EC platform demonstrated an inversely proportional relationship with target concentrations. It should be noted that the LAMP amplification prior to the CRISPR-based sample detection increases the sensitivity of the sensor, allowing for a consistently distinguishable signal from samples that are at or above the limit of detection (LOD) of the sensor (Fig. 4-9B).

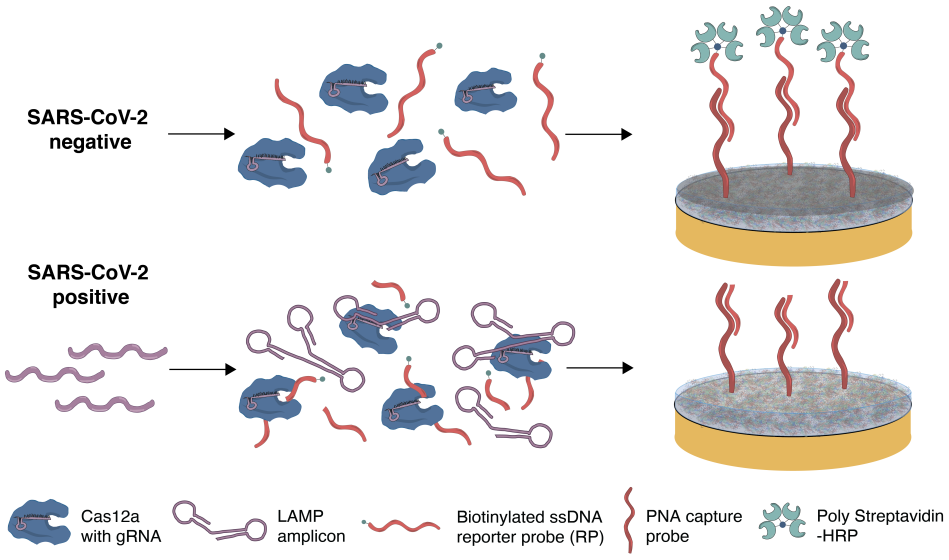
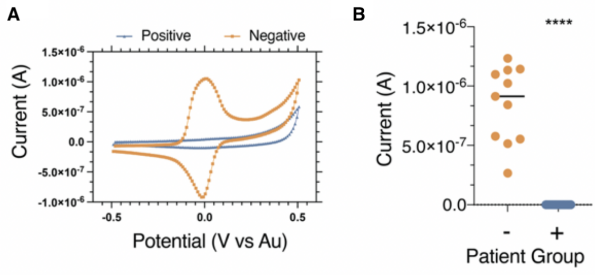



Figure 4-8: **Schematic illustrating the surface chemistry of the EC assay.** Without viral RNA present, the biotinylated ssDNA reporter probe is not cleaved; therefore, the poly-HRP streptavidin binds to the PNA/biotin-DNA duplex when added to the EC sensor chip and consequently precipitates TMB resulting in an increase in current. In contrast, the biotinylated reporter ssDNA is hydrolyzed in the presence of viral target RNA, cleaving the biotin group. Consequently, poly-HRP streptavidin does not bind to the surface of the chips, resulting in no TMB precipitation and no increase in current.

To optimize the binding efficiency of the PNA-based CRISPR-EC sensor platform, we varied the concentration and incubation time of the RP to obtain a rapid, high signal-to-noise ratio (Fig. B-4A, B-5). Among all the concentrations tested, 1nM RP and 5-min incubation produced a high signal with no background. Interestingly, the CRISPR-EC sensor platform gave a single molecule LOD of 0.8 cp/μL, which was nearly four times more sensitive than the initial fluorescence-based assays used to validate the primer and guide pairs (Supplementary Table S1, Figs. 4-6B and B-4B.

To determine the potential clinical value of the optimized CRISPR-EC sensor platform, we extracted RNA from 19 patient saliva samples that were positive for SARS-CoV-2 based on RT-qPCR with a range of CT values and 11 RT-qPCR negative clinical saliva samples (Fig. 4-7A). The current measured in the form of the output signal from the electrodes was clearly distinguishable (p-value $<0.0001$) when comparing the SARS-CoV-2 positive and negative samples (Fig. 4-9B). In addition, ROC curve analysis demonstrated an impressive correlation with RT-qPCR and CRISPR-based fluorescent detection, with 100% accuracy, and AUC=1 (Fig. 4-9C).




C

| ROC Curve analysis summary | |
|---|---|
| AUC | 1.00 |
| 95% conf. int. | [1.00-1.00] |
| P value | <0.0001 |
| Cutoff | 8859 |
| Sensitivity | 1.00 [0.83-1.00] |
| Specificity | 1.00 [0.74-1.00] |
| N pos. | 19 |
| N neg. | 10 |

Figure 4-9: **Performance of the CRISPR electrochemical assays using clinical samples.** A. Cyclic voltammogram showing the typical current peak signal achieved after incubation of samples from both SARS-CoV-2 negative (orange) and positive (blue) clinical samples. B. Clinical samples that contained SARS-CoV-2 viral RNA (+, blue) had low signals in our device and were clearly distinguishable from the high signals obtained for samples that did not contain viral RNA (-, orange). Student's t-test p value <0.001 (***). C. Summary table listing the numerical values of the receiver operating characteristic (ROC) curve analysis of the patient sample data collected for the SARS CoV-2 assay. The table shows a summary of the results from 19 RT-qPCR confirmed positive and 10 negative human saliva samples. AUC: area under the curve; 95% conf. Int.: 95% confidence interval; Sens: sensitivity; Spec.: specificity; N pos.: number of RT-qPCR SARS-CoV-2 positive samples; N neg.: number of RT-qPCR SARS-CoV-2 negative clinical samples.

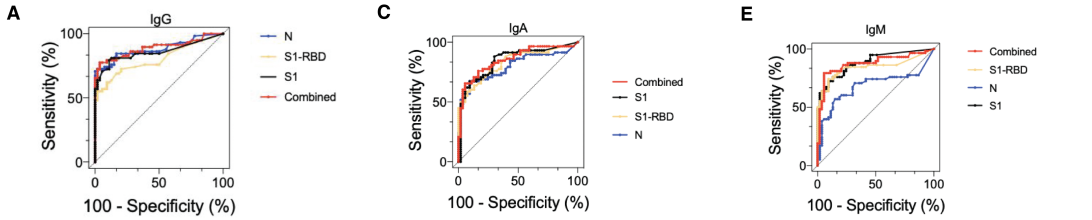
### 4.2.3 Integrated CRISPR-based molecular diagnostic assay

Multiplexed assays that diagnose disease by combining viral RNA with serology markers can lead to a more robust understanding of the progression of diseases, including SARS-CoV-2 [210]. The primary antigens that elicit antibodies during coronavirus infection are the N and S proteins [63]. The N protein is the most abundant viral protein and is highly conserved among the coronavirus family [188], with several studies indicating that IgG antibodies targeting the N protein may be more sensitive due to their early appearance and longer clearance rates during and post-infection [150, 222, 142]. While the S protein is less conserved than N, it is highly immunogenic due to the S1 subunit which mediates attachment to target cells [96]. Studies indicate IgG antibodies against the S1 subunit (S1) are more specific and strongly correlate with virus neutralization [171, 8].

Antibodies targeting the S1 subunit and the ribosomal binding domain within the subunit (S1-RBD) may also be used to ensure the effectiveness of vaccination responses and titer maintenance over time, with SARS-CoV-2 vaccine efficacy trials highlighting a direct correlation between the titer of antibodies targeting the RBD of the S protein, the neutralizing antibody titer, and vaccine efficacy [78]. Therefore, to maximize our assay's accuracy, we leveraged the multiplexing capabilities of our EC sensor and fabricated a multiplexed serology assay capable of measuring antibod-

ies against the S1, S1-RBD, and N proteins for detection of a patient's immunity, whether through prior viral infection or vaccination.

We began by using an Enzyme-Linked Immunosorbent Assay (ELISA) to optimize the reagents prior to assembling the multiplexed serology EC sensor chip (Fig. B-6, B-7, B-8, 4-10). Initially, a small set of high-titer and low-titer positive clinical plasma samples were used to optimize the assay over a broad range of IgG titers. The best performing capture antigens were S1, S1-RBD, and N, and the best performing detection antibodies were biotinylated goat anti-human IgG, rabbit anti-human IgM, and goat anti-human IgA. We validated the ELISA accuracy with 58 SARS-CoV-2 plasma samples from patients with a prior positive SARS-CoV-2 RT-qPCR result and 54 SARS-CoV-2 negative samples. Out of the 54 negative SARS-CoV-2 samples, 22 were collected before the onset of the SARS-CoV-2 pandemic. The ROC curve analysis of the ELISA results yielded areas under the curve (AUC) between 0.68 and 0.89 for IgG, IgM and IgA (Fig. 4-10).




B

| | IgG | | | |
|---|---|---|---|---|
| | S1 | S1-RBD | N | Combined |
| AUC | 0.86 | 0.80 | 0.88 | 0.89 |
| 95% conf. int. | [0.78-0.93] | [0.72-0.88] | [0.82-0.95] | [0.82-0.96] |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Cutoff | >0.018 | >0.022 | >0.041 | >0.064 |
| Sens. | 0.64 [0.51-0.75] | 0.57 [0.44-0.69] | 0.72 [0.60-0.82] | 0.72 [0.60-0.82] |
| Spec. | 0.98 [0.90-1.00] | 0.94 [0.85-0.99] | 0.98 [0.90-1.00] | 0.98 [0.90-1.00] |
| N pos. | 54 | 54 | 54 | 54 |
| N neg. | 58 | 58 | 58 | 58 |

D

| | IgA | | | |
|---|---|---|---|---|
| | S1 | S1-RBD | N | Combined |
| AUC | 0.85 | 0.83 | 0.80 | 0.85 |
| 95% conf. int. | [0.77-0.92] | [0.75-0.91] | [0.71-0.88] | [0.78-0.93] |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Cutoff | >0.110 | >0.115 | >0.260 | >0.376 |
| Sens. | 0.48 [0.36-0.61] | 0.50 [0.38-0.63] | 0.52 [0.39-0.64] | 0.60 [0.48-0.72] |
| Spec. | 0.98 [0.90-1.00] | 0.98 [0.90-1.00] | 0.98 [0.90-1.00] | 0.96 [0.88-0.99] |
| N pos. | 54 | 54 | 54 | 54 |
| N neg. | 58 | 58 | 58 | 58 |

F

| | IgM | | | |
|---|---|---|---|---|
| | S1 | S1-RBD | N | Combined |
| AUC | 0.89 | 0.84 | 0.68 | 0.87 |
| 95% conf. int. | [0.82-0.95] | [0.77-0.92] | [0.58-0.79] | [0.80-0.95] |
| P value | <0.0001 | <0.0001 | 0.0009 | <0.0001 |
| Cutoff | >0.007 | >0.067 | >0.080 | >0.128 |
| Sens. | 0.62 [0.49-0.73] | 0.55 [0.42-0.67] | 0.38 [0.27-0.51] | 0.79 [0.67-0.88] |
| Spec. | 0.98 [0.90-1.00] | 0.98 [0.90-1.00] | 0.96 [0.87-0.99] | 0.94 [0.85-0.99] |
| N pos. | 54 | 54 | 54 | 54 |
| N neg. | 58 | 58 | 58 | 58 |

Figure 4-10: **ROC curve for SARS-CoV-2 ELISA from patient plasma samples.** (A, C, E). ROC curve analysis of the patient sample data collected for the IgG, IgA, and IgM SARS CoV-2 assay, respectively. (B, D, F). Table listing numerical values of the ROC curve analysis. AUC: area under the curve; 95% conf. Int.: 95% confidence interval; Sens: sensitivity; Spec.: specificity; N pos.: number of RT-qPCR SARS-CoV-2 positive samples; N neg.: number of RT-qPCR SARS-CoV-2 negative clinical samples.

Next, we used the optimized reagents to develop a multiplexed EC sensor to measure the humoral response against SARS-CoV-2 from patient plasma samples. We fabricated an EC sensor with an antifouling coating composed of bovine serum albumin and reduced graphene crosslinked with glutaraldehyde (BSA/rGOx/GA) [235, 29, 217] 25-27, where each electrode was individually functionalized with S1 (electrode 1), S1-RBD (electrode 2), N (electrode 3), or BSA as an on-chip negative control (electrode

4) to perform a multi-antigen sandwich EC ELISA (Fig. 4-11).

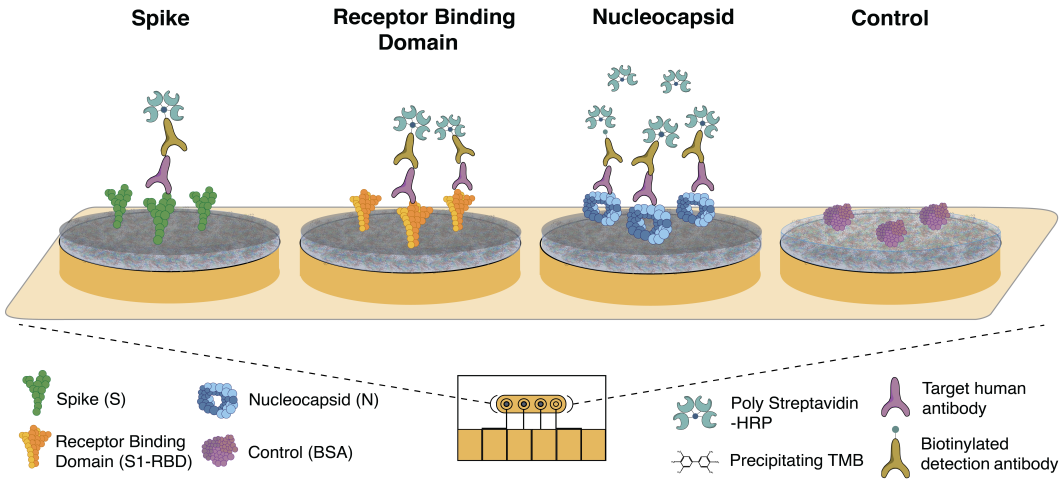



Figure 4-11: **Schematic of the multiplexed serology assay.** Schematic illustrating the multiplexed EC serological assay to assess host antibody responses on electrodes functionalized with SARS-CoV-2 antigens. Host antibodies bind to the SARS-CoV-2 antigens immobilized on the chips. Subsequently, biotinylated anti-human IgG secondary antibodies bind, followed by poly HRP-streptavidin binding and TMB precipitation on the chips.

We used an affinity-based sandwich strategy for the EC sensor assay, meaning that when SARS-CoV-2 antibodies were present, they bound to both the surface antigen and secondary antibody, leading to an increase in the EC signal. Each immunoglobulin isotype (IgG, IgM, IgA) was detected individually on the multiplexed EC sensor chips. Fig. 4-12 shows typical CV results obtained with the multiplexed EC sensor chips for detection of anti-SARS-CoV-2 IgG using SARS-CoV-2 positive and negative clinical samples.

We optimized the assay conditions of the serology EC assays to obtain the highest signal-to-noise ratios on both high- and low-titer clinical samples by varying factors such as plasma dilutions, sample incubation times, and TMB precipitation times (Fig. B-9, B-10, B-11). We found that a 30-min sample incubation at a 1:9 plasma dilution with a 3-min TMB precipitation time resulted in the highest sensitivity and specificity for detection of SARS-CoV-2 antibodies in clinical plasma samples (Fig. B-10C).

We then evaluated the accuracy of the EC serology platform using plasma samples from patients with a prior SARS-CoV-2 RT-qPCR positive result. We performed an ROC curve analysis utilizing the 58 SARS-CoV-2 positive plasma samples and 54 SARS-CoV-2 negative samples used to optimize the ELISA assay (Fig. 4-15A, 4-13A).

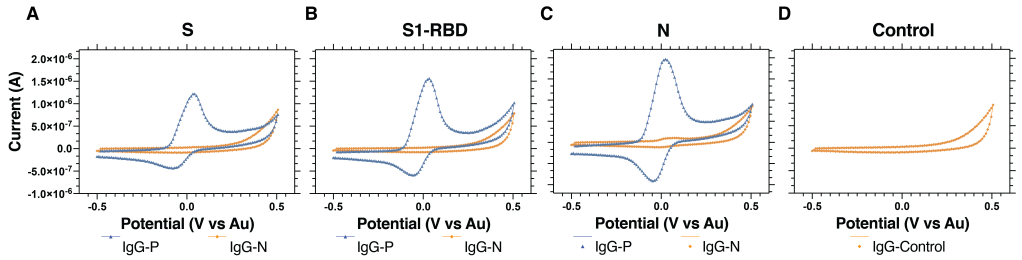
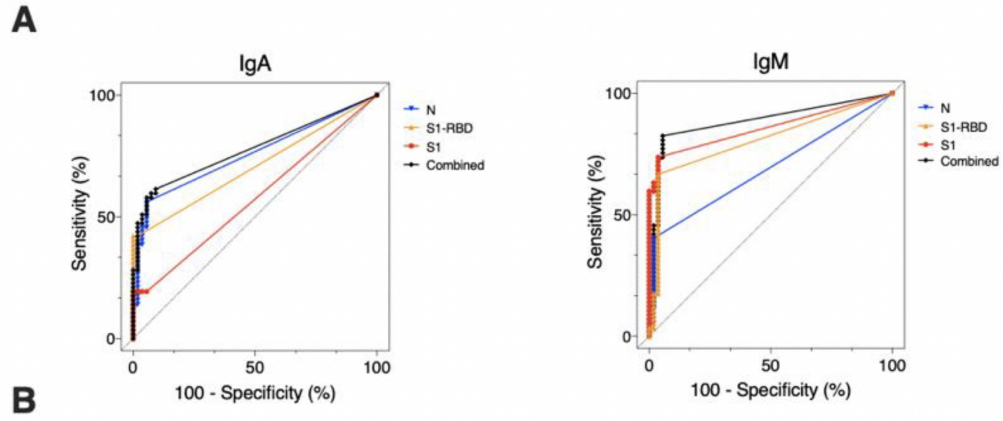



Figure 4-12: **Representative raw cyclic voltammetry data for the multiplexed serology assay.** Typical cyclic voltammograms for the four different electrodes that target host antibodies against A. Spike 1 subunit (S1), B. Spike 1-receptor binding domain (S1-RBD), C. nucleocapsid (N), and D. BSA negative control




B

| | IgA | | | | IgM | | | |
|---|---|---|---|---|---|---|---|---|
| | N | S1-RBD | S1 | Combined | N | S1-RBD | S1 | Combined |
| AUC | 0.71 | 0.76 | 0.57 | 0.78 | 0.80 | 0.86 | 0.70 | 0.89 |
| 95% conf. int. | [0.61-0.80] | [0.66-0.85] | [0.47-0.68] | [0.69-0.87] | [0.72-0.89] | [0.79-0.93] | [0.59-0.79] | [0.82-0.95] |
| P value | 0.0002 | <0.0001 | 0.18 | <0.0001 | <0.0001 | <0.0001 | 0.0006 | <0.0001 |
| Cutoff | >5.3e-8 | >4.1e-7 | >4.14e-8 | >4.5e-8 | >5.3e-9 | >2.1e-9 | >9.3e-9 | >5.3e-9 |
| Sens. | 0.42 [0.30-0.55] | 0.39 [0.27-0.52] | 0.19 [0.11-0.31] | 0.58 [0.45-0.70] | 0.67 [0.54-0.78] | 0.73 [0.61-0.83] | 0.40 [0.29-0.53] | 0.83 [0.71-0.90] |
| Spec. | 1 [0.93-1.00] | 0.98 [0.90-1.00] | 1 [0.27-0.52] | 0.94 [0.85-0.99] | 0.96 [0.87-0.99] | 0.96 [0.88-0.99] | 0.98 [0.90-1.00] | 0.95 [0.85-0.99] |
| N pos. | 54 | 54 | 54 | 54 | 54 | 54 | 54 | 54 |
| N neg. | 58 | 58 | 58 | 58 | 57 | 57 | 57 | 57 |

Figure 4-13: **Electrochemical serological assays can detect SARS-CoV-2-specific IgA and IgM antibodies in plasma samples.** Electrochemical serological assays can detect SARS-CoV-2-specific IgA and IgM antibodies in plasma samples. A. Receiver operating characteristic (ROC) curve analysis of the patient sample data collected for the IgA and IgM SARS CoV-2 electrochemical assay using results from 54 RT-qPCR confirmed positive and 58 negative human plasma samples. B. Table listing numerical values of the ROC curve analysis. AUC: area under the curve; 95% conf. Int.: 95% confidence interval; Sens: sensitivity; Spec.: specificity; N pos.: number of RT-qPCR SARS-CoV-2 positive samples; N neg.: number of RT-qPCR SARS-CoV-2 negative clinical samples.

The clinical samples had a wide range of IgG levels that were shown to have excellent correlation with our assay (Fig. 4-14). Overall, the AUC for anti-SARS-CoV-2 IgG (Fig. 4-15B) was higher than IgM (Fig. 4-14B), whereas IgA's AUC was lowest (between 0.57 and 0.78) and did not add diagnostic value (Fig. 4-14B). The specificity of each individual sensor modified with either S1, N or RBD was over 95% for IgG. Further analysis of the multiplexed assay's ROC curves showed S1-RBD was the most accurate capture probe (AUC=1 for IgG), followed by N and S1 (Fig. 4-15B, 4-13).

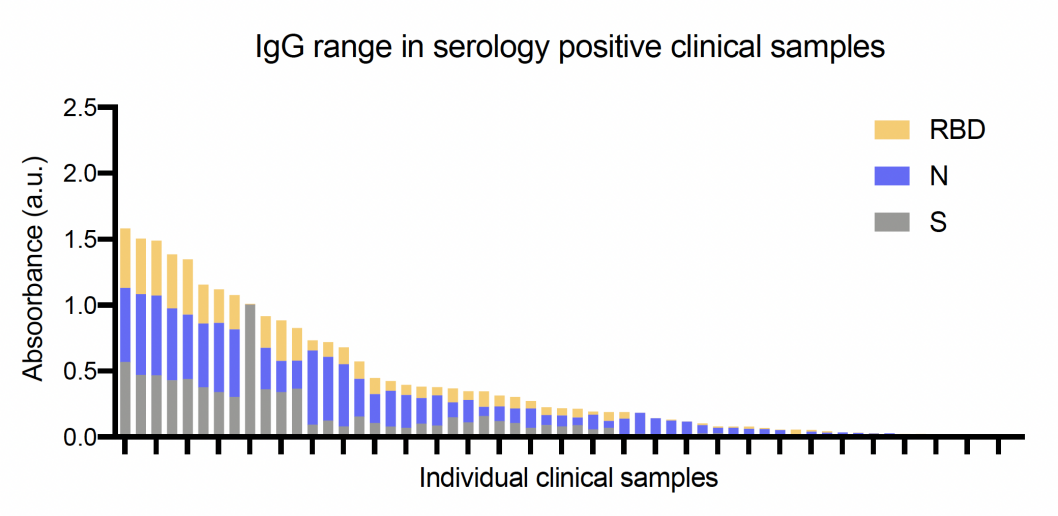
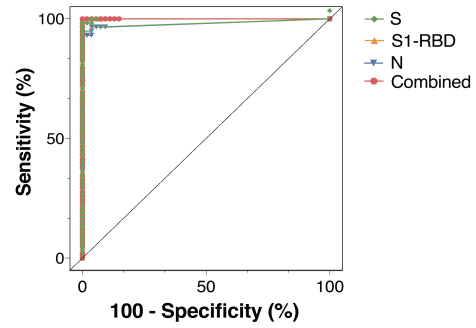



Figure 4-14: **IgG range for several SARS-CoV-2 antigens detected within 58 positive clinical plasma samples.** Levels of IgG directed against N (purple), S1 (grey) and RBD (yellow) detected by ELISA are plotted for each of the positive clinical plasma samples. The samples were collected were from convalescent patients at different stages post-disease and contained variable antibody titers, allowing us to validate that our multiplexed electrochemical sensors perform well across a range possible antibody levels.

In addition, we found that combining the results of the three antigens as a multiplexed readout resulted in a slight increase in accuracy when assessing a patient's immune status (AUC=1 for IgG) (Fig. 4-15). The combined EC IgG assay had an excellent correlation with prior SARS-CoV-2 infection and was shown to be 100% accurate (100% sensitivity and specificity); we therefore used this for subsequent experiments on the multiplexed EC sensor chips. We also found the EC sensor-based serology platform to be more accurate in detecting samples from patients with prior SARS-CoV-2 infection than the ELISA. The high specificity of the EC sensor platform may be attributed to the low non-specific binding on our nanocomposite-coated EC electrodes.

A




B

| | IgG | | | |
|---|---|---|---|---|
| | S | S1-RBD | N | Combined |
| AUC | 0.98 | 1.00 | 0.98 | 1.00 |
| 95% conf. int. | [0.95-1.00] | [1.00-1.00] | [0.95-1.00] | [1.00-1.00] |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Cutoff | >1.9e-7 | >1.7e-7 | >3.1e-8 | >2.4e-7 |
| Sens. | 0.95 [0.83-0.97] | 1.00 [0.94-1.00] | 0.95 [0.86-0.98] | 1.00 [0.94-1.00] |
| Spec. | 1 [0.93-1.00] | 1.00 [0.93-1.00] | 1 [0.93-1.00] | 1.00 [0.93-1.00] |
| N pos. | 54 | 54 | 54 | 54 |
| N neg. | 58 | 58 | 58 | 58 |

Figure 4-15: **ROC curve for IgG EC serology assay performance on patient samples.** A. ROC curves generated from the patient sample data obtained for the IgG EC serology assay.B. Table listing the numerical values of the sensitivity and specificity results. AUC: area under the curve; 95% conf. Int.: 95% confidence interval; Sens: sensitivity; Spec.: specificity; N pos.: number of gold standard SARS-CoV-2 positive samples; N neg.: number of gold standard SARS-CoV-2 negative clinical samples.

### 4.2.4 A Multiplexed viral RNA and antibody diagnostic in a 3D-printed LOC platform

We next worked to combine both EC sensors to create a multiplexed assay for simultaneous viral RNA and serological biomarker detection on-chip to facilitate increased sensitivity and specificity of SARS-292 CoV-2 detection [210]. Saliva is an excellent source of both viral RNA as well as host antibodies (IgG, IgM, IgA) in SARS-CoV-2 patients [202], and hence is an ideal sample for a multiplexed assay for viral RNA and serology. Unfortunately, the patient saliva samples used for this study had to be heat-inactivated before use as saliva from SARS-CoV-2 infected patients are of a highly contagious nature. Because high heat denatures the antibodies [98], we spiked a National Institute for Biological Standards and Control (NIBSC) SARS-CoV-2 IgG calibrant into heat-inactivated saliva samples at 1:20 dilution to simulate the ratio of IgG present in saliva. After confirming that the signal outputs from the spiked saliva samples were consistent with the signals generated from the plasma samples (Fig. 4-16), we modified the EC sensor chip so that the four electrodes could be used individually to detect the three antigens (S1, 301 N, S1-RBD) and PNA (Fig. 4-17).

To study the performance of our multiplexed EC sensor, we first conducted a two-step assay where antibody-spiked saliva was split into two volumes: 15µl was incubated on the chip for multiplexed serological detection of the host's anti-SARS-CoV-2 antibodies and 280µl was used for RNA extraction followed by the LAMP-CRISPR-based assay and incubation on the same chip. Following incubation of both samples on the electrodes, we simultaneously measured the SARS-CoV-2 viral RNA and host antibodies on-chip with the EC sensor readout of precipitating TMB following the addition of poly-HRP streptavidin. We validated the assay performance by testing the

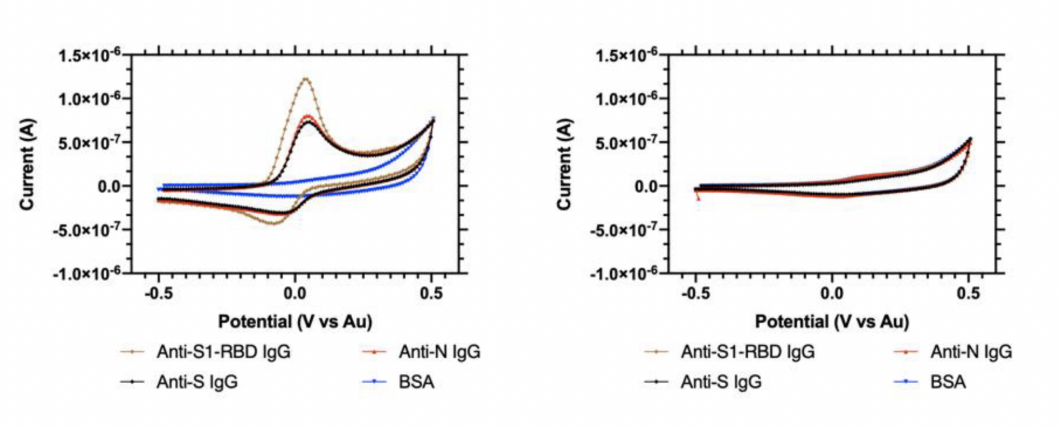



Figure 4-16: **Example of raw CV output for saliva sample.** An example of the raw CV output data when testing the serology EC sensor on saliva spiked with IgG calibrant.

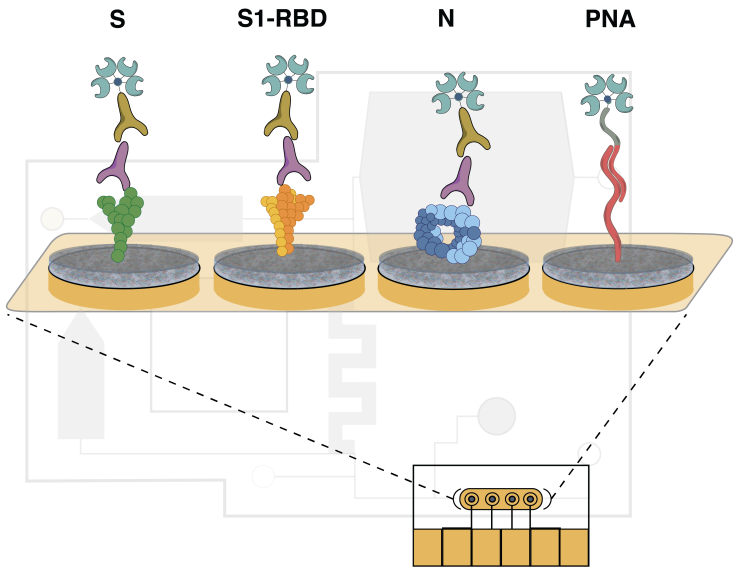



Figure 4-17: **Schematic of the multiplexed chip.** Schematic of the multiplexed chip surface conjugated with SARS-CoV-2 antigens: Spike (S1), S1-Receptor binding domain (S1-RBD), and Nucleocapsid (N); as well as peptide nucleic acid (PNA) for the detection of SARS-CoV-2 viral RNA on the LOC microfluidic system.

four possible combinations of serology and RNA-positive and RNA-negative clinical samples (Fig. 4-18). We achieved a very high probability (student's t-test $p<0.0001$) in distinguishing positive from negative samples with an ultra-low background signal in all the combinations tested. Taken together, these results showed excellent multiplexing capacity for SARS-CoV-2 viral RNA and host antibodies on the chips with

100% correlation in specificity and sensitivity.

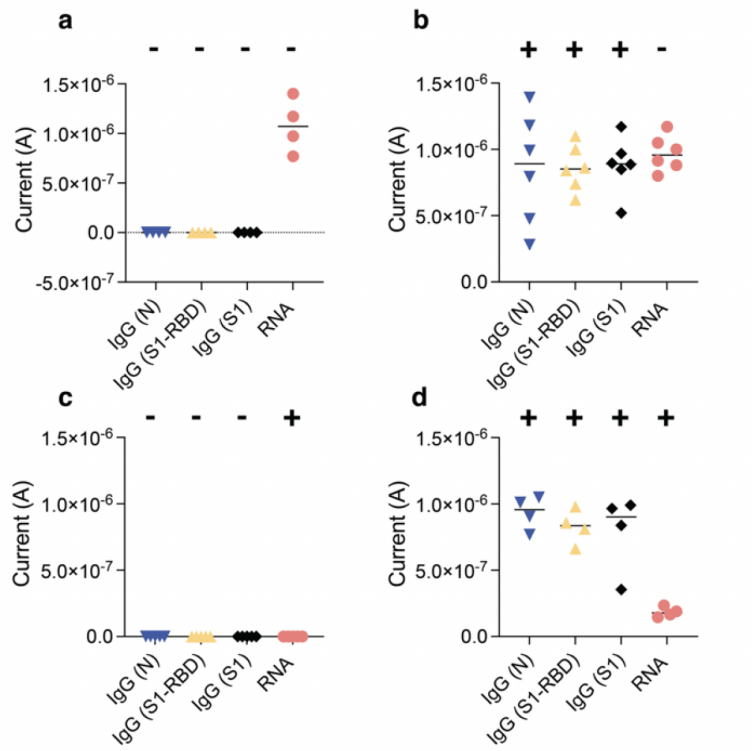



Figure 4-18: **Electrochemical platforms can be used for simultaneous detection of SARS-CoV-2 viral RNA and host antibodies against the virus.** Electrochemical readout for clinical samples that contain different host antibody and viral RNA combinations:A. Clinical samples negative for both serology and viral RNA. B. Clinical samples with negative host antibody levels and positive for viral RNA. C. Clinical samples that contain host antibodies against SARS-CoV-2 but are negative for viral RNA, and D. clinical samples with both positive host antibodies and viral RNA. The RNA signal significantly decreases (P value = 0.0006, Student's t test) when comparing IgG-positive and viral RNA-positive (d) and RNA negative samples (b and d).

Following these validation experiments, we integrated the multiplexed EC sensor chip into the compact LOC microfluidic platform to test the capabilities for automated viral RNA preparation prior to the EC sensor readout. Using the same functionalized chip assay as described above, each clinical saliva sample was split between the RNA and antibody reservoirs and the assay was performed as described in Fig. 4-2 and Supplementary Fig. B-3. Similar to the initial validation of the multiplexed assay, the assay performance of the LOC microfluidic system with integrated EC sensor chip was validated by testing the four possible combinations of serology and RNA-positive and RNA-negative clinical samples (Fig. 4-19).

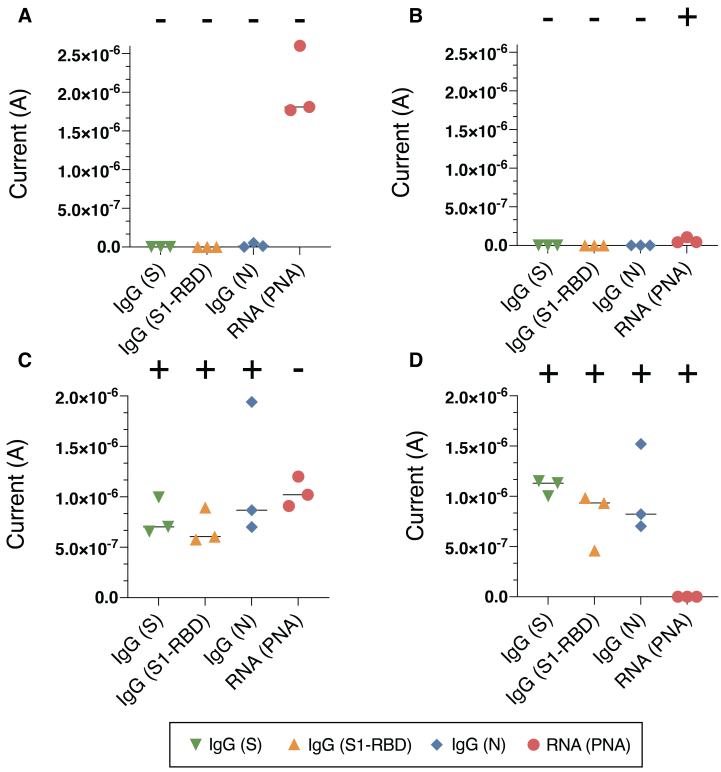



Figure 4-19: **EC sensor chip results using the LOC microfluidic platform.** Current (A) EC readout from the LOC microfluidic chip for clinical samples containing different host antibody and viral RNA combinations: A. Clinical samples negative for both serology and viral RNA. B. Clinical samples with negative host antibody levels and positive for viral RNA. C. Clinical samples that contain host antibodies against SARS-CoV-2 but are negative for viral RNA. D. Clinical samples with both positive host antibodies and viral RNA.

Clinical IgG negative samples showed no EC signal for the N, S1 and S1-RBD antigen-conjugated electrodes (Fig. 4-19A,B), whereas clinical samples from patients exposed to SARS-CoV-2 showed high IgG loads in all three antigen test areas (student's t-test for S1 and S1-RBD $p<0.0001$ and for N $p=0.0004$) (Fig. 4-19C,D). Moreover, we measured high signals on the PNA conjugated electrode for all clinical samples that were negative for SARS-CoV-2 viral RNA (Fig. 4-19A,C) and low or no currents on the PNA electrode for SARS-CoV-2 RT-qPCR RNA positive samples (student's t-test $p=0.0002$) (Fig. 4-19B,D).

Taken together, the results show that the LOC system can effectively prepare and deliver unprocessed saliva samples containing both SARS-CoV-2 viral RNA and host antibodies for on-chip simultaneous multiplexed detection with 100% correlation in specificity and sensitivity. A cost analysis of the LOC multiplexed diagnostic can be seen in Table B.4).

## 4.3 Discussion

Molecular diagnostics for detection of pathogen RNA and serological assays for assessment of host antibody responses are complementary tools that provide critical information to respond to disease outbreaks, assess vaccination status, and manage patient care and risks. While detection of SARS-CoV-2 RNA can be an important indicator of viral shedding during the infectious phase of the disease, factors such as sample types, virus variants, and infection severity [242, 239, 30, 61, 223, 79, 173] can impact diagnostic efficacy during the course of an infection and complicate patient management. Including serological assays can improve understanding of a patient's clinical status following viral infection and can also be used to assess a patient's response to vaccination. For example, currently available SARS-CoV-2 vaccines induce antibody production against SARS-CoV-2 S and S1-RBD proteins, and vaccinated individuals who have not been infected with SARS-CoV-2 are expected to develop measurable antibodies against the S, but not the N protein [53]. Therefore, multiplexed serology assays that target antibodies against several viral antigens can become key for seroprevalence studies to estimate the proportion of people in a population that have been infected, including asymptomatic infection, and/or immunized with vaccines. This information is key to estimate herd immunity and vaccine efficacy [78], which is critical for the decision to reopen economies [157, 49].

In the present study, we described a sample-in-answer-out diagnostic platform that integrates ultrasensitive and highly specific multiplexed EC sensors within an LOC microfluidic platform to rapidly and simultaneously detect both clinically-relevant quantities of SARS-CoV-2 viral RNA and antibodies within patient saliva samples. This platform incorporates multiple innovations including: (1) engineering of a high-specificity and high-sensitivity multiplexed EC sensor chip that enables detection of both proteins and nucleic acids with clinically relevant samples of biological fluids perfused through a single microfluidic channel; (2) development of a CRISPR-based detection assay optimized to function in a multiplexed EC assay in parallel with antibody-based detection assays that is amenable for POC applications; (3) development of methods for automated microfluidic extraction and amplification of RNA from raw patient saliva samples; and (4) fluidic integration of sample preparation process with the EC sensor chip.

While other CRISPR-based diagnostics (e.g., for SARS-CoV-2) have been described [28, 180, 11], they are typically limited to fluorescence and lateral flow readouts. Similarly, while there have been CRISPR-based EC sensors published [88, 240, 125], none have been sensitive enough to detect clinically relevant quantities of target nucleic acid. A key difference between our work and that of other recently published POC CRISPR-based electrochemical detection platforms is in our use of poly-HRP streptavidin/TMB-based reaction chemistry for readout, which enables further amplification of the EC signal for both the serological and CRISPR-based RNA sensors. Some EC-based sensors, such as a novel graphene FET-based EC biosensors for SARS-CoV-2, have displayed similar levels of sensitivity to the 0.8 ct/µl LOD shown here

[228], however those assays were only validated for VTM-based nasopharyngeal swab samples and more significantly, do not have multiplexing capabilities.

In this study, we also designed a customized molecular assay for nucleic acid detection that combines isothermal nucleic acid amplification with CRISPR-based enzymatic detection to target SARS-CoV-2 viral RNA in saliva. This was possible through the development and optimization of RNA-dependent Cas12a cleavage of a biotinylated ssDNA reporter, which interacts with the chip-based poly-HRP streptavidin/TMB-based reaction chemistry that allows for target detection more sensitively than fluorescence-based CRISPR-Cas12a assays, but with a similar time to result. The customizability of the EC chip's BSA/rGOx/GA-based antifouling surface chemistry allowed for a novel multiplexed sensing system that generates high conductivity and low nonspecific binding signals across a variety of biomolecular targets, aided by the sensitivity of the poly-HRP streptavidin/TMB reaction chemistry. Through optimization of reaction conditions, our EC assay has ultra-low background as well as improved sensitivity necessary for clinically relevant diagnostics that detect both nucleic acid and antibody targets [175, 213].

We validated the EC sensor platform for serology and obtained an accuracy of 100% (100% sensitivity and 100% specificity) for IgG as compared to traditional ELISA, with S1-RBD showing the highest accuracy in detecting IgG in clinical samples. This increased specificity for S1-RBD may be explained by the RBD domain being a highly immunogenic epitope for development of neutralizing antibodies to SARS-CoV-2 during the humoral response 47. In addition, we characterized a multiplexed EC sensor that detects antibodies against relevant viral structural proteins (S1-RBD, S1, and N) for a variety of antibody isotypes (IgG, IgM, and IgA), which enables a more robust understanding of the humoral response in patients. Importantly, we also found that simultaneous multiplexed detection of different viral antigens led to increased diagnostic sensitivity.

Finally, we designed and tested a novel microfluidic LOC platform that integrates with the multiplexed EC sensor for simultaneous detection of both RNA and IgG in clinical saliva samples. Saliva is an excellent alternative to nasopharyngeal swabs and nasal swabs for SARS-CoV-2 diagnosis, as it is simple to collect, does not require extensive collection equipment, and has been shown to provide both nucleic acid and serological data during and post infection or vaccination [98, 234, 122]. The microfluidic chip eliminates the need for RNA extraction kits by automating raw saliva sample preparation, RNA amplification, and CRISPR-based RNA detection steps. With its compact and sealed design, our LOC platform limits user steps to avoid possible sources of contamination or human-introduced error to allow for device use by untrained end-users and further increasing its potential as a POC testing system. With our integrated device, we are able to perform both serological and RNA detection from a saliva sample, which does not require specialized collection reagents or equipment, and the results are reported concurrently. To our knowledge, this is the first report of an EC diagnostic device that is multiplexed, highly sensitive, and

capable of processing raw biological samples such as saliva.

We tested our microfluidic LOC platform on four categories of SARS-CoV-2 clinical saliva samples: RNA negative and antibody negative; RNA positive and antibody positive; RNA negative and antibody positive; and RNA positive and antibody negative. Clear differentiation was seen between IgG negative and positive samples for N, S1 and S1-RBD antigen-conjugated electrodes (student's t-test for S1 and S1-RBD $p<0.0001$ and for N $p=0.0004$) as well as between SARS-CoV-2 viral RNA (student's t-test $p=0.0002$) within two hours after inputting the unprocessed saliva sample, highlighting the utility of our platform.

Limitations to our study include the small set of clinical COVID-19 saliva samples available due to the difficulty in acquiring saliva through biorepositories that do not routinely collect this sample type. However, the fact that we were able to obtain 100% accuracy with clinical samples that contained a wide range of viral loads strongly suggests that our CRISPR-based EC sensor platform could become a faster, simpler, and cheaper non-invasive strategy compared to RT-qPCR and traditional fluorescent diagnostics. Additionally, to assure that our multiplexed system was sensitive for clinically relevant levels of the biomarkers in question, we devised a system that would provide simultaneous qualitative results to generate a more robust and accurate estimation of a patient's infectious and immune status. Future work can be directed towards designing a multiplexed system capable of providing semi-quantitative or quantitative data on biomarkers of interest.

While our LOC platform is promising, there are bottlenecks that need to be addressed before this technology can be readily adopted for clinical POC settings. Currently, our system uses peristaltic pumping for fluid movement and a potentiostat for the readout of the EC sensor chips. Further integration of the electronics, peristaltic pumping, and potentiostat-based readout would allow for a robust reusable system that could make these multiplexed diagnostic useful for healthcare and clinical POC settings. Similarly, the 3-D printed microfluidic chips can be made at scale by transitioning to injection molding techniques, with the LAMP and the CRISPR detection reagents pre-measured, lyophilized, and sealed within the cartridge for streamlined assay use.

As the COVID-19 pandemic has shown, there is a critical need to rapidly adapt current testing strategies to more quickly and easily monitor both the infection and immune status of patients. Knowledge of infection stages can help curb disease spread, while insights on antibody titer levels can help with understanding how novel variants may affect individuals with immune protection through infection, vaccination, or a combination of the two. The streamlined workflow and multiplexing capabilities of our EC sensor represents important steps towards building the infrastructure necessary to provide this information to clinicians and members of the public alike.

## 4.4 Materials and Methods

### 4.4.1 Chip preparation

Gold chips were custom manufactured by Telic Company using a standard photolithography process with deposition of 15 nm of chromium and 100 nm of gold on a glass wafer. The area of electrodes was controlled by depositing a layer of 2 μm of insulating layer (SU-8). Prior to use, gold chips were cleaned by 5-min sonication in acetone (Sigma Aldrich, USA, no. 650501) followed by isopropanol (Sigma Aldrich, USA, no. W292907). To ensure a clean surface, the chips were then treated with oxygen plasma using a Zepto Diener plasma cleaner (Diener Electronics, Germany) at 0.5 mbar and 50% power for 2 min.

### 4.4.2 Nanocomposite preparation and activation

Nanocomposite coating was prepared using the previously described method43. Briefly, amine-functional reduced graphene oxide (Sigma Aldrich, USA, no. 805432) was dissolved in 5 mg/mL BSA (Sigma Aldrich, USA, no. 05470) in 10 mM PBS solution, pH 7.4 (Sigma Aldrich, USA, no. D8537), and ultrasonicated for 1 h using 1-s on/off cycles at 50% power. The solution was then denatured by heating at 105℃ for 5 min and centrifuged to remove the excess aggregates. The nanomaterial solution was then crosslinked by mixing with 70% glutaraldehyde (Sigma Aldrich, USA, no. G7776) at a ratio of 69:1, deposited on the glass chip with the gold electrodes and incubated in a humidity chamber for 20-24 h to form a conductive nanocomposite65. After nanocomposite deposition, gold chips were washed in PBS by agitation (500 rpm) for 10 min and dried with pressurized air. EDC (Thermo Fisher Scientific, USA, no. 22980) and NHS (Sigma Aldrich, USA, no. 130672) were dissolved in 50 mM MES buffer (pH 6.2) at 400 mM and 200 mM, respectively, and deposited on nanocomposite-covered gold chips for 30 min. After surface activation, chips were quickly rinsed with ultra-pure water and dried, and the capture probes were spotted on top of the working electrode area.

### 4.4.3 Clinical samples and ethics statement

De-identified clinical saliva samples from the Dominican Republic were obtained from Boca Biolistics under their ethical approvals. Because the saliva samples were collected from actively infectious patients, Saliva samples were heat-inactivated at 95℃ for 10 min to inactivate SARS-CoV-2 virions present in the samples, which denatured the antibodies present within the samples. RT-qPCR was performed by Boca Biolistics using the Perkin Elmer New Coronavirus Nucleic Acid Detection kit. De-identified clinical plasma samples were obtained from the Crimson Biomaterials Collection Core Facility at Partners Healthcare (currently Mass General Brigham). Additional de-identified clinical plasma and saliva samples were obtained through the

Massachusetts Consortium on Pathogen Readiness (MassCPR); these samples had been collected by Prof. Jonathan Li and Prof. Xu Yu. Additional pre-SARS-CoV-2 pandemic samples were obtained from the Walt Laboratory at Brigham and Women's Hospital. The Institutional Review Boards at the MGH, MGB, and Harvard University as well as the Harvard Committee on Microbiological Safety approved the use of the clinical samples in this study. All clinical plasma samples were collected from convalescent patients and were inactivated by heating at 65℃ for 30 min prior to use to denature any potential SARS-CoV-2 virions that might be present in the samples; this lower temperature denaturation method allows us to retain antigenicity of antibodies present within the samples.

### 4.4.4 CRISPR-based RNA assay with fluorescent reporter

CRISPR-based assays require the selection of both LAMP isothermal amplification primers and gRNAs to detect the LAMP amplicons. LAMP amplification primers (Table B.3) were selected after testing a range of LAMP primers, including some from the literature [175, 138, 186, 156, 237]. Cas12a gRNAs consist of two parts: the handle region (UAAUUUCUACUAAGUGUAGAU) that the Cas protein recognizes and binds, and a user-defined region at the 3' end of the handle that determines the specificity to the target. Spacer regions were selected following established guidelines [74]. To synthesize the gRNA (gRNA sequence: UAA UUU CUA CUA AGU GUA GAU GGU GAA ACA UUU GTC ACG CA), synthetic DNA with an upstream T7 promoter sequence (5' GAAATTAATACGACTCACTATAGGG 3') was purchased from Integrated DNA Technologies (IDT) and in vitro transcribed using the HiScribe T7 High Yield RNA Synthesis kit from New England Biolabs (NEB). Reactions were incubated for 16 h at 37℃, treated with DNase I (NEB), and purified using the RNA Clean & Concentrator-25 kit (ZymoResearch). gRNA was quantified (ng/μL) on a Nanodrop 2000 (Thermo Fisher Scientific).

Simulated SARS-CoV-2 samples were prepared by serially diluting full-length SARS-CoV-2 viral RNA (Twist Biosciences, MT106054.1) in nuclease-free water. Viral RNA extracted from saliva samples was used after purification via the QiAmp viral RNA extraction kit, as explained above. RNA was then amplified by LAMP and further detected by collateral cleavage of the fluorophore-quenched ssDNA reporter probe. Briefly, 5 μL of the diluted genomic DNA, or clinical sample RNA extract was added to 2.5 μL of the 10X primer mix (Table B.3), 12.5 μL of the LAMP master mix (NEB), and 5 μL of water. LAMP mixtures were incubated for 30 min at 65℃. After LAMP amplification, 4 μL of the amplified LAMP mixture were mixed with 11 μL of nuclease-free water and 5 μL of the CRISPR mixture, which contained 1 μM ssDNA fluorophore-quencher reporter (sequence: 6-FAM/TTATT/IABkFQ), 100 nM Cas, 200 nM gRNA in 10X NEB 2.1 buffer. Reactions were incubated at 37℃ for 20 min and fluorescence kinetics were measured using a BioTek NEO HTS plate reader (BioTek Instruments) with readings every 2 min (excitation: 485 nm; emission 528 nm).

### 4.4.5 Chip functionalization for the CRISPR-based RNA assay with electrochemical reporter

For CRISPR sensors, custom-synthesized amine-terminated peptide nucleic acid (AEEA -ACAACAACAACAACA) where AEEA is an O-linker was obtained from PNAbio, USA. PNA is a synthetic analog of DNA with a backbone utilizing repeating units of N-(2-aminoethyl) glycine linked through amide bonds. PNA contains the same four nucleotide bases as DNA (adenine, cytosine, guanine, and thymine) but are connected through methylene bridges and a carbonyl group to the central amine of a peptide backbone66. Stock PNA was diluted to 20 μM in 50mM MES buffer and spotted on the working electrode. One electrode was spotted with 1 mg/mL BSA as a negative control. The spotted chips were incubated overnight in a humidity chamber. After conjugation, chips were washed and quenched in 1 M ethanolamine dissolved in 10 mM PBS, pH 7.4 for 30 min and blocked with 1% BSA in 10 mM PBS containing 0.05% Tween® 20.

### 4.4.6 CRISPR-based RNA assay with electrochemical reporter

The reporter sequence for CRISPR-based EC assays was a ssDNA (sequence: /5Biosg/ AT TAT TAT TAT TAT TTG TTG TTG TTG TTG T) conjugated to a biotin that bound to poly-streptavidin-HRP. Upon Cas12a activation, the ssDNA-biotin reporter is cleaved in solution, thus preventing binding to the complementary PNA sequence on the surface. Poly-streptavidin-HRP is then added and able to bind to the ssDNA reporter-biotin. The concentration of HRP bound to the electrode was read by HRP-dependent oxidation of precipitating TMB (TMB enhanced one component, Sigma Aldrich, US, no. T9455). TMB precipitation forms an insulating, non-soluble layer on the electrode surface. Full-length genomic RNA (Twist Biosciences, MT106054.1) was serially diluted and amplified with 2X LAMP master mix (NEB) for 30 min at 65℃. Viral RNA was extracted from saliva via purification with the QiAmp viral RNA extraction kit. Similar to the protocol explained above, 5 μL of the viral RNA was added to 2.5 μL of the 10X primer mix (Table B.3), 12.5 μL of the LAMP master mix (NEB), and 5 μL of water. LAMP mixtures were incubated for 30 min at 65℃. After LAMP amplification, 4 μL of the amplified LAMP product was mixed with 10 μL of nuclease-free water and 5 μL of the CRISPR mix, which contained 4 nM reporter, 100 nM Cas, 200 nM gRNA in 10X NEB 2.1 buffer. Mixtures were incubated for 20 min at 37℃, during which time the ssDNA biotinylated reporter was cleaved. After that, 15 μL of the LAMP/reporter/Cas mixtures was deposited on the chips for 5 min. Thereafter, the chips were washed and incubated with poly-HRP streptavidin and TMB for 5 min and 1 min, respectively. Final measurement was then performed in PBST using a potentiostat (Autolab PGSTAT128N, Metrohm; VSP, Bio-Logic) by a CV scan with 1 V/s scan rate between -0.5 and 0.5 V vs on-chip integrated gold quasi reference electrode. Peak oxidation current was calculated using Nova

1.11 software. Cyclic voltammetry allowed us to measure attomolar concentrations of SARS-CoV-2 target RNA.

### 4.4.7 Serology ELISA assay

ELISA assays were optimized in a 96-well plate format. 100 µL of 1 µg/mL antigens: Spike S1 (SinoBiological, China, no. 40591-V08H), Nucleocapsid (RayBiotech, US, no. 130-10760) and Spike (S1) RBD (The Native Antigen Company, UK, no. REC31849) were prepared in a 10 mM PBS buffer at pH 7.4 and added to Nunc™ MaxiSorp™ELISA plates (BioLegend, no. 423501) and immobilized on the plates by overnight incubation at 4℃. The plates were washed three times with 200 µL of PBST followed by the addition of 250 µL of 5% Blotto for 1 h. After washing the plates, 100 µL of the clinical plasma samples diluted in 2.5% Blotto was added and incubated for 1h at RT. Plates were further washed and HRP conjugated anti-human IgA/IgM or biotin-conjugated anti-human IgG detection antibodies were added for 1h. The secondary antibodies used were: HRP conjugated anti-human IgM (Human IgM mu chain rabbit Antibody, Rockland, us, no. 109-4107) or IgA (AffiniPure Goat Anti-Human Serum IgA, $\alpha$ chain specific, Jackson ImmunoResearch, US, no. 109-005-011) or biotin-anti-human IgG (AffiniPure Goat Anti-Rabbit IgG, Fc fragment specific, Jackson ImmunoResearch, US, no. 111-005-008). The IgG plate was further mixed with 100 µL of Streptavidin-HRP (1:200 dilution in 2.5 % Blotto) and washed. 100 µL of turbo TMB (Thermo Scientific, no. 34022) was added for 20 min followed by the addition of 100 µL of sulfuric acid (0.1M $H_2SO_4$ in Water) to stop the reaction. The absorbance of the plates was immediately read using a microplate reader (BioTek NEO HTS plate reader, BioTek Instruments) at 450 nm.

### 4.4.8 Electrochemical serology assay

To translate the ELISA assays to EC readouts, Spike S1, Nucleocapsid and Spike-RBD were diluted to 1 mg/mL in the PBS buffer and spotted in three electrodes of the EC chip. An additional electrode was spotted with 1 mg/mL BSA as a negative control. The spotted chips were incubated overnight in a humidity chamber. After conjugation, chips were washed and quenched in 1 M ethanolamine (Sigma Aldrich, US, no.E9508) dissolved in 10 mM PBS, pH 7.4 for 30 min and blocked with 5% Blotto (Santa Cruz Biotechnology, US, no. sc-2324) in 10 mM PBS containing 0.05% Tween® 20 (Sigma Aldrich, US, no. P9416). The fabricated sensor was then used to detect immunoglobulins from clinical samples. Each sensor was either used to detect IgG, IgM, or IgA against the three antigens that were immobilized on the chips. 1.5 µL of clinical plasma samples were mixed with 13.5 µL of 2.5% Blotto and incubated on the chips for 30 min at RT followed by a rinsing step. After that, HRP conjugated anti-human IgM /IgA/ biotin-anti-human IgG was added for 30 min at RT. 1 µg/mL of poly-HRP-streptavidin (ThermoScientific, US, no. N200) diluted in 0.1% BSA in PBST was added to chips with IgG for 5 min. The chips were rinsed and

precipitating TMB was added for 3 min followed by final rinse and EC measurement using a potentiostat by cyclic voltammograms with a scan rate of 1 V/s between -0.5 and 0.5 V vs on-chip integrated gold quasi reference electrode. Additional antibodies were screened, including: $F(ab')_2$ Goat anti-Human IgG-Fc Fragment Antibody Biotinylated (Bethyl Laboratories, no. A80-148B), Goat anti-Human IgG Fc Secondary Antibody, Biotin (ThermoFisher Scientific, US, no. A18821), Purified anti-human IgG Fc Antibody (BioLegend, no. 409302), Purified anti-human IgG Fc Antibody (BioLegend, no. 410701), and AffiniPure $F(ab')_2$ Fragment Goat Anti-Human IgG, Fc$\gamma$ fragment specific-biotin (Jackson ImmunoResearch, US, no. 109-006-170).

### 4.4.9 Multiplexed electrochemical serology and CRISPR-based RNA detection

Multiplexed sensors for both nucleic acid and host antibody detection were prepared by spotting three electrodes of the EC chip with proteins: Spike S1, Nucleocapsid and Spike-RBD; and spotting the amine-terminated peptide nucleic acid (AEEA-ACAACAACAACAACA) reporter on the fourth electrode. The spotted chips were incubated overnight in a humidity chamber. After conjugation, chips were washed and quenched in 1 M ethanolamine (Sigma Aldrich, US, no.E9508) dissolved in 10 mM PBS, pH 7.4 for 30 min and blocked with 1% BSA in 10 mM PBS containing 0.05% Tween® 20 (Sigma Aldrich, US, no. P9416). The fabricated sensor was then used to detect IgG's as well as viral RNA from clinical samples.

Multiplexed chips were used to detect viral RNA as well as IgG against the three antigens that were immobilized on the EC sensor chips. Negative control saliva (RT-qPCR negative) was heat-inactivated and spiked with plasma at a ratio of 1:20 to simulate IgG concentrations in saliva. Two experiments were done in parallel on each chip, as follows: (1) 15 µl of plasma-spiked saliva was used for the serology assays as explained above. Briefly, 0.75 µl of the plasma sample was mixed with 14.25 µl of control saliva and incubated on the chips for 30 min at RT followed by a rinsing step. After that, biotin-anti-human IgG was added for 30 min at RT. Chips were then rinsed. (2) In parallel, RNA extracted from RT-qPCR positive and negative clinical samples was amplified by LAMP for 30 min at 65℃ as explained above. Then, 4 µL of the amplified LAMP product was mixed the CRISPR mix, which contained the biotinylated reporter and incubated for 20 min at 37℃. 15 µL of the LAMP/reporter/Cas mixtures was deposited on the chips after the chips had been exposed to SARS-CoV-2 IgG. Thereafter, the chips were washed and incubated with poly-HRP streptavidin and TMB for 1 min. Final measurement was then performed in PBST using a potentiostat (Autolab PGSTAT128N, Metrohm; VSP, Bio-Logic) by a CV scan with 1 V/s scan rate between -0.5 and 0.5 V vs on-chip integrated gold quasi reference electrode. Peak oxidation current was calculated using Nova 1.11 software. Cyclic voltammetry allowed us to measure both the presence of IgG antibodies as well as attomolar concentrations of SARS-CoV-2 target RNA.

### 4.4.10 Microfluidic chip for multiplexed electrochemical serology and CRISPR-based RNA detection

A microfluidic chip for a point-of-care diagnostic that integrated the multiplexed chip was designed using Autocad software and printed on a Formlabs Form 3B 3D SLA printer in grey resin, Version 4 (Formlabs RS-F2-GPGR-04). The layer thickness chosen for the 3-D printing parameters is 50 μm. The chip is cleaned in an isopropanol bath for 20 min, then cured in a heated UV chamber for 1 h at 60℃ (Form wash and Form cure, Formlabs). The chip was designed with chambers for saliva, LAMP reagents, and CRISPR reagents, as well as a reaction chamber lined with a PES membrane and a waste port. The PES membrane (Millipore, catalog no. GPWP04700) was laser cut using Epilog Fusion edge 24 laser cutter to fit within the serpentine reaction chamber. The chambers were designed to be connected to the reaction chamber through channels that are closed when under vacuum. Additionally, the chip has a channel and waste port which allow for the integration of the EC sensor chip. The chip was sealed using a clear delayed tack adhesive tape (3M 9795R) to prevent evaporation. Two high-power resistors were used as heating elements for the sample preparation camber and reaction chamber (Digikey 355018RJT) and placed in a custom-printed backing to assure accurate and repeatable heading was provided. A full list of parts along with cost analysis can be found in Table B.4. The voltages used for each temperature are provided in Supplementary Fig. S5 and were generated using an Agilent E3631A DC power supply. A complete view of the microfluidic chip can be seen in 4-2 and B-3.

A small DC pump (Takasago RP-Q series) connected to an Arduino Uno was used for the pumping. The workflow for the multiplexed EC assay on the microfluidic chip is as follows: (1) The microfluidic chip was sealed with clear delayed tack adhesive tape. (2) Saliva from patient samples (BioIVT) was inactivated at 95℃ for 5 min and mixed with 10% by volume of Proteinase K (NEB P8107S) that was diluted 1:10 in nuclease free water. (3) 450 μL of the saliva mixture was added into the sample preparation chamber and the chamber was heated to 55℃ for 15 min, followed by 95℃ for 5 min. (4) The sample was pumped over the PES membrane within the reaction chamber at a speed of 50-100μL/min. (5) The reaction chamber was heated to 95℃ for 3 min. (6) 25μL of LAMP reaction (12.5 μL of NEB 2X WarmStart Mastermix, 10 μL of nuclease-free water, and 2.5 μL 10x of LAMP primer mix) was pumped into the reaction chamber and incubated at 65℃ for 30 min. (7) The heat was decreased to 37℃ and 22 μL of the Cas mix (15 μL nuclease-free water and 7 μL of the CRISPR mix, which contained 4 nM reporter, 100 nM Cas, 200 nM gRNA in 10X NEB 2.1 buffer) was pumped into the reaction chamber and incubated for 30 min. (8) The LAMP/CRISPR reaction was pumped over the EC sensor chip at 4 μL/min for 5 min followed by 20 μL of PBST at 15 μL/min. (9) 20 μL of saliva sample spiked with NIBSC antibody calibrant mixed with anti-IgG detection antibody linked with

biotin was flowed through the EC sensor at 4 μL/min for 5 min. (10) The EC sensor chip was washed with 20 μL of PBST at 15 μL/min followed by addition of 20 μL streptavidin-poly HRP at 6 μL/min. (11) The EC sensor chip was washed with 20 μL of PBST at 15 μL/min and 20 μL of TMB at 8 μL/min. (12) The EC sensor chip was washed with 30 μL of PBST at 15 μL/min. (13) The EC sensor chip was read on a potentiostat and the cyclic voltammogram was generated.

### 4.4.11 Data analysis

Fluorescence values are reported as absolute values for all experiments used for CRISPR-based fluorescence assays. Absorbances for the ELISA assays are reported as background-subtracted values to normalize for plate-to-plate variability. Peak oxidation current for EC CRISPR and serology assays was calculated using Nova 1.11 software. All data were plotted and statistical tests were performed using GraphPad Prism 9. Gold standards for ROC curve analysis: the individual samples for both serology and viral RNA detection were validated using SARS-CoV-2 RT-qPCR. For molecular assays, SARS-CoV-2 positive saliva samples were RT-qPCR positive at the time of saliva collection. For serology assays, SARS-CoV-2 positive plasma samples were RT-qPCR positive at the time of plasma or at an earlier date. Receiver operating characteristic (ROC) curves were used to evaluate the performance of diagnostic assays as a function of the discrimination threshold, plotted as sensitivity (%) versus 100 specificity (%). The areas under the ROC curve (AUC) are a proxy of test performance, where 1 represents a perfect test and 0.5 represents a random predictor. ROC curve analysis was done in GraphPad Prism 9 using a 95% confidence interval and the Wilson/Brown method. Figures were created using Prism 9 and Adobe Illustrator.

**Author contributions:** D.N., J.R., S.S.T., P.J., H.d.P., M.Y., N.D. and H.S. conceived the study under the guidance of J.A.P., P.E., J.J.C. and D.E.I. Experiments were performed and validated by D.N., J.R., S.S.T., P.J. and H.d.P. H.d.P. and P.J. formulated the idea, organized the experiments and managed the project. M.Y., N.D., D.R.W., G.A., J.Z.L. and X.G.Y. contributed to the collection and characterization of clinical serum and saliva samples. All authors contributed to manuscript preparation and editing.

# Chapter 5

# Sensors in the Field: How Data Figures into Local Monitoring

## 5.1 Introduction

In recent years, researcher have published much work in the field of biology and bio-engineering that claims to be intended for "low-resource settings," though much of this work is never actually field-tested in the settings that they claim to be built for. My work is no exception, with much of my biological research concerning around biotechnology built for a lab-free environment and a non-scientist user but containing no user testing by non-scientist users. While the primary reason why my work was not more thoroughly validated due to a global pandemic and consequently the time-sensitive nature of the work that was being developed, it left me feeling part of the problem. I want biosensors that are validated in the resource-limited settings that are discussed in these papers to confirm the claims of robustness, ease of building, ease if use, and ease of interpretation. But the truth is, most researchers do not want to use systems designed for non-laboratory equipment, they want to design low-cost technology in papers while providing data that primarily comes from from high-cost specialized instruments. This is because data, and data quality, is typically of the utmost priority (and of primary concern when meant to be performed on non-proprietary instruments). Unsurprisingly, it is the output that people care most about. Data, what it is, how much of it we have, how much of it we want, and what we can do with it, is the topic of this section. Within environmental monitoring in particular, we are building more and more tools that provide data at a variety of different resolutions and are attempting to (somehow) find ways to make sense of it all.

The focus on data seems to have reached a fever pitch in the past two decades. It has pushed us into the realm of "Big Data," which Kitchin and McArdle describe as being a fairly amorphous category, albeit one that should have data sets which possess the three V's of: volume, velocity, and variety (although that is not always the case) [115]. The Big Data boom can be attributed to computing and storage, both of which have become increasingly cheap, fast, and ubiquitous. This desire for

data has certainly made its way to environmental monitoring, with long-term sensing research projects like the Center for Embedded Networked Sensing (CENS) [60] and National Ecological Observatory Network (NEON) [111] collecting longitudinal data sets meant for understanding our surroundings, countless environmental monitoring-related citizen science projects designed to aggregate large quantities of data through grassroots methods [40, 148], satellite networks that aim to image the entire planet every day, and governments that work internet of things-powered devices into urban planning projects. All of this builds into a planet full of data (hoarded on server farms at a significant environmental cost [55]), more data than anyone knows what to do with, in the hopes that it can someday be useful.

This is not to say that the aggregated data is not useful, because more and more machine learning has attempted to massage and process these large data sets without human oversight in an attempt to discover patterns that can aid in decision making, predictive modeling, and management. However, peeling back the veneer of utility, it is clear that the underlying desire to amass larger and larger data sets likely goes a bit deeper than utility. In her essay *The secret life of big data*, Genevieve Bell asks:

> *"What is going on right now that makes us say we need data, more and more and more of it? What is the anxiety, the fear, the instability, the place that the world has moved in such a way that the thing that we think will be comforting, although we would never use that language, is more data. Why is this so comforting and so seductive, or, why is it that this is the moment in which more data seduces?"* [19]

I think the answer, in addition to the usual desire for technological and societal advancements that allow for these data streams and the demonstrations thereof, as well as the promise that this information may be useful at some point in some way, is a heightened sense of urgency due to an increasingly globalized world in crisis. At present, I am writing this dissertation during the third year of a global pandemic as temperatures have reached record-breaking heights in many parts of the world. There are forest fires raging from California to Spain, events that are somehow a bit anti-climactic now that they have been happening more and more frequently in the past few years. There are already climate refugees and those who are at risk of dying due to heat, draught, storm, and poor public health infrastructures. The species on our planet – from the microscopic plankton to large land mammals – are going extinct, some before we even have the chance to catalog them (the ethics of our desire to catalogue things being a discussion for another time).

I believe that the fields of biology and environmental studies are searching for more and more data in a hope that it will allow them to better decode the world around us as quickly as possible. Time is of the essence, this is a moment for quantity over quality, if one had to choose. Researchers are rushing to collect as much data as

possible to feed into computational model hungry for data in the hopes that a generalizable trend can be established and a plan devised. As Prof. Jennifer Gabrys notes in Practicing, materializing and contesting environmental data, "...the state of the planet in crisis is often met with the response to gather more data." [70].

### 5.1.1 Political ecology and environmental data

The question of data and its utility can be explored in countless ways, however for the purposes of this chapter I will be exploring it through the lens of political ecology, which works to understand our environment within the context of the political and economic structures that our environment has become embedded in. Political ecology takes a high level view of the systems that overlay our environment – from the local to the global – and underlie the decisions that are made with the data collected from those systems. As digital technologies make it possible to gather data sets previously considered unfathomable and unprocessable, the question becomes: what *do* we do with all of it? Is it necessary? Is it useful? Is it helping?

In "A political ecology of data," Eric Nost and Jenny Goldstein consider these questions, and address the governance, materialization, and practice of environmental data collection as well as the development of infrastructure that takes the above into consideration when developing data or data analysis protocols for future systems [161]. They state that "Part of the problem with 'data technologies will help us save the planet' discourses is that they rarely consider how data will actually be developed, deployed, maintained, or otherwise circulated." Data is not neutral, not free of context, and can be dangerous when used wrong (or even lazily). Data collection is therefore a process that needs to be considered in the environmental observation-to-action pipeline.

Data collection is critical to the process of assembling data infrastructures that provide the *right* kinds of information, which is not necessarily predicated on volume or sensitivity or locality, to assist with decision making at all levels, from the personal to the governmental. Sentiments such as these are what drove me to more closely study how the biosensors I develop could work within the local monitoring framework I claim they are best suited for. In particular, this work focuses on the role that materiality and context play in local environmental monitoring and citizen science, and how the resulting data may be made useful within the large deluge of environmental data.

### 5.1.2 Considering Big Data and datadiversity

Designing data infrastructures is not only an issue in political ecology, because the datafication of most every part of life has already been underway for decades. Datafication is a topic shared between city planners, technologists, and social scientists,

who question the relationship between data and cities, working to understand if, and how, data can bring about utility, equity, democracy, governance, and accessibility within civic processes [18, 179, 43, 24, 92]. One of the reason that datafication and data infrastructures is most relevant to the field of environmental sciences is due to the aggregation of data sets under the roof of "Big Data," which has become one of the most recent tools that public and private actors use to try and get a handle on the changing landscapes (literally). However, as Geoffery Bowker notes in Biodiversity Datadiversity, though archiving data is important as a way of contributing to science and future research inasmuch as "the working archive is a management tool," these archives also need a level of historical and ontological complexity that is commensurate with the worlds that it is mapping [26].

Bowker focuses on the state of biological and environmental data, which happen to also be the focus of this dissertation. Within the larger conception of biodiversity, the realms of biology and environmental studies are distinct fields of study, albeit ones that overlap. Both are quite broad, with biology encompassing all kingdoms of life – microbes to megafauna, and environmental studies lying at the intersection of the natural and built world, including all of the abiotic and biotic relations that continue to build and rebuild them. They are both interested in data sets that are longitudinal and implement highly standardized protocols and precise instruments (oftentimes the same instrumentation and protocols for chemical and biological studies). They are both interested in gathering data not only on their object of interest, but also on the conditions that the object of interest exists within and how that information may add dimensionality to the overall study. They are both also dealing with difficulties in reproducing results, often because the conditions within biology and environmental sciences are difficult, if not completely impossible, to replicate exactly [14, 149]. But more importantly, Bowker notes that both fields have undergone a transformation in an era that heralds collection and analysis of large-scale data science, whereby much of the work published in these fields are focused more on gathering data not just as a means to an end, but as the end itself. Data has become the scientific product, rather than the raw material upon which to produce a paper or theory within these fields.

This sentiment is emphasized by Sabina Leonelli, who further probes this desire for researchers to pursue data as a means to an end in the era of "Big Data." Leonelli describes the "data journey" that goes into turning a data set into one that is part of the world of Big Data, which involves three steps: (1) data de-contextualization, where the data is integrated into the larger database, (2) re-contextualization, where the decontextualized data is given a new life by being used in a different context from the initial experiments that collected the data, and finally (3) reuse, where the recontextualized data is used for novel discoveries that would be otherwise impossible if the data set had not been assembled [124]. This conceptual framework is in agreement with Leonelli's previous claims that data are not necessarily "local", or embedded in their context, because of the ability of researchers to "free" data sets collected in a different research context and use it for their own research, as long as the data set is packaged in a way that provides a way for researchers to gauge the relevance and

reliability [123].

Leonelli not only highlights the difficulties present in the data journey, of which there are many specific to Big Data in the context of biology, such as successfully standardizing and integrating the protocols and methods of data collection and metadata needed to properly de-contextualize and re-contextualize the data, but further notes that the pursuit of large-scale data sets and Big Data as a means to an end can ultimately skew the kind of data collection and context-specific research that is most useful to fields for drawing unbiased results and correlations that can allow for these scientific data sets to be useful in the first place. In their book *Big Data: A Revolution That Will Transform How we Live, Work, and Think*, Viktor Mayer-Schonberger and Kenneth Cukier point out that in the field of Big Data, "...we'll often be satisfied with a sense of general direction rather than knowing a phenomenon down to the inch, the penny, the atom." [144]. The scaling up of data creates an environment where correlations drawn from Big Data are not necessarily capable of fine-grain detail, but only a general idea of things. More often than not, researchers in the field of biology will not take the generalized correlations found in the re-contextualization and re-use process as publishable results, but as the basis for the development of a new set of experiments that will collect yet another data set, which will be compared with the results seen in the correlation from the Big Data set.

Indeed, this is a common critique of Big Data found in Critical Data Studies. As Dalton and Thatcher state, "...a reliance on volume to justify import plays into the mythologies of 'Big Data' wherein larger is always more clear, more comprehensive, and generally better, furthering a seductive pseudopositivist research orientation in which 'raw' data becomes epistemic reality." [45]. Which is to say, not only is it a potentially useful but logistically difficult and energy consuming process to turn data into a product that can be used within a larger data infrastructure, but it is also not necessarily something that is going to end up even being useful for the task at hand. In fact, for biological and environmental data sets, just properly contextualizing a data set and providing the requisite metadata is often as useful, if not more useful, than collecting a much larger data set which is not properly logged or organized.

This brings us back to the question of data collection, contextualization, and infrastructure. In a world where more is not better, which is to say that the size of data and the utility of data are not linked, designing experiments or a protocol should be most focused on answering the research question at hand as best as it can, rather than prioritizing that the data generated be easily integrated into the Big Data network. But how does one try and design a system of data collection that is at the appropriate scale and resolution?

### 5.1.3 Multi-modal local sensing and "just-good-enough" data

To begin answering the questions of data collection, which includes the scale and resolution of that data, requires scoping. It is a particularly difficult endeavor in the age of endless data streams to work on methods that provide a more sparse data set, or sets of interrelated more sparse data sets, albeit one(s) that may more easily or effectively answer a question. There are a few parts to this equation: (1) the method(s) used to generate the data (2) the method(s) used to process the data (3) the data product and its format.

The area of focus for this research is in developing new tools that can be used for biological sensing, in specific sensing nucleic acids, ideally in a way that can be amenable for use outside of a highly instrumented and staffed laboratory. Though the fields of molecular and synthetic biology have developed highly sensitive protocols that can allow for truly unprecedented levels of resolution on the invisible microscopic worlds that exist alongside and within us, these protocols operate on clean (or cleaned up) samples that are run on calibrated and standardized pieces of technology that ultimately generate computationally-intensive data streams. Which is to say, biological sensing is, as of yet, fairly unprepared for the field. (The COVID-19 pandemic has only highlighted the issues of biology remaining centralized in laboratories, with the need for testing and analysis of viral samples almost immediately overwhelming the system).

Biological organisms, and microbes in particular, are the foundation of our biosphere [200]. These microbes are key nutrient cyclers and chemical regulators, as well as biological sensors in and of themselves. But more broadly, there are a range of micro-to-macro bioindicator species that scientists have incorporated into their study of environments [76]. Just focusing on a bioindicator alone can provide for a holistic view of the chemical and physical environmental changes that are occurring in an area. Linking these together highlights the complicated and often sensitive relationship between organisms and their surroundings. These relationships have the ability to teach us not only about how reliant living things (like humans) are on other living things, but just how critical those other living things are in making sure that ecosystems continue to function.

Similar to environmental sensing, biological sensing often requires the participation, indirectly or directly, in multi-species encounters that shift perspectives and ways of relating and being in the world. Importantly, it allows for encounters that allow humans to question the roles we play within our ecosystems; it is one of opportunities where we have the chance to reach across the aisle of the kingdoms of life. Therefore, designing tools that can be used for the direct identification of these microorganisms in-situ can, and should, be able to collect data streams that are as complex as the relationships being established through the sensing.

A multi-modal approach to sensing, where information from multiple overlapping

data streams are combined, is not new within the fields of environmental sensing or biology. As fields of inquiry, both contain research areas that focus on the merging of people and place, and the things that govern those relationships. To get at these relationships, there are a range of research methods that can be used to incorporate the scientific with the social. These research areas may touch on biology, chemistry, physics, and geology, as well as history, archaeology, sociology, and politics. Multi-modal sensing allows for the consideration of the human and more-than-human interactions and experiences that can contribute to ways of knowing and understanding a place over time, and how that colors the sensing data that is collected.

These thoughts can be summed up as a desire and need for *context* to be incorporated into environmental data generation methods, or as Geneveive Bell writes, "data has a country." [19]. Nost and Goldstein further emphasize that *"context—inclusive of history, place, and scale—matters, in contrast to the universality and novelty often associated with data technologies."* [161]. But context in the larger, non-Leonelli sense, which is to say, not just providing a metadata trail that can allow for the use and re-use of the data, but multiple streams of data that can overlay a data set and provide critical information which helps to understand and untangle relationships. These streams of multi-modal data are difficult, if not completely impossible, to capture through the deployment of autonomous sampling systems which cannot provide data beyond the qualitative or quantitative measurements it is programmed to capture.

Two areas of research that have collected multi-modal data streams are the fields of citizen science and participatory monitoring. Citizen science is the crowdsourcing of data, typically relating to the natural world, by the general public. It is often done in collaboration with a research scientist or research group, who generate tools, protocols, and data collection methods that support the large-scale effort [137]. Participatory monitoring is another method that engages locals as stakeholders in a process to gather data that can be used to improve management and practices within an area [59]. Though locals should to have more oversight in setting the research agenda and methodology in a participatory monitoring process, there is also the spectre of other stakeholders, be it governments, NGOs, or funders, that loom and can impact the program and the type of data and monitoring in performs.

Having samplers who have lived experiences with the environment being monitored allows for the addition of their own observations, often informal or qualitative, which can be used as metadata to further situate the data collected by these samplers. Clifford Geertz bring up Heidigger's question, "who knows the river better, the hydrologist or the swimmer?" [75] He refers to the swimmer's knowledge as "the local variety," and emphasizes that it certainly holds its own against that of the knowledge that the hydrogeologist has (or claims will develop). Citizen Science and participatory monitoring is not only collecting data from the point of view of the swimmer, it is also, at times, giving those swimmers the instruments and methods of the hydrologist and having them use it towards their goals.

Researchers who engage in participatory monitoring or citizen science programs can sometimes emphasize how these methodologies allow for the creation of "data stories" [71] and "data narratives" [54], ways of acknowledging that data alone does not have the ability to interpret itself. Data often needs to be in proximity to other data in order for meaning making to happen. Emphasizing an ethnographic approach to data gathering not only allows for easier re-contextualization and re-use of the data, it also allows data that might be otherwise considered too low quality to become useful. The latter type of data collection has been termed "just good enough" data, whereby "...citizen gathered data is often 'just good enough' to establish patterns of evidence that can mobilize community responses in terms of communicating with regulators, requesting follow-up monitoring, making the case for improved regulation and industry accountability, and keeping track of exposures both on an individual and collective level." [71]). "Just good enough" data merges the data set with the "data setting" [136], making the collected information compelling for further study. Opening up monitoring and sensing to data streams that take advantage of lower cost and possibly lower quality ad-hoc tools for collection purposes opens up the accessibility of environmental monitoring to communities who otherwise may feel the barrier of entry is too high.

However the benefits of multi-modal local sensing go beyond just a more highly-contextualized data set, exposing how "different modes of sensing and experience provoke not just different ways of being in and for worlds" as well as "that different worlds and environmental relations are at stake" [69]). These "different ways of being in and for the world" [70] that stem from local participatory data makes room for more generous, creative, and expansive solutions. It also highlights the ways in which changing environments and social relations are always intertwined and often co-produced [199] Not only does participatory sensing allow for richer data sets, it also has real and obvious impacts in how those local monitors think about and act in their surrounding environment on an individual and collective basis.

### 5.1.4 Decolonized research practices

Often with citizen science and participatory-led monitoring programs, there is something that is changing or broken within the community that galvanizes them to study it deeper. The changes being monitored are ones that can impact, or threaten to impact, the community's way of being physically and culturally. Taking a local an participatory approach to sensing, one that allows for the opportunity for community members to set the research agenda through the lens of their lived experiences and historical connection, is (hopefully) a move away from parachute science, where a scientist drops into an unfamiliar community and sets a research agenda without listening to or engaging with locals. Participation in local monitoring programs offers the opportunity for people to decolonize their conception of what science can be and engage in an ecological intimacy that can, and should, be a call to bring all

parts of oneself to the process. This is echoed in the perspective of Robin Wall Kimmerer, who notes that "ecological restoration is inseparable from cultural and spiritual restoration, and is inseparable from the spiritual responsibilities of care-giving and world-renewal." [113].

A decolonized research methodology emphasizes bringing all parts of yourself to a process – body, mind, spirit, emotion [113]. Many who participate in citizen science and participatory monitoring are already equipped with tools that can help them with their work, because they are the ones who have been swimming in the river (literally and metaphorically). Monitoring processes have the opportunity to physically heal the natural landscape as well as the opportunity to heal much deeper relationships between locals and the land, which can span generations and centuries. Within indigenous traditions, there is no clear separation between humans and the rest of the natural world. We are all one large organism, bound by a culture of reciprocity. Reciprocity – between researchers, communities and the areas they are monitoring, is critical to the work. Environmental monitoring, like research, works best when there is generosity of time and spirit involved, as well as a sense of gratitude for that generosity. This is why designing a monitoring project that incorporates narratives and stories in the data gathering process is so necessary; data gathering should always grapple with and respect the ways in which culture is embedded within every measurement collected [54]

However, there is more to learn when aiming to decolonize research practices than acknowledging that we should be taking a human-centered approach that demands (and celebrates) that multiple ways of knowing and monitoring be respected and that generosity of time and spirit be a given. In the seminal text *Decolonizing Methodologies*, Prof. Linda Tuhiwai Smith provides an indigenous perspective towards research and knowledge production that works to decouple the colonial and imperial underpinning of much of the academic and formalized research has been performed. She provides guidelines that researchers can follow in their work, a framework that I believe can also be applied to community-guided research. She has a set of questions that she considers important for those engaging in research to as themselves, namely:
*who is participating in the process of this work?*
*Who is this process serving?*
*What is the process serving?*
*Who will benefit from the work being done?*
*What are the potential risks?*
*Is the process supportive to those who are participating and who are at the center of this work?* [197].

These questions get at the root of a research endeavor by trying to understand who stands to benefit from the work and who is most at risk. It works to establish a project's values, its desires to protect or expose its subjects, and the power dynamics that will govern who makes decisions and how.

With that in mind, I will provide what amounts to an ethnographic overview of my time shadowing and learning from communities who are participating in local monitoring of water bodies. The first is an attempt to understand how citizen science and human sampling allow for the Mystic River Watershed Association to gather a data set that monitors have been volunteering to contribute to for 20 years. The second is an attempt to understand how the Mahia Māori Committee have been working to establish a research agenda and monitoring goals when forming their monitoring program, making sure that indigenous and traditional values are at the forefront of their work. I would like to make it clear that though I have engaged in water monitoring and sampling with both communities, I have not deployed any of the CRISPR biosensors described in the earlier chapters or chapters to come within this dissertation. The goal of my research with these communities was not to try out my technologies and see how it would work, but to better understand the places that my technology may eventually live in. Understanding how these groups decide what data is important or useful to them, as well as the methodologies that they are currently using and may want to use to collect that data, was the goal. The ability to learn from these communities over the past few years not only impacted the devices and methods that I worked on within my biological engineering research, but it gave me continued support that trying to get these sensing systems to work outside the lab was useful and important to many, and that it was a goal worthy of further study.

## 5.2 Mystic Monitors: The Production of Citizen Science

### 5.2.1 Mystic River Watershed Association

The Mystic River (from the Native American MissiTuk, meaning "great tidal river") watershed covers 76 square miles and 21 communities in the Greater Boston area [1]. The headwaters begin in Reading, MA with the Aberjona River and flow from Upper Mystic Lake to Lower Mystic Lake, where the Mystic River winds and eventually reaches Boston Harbor. Between the eighteenth and twentieth century, the Mystic River ran through a heavily industrialized area, which left heavily polluted waters that lingered even after the industry was long gone[56]. In the mid-1960s, there were two large public infrastructure developments that further impacted the water quality of the Mystic: the construction of I-93 dredged and re-aligned much of the lower river while the new Amelia Earhart Dam separated the once-tidal Mystic River into an upper freshwater basin and a lower tidal body [192].

Around this time, the formation of the Environmental Protection Agency and the passing of the Clean Water Act (1972) created a push from community organizations to monitor water quality in the Boston area. MyRWA was one of the community groups that began organizing for improved quality and monitoring in the area beginning in the 1970's with members of the Tufts University community beginning

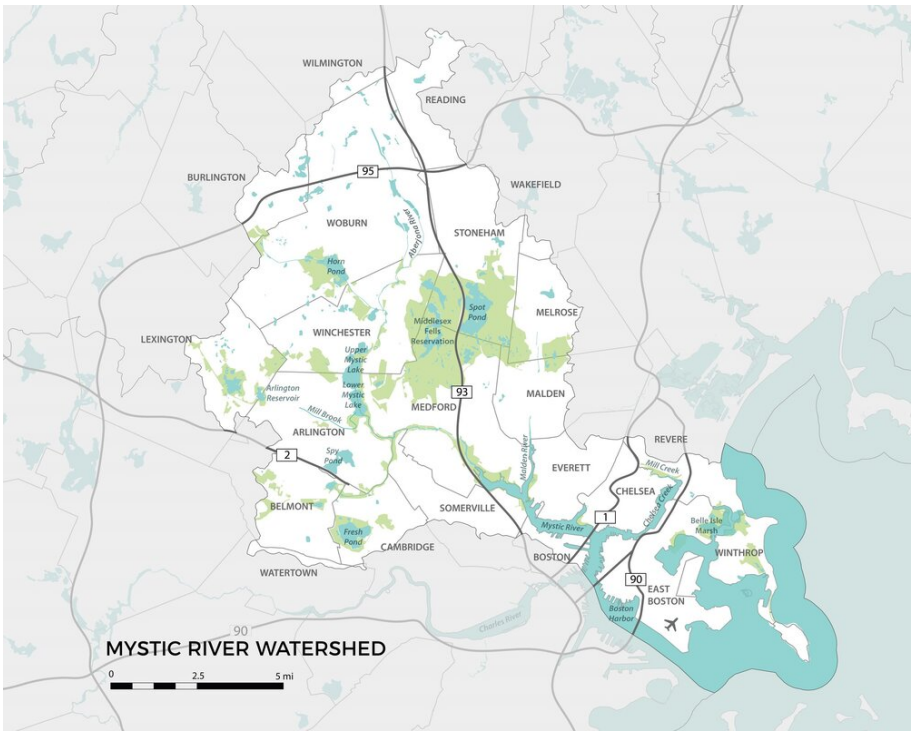



Figure 5-1: **Mystic River Watershed.** A map of the Mystic River Watershed (https://mysticriver.org/maps).

an informal sampling process to test water quality [129]. After the court-mandated cleanup of Boston Harbor, the Massachusetts Water Resource Authority (MWRA) began a compulsory water monitoring program set up by the EPA for all water bodies that lead into Boston Harbor [189]. To complement this data, MyRWA organized their own volunteer-based monitoring within the Mystic River and its tributaries (Fig. 5-1) to add further data that would add dimension to MWRA's readings. MyRWA coordinated with the MRWA to deliver the water samples collected by their monitors so that the samples from the mandated sampling sites and the additional monitoring sites could be processed using the same protocols. There have been additional efforts to collect water quality data from the Mystic, such as an EPA buoy near the Blessing of the Bay Boathouse, which is meant to provide real-time water quality data, but crucially does not have the ability to provide real-time bacterial information (at the time of this dissertation's writing, the data can be accessed through https://www.epa.gov/mysticriver/live-water-quality-data-mystic-river).

Currently, the way that the watershed can provide insights to the public on current bacterial levels is through algorithms that have been developed through collecting water samples periodically, but particularly before, during, and after rain events. These algorithms are currently used to understand the safety in boating on various water bodies (it should be noted that the EPA has different water quality requirements for boating vs. swimming, with boating allowing for relatively higher levels of contamination) but become less accurate as climate change significantly impacts

water qualities such as water levels due to drought, water temperature, pH, and the impacts of wet weather events. For example, in the summer of 2021 there were so many wet weather events that it was difficult to find a day through July and August where the algorithm didn't indicate that there were likely high levels of bacterial contamination. Conversely, the summer of 2022 has seen drought conditions throughout Massachusetts, which made it impossible for the algorithm to provide data on the predicted bacterial water quality.

MyRWA has been running a monthly water sampling program, referred to as the Mystic Monitoring Network (MMN) since 2000, with some of the volunteer water monitors having participated for as long as the program has run. The sampling sites were chosen based on the different tributaries that feed into the Mystic and have varying flow rates, depths, and levels of accessability depending on the tributary. For monitors, their monthly encounters with the river consist of waking up at 6AM to get ready and meet your sampling partner between 6-8AM at your site or 2 hours after low tide for the sampling that occurs in the lower Mystic; taking the samples with either a pole given by MyRWA or a sampling device of their own design; taking both the air and water temperature; placing all of the collected samples in a cooler on ice while performing the Winkler test for dissolved oxygen, assigning a water color and smell based on predetermined options; driving the sample over to the MyRWA offices in Arlington or to a meet-up spot near the Lower Mystic; and signing off on the chain of custody. The collected samples are then delivered to the MWRA laboratory on Deer Island, which tests for bacteria, salinity, total suspended solids, nitrogen, and phosphorus. The full water monitoring manual can currently be found at https://static1.squarespace.com/static/563d6078e4b0396c216603c8/t/58505b87b8a79b7128bfdfc1/1481661332133/2014_MyRWA_WQ_Manual.pdf.

I was interested in studying the MyRWA monitoring group because the data set is a great example of effective longitudinal citizen engagement with monitoring which has had directly positive impacts on the watershed it is monitoring. In particular, I was interested to learn whether the monitors used or developed any knowledge specific to their sampling site as they returned month after month to perform their sampling, a place-based knowledge that they used as a tool to better understand and analyze the site. I was curious, did those qualitative insights made it into the data set they were collecting? Did it change the way that they see their site? The watershed? Their towns? Did it change where or how they wanted to do their sampling? Did looking at the quantitative data collected at their site further refine their ability to monitor their site over time? My inquiry can be more or less described as at the intersection of datafication, water, and citizen science and participatory monitoring.

Water resources have a long history within the social sciences, where it has been studied as a body of politics and culture [7, 107, 177] as well as community resource [164, 15, 190], a slippery thing that is always moving and changing. Citizen science, where volunteers provide time, expertise, and momentum behind civic researcher efforts, has been studied by researchers who have an interest in, how they are forms,

how citizens participate, and the quality of data the programs generated [40, 176, 158]. There have been some who have tread in the specific waters of citizen water monitoring [3, 101, 114], analyzing motivation, participation, data generation, and outcomes from citizen science water monitoring efforts. I hope that my approach to understanding this local water monitoring effort from the perspective of the monitors can contribute some small insights into this larger effort.

### 5.2.2 Research Methods

In early April 2020 I interviewed M1, who is a water scientist for the MyRWA and responsible for this volunteer program. In completing this interview, I was able to better understand the scope and context of the monitoring program and generate a set of semi-structured interview questions which became the centerpiece for all later interviews which were conducted between April and May 2020 by telephone (with the exception of M5, which was conducted via email). Each interview typically lasted between 30 to 60 minutes. I contacted all 32 current volunteer monitors, received a response from 17, and interviewed 15, 13 of whom have an assigned site and two of whom are substitute monitors. Sampling is completed in pairs; I interviewed four pairs of monitors separately as part of the interview process. In total, the respondents covered ten sites within the watershed out of the 15 sites covered in the sampling program. The transcriptions of these interviews are the basis of this paper.

Due to the COVID-19 pandemic, April 2020 was the first month in twenty years of sampling that the Mystic River Watershed Association (MyRWA) will not have data points for the 15 sites they monitor. April was also the month that I began my research with the volunteer monitors, some of who have gathered water samples for as long as the program has existed. During my days indoors, interviews with monitors led me to Google Maps, a tool I used to take virtual tours of the watershed and sites they mentioned. In this virtual environment, I attempted to retrace the steps of the monitors, locating the bridges and banks used as entry points for the Mystic and its tributaries. I reflected on the different experiences I had in the past within the watershed: walking down to the muddy banks of the Mystic River and looking out to the newly built Encore Casino before heading to a movie at Assembly Row; walking along the reeds at Alewife Brook Reservation and watching ducks and great blue herons enjoying the water; helping a friend test out a boat prototype at Chelsea Creek; watching students in a Sunday program I volunteer with collect water samples from the detritus-filled Mill Brook for analysis.

I began each interview by asking the respondent how long they had lived in the Mystic River watershed and what initially brought them to the area. Following this, I asked when and why they initially got involved with the monitoring program to better understand their interest in the program and their decision to be trained. I then inquired about the length of time they had been a volunteer water sampler, their site location, as well as the routine they follow when they are sampling their site. These

questions were to allow the respondent to choose what they felt was mundane and interesting in their sampling site and process, as well as to allow them to reflect on the instrumentation, protocols, and sampling parameters that constitute this monthly activity. I further asked whether there are ever times they wish they deviated from the sampling schedule based on any observations they made in their daily lives, whether they incorporate any qualitative observations into the their monitoring protocol that are useful for the understanding and interpreting their site, whether there are any parameters or locations they would like to add to the program, and whether they look at the quantitative results of their samples once they are analyzed by the laboratory. The semi-structured interview questions can be found in Appendix C.

Through these questions I sought to understand the boundaries of the program; that is, whether the respondents felt at all constrained by the focus on repeatability and consistency as part of their monitoring activities and whether they see themselves as having agency in the water monitoring process, or merely collectors or keepers of data? Finally, I asked whether there was a specific title that they would assign to themselves based on the work that they do, whether it be a citizen science, a field technician, a volunteer, or otherwise. I typically pushed the respondents to further explain why they chose the title they did in order to get a better sense of how they see themselves and their work situated within the larger picture of the watershed and the work that they do to generate data. After respondents answered the questions, I let them know I was complete but would love to hear a sampling story they felt was memorable or answer any questions they had on my research within water monitoring technologies. This open-ended statement was to allow a non-structured environment for the respondent to share anything that they found important or interesting, but that might be out of the scope of the questions that were asked.

Once I completed the interviews, I read through three transcripts for themes and wrote a codebook which I used to code the rest of the interviews, which can be seen in Appendix C. These codes included references to water sampling, water contamination, bacteria, nutrients, following protocols, urbanization, development, rain, the watershed, consistency, convenience, quantitative, qualitative, the data set, citizen science, as well the reason that the monitor began sampling. These codes were added to the interview text in a computer file and then these coded pieces of text were extracted and organized by code for analysis. The results of this analysis will now be discussed.

### 5.2.3 In service of "The Data Set"

At the heart of MyRWA's monitoring program is, unsurprisingly, the data. The monitors had much to say on the data set, as well as the specific water quality measurements that are processed each month. M1, the watershed scientist and the sample coordinator for the program, provided insights on how the sampling program was set up, which reflected the close ties that the program has to the EPA-mandated sampling (involving the collaboration between MWRA, EPA, DEP, and USGS among

others). He refers to the sampling work that MyRWA does as one that attempts to "add value" to the data set being gathered on Boston Harbor, and mentioned that the water quality metrics were chosen to correspond with the Boston Harbor tributary sampling. Many of the water samples are even processed on the same instrumentation as the EPA-mandated samples by the same water scientists.

I found it interesting that focusing on the water sampling within the context of the larger data set put limitations on the sampling processes of the MyRWA volunteers, even though there are no formal regulatory requirements on their sampling. For example, I brought up to M1 that I was surprised to see that pH was not part of the water quality panel done, because this measurement is one of the most common done in water monitoring context. He responded that he didn't know why pH was not a part of the baseline work, and followed that by saying *"...there may be some pH data that's acquired over the years, but it's not part of that main data set" (M1: 189-190)*

There is a distinction being made between data and "that main data set." Which is to say, even if a MyRWA monitor has an interest in collecting specific types of data that would be relevant to the health of the watershed, that data would be categorized and considered differently (likely less valuably) because it does not have a neat place within the larger data set. This, of course, falls into the larger discourse around the "datafication" of the watershed, and how the goal of generating data that can neatly fit into the Big Data set can distract. I will offer a qualification though, which is that "datafication" and "big data" are often terms applied to things like apps, cell phones, Fitbits, and latent air quality monitors, i.e. things that can produce data points every second or minute. The entire data set of monthly water sampling measurements over twenty years is actually quite sparse when compared to the standard data sets discussing Big Data. However, I believe that the generalized procedure of social-action-to-quantitative-data pipeline for the optimization of future actions and outcomes does still hold in this case.

Indeed, I saw this mindset throughout the responses of the water monitors when asked what kept them sampling from month to month as a volunteer. The monitors consistently mentioned the data set, praising the value of the data they were collecting in the context of the longitudinal record that could be used for future benefit.

> *"Yeah. And so the data's the a kind of witness or like, you know, the sort of like the history and it's, it's as you were saying this like impartial, non-politicized and I mean you can obviously politicize it and people can decide whether or not they want to accept it, but is this like very quantitative, um, sort of like, you know, documentation witness for whatever's happened and you know, anyone at any point can go back and sort of like look at it and sort of listen to what it's saying and see what, where to go from there."* (M4: 355-361)

*"...being a biologist or an ecologist, I sort of know the value of these long-term data sets. You can't recreate it. I mean, people can try doing some, you know, taking an algorithm and plotting everything, but it's kind of amazing to have it."* (M6: 447-450).

*" I think the critical part is to have verified record that when you need to act, give you a basis to act successfully"* (M7: 174-177)

*"MyRWA has been around for a while now, so they have a data set that the state or the cities would otherwise not have and it shows you progressions that you can then maybe relate to what's going on in the community and see effects of changes in regulations."* (M8: 89-93).

*"It's always great to have these longitudinal records, you know that we have established through methodology that you just keep doing and then you can see, that is really one good aspect that we have a long track record now.* (M13: 121-123)

*"Yeah. It's incredible. I was talking to somebody who is from a different town and he's also a sampler. He was saying how he doesn't have a lot of faith in his town to do anything. But, just knowing that this database exists and having it as sort of impartial, maybe like some sort of thing that's not caught up in politics is so important for him because even if he doesn't feel confident in the town, he knows they're still gathering all this data that can then be used into the future. Even if right now people aren't making use of it.* (M10:316-323)

I realized quickly that the data set had become to the monitors, as Bowker and others had warned, the *product* of this volunteer monitoring endeavor. This was seen not only in the ways that the samplers viewed the process of sampling, but also in the ways that they described the data set itself as the lever for change within governmental and local spheres.

*"We can use [the data] to influence public opinion and public decision making, and... you [EPA] can use that information to find out where you should apply your resources to investigating pollution sources and getting rid of them with your powers as a government agency."* (M1: 213-222)

*"...the fact that we had this consistent sampling record going back far that said that the waterways were a problem and we kind of knew what was wrong with them but nobody would fix them, created that pressure to*

*fix them. And, and so it's a very powerful policy tool."* (M10: 208-211)

*"I mean, I guess I, I'm very glad that they have that amount of data because I think it, it definitely helps inform various policies throughout the watershed and how to go about cleaning up particular areas"* (M14: 165-167)

What I found especially interesting was that even with the admission by many that the goal of their monthly volunteering was the collection of data for a data set, they also readily admitted to the fallibility inherent in collecting and relying on one set of measurements per site per month as an accurate picture of the watershed. Data quality was often discussed in the context of attempting to collect a representative sample (this will be discussed at length in the next section on citizen field techs) and in the context of rainfall. Within the watershed, a storm event (wet weather event that sees 0.25 inches or more of rain within 48 hours, causing a large influx of stormwater into local water bodies) impacts the baseline data collected.

These wet weather events are so problematic that MyRWA has a special hot spot monitoring program where they try and collect water samples after these storm events, typically to have an upper bound estimate of the levels of contamination within the watershed. Significant water contamination during a wet weather event is fairly common within urban watersheds, where a significant portion of the ground is nonporous. Stormwater runs on pavement and asphalt, picking up whatever contaminants are present (fertilizer, dog feces, trash etc.), before entering storm drains and urban water bodies. On top of that, Massachusetts' aging sewer infrastructure, which includes cracked underwater pipes and combined sewers, add further contamination to these urban water bodies. Simply put, unpredictability of weather as a factor in the consistent monthly timing of water sampling makes parsing the data *"a signal to noise problem"* (M4: 227), which in turn makes it difficult to make qualitative called like whether things are improving within the watershed.

While there are certainly databases that researchers can cross-reference to try and understand why some data points were so vastly differ from others, I think this does raise a question about how these data sets are contextualized and used. M4, who has been sampling for 20 years, mentioned in regard to the data collection process:

*"...for 20 years I had a feeling it was useless, but I was going to do it anyway because I thought it was going to be useful someday, and I still haven't seen that day."* (M4: 267-269)

There was further uncertainty both in the measurements that were being collected as well as the utilization of qualitative data streams (assessing weather as well as the

color and odor of water) within the monitoring process expressed by M15.

> *"The fact is, it's a bit of a black box to me. We collect these samples, we write down the data, our observations, you know, is it cloudy? Is it partly cloudy? Clearly that's a subjective call. But, um, but we write them down and we submit them, and they go away and I don't know what happens to them."* (M15: 123-127)

I found it quite interesting that there were volunteers who had little idea of the results from the water that they were collecting (for the first 15 or so years before M1 became the watershed scientist, the results weren't aggregated or sent out to the monitors) and still participated, month after month, some for two decades. There were even monitors who doubted the utility of the data that they were collecting and still participated, month after month, some for two decades! And yet, the data set persists through the will of these monitors and their hope that there will be an ultimate use for the data; that sometime in the future, someone will be able to find this data set and improve the watershed through it. (Some of the water quality data can be seen aggregated in the EPA's report card for the Mystic River, which at the time of this dissertation's writing can be found through https://www.epa.gov/mysticriver and the rest is aggregated by MyRWA in their water quality database).

There were even questions as to why humans were participating in the sample process at all (this quote is quite long, but it is too relevant to the topic at hand to cut shorter):

> *"Periodically at MyRWA, I go over and wave my arms and advocate for automation of some of this stuff. One of my pet peeves is, so we've got a water temperature reading to do. We collect a sample, we drop the thermometer in the sample, by now the sample is sitting in a bottle on the sidewalk and it's either warming up or cooling off, depending on the difference in the water temperature and the air temperature of the surroundings. But the thermometer's in there. And while it's in there, the thermometer is getting closer to the real temperature that the water was when it was collected. And then we pull out the thermometer and read it and we pull it out of the water and hold it in the air and the water, the temperature now is changing to the air temperature and the quicker we can read it the better. So we read the water temperature and there's an opportunity for misreading it. And then we write it down on the data sheet and there's an opportunity for mis-transcribing there. And then we turn in the data sheet and someone looks at it and there's an opportunity for them to misread my handwriting that could have been shaky because it was too cold to be getting water samples. And they take that water temperature and they transcribe it and put it into some computer program and there's*

*an opportunity for them not only to misread it but also to mis-enter it. And then it goes into a computer database. So I, I wonder what is the likelihood that, that, that datum, uh, bears close resemblance to the reality. But that's a depressing thought. So I prefer not to prefer not to think about it very much. So I wonder in this age of technology, why is it that my phone hasn't taken the temperature of the water sample and directly entered into the database without any opportunities for misreading, mistranscribing, misentering, and so on and so on?"*(M15:222-246)

It's hard to imagine that someone else might take these measurements seriously if the samplers themselves don't, but it's also possible that they're skeptical because they're watching the sausage get made. But considering the participation of volunteers in this monitoring program became a significant fascination after realizing that the participation was not contingent on some sort of scientific product or instant gratification due to a result. I decided to probe more deeply into the concept of monitoring and the production of "science" to try and understand more deeply where the monitors saw themselves in the pipeline from nature to data.

### 5.2.4 Citizen field techs and (citizen) science as a vocation

I had always felt that the MyRWA volunteer monitoring program was a clear example of citizen science at work. Citizen science brings together members of the public who engage in scientific work, typically in collaboration with or under the direction of scientists or scientific institutions. The MyRWA program, supported by local nonprofits and aided by city infrastructure, centers around the monthly citizen volunteer sampling at a variety of sites around the watershed. However, when I brought up the role of volunteers in the monitoring process to M1, the watershed scientist, particularly whether their experiences during their sampling in the watershed ever impact the protocols for sample collection. He provided the following insights:

*"We call it a citizen science program in, and it certainly is. This is volunteer citizen engagement in data collection. But I sometimes think of it as citizen field tech rather than citizen scientists. I happen to be a birder so I take part in this big database at Cornell called eBird where people contribute data and it's a citizen science project, but there, the observer in the field is making judgment calls, you know, how many, how many starlings, where at that location, was that a black crowned or yellow crowned, like heron or whatever. There, the observer is being asked to add, to contribute expertise at the site...and they're not making, well, they're making some qualitative judgments about whether the sample has an odor or whether it has, what it's appearance is, but the goal of the protocol for our samplers is to acquire a repeatable representative sample of the stream, of what's in*

*the stream."* (M1: 310-330)

M1's assessment of what constitutes a "citizen scientist" is someone who has specialized knowledge that can be used to make real-time judgment calls, though he did acknowledge that the MyRWA monitors do, in fact, make qualitative calls each month on the status of the water (though the qualitative data collected seems to be secondary to the quantitative data from the collected water sample is obvious). But how does collecting samples (and collecting it well – which is to say a repeatable, representative sample) count as less "citizen science" than the technicians or scientists at Deer Island who are simply receiving labeled bottles and processing the samples? One of the difficulties with water monitoring that may be a part of M1's separate conception of a birder and a water monitor is that unlike a birder, the subject of monitoring in the watershed is a set of characteristics that are invisible to the naked eye. The water monitoring efforts therefore require additional steps through off-body sensors to move from from sampling/observing to result. It seems that when off-body sensing is required, those who participate in the sample collection are not considered to be where the "science" is happening – that designation is saved for the sensor that is generating the data. Undervaluing the work of technicians and others who, with their embodied and contextual knowledge, are critical in the production of science, is unfortunately common when held up against the formal knowledge held by the "scientist" [17].However I was curious how the monitors themselves chose to identify their role in the monitoring program and whether that had any impact on the work they do.

The monitors each had different ways that they related to their roles, some feeling very close to the term "citizen scientist," others feeling it was too grandiose for the type of quick work they do once a month in the morning before going about their day. However, there was always a great deal of pride in the work performed and a fastidious attention to detail and consistency when describing the sampling process that I believe underscores the type of intimate knowledge and deep connection that these monitors have with their site and with the data that they are collecting. Many of the volunteers happen to be active or retired scientists with familiarity on the processes or the type of data that the program is trying to collect. It is possible that their knowledge of (or at minimum idea of) the technical definition of what a formal "scientist" is may be a significant factor in the definitions they provided. Often the distance from the term "citizen science" by the monitors stems both from the distance from the data that results from the sampling, as well as the lack freedom to choose what is sampled, where, and when.

> *"Well, we don't have to be scientists to be sampling the river... but it can be a baker who does the sampling, you know, once you're trained you're good. You don't have to be a scientist to do the sampling." (M2: 51-59)*

> *"Sure, it's citizen science, but I'm not engaged in any way in changing*

*anything or rethinking, you know. . . .citizen science for sure, but not engaged in changing anything about how anything's done. (M3: 71-81)*

*"I sample the water, I deliver the samples and log information with the samples, I pick up next month's supplies. I like the term Water sampler or water quality monitor volunteer. No point in making it fancy, it is what it is.* (M5)

*"Oh yeah, I think it's citizen science. Yeah. You know citizen science with a fairly, I mean it's a good citizen science project cause is pretty good quality control."* (M6: 278-280)

*"I don't, you know, I'm not doing the upstream portion, I'm only taking the samples, but I would call it citizen science, it's pretty cool. I mean, for me it's satisfying to get out and do and I feel like it's a good cause, especially when I see how well the folks manage it, and they manage the data well too, it's pretty impressive.* (M12: 136-140)

*"I don't know if I would say it's citizen science because we're not doing the lab work. Um, we're just collecting. So I guess I personally view this as closer to a field tech kind of thing and it just happens to be done by a bunch of volunteers. So in terms of citizen science, no, but maybe citizen field tech guy, I don't know. Yeah, I guess I don't view this as like hard science per se, but it is definitely a crucial part to be able to get the, um, get the samples necessary so that you can take them off to the labs.* (M14: 183-189)

*"You know I guess I think the term citizen science is a bit of a grandiose term anyway, knowing what I do."* (M15: 112-113)

*"I know they like to use the word citizen scientist. I feel, um, tech is like an interesting word. I'm, I mostly feel like I'm a volunteer that shows up once in a while and I got trained on doing the procedure properly, but mostly I feel like a volunteer. Um, but tech is kind of an interesting, interesting word for it. . . most of the science is happening, most of the testing, the actual testing and things like that are happening, you know, once we return those vials. So I think tech makes more sense because we're doing kind of the, the, the, the collecting bit of it and not necessarily the testing and analysis of it."* (M16: 123-133)

In his lecture "Science as a vocation," Max Weber is lamenting what he sees as the current state of science (in this case he is referring to the social sciences and the

humanities in 1917). Of course, a citizen science program is not the target audience for this lecture – the monitoring program in question is certainly not a vocation (it is all volunteer), nor is it really the type of "science" that Weber is referring to when he uses the term "science." However, I do believe that some of the points raised are relevant and reflect some of the questions that monitors had with their monthly work. When describing the importance of passion and inspiration, Weber notes:

> *"Among young people nowadays the idea is very widespread that science has become a question of simple calculation, something produced in laboratories or statistical card indexes, just as 'in a factory,' with nothing but cold reason and not with the entire 'soul.' Though of course we should note in passing that for the most part there is very little understanding of what actually goes on in a factory or a laboratory. In both places it is necessary for something, and the right thing at that, to occur to people if they are to achieve anything worthwhile."* (Weber, 8)

Weber feels that "scientists" are responsible for the technical *what* that they are working on and considers whether they should also be on the hook for the conceptual *why*. I think this framing of "science" is quite interesting when applied to the MyRWA water monitoring program, where citizens distance themselves from the term because they don't have full oversight in the process. However, from the Weberian standpoint, it seems that involvement in the production of data (i.e. the water monitors) would fall into the category of a scientist, performing the daily tasks of consistently collect samples that lead to results (the "scientist" term would also extend to those at the end of the chain of custody who process and log the results for the samples). Within the routine work that the monitors perform, there was a significant amount complexity, rigor, and consistency that is described which is a large part of the resulting high-quality data that is generated.

All of the monitors have developed a sense of how to get the *right* kind of data, which is to say, a sample that is representative of the body of water they are tasked with sampling. From the MyRWA side, this is somewhat coordinated through training monitors on best practices for collecting, measuring, icing, and delivering the samples. But for many of the monitors, the collection process is informed not only by the training, but by the other experiences in the lab and in the field in which these monitors have participated. Almost all of the volunteers that I spoke with (14/15) have at some point been involved in the sciences, if they were not still active. Monitors' background ranged from PhD's in environmental studies and biology to Masters degrees in hydrogeology and ecological management. Many mentioned that they were modeling their actions off of a combination of previous experiences (both from their previous professional backgrounds, the training, and years at their site). M12 described how within the sampling, *"the protocols that we follow are pretty much what I followed professionally."* (M12: 223-224). M15 acknowledged *"[he's] an engineer by trade and*

*by training, and [he] think[s] [he] bring[s] an approach that is more scientifically oriented just out of habit.* (M15: 114-115). Below is an example of the descriptions that the samplers provided as they described sampling at their sites:

> *"The water from the stream tends to be clearer than the pond. The pond is turbid. So you can actually see where you're sampling the actual stream. Instead of having the stream contaminated, just be sampling the water of the pond. However, you're not ensured that some of that water isn't actually swirling over and getting in, I mean it's not a perfect flow. But to do the best you can, we do duplicates and they came back consistently So whatever we're doing is reproducible." (M4: 117-123)*

> *"We, in fact when I started sampling, I said we really need to be sampling upstream of the bridge rather than downstream because the person who had been sampling it couldn't tell what was up or down stream because it often depends on which way the wind is blowing, rather than how much water is going down the Malden, so it's pretty stagnant." (M10: 247-251)*

> *"We actually, we used to use a pole and we used to do one bottle at a time with the sampling poles, but in our site, since we're over a dock, we started using a shower caddy" (M11: 74-76)*

> *"At some point the real scientists at MyRWA decided that it might be adequate to take samples from the bridge itself. I gather they've been doing that sort of thing on the Charles River for some time and had decided that it was not a risk to the integrity of the samples. So, they came up with this contraption, that I call SC-1: that stands for shower caddy, something originally intended to be used in the shower to hold your shampoo and whatnot. So, they showed me that and suggested that we consider using that in Alewife Brook. I looked at it and decided that I could do better. And that's when I started tinkering around in my garage with PVC pipe and various exercise weights. Then I created RS-1, which stood for ring stand, because I thought that looked like that kind of an instrument." (M15: 193-205)*

> *"I mean, I think my technical background helps with my awareness of why consistency and precision are necessary when doing this. I think it helps with my awareness of why we're doing the same sampling, same places, same time of the month, same time of the day, every single month and we, you know, you don't go scoop it out of the muddiest part of the bank. You try to get the water that's flowing and stuff like that. I think it helped with my understanding for why's" (M16: 136-141)*

(a) (b)

Figure 5-2: (a) Alewife Brook Reserve monitoring site. (b) A monitor using a custom sample collector over the bridge at the Alewife Brook Reserve site.

It should be noted the M15 and his sampling partner have a particularly tricky site to sample, which is either accessed through walking down a steep muddy bank that is often strewn with trash and debris, or by dropping a sampling device from a bridge that is above the stream (Fig. 5-2a, 5-2b). The two monitors constructed multiple rigs that they tested out before settling on their current set up, which allows for a representative sample that is safe and easy to collect. I spent time with M16 and his monitoring partner, watching how they use their device. When at their site they are comfortable, brisk, and efficient in their work. They first collect some water that they drop a thermometer in to get the relative temperature. Then, they look at the kit picked up from the MyRWA offices that provides the containers necessary for sending to MWRA facilities and they add those to their rig. They drop the rig down into the water and take a moment to do a holistic survey of the site for the color and odor of the water, as well as to notice anything else that might be out of the ordinary, before pulling the rig back up. Then they cap up the water samples and place them on ice, fill out their chain of custody slips, note the weather and are off to deliver the samples to M1 before returning to their day.

I would consider the level of detail and thought put into the sampling that these monitors go through to be paramount to the success of the data, having a far greater impact on the quality of the data that then the MWRA scientist on Deer Island ever could. The very success of the monitoring program is, in fact, contingent upon this

understanding, and the performance of science. That *is* science, at its core. The monitors are the ones that make sure the downstream data are representative, reproducible, and accurate (critical to a longitudinal data set) not the MWRA laboratory staff and scientists. However, as mentioned above, it is common for work done by technicians to be undervalued within the larger scientific endeavor. The mentality of not putting the label of "science" on earlier in the data generation pipeline could be correlated to the fact that the overwhelming majority of the volunteers come from quantitative science backgrounds. It is almost as if the scientists are too familiar with the idea of "science" to feel that this counts. Personally, I believe that there is no less skill and training required in the proper collecting and handling of the samples in the field as there is to run a reaction in a laboratory, having tried both. It would be interesting to have an opportunity to speak with those who work in the laboratory on Deer Island and inquire whether they felt they had any ownership of the data.

However, the fact that the sensor technologies used for MyRWA's water quality measurements were unavailable to these monitors and therefore the resulting perception that they were not the ones performing "science" in the same ways as those with the sensors certainly was surprising and was an idea that I thought a lot more about when designing sensors that try and cut out the laboratory middle man. Providing low-cost or easy-to-use sensors that local monitoring can use in the field is something that not only keeps the the sample-to-answer pipeline short and the data close, but it also gives the monitors a greater feeling of ownership over their work and a feeling that they are engaging in "citizen science".

### 5.2.5 The Structuration of Citizen Science

However, beyond the ontological and epistemological complexity of "science" and how citizen science figures into the may ways of engaging in a monitoring program, there is still the question brought up by monitor M15, namely, *why is it that my phone hasn't taken the temperature of the water sample and directly entered into the database without any opportunities for misreading, mistranscribing, misentering, and so on and so on?*). Why *does* it matter that citizens figure into the monitoring at all? There are many examples of citizen science projects where there are volunteers who are "collecting data," wherein the data is actually autonomously collected by a device that saves the information or else is connected to a nearby computer an uploads the findings ([41, 25, 193, 231, 140], among countless examples). Is there something that is gained from having humans into this process in such an intimate way? Perhaps something unique that the monitoring program wouldn't otherwise have?

I would like to further this question through the lens of a comment that M1 made, one that most any manager for a participatory monitoring program may be able to resonate with, that speaks to the difficult nature of trying to build a quality data set through a decentralized process.

*"... that it reminds me of something that you often say that like, you know it's my job to go and get data, but sometimes I'm in the field and I say Oh, that's interesting. And I kick myself and say, [M1], it's not your job to get interesting data, It's your job to get useful data. If it's not useful, please move on." (M1: 540-545)*

Though I do not like the framing of data as either "useful" or "interesting" (a false dichotomy), I do think that it is a really good case for why having a human in the monitoring process can go above and beyond an automated sampling, even if brings the risk of data consistency along with it. This is because when humans are engaging in a data collection process, especially one that is outdoors, there is inevitably what Mike Michael refers to as "overspill" that occurs, something that participants engage in that is outside the prescribed intention of the event [151]. I believe that the "overspill" is similar to the "interesting" data M1 referred to, information that is outside of the purview of the data set that the monitors are collecting but nonetheless something worth noting and useful it its own right.

MyRWA's water monitoring program, in the absence of cheap and easy sensor technology that can be deployed for autonomous sensing in the ways of air or noise pollution, still requires bodies in a space. Once at the site, the monitors can't help but look around. These monitors, even when collecting their "useful" data before going about their day, also happen upon all sorts of "interesting" data, which can be part of an ad-hoc monitoring process that goes on at each site at the same time as longitudinal samples are being collected. As mentioned previously, some of these volunteer monitors have been returning to the same site for upwards of two decades. They have an intimate knowledge of their small piece of the watershed, the type of local knowledge that is impossible to put in a data set but is easily activated when the landscape shifts in some way. Month by month, they are calibrating their own on-body sensors to their surroundings, increasing their sensitivity with each passing month. Monitors mentioned noticing anomalies like illegal pipe connection, gas and oil spills, hydraulic fluid, fish kills, sewage overflows, laundry suds, and a variety of garbage when doing their monthly monitoring. As M10 noted,*It's two sets eyes watching a river on a regular basis, even if it's just monthly. And I think that has an impact."*(M10: 72-79). Even M15, who was the monitor most skeptical of the role of humans within the sampling program, admits *"There are dead pipes that are visible from the place where we stand when we go to collect our samples. And if all of a sudden one of those pipes started flowing, both [M7] and I would notice it. It wouldn't be until the next time we went for sampling, but we would notice it."* (M15: 416-421).

I was concerned, however, that this "interesting" data may fall through the cracks, because it is outside the structure of the monitoring program. However, it seems that just by participating in the monthly monitoring program, the participants also learned more about the larger municipal infrastructure that exists within their city and felt more comfortable reaching out to report events when they encountered "in-

teresting" overspill data. This means the monitoring program is not only producing a data set for water quality, it is also producing informed and engaged locals who have the ability to perform ad-hoc monitoring in their local environment and know how to report this data to those who can make use of it. These ad-hoc observations often fall into the category of "just-good-enough" data, something that can push for further monitoring or specialized analysis within the watershed.

Considering how intertwined the ad-hoc monitoring is with the monthly monitoring program, as well as how prevalent the different types of life experiences and knowledge are to the monitoring, I was reminded of social scientist Anthony Giddens' structuration theory [77], which emphasizes that social phenomena are controlled by a combination of social structure and the agents within that society (not through either one imposing a force on the other, as postulated by various sociologists). Social structure and agents are in a dance together, leading them both to evolve due to pressures that they exert on one another. This pressure is what allows for the constant making and remaking of our social existence. I will briefly explain a bit about structuration theory as I feel it applies to this monitoring program, though this is by no means an exhaustive summary of the theory. It should also be noted that this theory fits well because it is, as some have noted, more of a meta-theory (a way of thinking about the world, than a theory that could tested and proven [102]. I use it here as a way of acknowledging and attempting to understand the ways in which the monitors (agents) and the monitoring program (social structure) have had influences on each other and how that, in turn, has influence on the watershed.

As monitors go about their daily routines, they are building "memory traces." There are typically three types of memory traces: "legitimation," which provide information on what should happen in a given situation, "signification", which is how to make meaning or interpret an event, and "domination," which refers to the means (people or object) needed in order to accomplish something [77]. All of these memory traces provide a level of "capability," wherein the agent has the ability to act on, ignore, or change their behavior within the structure (even while limited by it) and "knowledgeability," which is an actor's awareness of their actions. Knowledgeability is informed by both "discursive consciousness", which is when an agent can verbally express why they are doing something, as well as "practical consciousness", which is knowledge that is so ingrained due to everyday life that they hardly notice it until someone brings it up. Finally, Giddens emphasizes both time and space as critical components within structuration, as they allow for what he refers to as "regionalization," or "the movement of life paths through settings of interaction that have various forms of spatial demarcation." [218].

For many of these monitors, the knowledgability that they bring to their monitoring work comes not only from their history with the site and interactions with the watershed scientist who leads the sampling program, but from their previous experiences as scientists, researchers, and professionals collecting high quality data sets. Sometimes they have ways of verbalizing how they know what to do when they arrive at a site,

whether it be the type of sample they are trying to gather or the way in which they can assure there is repeatability, but others times it is something they just seem to know how to do without even thinking about it. They is also a regionalization specific to their specific sampling site that allows for a deeper, intuitive understanding of the locale and the ways in which the monitors can understand an engage with the space. All of this comes together into a type of "practical consciousness," things the samplers and the watershed scientist might even take for granted (such as the importance of replicability and reproducibility of data, consistency between measurements, and the utility of data sets as a lever for policy and regulatory change).

M14, a substitute monitor and therefore rotates sites and sampling partner (rare within the monitoring program), provided a useful description for the type of knowledgeability and regionalization within the volunteer water monitors:

> *"Yeah, so in general, most of the people, at least the ones who had been there for longer than a year, they generally have a sense of whether or not, based on the recent weather conditions or whatever was going on in that area, they would have a good sense about whether or not there'd be high levels of DO for example, or something like that. So they would kind of get a sense of how to handle that. There's other people who have a good sense of how to actually handle the samples better. For example, like with the DO [dissolved oxygen] samples, when we're fixing them in say January, it's so, so cold. There's one guy who taught me this trick that in order to get it to actually participate out really quickly, you actually need to hold it in your hand or in a pocket or something, it needs to be a little bit warm, it's going to take forever or possibly never separate if it's in the air or just sitting there. So, just little tips and tricks like that I've learned from some people. And then in terms of the clarity of the water, sometimes what's been going on in the watershed, like there's a couple where people just had a good institutional knowledge of airport operations or the container ships or something that go up, down through Chelsea Creek. So that kind of stuff some people have definitely paid attention to. But I think that usually it takes people, I don't know, a year or so to kind of get a handle on what are the various patterns of their site."* (M14: 113-130)

While I initially viewed the structure of the monitoring program as something that limited the monitors, the structure and consistency of the program is likely what enables them to develop the type of sensitivity towards their site that allows for interesting ad-hoc monitoring. For each monitor, their assigned sampling location and sampling time (7AM on the third Wednesday of the month for freshwater, 7AM on a tidal-dependent date for the saltwater) creates a type of regionalization and a routinization. There are memory traces, knowledgeability, and capability stemming from their educational and professional background, the trainings they participated

in, the observations of their sampling partner performing tasks, and their own personal experience of engaging in the sampling on a monthly basis. All of these create a complicated social practice which cannot, now that I think about it, be properly conveyed through a label like "citizen scientist."

From within this structure, though, monitors can be flexible. The very structure of the program allows at least two trained sets of eyes to engage with the landscape and react or respond to it as needed each time they sample. While monitors collect their water samples in labeled containers to generate the consistent quantitative data set, they also leave room for spontaneous qualitative observations and analysis by the monitors that can allow for new interventions within the landscape, and possibly even the sampling process. They can push on the structure each and every time they return to the site both because they know more each time, but also because the site is always the same and always changing. Below are a few choice quotes from a monitor that reflects these ideas:

> *"One is we're getting a good record of how these changes are impacting the water quality in the Malden. Um, I'm not sure if it shows up so much as the bacteria, but there's all the different indicators we have. And so some of this development may actually make water quality worse if we're seeing a lot more landscaping. Right. But we're also a set of eyes on the river in that, in the past I have, I caught several spills of things like hydraulic fluid and the washing of the buses and all of these things that were put into the river that we might have missed it, but because I was there every month and I could see what was going on and I would take pictures, and we had a fish kill one year too, I mean, it's just stuff like that . Um, it's, you know, two sets eyes watching a river on a regular basis, even if it's just monthly. And I think that has an impact. And I think that also just that we have data for the river going back 20 years is pretty important as well."* (M10: 68-80)

> *"You know it's funny though cause I like I said, I've been working with these guys a long time as a volunteer and I have noticed a few things cause I do know the watershed pretty well cause I, part of my professional career I actually walked entire Upper Mystic watershed doing, we were doing mapping, so I know the watershed pretty well, certainly right near my hometown near Arlington. So I did notice a couple of times where there was like suds and things, it Look like someone's laundry detergent essentially, and it was where there's an illegal connection from a sewer, it's highly industrialized in this part of the world and so looks like someone had somehow inadvertently connected their piping from their washer to this small creek that runs around town. So I called the guy who's the monitor now and he said okay, we'll check it out. "* (M12: 189-197)

> *"You're plugged in, you know what's going on and if something happens, I know the people, my local representative and local Senator, or know those people. Some cases you personally, so if there's any issue with the river, and there have been some issues in the past, I can just call them up and say, Hey, you know, are you aware of this? You know, something needs to be done and you know, they, they will respond. So, you know, being, being part of this whole network of people that are involved around the river or you know, makes you do something, but it's kind of this chicken and egg problems, you know?"* (M13: 175-184)

In summary, whether or not monthly monitors refer to themselves as citizen scientists or a field tech or monitor, all of these terms are likely unable to hold the epistemological overspill that comes from personal and historical experiences both in and out of the field. Regardless, locals are participating in a highly structured program that, while somewhat limited in scope, provides a social scaffold that they can consequently stray from and identify curious things at their sites. Robotic and automated samples are forever limited to attempting to gather their pre-programmed "useful" data, not able to reach the more interesting, and anomalous, data that relies on humans, with all of their idiosyncrasies, schedules, subjectivity, and fallibility. Robots are also less flexible than humans, unable to change their sample methodologies or location depending on qualitative characteristics of the water, which is critical to the overall integrity of the site monitoring. The Mystic monitors have gone above and beyond the definition of their position, contributing to a long-standing, rigorous, and highly standardized set of data that are used by municipalities and governmental agencies while picking up local anomalies along the way.

## 5.3 Horokaka Mātaitai: Multi-species monitoring

I will now spend some time talking about a very different monitoring program that is at a much earlier stage within its process. The world's oceans are some of our planet's most biologically diverse and complex ecosystems, with most being managed by indigenous communities [10, 221]. Unfortunately, these ecosystems are currently under threat due to to changes in climate, overfishing, and pollution among other stressors [89]. The second monitoring group that I have been learning from is the Horokaka mātaitai on the Mahia Peninsula, under the control of the Mahia Māori Committee (Fig. 5-3). I was initially connected with members of the Mahia Māori Committee through a collaborator who led Te Herenga, which is the national Māori network and a forum for Māori engagement with New Zealand's EPA. The Māori of Aotearoa New Zealand are the indigenous people of the land and have a history of environmental monitoring and management that long predates modern environmental protection laws through the implementation of *mātauranga māori*, or the Māori indigenous wisdom that has been passed down through generations [12, 38, 146, 226]. Being invited to learn about how the program was founded and their monitoring

goals, I hoped for the opportunity to learn how the *tangata whenua* (people native to the land) have managed their environments for hundreds of years and expand my understanding of the types of cultural tools a community-led water monitoring could implement alongside modern sensing strategies.

Figure 5-3: **Horokaka mātaitai** This is the indicator that delineates the boundaries of a mātaitai, where *tangata whenua* locally manage the fishing (Credit: Avery Normandin)

Working with this program was exciting not only because of my deep interest in local monitoring practices, particularly ones that emphasize the type of cultural and historical sensing that occurs over time (years to generations) which allow for connections that may otherwise be missed, but also because I felt a bit frustrated by how set in their ways the MyRWA monitoring program was. Though I acknowledge the importance that a long-term data set has for legacy, legitimacy, and substantive outcomes that can improve an area undergoing environmental and ecological changes, I was also curious how a program might decide which directions they wanted to go in and what parameters they might choose if not bound to an existing data set or attempting to replicate a Western science-based approach.

(A note: Within this section, words that are in Te Reo Māori will be italicized, with an approximate definition in parenthesis. As a *pakeha* (foreigner), I acknowledge that I have only a crude understanding of these concepts and will never fully understand the breadth and depth that these words carry. I am deeply grateful to the time and patience that my Mahia *whānau* has shown me in explaining these concepts and I recommend spending time researching the the work of Māori scholars who have spent their careers on these topics.)

### 5.3.1 The Mahia peninsula and Rongomaiwahini Iwi

The Mahia Māori Committee is made up of members of the Rongomaiwahine iwi, who *whakapapa* (trace their ancestry) to the Mahia peninsula and are considered the *tangata whenua* (indigenous people to the area), a small piece of land that protrudes from the southeast side of the North Island of Aotearoa New Zealand. Rongomaiwahine iwi members are descendants of Rongomaiwahine, who was a navigator, and a commander, and one of the few women who had more *mana* (cultural and spiritual authority or influence) than her husbands, one of which being Kahungunu, an iwi in the same region but that Rongomaiwahine iwi members do not consider themselves to be a part of. The Rongomaiwahine iwi has hundreds of years of deep physical and spiritual connection to their land, which traditionally holds great spiritual significance as the site of the Takitimu waka landing as well as the site of a school of sorts where people could come to spiritually connect and learn *karakia* (prayers and rituals for guidance and protection) from their ancestors.

Geologically, the peninsula is an incredibly active site. It is on the eastern edge of the Zealandian continental fragment, in an area that is being actively subducted [23], causing frequent earthquakes in the area that result in the peninsula growing higher and higher over time. The soil is rocky and prone to erosion, particularly following the clearing of dense forest during the introduction of farming to the area. When an earthquake does come, sometimes in the form of "slow slip" events where the earthquake happens over the course of weeks, it brings soil down into the water, which can be risky for the survival of intertidal species [20, 153]. The waters in the area are a mix, oceans that lead inland and turn to freshwater rivers and streams, teeming with clams, mussels, urchin, rock lobster, abalone, and others.

For Māori, ancestry and heritage is tied up not only in people, but in the nonhuman species and natural landmarks that they have developed with over the years [182]. These relationships are built on the reciprocity between humans and the land, forever indebted and bound to protect. The key species within an area which are culturally important are know as *taonga* species, and there is a concept of guardianship, *kaitiakitanga*, that exists within the whole community and particularly among the chosen *kaitiaki*, or guardians, who are appointed to protect these *taonga* so they can continue to thrive [183]. One of the many methods that Māori and *kaitiaki* have implemented to protect the waters has been through restrictions and special monitoring status for water bodies. Two common methods are a *rāhui*, which is a temporary restriction on accessing an area either for political control, to separate from something that is *tapu* (typically if someone died in an area), or more commonly for conservation purposes, and a *mātaitai*, which is a an area of water whose fishing is under the control of *tangata whenua*, who can set restrictions on the species and numbers that can be fished within an area. There are currently processes within the New Zealand Ministry for Primary Industries (who control the commercial fishing in the country) for iwi to apply to designate an area of water a *mātaitai* and consequently be allowed to have control over all of the non-commercial fishing, using *tikanga māori* (māori cultural

customs and practices based on mātauranga māori) as their guide for creating the bylaws.

### 5.3.2 Taking a species-centered approach to monitoring

I first visited the Mahia Peninsula with my colleague Avery, a fellow graduate researcher in the Media Lab who I have collaborated with on bioengineering project and educational programming, in October 2018 to learn more about the community's recent plans to develop the *mātaitai* and monitor the waters around the peninsula. The visit came just as the Mahia Maori Committee was finalizing its bylaws. The reason for designating the *mātaitai* was to protect the culturally important *taonga* species such as *karengo* (a type of seaweed), *pāua* (abalone), *kina* (sea urchin), and *koura* (rock lobster) among others that were being depleted due to overfishing, erosion, and climate change [147]. The main goal of this visit was in to learning more about how the community set their goals for the monitoring of their mātaitai and what more they hoped to do.

Figure 5-4: **Culturally important species)** (L-R) seaweed, sea urchin, abalone shell.

We attended community discussions at the *marae* which focused on the *tangata whenua*'s long history on the Mahia peninsula and the ways that the peninsula's beaches and intertidal waters have been used over time. Group discussions such as these, where there were multiple generations from the iwi sharing thoughts and insights both on the historical changes to the reefs together with their hopes for replenishment through implementing the mātaitai, provided level of detail and nuance that feels impossible to put in any technical document describing the bylaws. We were able to learn more about the community's current monitoring program and technical capacity with respect to visual monitoring approaches such as transect- and scuba-based, as well as watershed-related concerns due to changing land use (primarily to agriculture) which has caused increased sedimentation in the intertidal zone in recent years.

During the discussions, the community expressed a desire to to design a community-led marine monitoring program that used on-body and off-body sensing technology

(a) (b)

Figure 5-5: **Learning about the mātaitai.** Spending time on the reefs with (a) Wikitoria and (b) Sophie.

that prioritizes their values, goals, and vision for their marine ecosystems. We also spent time discussing methods of storing and analyzing the data that are generated and making sure that it was accessible locally within the Mahia community. We also mentioned the possibility of organizing *wānanga* (a traditional Māori method of generating and sharing knowledge and provides a space for garnering diverse, iterative, and culturally-sensitive community input [208]) for the program that focused on both Western-based electrochemical sensor systems together with traditional monitoring techniques and would be open to the entire Mahia community. The goal that was mentioned often was to collect data with the express purpose of designing a species reintroduction program for *paua* (abalone) from the peninsula that are now rare to find in the waters on the East Coast. There was, of course, a *kai* (a meal) as a break during the discussions and/or once we wrapped up, which gave more time for informal conversations with members of the iwi around these topics.

When not at the *marae*, Avery and I organized one-on-one meetings with community members, typically through day trips to the areas in their watershed that were most meaningful to them. These meetings allowed us to develop a better understanding of individual goals and interests for the monitoring, as well as to learn more about life on the peninsula then and now and how local and regional politics could impact the *mātaitai* and programs overall ((Fig. 5-5a, 5-5b). We saw childhood swim spots now almost entirely filled with sand only a few decades later, chitinous middens exposed on the beach from eroding soil that speak to the centuries of large Māori meals on the peninsula, sacred beaches where whales would wash up and a special *karakia* was performed to determine an appropriate burial location, as well as favorite overlooks

and diving spots. We spent a significant amount of time on the *mātaitai*, where we heard wave dynamics, tidal conditions, sedimentation, seaweed populations, and more on possible locations for monitoring and future transplantation.

One of the key insights on this initial trip was understanding that this monitoring program would be one that does not center human health or water quality directly, but the health of key *taonga* species as proxies for health (Fig. 5-4. This methodology is in line with the Cultural Health Index, a method first developed to assess freshwater streams and waterways within Aotearoa New Zealand by Gail Tipa and Laurel Teirney [214]. It was one of the first programs that the New Zealand EPA developed alongside Māori that prioritized the historical and cultural knowledge that the Māori have within environmental monitoring. The cultural assessment considers the long-standing spiritual and importance of a place, the historic activities that Māori did within these waters, most importantly the health and presence of culturally-important aquatic species within the area.

Specifically, the Cultural Health Index considers the health of waterways through the lens of (1) whether the site holds special importance within Maori culture, (2) traditional use of the waterway, and (3) whether it can still be used for traditional activities. It prioritizes cultural health indicators just as much as scientific health indicators, recognizing that the environmental monitoring is not isolated, so water monitoring necessarily involve monitoring the larger environment, from farming practices to erosion. This resonates with experiences on the *marae*, where *kaitiaki* made it clear that you need start at the highest point of the land and end in the sea when studying issues related to water. What is on land ends up in the waters, and if that is not incorporated into the monitoring and considerations of watershed planning, it will be impossible to get to the root of the problem.

Building a water monitoring program on understanding how changing the environmental conditions impact culturally important species in the area, with a goal of transplanting those species, moves the community from a mind set of depletion towards one of potential abundance. Centering these species also highlights the important ecological functions that keystone species perform [152], so taking a species-centered approach to monitoring and rehabilitation can have more of a significant impact in the overall health of an ecosystem then just focusing on water quality alone. In the words of Donna Haraway, "Co-constitutive companion species and co-evolution are the rule, not the exception" [90]. The approach that *mātauranga māori* and *tikanga māori* take by emphasizing the deep connection between the humans and natural world, be it nonhuman animals, plants, streams, or mountains, means that it is a priority to live in a way that honors and protects those connections and the reciprocal relationships.

### 5.3.3 Intergenerational sensing: developing the next generation of *kaitiaki*

After our first visit to the peninsula, the community was excited to try out water monitoring methods that integrated electronic and chemical sensing to add another dimension to their monitoring program (multi-modal sensing!). During a series of meetings with Paul Ratapu, the chairman of the Mahia Maori Committee, we discussed the evolving needs to of the community and how that should determine the type of sensors and events we would plan. We came to realize that it was best to have a sensor that could be easily moved to many locations to allow for a survey of the reefs and allow for residents to get hands-on experience with the probes. In addition, Paul was also interested in sensor systems that could allow for a more holistic understanding of what the sensor is doing, maybe having an output that was more descriptive than a digital output. I reached out to a fellow Media Lab PhD student, Laura Perovich, who focused her research on designing a remote-controlled boat with low-cost water sensors connected to programmed LED sensors that changed colors to provide real-time visual water quality outputs. [169].

In preparation, I received further funding to buy materials and assemble the sensor systems. The sensor for reef surveying used a set of Atlas Scientific sensor probes together with a Raspberry pi and a mountable 7" touch screen which could be easily calibrated and moved from location to location to take scientifically relevant and accurate measurements of temperature, dissolved oxygen, conductivity, and pH. The interactive probes for water monitoring education in collaboration with Laura Perovich was driven by an Adafruit feather that had probes for temperature and conductivity connection to the LEDs and housed in a waterproof sphere.

The second visit to Mahia happen took place in July 2019 and was framed as a set of *wānanga* led by *kaitiaki* and other community members. These *wānanga* were meant to be the first in a series of community-led prototyping activities, but due to the COVID-19 pandemic this was the only one that we were able to organize (the second set had been schedule for late March 2020). The sensors that were part of the first workshop included tablet-based kits for testing nutrient, pH, dissolved oxygen, and coliform bacteria, the analytical-grade probes for pH, dissolved oxygen, temperature, and conductivity, and the visual-based sensors that were attached to LEDs and changed color in response to changing water quality (Fig. 5-6).

The first *wānanga* focused on understanding the history of the *tangata whenua* on the peninsula and using these stories as a way to choose sites for water sampling later in the day. The participants ranged from 10-80 years old and even included a member of the Regional Council and an employee of the Department of Conservation who came to the peninsula specifically for the *wānanga*. The structure of the event allowed for those who had never been to the *marae* to be welcomed (which is required before any new person enters into the building), followed by a *karakia* from *kaumatua* (an elder in the community) once everyone was inside. After this, we spent time on intro-

Figure 5-6: **Sensors at the wānanga.**The chemical, probe, and LED based sensors set up on the table in the *marae*.


Figure 5-7: ***Kaumatua* leading the wānanga.** Sophie Dodd (standing) sharing thoughts during the morning session of the wānanga.

(a) (b)

Figure 5-8: **At the wānanga.**(a) Collecting samples on the mātaitai (b) Testing the water quality in the marae.

ductions, which involves stating iwi associations, heritage, connections to the Mahia peninsula, and the reason for attending the meeting. Once this was completed, the *kaumatua* led the conversation around changes to the coast and the intertidal and subtidal zones that they have seen over the decades, including customs and practices that their parents and grandparents implemented while they were children and explaining the purposes and meaning behind some of the practices as they pertained to protection of the marine life. They took time to try recalling or understanding the reasons for stopping some of the practices that were traditionally kept, which sometimes had to do with the New Zealand government's role as a regulator of primary industry and can also be attributed to the increase in demand for seafood such as abalone and sea urchin in the country.

We also discussed the changes to the local population, with many residents no longer *tangata whenua* and therefore not bound by *tikanga māori* as it pertains to the environment. Many of the *kaumatua* were over seventy years old, so their history on the peninsula stretched back to at least the 1940's, sometimes earlier. Following this discussion, we took a break for a *kai*. Following lunch, Avery and I gave a presentation on the sensors that we brought over and provided context for how the sensor data can relate to water and reef health. Following this, the community identified spots within the surrounding area where they thought the water sampling would be useful, and we left the *marae* (which is conveniently across the street from the ocean) to go collect the samples (Fig. 5-8a, 5-8b).

After returning with the water samples, we took probe-based readings, performed tablet-based colorimetric water monitoring tests for nitrates, phosphates, and coliform bacteria, and played around with the LED-based sensors using the temperature probes. After collecting and consolidating all the probe-based data for each water sample, we came together as a group and discussed the results for each site and make sense of it (Fig. 5-9a, 5-9b).

(a) (b)

Figure 5-9: **Collecting and discussing sensor data.**(a) Wikitoria using the color-changing sensor (b) Paul leading a discussion on the probe-based sensor results

The color-changing LED water sensors were the most exciting and engaging of all the sensors tried during the workshops. The immediate color change during readings made it easier to communicate monitoring principles with participants who may not find the topic very engaging. The sensors were also useful when discussing how one body of water can have areas with locally different water qualities depending on factors such as pipes, runoff sites, sun exposure, and plant life. Playing around with the LED sensors also naturally jump-started conversations by community members on locations they would want to put long-term monitoring probe and the types of results they might see there over time. I didn't expect it to be as useful for planning monitoring locations, because the sensor itself is not meant to be left out long-term due to its battery life, but the intuitive conceptual understanding it provided for participants and visual appeal made it easier to approach.

Between the first and second *wānanga*, Avery and I spent more time with community members discussing the changes within the community, taking a tour of the Whangawehi river clean-up effort, which is a fresh water restoration project centered on developing riparian zones, in the area, as well as surveying the reefs.

The next *wānanga* was also held at the *marae*, where we discussed data management and analysis methods for the sensor data and discussed changes that could be made to have future workshops run more smoothly. We then continued our brainstorming on areas that were best suited for more frequent monitoring with the probe-based sensors, as well as monitoring methods to prepare for a potential *paua* transplantation. The *kaumatua* attending gave their input on site selection, culminating in the selection of two sites in the *mātaitai* that are going to be sampled more frequently to prepare for a transplantation program.

Intergenerational monitoring was a priority for the *mātaitai* leaders and informed the

schedule of activities for the *wānanga* that we organized. For the *kaumatua*, their children and grandchildren are the inheritors the role of *kaitiaki* and must be prepared for when that time comes. It is even more critical now, as more iwi members are moving away from peninsula towards bigger cities within New Zealand and beyond, meaning that many of the children within the iwi do not have the type of lived experience and long-term monitoring knowledge that the older generations hold. However, as New Zealand has worked to right the historic wrongs of the Crown through acknowledging Māori *rangatiratanga* (right to authority), the younger generation has access to schooling that teach Te reo and embrace *kaupapa māori* principles, which allows them to proudly step into their heritage in a way that was not possible for previous generations.

The *kaumatua* have great hope that the future generations will be the ones to find balance between the worlds of science and tradition. They learn the ways of western science as well as *tikanga māori*, which provides them will a rich a varied set of tools that they can use in an effort to be better manage and protect their lands. One comment we heard at the second *wānanga* was to incorporate more monitoring and sensors into the curriculum in the local *kura kaupapa* (māori-led school), because if they do not learn how to manage their lands someone else will come in and do it for them, and not on their terms. Rather than a *pakeha* scientist parachuting into the community and dictating what is best for their lands, they hope for Māori scientists who can embody the best of both, cutting out the middle man. In her book *Braiding Sweetgrass*, Robin Wall Kimmerer described reaching a similar point within her journey in ecological studies:

> *There was a time when I teetered precariously with an awkward foot in each of two worlds—the scientific and the indigenous. But then I learned to fly. Or at least try. It was the bees that showed me how to move between different flowers—to drink the nectar and gather pollen from both.* ***It is this dance of cross-pollination that can produce a new species of knowledge, a new way of being in the world. After all, there aren't two worlds, there is just this one good green earth.***(p.47)" [113]

The structure of the first *wānanga*, where the *kaumatua* were able to share their experiences and the history of the waterways before the younger participants took the lead in sampling and testing the water, reflected the ways in which the community hope to keep the traditional ecological knowledge within the community add have a monitoring and management program that can have benefits for generations to come.

## 5.4 The Citizen Science of Showing Up

Both experiences with monitoring programs were hugely informative and meaningful as a technologist interested in understanding the inner workings of community-centered monitoring efforts. Outside of the change of scenery and break from the tedium that repetitive wet lab work can induce, there were also insights that I gain from these experiences that would have been difficult to get in any other way. Some of them seemed obvious to me, while others I wasn't necessarily cognisant of until I found myself in meetings with my bioengineering colleagues, disagreeing with their suggestions over a user-related concern. Below are some of the major takeaways that I would like to highlight.

1. **You don't need a lot of data, but you do need consistent, trustworthy data**. Data for data's sake is not very useful, particularly because people are not necessarily looking at it (or at least not immediately). What is important is the trust that people (monitors, community, municipality, etc.) have in the data and the program and in the consistency inherent in the measurements. Data is not created in a vacuum, and it is not much use unless it can be acted upon in some way. This requires trust in the measurements, in the people who are measuring, and in the analysis of the measurements. This trust needs to work in both directions, from the monitors to the program managers and the managers to the monitors, to make sure that the methodologies used for monitoring are appropriate for the program and that they are being correctly implemented in the field. Depending on who this data is meant for – communities, government organizations, universities, etc. – it will be important to effectively communicate these procedures and create materials or databases that are easy to access and interpret. Different monitors and programs will want to access and understand the data differently – be in conversation as a group on the cadence that seems to be best fit for the program. If participants feel that the overall data collected by the program is quality, it can motivate them to keep collecting quality data and participating in the program, even if the utility is not immediately apparent (which it often isn't for long-term monitoring).

2. **Multi-modal monitoring can be more than the sum of its parts**. Bringing together data streams, especially qualitative and quantitative data streams, can provide a more in-depth picture than collecting streams of data in isolation. It provides contextualization, which can make data analysis and interpretation easier. However, it is worth taking the time to explain to citizen scientist/monitors what the point of the different data streams are and how they are logged/taken into consideration, to make sure that there is a clear understanding of how to collect the data and what it is used for.

3. **Short-term and long-terms data can both be important.** It is helpful to take time discussing both short-term and long-term monitoring goals, as well as the sensors and methods that would work best to serve reach them. It's possible two different tools or methods are needed to produce data at those two different resolu-

tions, because there is no one-size-fits-all approach to data collection. For instance, I can imagine a world where a CRISPR biosensor (which provides qualitative yes/no data) can be useful for for short-term or ad-hoc monitoring applications, while in the same monitoring session a sample can be taken and sent to the MWRA laboratories for traditional culture-based assay which is useful for a long-term data set. Both provide data about the water, one in an hour and the other in two days, and there are reasons and benefits to having both. Don't fall into the trap that a monitoring program should only do things one way. However, it is also important to be mindful of the capacity of the participants in the programs and make sure that all are aligned on what is being collected, why, and how.

4. **Monitors like devices, but not complication.** If you are going to have volunteers participating in data collection or analysis, it is best to try and minimize steps between sample and answer. This gives monitors further confidence that there aren't errors being generated (either by them or other participants) that are impacting the quality of the data set (which goes back to point 1) and gives the program the best chance of collecting error-free data. I have seen an increase in confidence when electronic-based sensors are used in a monitoring process from both of the communities that I learned from. While initially a bit disappointed (I am uncomfortable with the ways in which people defer to outputs of electronics and technology when it is placed in their hands), it is also important to acknowledge that there is something exciting about a new tool that has technical abilities we don't carry ourselves. People also just like cool gadgets! Many of the people who volunteer to participate are excited about new monitoring tools and want to try them out. Finding ways to incorporate tools that are easy to use and produce reliable data can be a motivating factor for volunteers. Having sensors that perform automated or electronic analysis also takes some pressure off of the monitor and make them feel like there are other capable things sharing the burden of data collection and processing, though I would argue in a way that keeps more agency over the data and "scientific" process with the monitor than a method where samples are sent to a lab for analysis.

5. **Controls and replicates when possible!** It's good practice to collect replicates (in biology, triplicate is standard) and to have positive and negative controls. Sometimes the kit you're using is a dud! Sometimes things get contaminated! Sometimes the sample you collected is not – as hard as you tried – representative. For programs that are only collecting a few data points a year (like monthly programs) it is worth a bit of extra effort to make sure that the results of you data are coming from the site itself and not from a broken sensor or a bad batch of kits. Having sanity checks, particularly when readings are not what was expected, are helpful when the goal for monitoring is understanding fluctuations and changes. This is especially critical if the data set is to be used for policy decisions.

6. **Visual and experiential aids whenever possible.** Visual aids are typically useful for demonstrations regardless of the educational endeavor. Want to explain a monitoring or scientific concept to a group of volunteers? Visual aids can be hugely

helpful is trying to communicate concepts that otherwise might be a bunch of unfamiliar words, equations, or lines of code. Want to make an easy sensor for anyone to read? Try using a visual-based output (colorimetric or fluorescent) to help guide intuition on the result (and avoid tests where the results are producing non-intuitive outputs!). Want to teach a group of monitors how to use a new tool? Plan a workshop and go through it step-by-step and providing written instructions for them to reference after.

7. **People belong in monitoring programs!** As you might gave gathered from the points mentioned above, I believe that people absolutely belong in monitoring programs! I admit I believed this before I started working with these communities, but now I believe it even more strongly. Science is in the business of developing tools, but it is people who direct those tools. Involving locals and incorporating traditional and historic knowledge into a monitoring program is a great way of learning what or where to direct these tools towards, because often programs are constrained by budget and equipment. Additionally, monitoring programs are not only collecting data sets (which is, to me, the boring part), they are also generating informed, engaged, and skilled citizens (who are all walking sensors in their own right), who have developed (or are developing in the case of monitors who are assigned sites that are outside of their local area) a nuanced and intimate knowledge of their surroundings just by going about their daily lives. Biosensors and microbes are ecological sentinels – and humans can be too! We all can do our part. Someone paying close attention, even for ten minutes, in a place once a month seems like a very small thing – but it's not! That's what it takes! That's what it's all about!

8. **Monitor the monitors.** The citizen scientists and monitors within a program are collecting a lot of extra data – the overspill data that was discussed previously. It behooves a program to find ways of accessing that data, be it through a periodic check-in at a local meeting, a form that's sent out, a running log that people have access to, a paper on a bulletin board, whatever works best for the monitors. Regardless of the method, don't sleep on that data! Checking in with monitors is a way of not only learning how the program is going, but also whether there are any adjustments that may be needed with the methodology or practices.

9. **Automated monitors are not the enemy!** While I call into question the utility for the endless data streams that technology promises us, automated and latent sensing can be tremendously useful for environmental monitoring, particularly programs that do not have access to volunteer or who work in locations where it is difficult to collect samples. Collecting baseline data on water quality (such as pH, temperature, dissolved oxygen, turbidity, and conductivity) with sophisticated or or even DIY sensors can help contextualize a watershed and the data that comes out of it. While more is not always more (as discussed previously) with data, there is a balance that can be struck between automated and in-person monitoring methods. I wrote this chapter primarily because I fear there may be an *over-reliance* on automated data collected through latent sensing methods, which I find to be lacking in the type of multi-dimensionality that is available for in-person citizen and participatory

monitoring. In their work, sociologist speak about developing "thick description" (a phrase coined by Gilbert Ryle and made popular by Clifford Geertz) within their areas of study – I find that sensor data is fine, but it tends to be thin. However, I *do* think that an integrated methodology, which can combined automated monitoring with in-person measurements and observations, can allow for the best of both world. Citizen and participatory monitoring provide sanity-checks for automated sampling and together, these methods can allow for high-quality and rich data streams.

# Chapter 6

# Next Steps

It seems as if is never enough time to accomplish what one sets out to do. It feels doubly true when living through a time where being around others is considered a high-risk endeavor. In the summer of 2019, I hoped to work on a dissertation that went from the lab to the community and back, designing and testing sensing technologies that were useful to people who did not have laboratory backgrounds and who wanted to get a better sense of the types of organisms within their environment. What I got was something a bit different – albeit something that I am grateful for and immensely proud of. COVID-19 was an unfortunate wake up call which emphasized how inadequate and unsustainable centralized testing is for pathogens detection and microbial monitoring. It also forced our small team of researchers to work in ways that were faster, more focused, and more collaborative than I could have ever dreamed of. If I could continue my PhD alongside my current lab mates for another six years, I can't even imagine how much more we could do. There were so many lessons learned through this process and so much that can be done to improve on this work. I hope that this dissertation provides a strong conceptual and technical foundation for others to use for their biosensing and monitoring efforts.

## 6.1 Next Steps

I imagine two different use cases for the two CRISPR biosensor technologies that were described in this dissertation. In both cases, it would be interesting to enable semi-quantitative or quantitative results for the assays. In many of the use cases, such as pathogen monitoring, a qualitative yes/no result is sufficient. However, there are some pathogens of interest where pathogenic load is relevant, therefore it would therefore be useful to know the relative abundance of the pathogen of interest. This would require modifications to the amplification steps for either reaction, which is what currently allows for the high sensitivity of these assays but complicates quantitation.

With regard to miSHERLOCK described in Chapter 2 and 3, it would be of great

benefit to design versions of these sensors that could perform without the use of any heating elements, and even better if the assay output does not require any electronic components in order to be read. This could be done through modifying the amplification methods and attempting to develop colorimetric assays (ones that are at least as sensitive as what we can generate through fluorescent detection). These sensors could be useful as a more sensitive and specific alternative to a lateral flow diagnostic, one that can be simpler to design and manufacture because it relies on nucleic acids rather than antigen-antibody pairs and performed quickly and easily by a user before being disposed.

With regard to MELODe described in Chapter 4, it would be of great benefit to determine whether the electrode surface could be washed and the PNA unbound from the cleaved product after the assay is run, which could allow for a reusable electrochemical biosensor. This could bring the cost down on the system and differentiate from current single use assays available. Versions of the lab-on-a-chip microfluidic system that implement a large array of these electrochemical sensors could also open the door to an automated sensing system that has the capability to perform latent monitoring. It could be a system set up that is an early sentinel for future pathogens, as samples like city wastewater have been shown during the COVID-19 pandemic to be a strong predictor of disease spread. It would also be a welcome addition to current large-scale monitoring and sensing projects which often do not have the ability to sense biological species as a measurement. For example, the Living Observatory at Tidmarsh has embedded sensor nodes to collect audio, visual, and sensor data on environmental conditions, but does not yet have biosensing capabilities. While there are plans to deploy a machine-learning powered analysis and visualization of microscopic aquatic organisms such as plankton([241], there have been no plans for nucleic acid sensors. The array of MELODe electrochemical sensors could be a promising option to allow for microfluidic-based microbial and nucleic acid monitoring at a remote site such as Tidmarsh.

With regards to participatory engagements with the Mystic water monitors and Mahia Māori Committee, it would be of great benefit to leverage the ethnographic understanding of the water programs and the data streams that are of interest to them to further tailor biosensor technology. Ideally, this would include at least one workshop that explains how the biotechnology works and provide physical prototypes that the participants can try out to get a feel for the devices. There could be focused discussions on the current and potential features of the device, such as data logging, power and batteries, size, accessibility, and legibility, to make sure that the device can be the best fit for the community's needs and use cases. There could then be a follow-up workshop where the finalized prototype is demonstrated to participants and used on samples that are brought by participants or collected during the workshop. Ideally, the devices used at the workshop would stay with the community, along with a written and/or interactive manual, where they can continue to try it out and collect data with it. Following this, there can be another set of semi-structured interviews with participants to understand how they felt during the workshop, how they feel about the

sensing modality and sensor development process, ways in which they would change the design process or the the device itself if given the opportunity, and thoughts on any use cases for the device within their daily lives. It would also be interesting to understand how these novel biosensors integrate into existing monitoring regimes that these programs implement.

## 6.2 Lessons Learned

There is so much that I learned both in the wet lab and in the field that has improved the work that I do in both environments. First and foremost, things take time – sometimes for reasons you expect, and other for reasons you can't prepare for (pandemics, broken equipment, reagent delays, discontinued products, etc.). Make sure you allow yourself enough time to accomplish your goals, but more importantly, make sure that if you are going to be bringing materials to a workshop that you have enough time to prepare everything! Community events can take months to organize, the last thing you want is to not have materials available when the event arrives! If possible, don't aim to begin planning a workshop until all of the systems intended for it are designed and tested (though this is a large and often unrealistic ask)! It is a way to respect both the community and your work.

Be open to the changes that come with the research process. Even with lots of plans made, there are opportunities and collaborations that arise that might be worth veering off-course for. I did not expect to get as involved as I did in the electrochemical-based CRISPR sensors, but I followed my interest in lab-on-a-chip design and was able to design and validate a sensing system that, although not in the initial scope of my PhD, was an incredible learning experience that allowed me to re-think the types of biosensing systems that could be deployed. At the same time, if you don't make time for the projects you are passionate about, no one else will! I have been interested in water monitoring for many years and it took time and effort to continually bring myself back to that work during the pandemic-related research, but I am so happy I did.

Seek out new skills from people who can teach you, even if they might not seem immediately relevant. Learning a new skill during my PhD studies kept me inspired and prevented me from reaching a plateau state, skill-wise. Taking time to improve my computational modeling and soldering skills for the probe-based sensors and initial prototyping with early iterations of the microgravity biosensing system prepared me to work quickly once we set our sights on a miSHERLOCK device. Experience in electronics that I picked up from my lab mates and TA-ing classes in the Responsive Environments group provided me with necessary background that I was able to activate for the eRAPID-related electrochemical sensors. I never regretted taking time to learn new skills and was surprised to see how immediately they were useful for projects .

Collaboration is key! I can't imagine being able to accomplish even a small piece of what I did without all of the lab mates and collaborators who worked on these projects. To perform rigorous molecular biology requires huge amounts of time, effort, and experimentation. Having collaborators to help shoulder the burden, troubleshoot issues, provide expertise, and push it to the finish line is sometimes the difference between a project that is published versus one that sits on the shelf indefinitely. It is a point of great pride that I am co-first author on both of the published papers that are in this dissertation – I strongly believe in celebrating the ways that multiple researchers bring their expertise to a single project and hope that others will do so as well!

# Appendix A

# miSHERLOCK

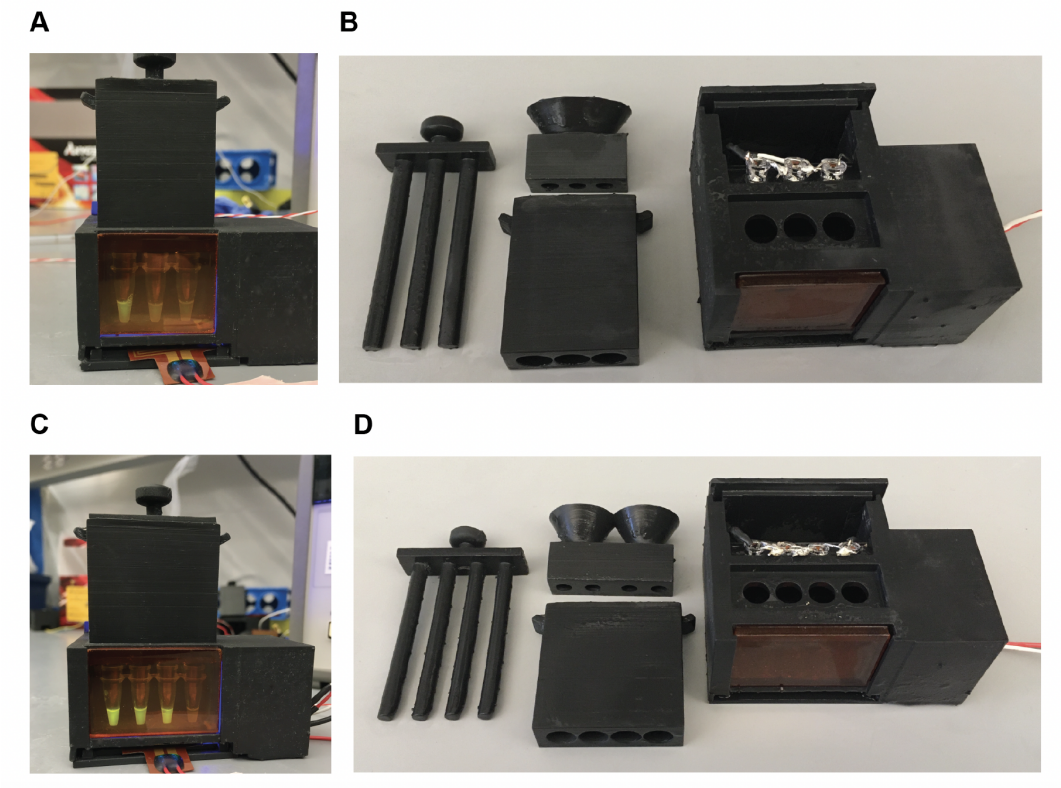



Figure A-1: **miSHERLOCK triplex and quadruplex 3D printed devices.** A. An assembled 3D printed triplexed miSHERLOCK device with a representative reaction of SARS-CoV-2 negative saliva showing one positive test (RNaseP positive control) and two negative tests (SARS-CoV-2 NP and N501Y). B. The components of the triplexed miSHERLOCK reader (L-R): plunger, saliva collector, sample preparation column, reaction chamber with heater, LEDs, and orange acrylic transilluminator filter. C. An assembled 3D printed quadruplexed miSHERLOCK device with a representative reaction showing three positive samples and one negative sample. D. The components of the quadruplexed miSHERLOCK reader (L-R): plunger, saliva collector, sample preparation column, reaction chamber with heater, LEDs, and orange acrylic transilluminator filter.

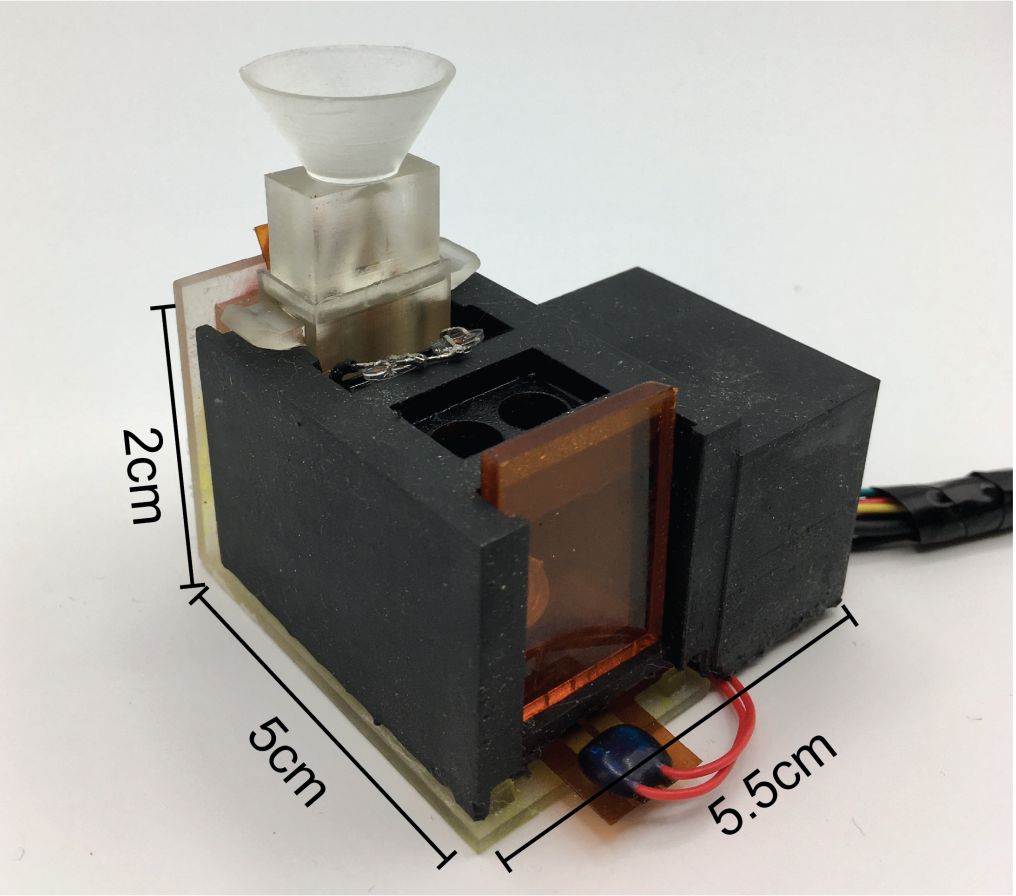
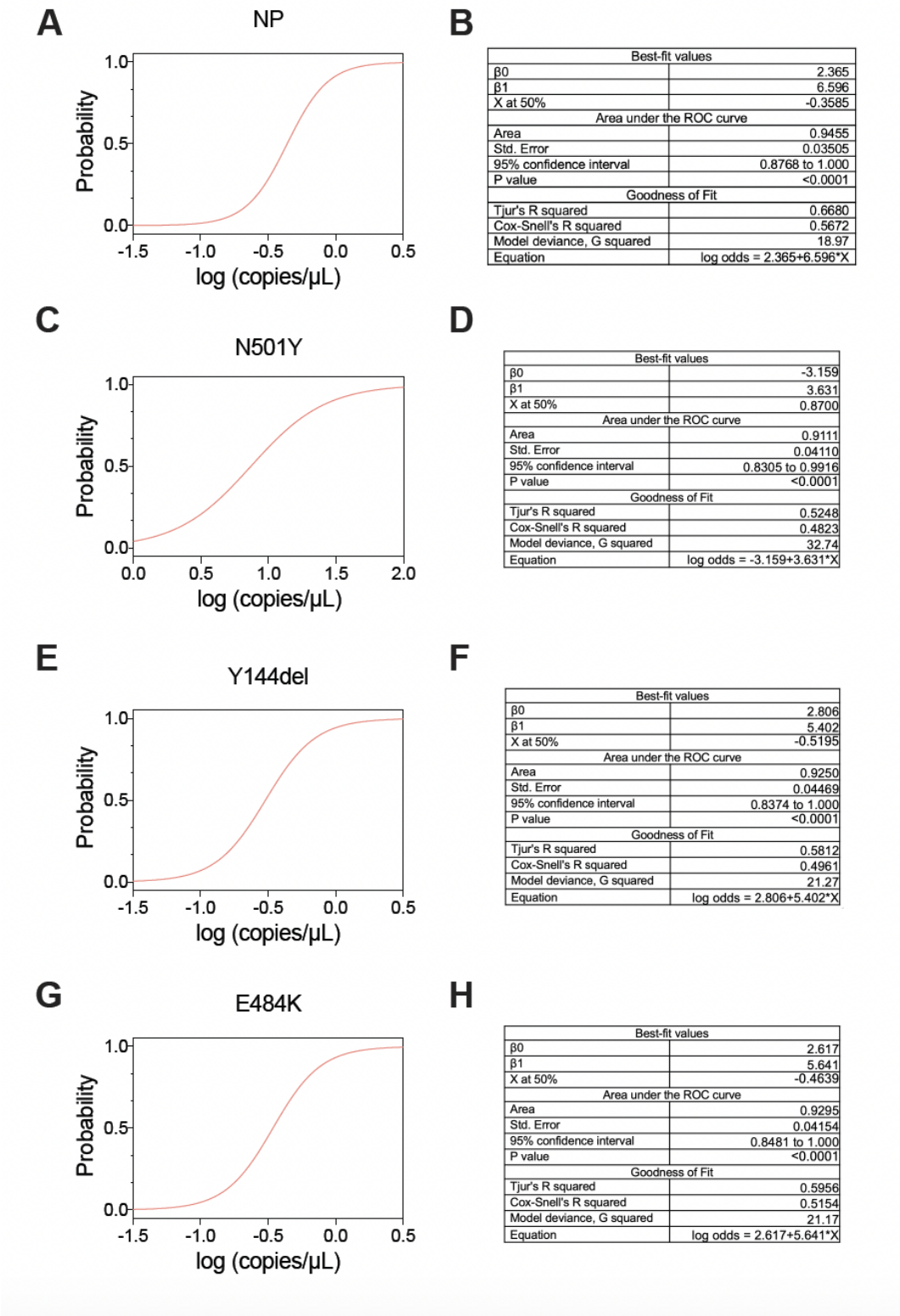



Figure A-2: **Oblique view and measurement of miSHERLOCK device before use.**


Figure A-3: **Adjunctive mobile phone application supports automated quantitation and result interpretation.**




| Best-fit values | |
|---|---|
| β0 | 2.365 |
| β1 | 6.596 |
| X at 50% | -0.3585 |
| **Area under the ROC curve** | |
| Area | 0.9455 |
| Std. Error | 0.03505 |
| 95% confidence interval | 0.8768 to 1.000 |
| P value | <0.0001 |
| **Goodness of Fit** | |
| Tjur's R squared | 0.6680 |
| Cox-Snell's R squared | 0.5672 |
| Model deviance, G squared | 18.97 |
| Equation | log odds = 2.365+6.596*X |

| Best-fit values | |
|---|---|
| β0 | -3.159 |
| β1 | 3.631 |
| X at 50% | 0.8700 |
| **Area under the ROC curve** | |
| Area | 0.9111 |
| Std. Error | 0.04110 |
| 95% confidence interval | 0.8305 to 0.9916 |
| P value | <0.0001 |
| **Goodness of Fit** | |
| Tjur's R squared | 0.5248 |
| Cox-Snell's R squared | 0.4823 |
| Model deviance, G squared | 32.74 |
| Equation | log odds = -3.159+3.631*X |

| Best-fit values | |
|---|---|
| β0 | 2.806 |
| β1 | 5.402 |
| X at 50% | -0.5195 |
| **Area under the ROC curve** | |
| Area | 0.9250 |
| Std. Error | 0.04469 |
| 95% confidence interval | 0.8374 to 1.000 |
| P value | <0.0001 |
| **Goodness of Fit** | |
| Tjur's R squared | 0.5812 |
| Cox-Snell's R squared | 0.4961 |
| Model deviance, G squared | 21.27 |
| Equation | log odds = 2.806+5.402*X |

| Best-fit values | |
|---|---|
| β0 | 2.617 |
| β1 | 5.641 |
| X at 50% | -0.4639 |
| **Area under the ROC curve** | |
| Area | 0.9295 |
| Std. Error | 0.04154 |
| 95% confidence interval | 0.8481 to 1.000 |
| P value | <0.0001 |
| **Goodness of Fit** | |
| Tjur's R squared | 0.5956 |
| Cox-Snell's R squared | 0.5154 |
| Model deviance, G squared | 21.17 |
| Equation | log odds = 2.617+5.641*X |

Figure A-4: **Probit curves for miSHERLOCK SARS-CoV-2 assays.** Sensitivity of miSHERLOCK diagnostic for universal SARS-CoV-2 (nucleoprotein or NP assay), (A-B), N501K (C-D), Y144del (E-F), and E484K (G-H) mutation detection by comparison of probit regression curves and fit characteristics.

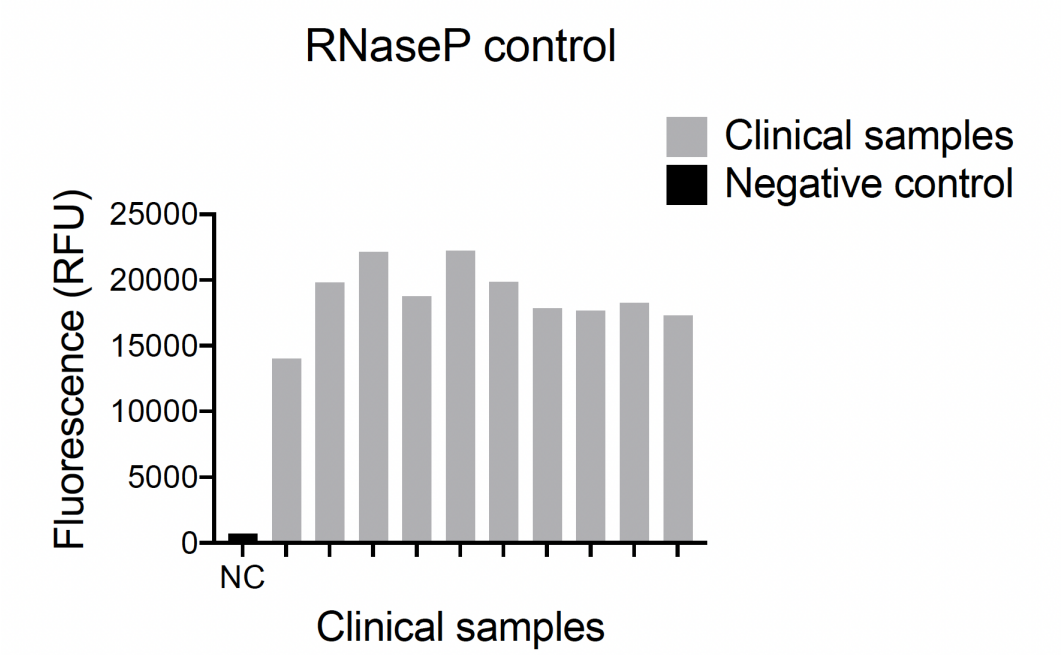



Figure A-5: **Clinical saliva samples tested with miSHERLOCK Human RNaseP assay.** The RNaseP assay performed well when tested with clinical saliva samples. NC indicates negative control (water only, as there were no non-human saliva samples available).

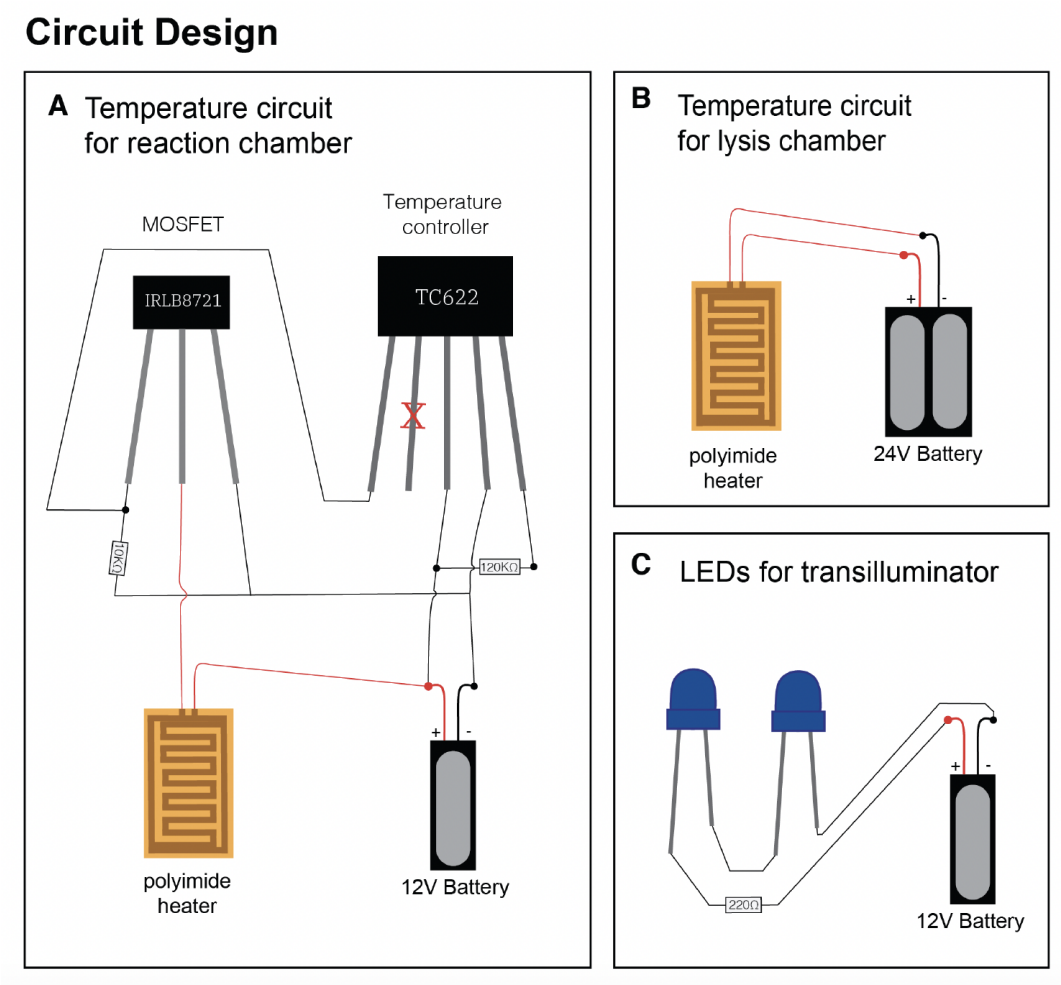



Figure A-6: **Circuit diagrams for temperature and LED control.** A. The reaction chamber temperature circuit consists of a MOSFET transistor that is attached to the polyimide heater and the trip-point temperature sensor. The temperature sensor is programmed to 37℃ by placing a 120 kΩ resistor between pins 3 and 5. The heater and trip-point sensor are both attached to a 12 V battery source. B. The lysis chamber temperature circuit consists of a polyimide heater attached to a 24 V battery source. There is no additional temperature control needed, because the polyimide heater was shown to maintain the high temperature necessary to lyse viral particles and denature potential nucleases with an appropriate voltage source alone. C. The LED circuit consists of two royal blue LEDs in series attached to a 12 V voltage source by a 270 Ω resistor to keep the current within the optimal range of 20 mA.

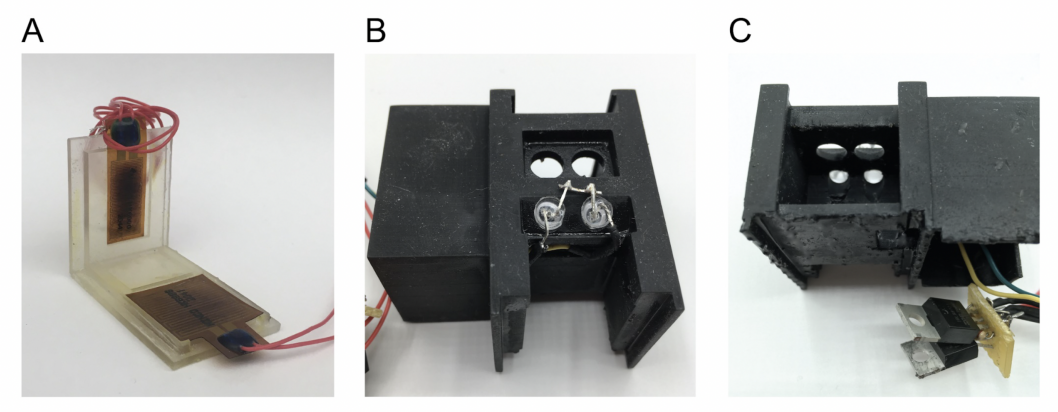



Figure A-7: **MiSHERLOCK electronics integrated into a 3D printed device.** A. An assembled 3D printed reusable temperature control housing containing two heaters for the two distinct temperature zones needed. B. The optics housing contains two LEDs to excite the fluorophores released in the SHERLOCK reaction. C. The assembled temperature controller circuit beneath the electronics box provides temperature regulation for the reaction chamber of the miSHERLOCK platform.

Table A.1: Genomic analysis of gRNAs and RPA primers used to detect key SARS-CoV-2 variants.

| **Sequences** | **All (NCBI 43k genomes** | **B.1.1.7 (GISAID 50k genomes)** | **Short heading** | **B.1.351 (GISAID 577 genomes)** | **P.1 (GISAID 78 genomes)** |
|---|---|---|---|---|---|
| NP gRNA | Universal | 90.7% | 99.7% | 100% | 100% |
| NP RPA Forward | Universal | 94.7% | 99.8% | 99.5% | 100% |
| NP RPA Reverse | Universal | 97.0% | 0.06% | 100% | 100% |
| Y144del gRNA | B.1.1.7 | 0.33% | 98.0% | 0.00% | 0.00% |
| Y144del F2 | B.1.1.7 | 99.2% | 99.7% | 100% | 100% |
| Y144del R2 | B.1.1.7 | 99.7% | 99.8% | 100% | 100% |
| E484K gRNA | B.1.351, P.1 | 78.3% | 99.8% | 100% | 100% |
| E484K F2 | B.1.351, P.1 | 99.6% | 99.8% | 100% | 100% |
| E484K R1 | B.1.351, P.1 | 99.1% | 98.8% | 100% | 100% |
| N501Y gRNA6 | B.1.1.7, B.1.351, P.1 | 0.49% | 99.8% | 98.8% | 100% |
| N501Y F8 | B.1.1.7, B.1.351, P.1 | 98.7% | 98.7% | 0.17% | 0.00% |
| N501Y R1 | B.1.1.7, B.1.351, P.1 | 97.9% | 99.6% | 95.3% | 100% |

Table A.2: Oligonucleotide sequences used in this study.

| **Name** | **Sequence (5' to 3')** |
|---|---|
| SARS-CoV-2 Nucleoprotein Fwd Primer (44) | CGGCAGTCAAGCCTCTTCTCGTTCCTCATC |
| SARS-CoV-2 Nucleoprotein Rev Primer (44) | CAGACATTTTGCTCTCAAGCTGGTTCAATC |
| N501Y Fwd Primer | GGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATC |
| N501Y Rev Primer | TTTAGGTCCACAAACAGTTGCTGGTGCATGTAGAAGTT |
| Y144 Fwd Primer | AAGACCCAGTCCCTACTTATTGTTAATAACGC |
| Y144 Rev Primer | AAAGTGCAATTATTCGCACTAGAATAAACTCTGAACTC |
| E484K Fwd Primer | CCTTTTGAGAGAGATATTTCAACTGAAATCTAT |
| E484K Rev Primer | ACCATATGATTGTAAAGGAAAGTAACAATTAAAAC |
| Human RNaseP Fwd Primer | GGAGACAGCCGCTCACCTTGGCTATTCAGTTG |
| Human RNaseP Rev Primer | GAAGAAGTTGCTCTCAAAACATTGCAGTGAGATGG |
| Nucleoprotein gRNA | UAAUUUCUACUAAGUGUAGAUUUGAACUGUUGCGACU ACGU |
| N501Y gRNA | UAAUUUCUACUAAGUGUAGAUCAACCCACUUAUGG UGUUGG |
| Y144del gRNA | UAAUUUCUACUAAGUGUAGAUUGUUUUUGUGGUAAAC ACC |
| E484K gRNA | UAAUUUCUACUAAGUGUAGAUACACCAUUACAAGGUG UGCU |
| Human RNaseP gRNA | UAAUUUCUACUAAGUGUAGAUCCAAUUGUACAGGGA AAAUC |
| Reporter FQ Reporter | 6-FAM/TTATT/IABkFQ |

Continued on next page

| Name | Sequence (5' to 3') |
|---|---|
| N501Y-Mutant gblock | GCGCTAATACGACTCACTATAGGGAGGTTGGTGGTAAT<br>TATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTC<br>AAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCA<br>GGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTA<br>ATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCC<br>ACTTATGGTGTTGGTTACCAACCATACAGAGTAGTAGT<br>ACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTT<br>GTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAA<br>TGTGTCAATTTCAACTTCAATGGTTTAACAGGCACAGG<br>TGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCC<br>AACAATTTGGCAGAGACATTGATGACACTACTGATGCT<br>GTCCGTGATCCACAGACACTTGAGATTCTTGACATTAC<br>ACCATGTTCTTTTGGTG |
| Y144del-Mutant gblock | GCGCTAATACGACTCACTATAGGGTTGCTTCCACTGAG<br>AAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTAC<br>TTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAA<br>CGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAAT<br>TTTGTAATGATCCATTTTTGGGTGTTTACCACAAAAAC<br>AACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTC<br>TAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGC<br>CTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTC<br>AAAAATCTTAGGGAATTTGTGTTTAAGAATATTGATGG<br>TTATTTTAAAATATATTCTAAGCACACGCCTATTAATTT<br>AGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAAC<br>CATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGT<br>TTCAAACTTTACTTG |
| E484K-Mutant gblock | GCGCTAATACGACTCACTATAGGGTTAGAGGTGATGAA<br>GTCAGACAAATCGCTCCAGGGCAAACTGGAACGATTGC<br>TGATTATAATTATAAATTACCAGATGATTTTACAGGCTG<br>CGTTATAGCTTGGAATTCTAACAATCTTGATTCTAAGG<br>TTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGA<br>AGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACT<br>GAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGT<br>TAAAGGTTTTAATTGTTACTTTCCTTTACAATCATATG<br>GTTTCCAACCCACTTATGGTGTTGGTTACCAACCATAC<br>AGAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACC<br>AGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGG<br>TTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTA<br>ACAGGCACAGGTGTTC |
| Continued on next page | |

| Name | Sequence (5' to 3') |
|---|---|
| N501N-Wildtype gblock | GCGCTAATACGACTCACTATAGGGAGGTTGGTGGTAAT<br>TATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTC<br>AAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCA<br>GGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTA<br>ATTGTACTTTCCTTTACAATCATATGGTTTCCAACCCAC<br>TAATGGTGTTGGTTACCAACCATACAGAGTAGTAGTAC<br>TTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGT<br>GGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATG<br>TGTCAATTTCAACTTCAATGGTTTAACAGGCACAGGTG<br>TTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAA<br>CAATTTGGCAGAGACATTGCTGACACTACTGATGCTGT<br>CCGTGATCCACAGACACTTGAGATTCTTGACATTACAC<br>CATGTTCTTTTGGTG |
| Y144-Wildtype gblock | GCGCTAATACGACTCACTATAGGGTTGCTTCCACTGAG<br>AAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTAC<br>TTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATA<br>ACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAA<br>TTTTGTAATGATCCATTTTTGGGTGTTTATTACCACAA<br>AAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTT<br>TATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTC<br>TCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTA<br>ATTTCAAAAATCTTAGGGAATTTGTGTTTAAGAATATT<br>GATGGTTATTTTAAAATATATTCTAAGCACACGCCTATT<br>AATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTT<br>AGAACCATTGGTAGATTTGCCAATAGGTATTAACATCAC<br>TAGGTTTCAAACTTTACTTG |
| E484E-Wildtype gblock | GCGCTAATACGACTCACTATAGGGTTAGAGGTGATGAA<br>GTCAGACAAATCGCTCCAGGGCAAACTGGAACGATTGC<br>TGATTATAATTATAAATTACCAGATGATTTTACAGGCTG<br>CGTTATAGCTTGGAATTCTAACAATCTTGATTCTAAGG<br>TTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGA<br>AGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACT<br>GAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGT<br>TgAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGG<br>TTTCCAACCCACTTATGGTGTTGGTTACCAACCATACA<br>GAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACCA<br>GCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGT<br>TAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAA<br>CAGGCACAGGTGTTC |
| Continued on next page | |

| **Name** | **Sequence (5' to 3')** |
|---|---|
| Human RNaseP gblock | GCGCTAATACGACTCACTATAGGGATGGGACTTCAGCA<br>TGGCGGTGTTTGCAGATTTGGACCTGCGAGCGGGTTC<br>TGACCTGAAGGCTCTGCGCGGACTTGTGGAGACAGCC<br>GCTCACCTTGGCTATTCAGTTGTTGCTATCAATCATATC<br>GTTGACTTTAAGGAAAAGAAACAGGAAATTGAAAAAC<br>CAGTAGCTGTTTCTGAACTCTTCACAACTTTGCCAATT<br>GTACAGGGAAAATCAAGACCAATTAAAATTTTAACTAG<br>ATTAACAATTATTGTCTCGGATCCATCTCACTGCAATGT<br>TTTGAGAGCAACTTCTTCAAGGGCCCGGCTCTATGATG<br>TTGTTGCAGTTTTTCCAAAGACAGAAAAGCTTTTTCAT<br>ATTGCTTGCACACATTTAGATGTGGATTTAGTCTGCAT<br>AACTGTAACAGAGAAACTACCATTTTACTTCAAAAGAC<br>CTCCTATTAATGTGGCGAT |

Table A.3: Total cost analysis of miSHERLOCK duplexed diagnostic device and assay.

| **Sequences** | **Unit Cost ($)** | **Quant -ity** | **Cost ($)** | **Vendor** | **Product Number** |
|---|---|---|---|---|---|
| **Housing** | | | | | |
| PLA filament | 20.99 | 0.055 | 1.15 | Prusa | FLM-PLA-175-BLK |
| Polyimide heater (2pc) | 0.2 | 2 | 0.40 | Alibaba | |
| Transparent orange acrylic | 5.49 | 0.01 | 0.05 | Inventables | 24164-03 |
| Battery holders (12V, 9V) | 0.13 | 3 | 0.39 | Alibaba | |
| PES membrane | 0.02 | 2 | 0.04 | Millipore | GPWP04700 |
| **Temperature Sensor** | | | | | |
| N-channel MOSFET | 1.02 | 1 | 1.02 | Digikey | IRLB8721PBF-ND |
| 10k Ohm resistor | 0.1 | 1 | 0.10 | Digikey | S10CACT-ND |
| Temperature Sensor | 1.95 | 1 | 1.95 | Digikey | TC622VAT-ND |
| 120k Ohm resistor | 0.1 | 1 | 0.1 | Digikey | S120KCACT-ND |
| Protoboard | 7.5 | 0.03 | 0.23 | Digikey | 438-1037-ND |
| **LED** | | | | | |
| Super Bright Blue 5mm LED | 0.32 | 2 | 0.64 | Adafruit | 301 |
| 220 Ohm resistor | 0.1 | 1 | 0.10 | Digikey | S220CACT-ND |
| **Biochemical Assay** | | | | | |
| RPA Basic pellet | 327 | 1 | 3.40 | TwistDx | TABAS03KIT |
| Cas12a | 250 | 5 pm | 0.63 | New England Biolabs | M0653T |

Continued on next page

| **Sequences** | **Unit Cost ($)** | **Quant -ity** | **Cost ($)** | **Vendor** | **Product Number** |
|---|---|---|---|---|---|
| Reverse Transcrip-tase | 656 | 250 units | 4.10 | New England Biolabs | M0368X |
| RNase H | 293 | 2.5 units | 0.73 | Ambion | AM2293 |
| gRNA, reporter, buffer | | | 0 | | |
| **Total** | **$15.03** | | | | |

Table A.4: Cost estimate of miSHERLOCK duplexed diagnostic device components at 10,000+ scale.

| **Sequences** | **Unit Cost ($)** | **Quant -ity** | **Cost ($)** | **Vendor** | **Product Number** |
|---|---|---|---|---|---|
| **Temperature Sensor Circuit** | | | | | |
| N-channel MOSFET | 0.144 | 1 | 0.144 | Digikey | TSM240N03CX6RFGTR |
| 10K Ohm resistor | 0.055 | 1 | 0.055 | Digikey | 2019-RK73H2BTTD1002FTR |
| Temperature Sensor | 0.580 | 1 | 0.580 | Digikey | 2156-LM26CIM5-YPE/NOPB |
| 120K Ohm resistor | 0.055 | 1 | 0.055 | Digikey | 2019-RK73H2BTTD1203FTR |
| **LED** | | | | | |
| Blue 5mm LED | 0.148 | 2 | 0.296 | Digikey | 754-1807 |
| 220 Ohm resistor | 0.055 | 1 | 0.055 | Digikey | 2019-RK73H2BTTD2200FTR |
| **Total** | **$1.19** | | | | |

# Appendix B

# Multiplexed Electrochemical sensors

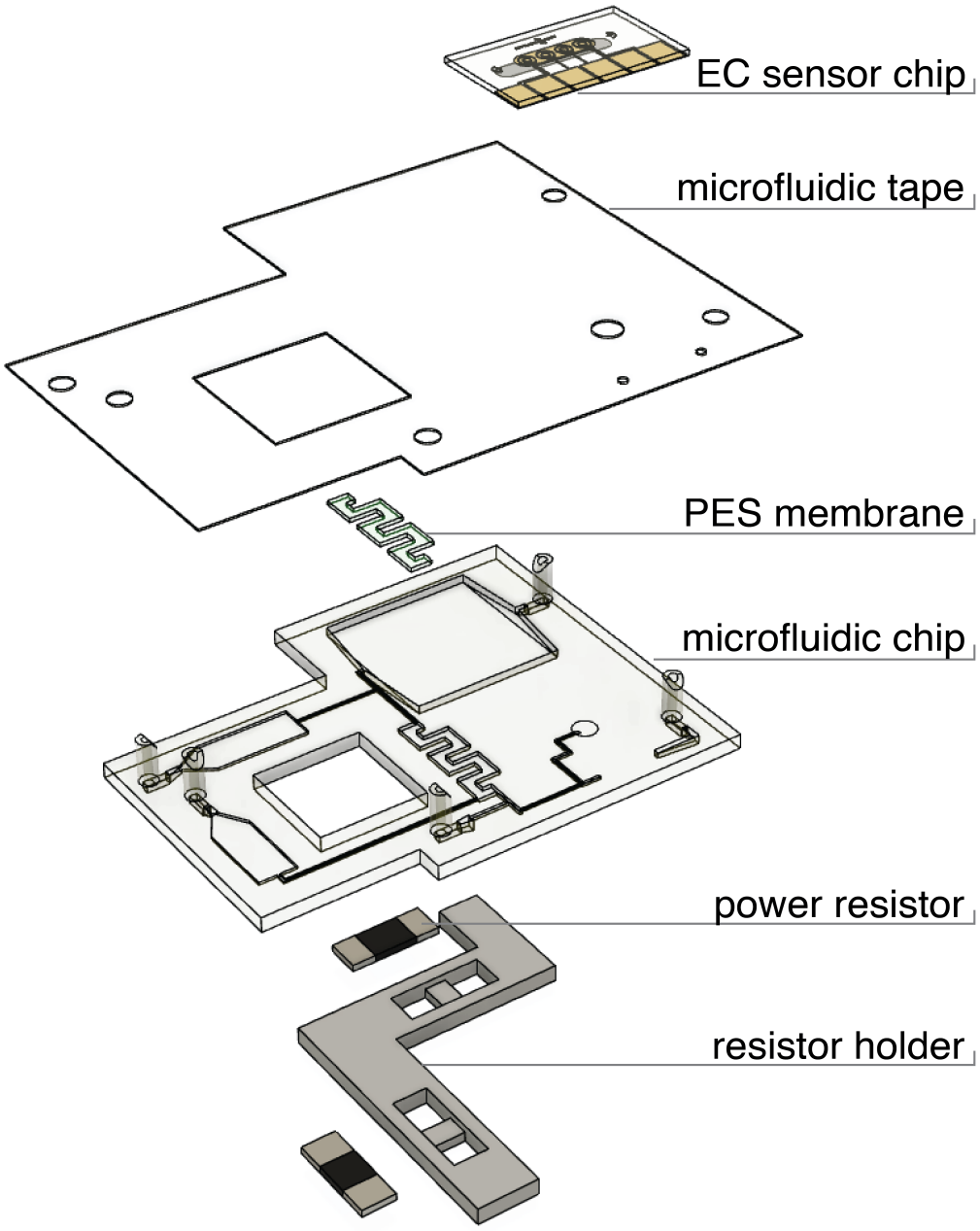



Figure B-1: **Exploded view of LOC microfluidic platform.** An exploded view of the multiplexed system, which includes a heater system, a sealed microfluidic chip, and a multiplexed EC sensor chip.

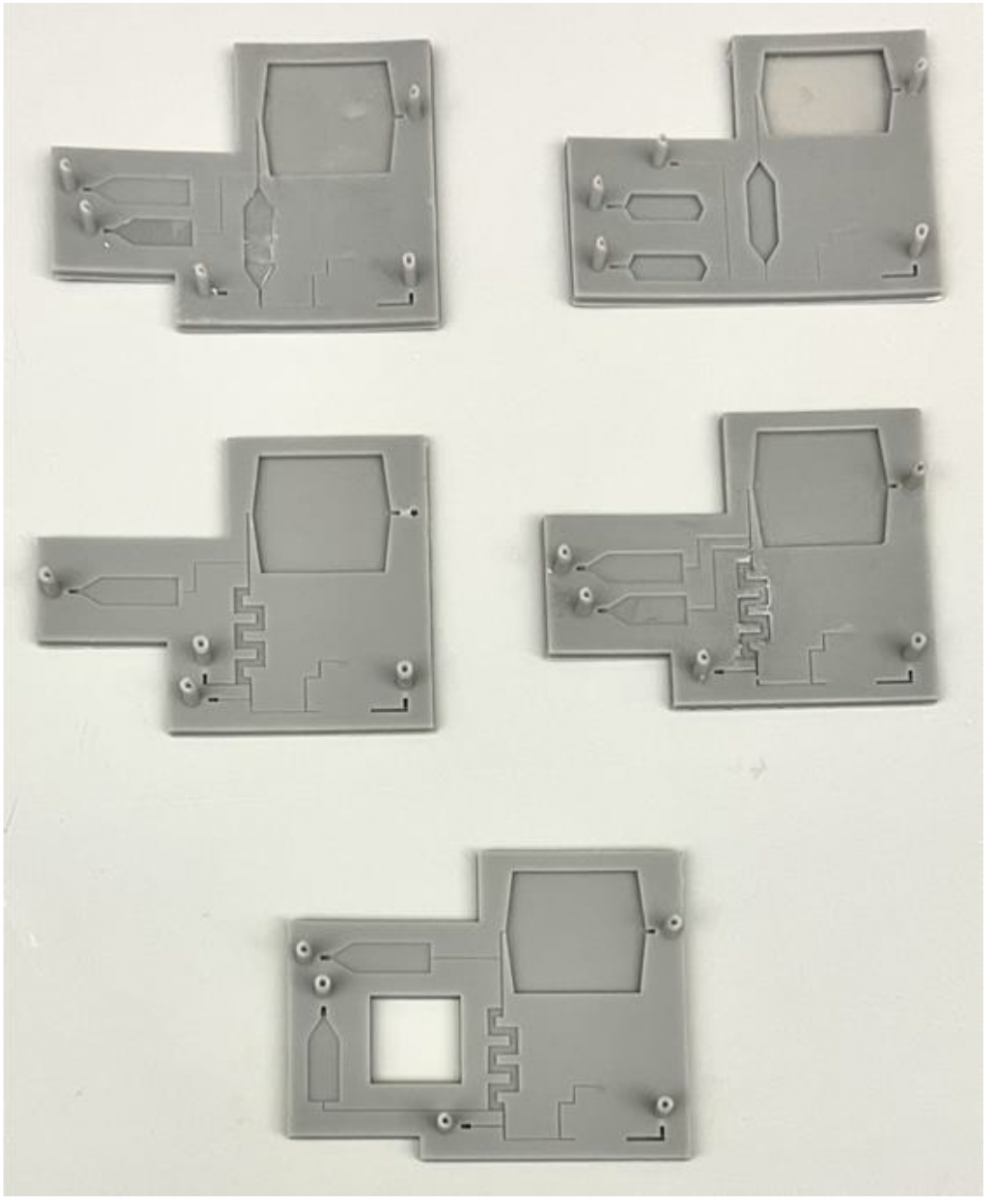


Figure B-2: **Iterations of the microfluidic chip.** Five designs with varied reservoirs, reaction chamber, channel sizes and layouts that were among the microfluidic chips tested.

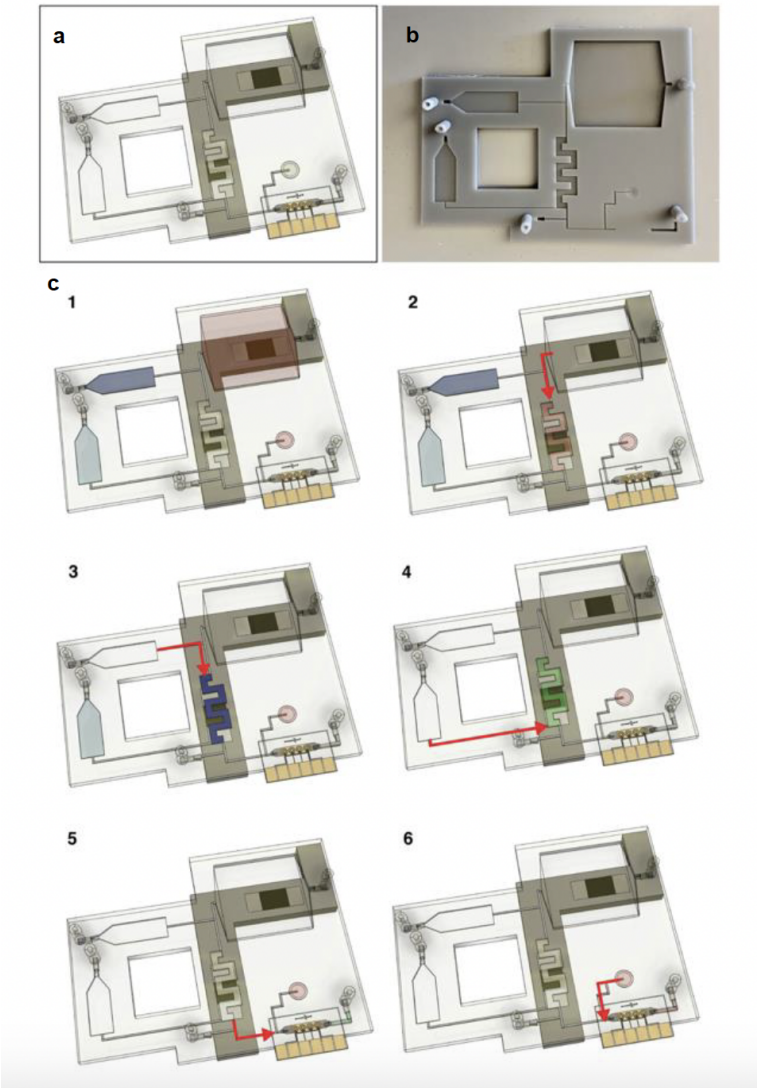



Figure B-3: **Microfluidic chip workflow.** a. Labeled image of the microfluidic chip. b. Printed microfluidic chip. c. A detailed description of the movement of liquid through the device. The workflow is as follows: (1) Add 25 µL of nuclease-free water to the LAMP reservoir, and 22 µL nuclease-free water the CRISPR reservoir, 20 µL saliva to the saliva-based antibody reservoir and 400 µL saliva to 50 µL of a 1:10 dilution of proteinase k to the saliva-based RNA preparation reservoir. Incubate in RNA preparation reservoir at 55℃ for 15 mins followed by 95℃ for 5 min. (2) Saliva flows through the reaction chamber and after, the PES filter incubates at 95℃ for 3 min to assure there is no more active proteinase k or inhibitors within the saliva. (3) The rehydrated LAMP mix is pumped into reaction chamber and incubates at 65℃ for 30 min. (4) CRISPR mix is pumped into reaction chamber containing the LAMP reaction and incubates at 37℃ for 30 min. (5) The LAMP-CRISPR mix is pumped over the EC sensor chip. (6) The 20 µL saliva sample from the antibody detection reservoir is pumped over EC sensor chip.

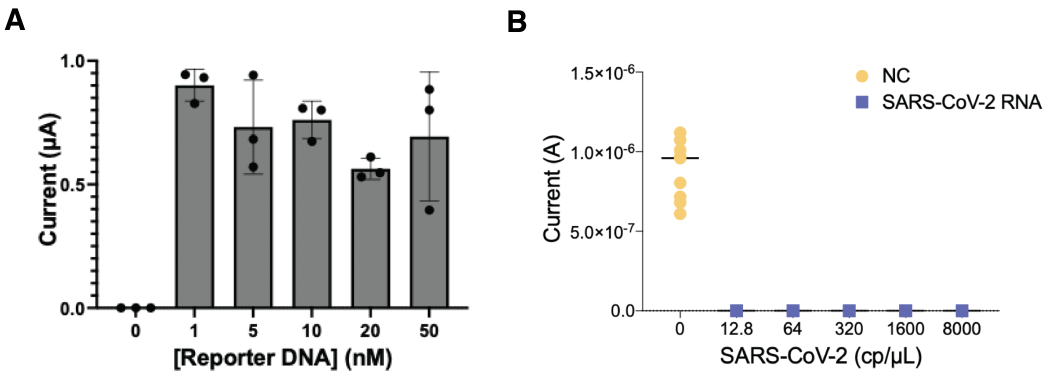



Figure B-4: **Optimization of CRISPR assay for EC-based SARS-CoV-2 RNA detection.** A. Optimization of the reporter concentrations in the CRISPR-electrochemistry RNA assay. We optimized the concentration of the biotinylated reporter probe for the electrochemical SARS-CoV-2 CRISPR-based RNA assays. We tested reporter concentrations between 0-50 nM and obtained optimal performance of the assays at 1nM final reporter concentration. Error bars represent the standard deviation of biological triplicate experiments.B. Limit of detection of the CRISPR-based electrochemical assay. Serially diluted full-length SARS-CoV-2 RNA (purple squares) was spiked in water and amplified by RT-LAMP. Results show that dilutions down to 12.8 cp/µL of viral RNA had a clear negative signal in our electrochemical assays, indicating no deposition of TMB. Unpaired Student's t-test $p<0.0001$ for differences in current (A) between 12.8 cp/µl and 0 cp/µl viral RNA. NC: negative control (yellow circle), 0cp/µl viral RNA. Error bars represent the standard deviation of independent triplicate experiments, biological replicates.

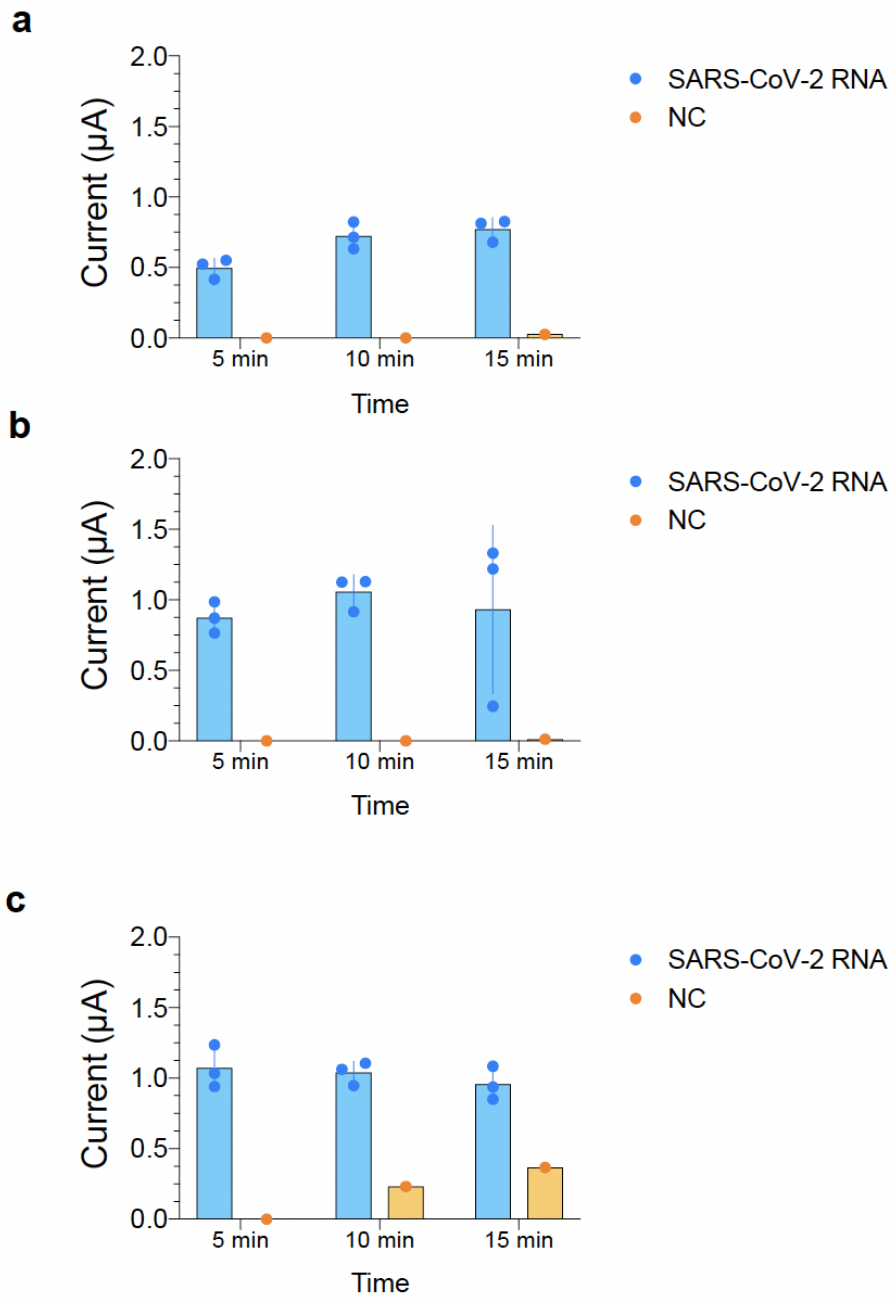



Figure B-5: **Optimization of the CRISPR-based electrochemistry assay time.** We optimized the reaction times by incubating (a) 0.5nM, (b) 1nM and (c) 5nM reporter probes over 5, 10 and 15 min. We selected 1nM reporter probe concentration and 5-min incubation times for further experiments. Error bars represent standard deviation of biological independent triplicate experiments.

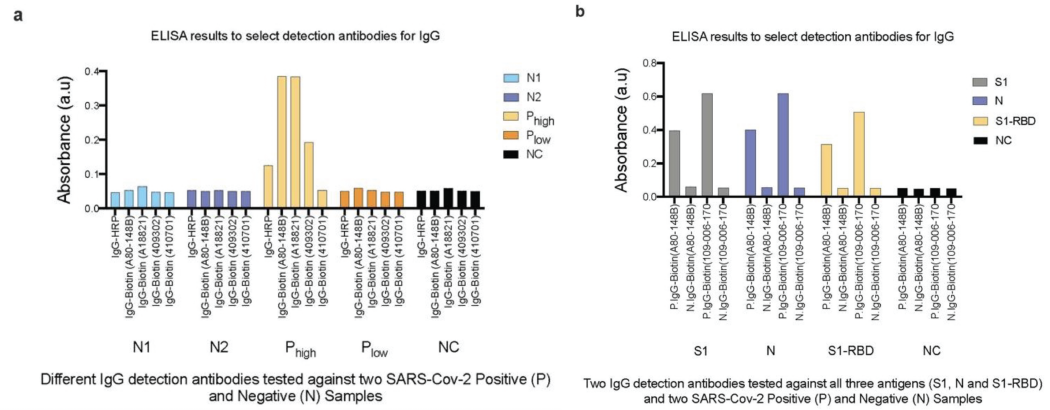



Figure B-6: **Selection of the SARS-CoV-2 anti-human IgG detection antibody against Spike, NC, and RBD protein in an ELISA format using SARS-CoV-2 positive and negative samples.** A. Biotinylated goat anti-human IgG Biotin (A80-148B) and (A18821) had high signals for positive sample $P_{high}$. All detection antibodies gave a similar low signal for negative clinical samples and PBS negative control background (Fig. B-4B). Overall, the anti-IgG Biotin (A80-148B) had a higher signal-to-noise ratio. B. Comparison of the performance of anti-IgG Biotin (A80-148B) with anti-IgG Biotin (109-006-170) using antigens S1, N, and S1-RBD. Anti-IgG Biotin (109-006-170) was selected for further experiments because it had higher signal-to-noise ratios for all antigens.

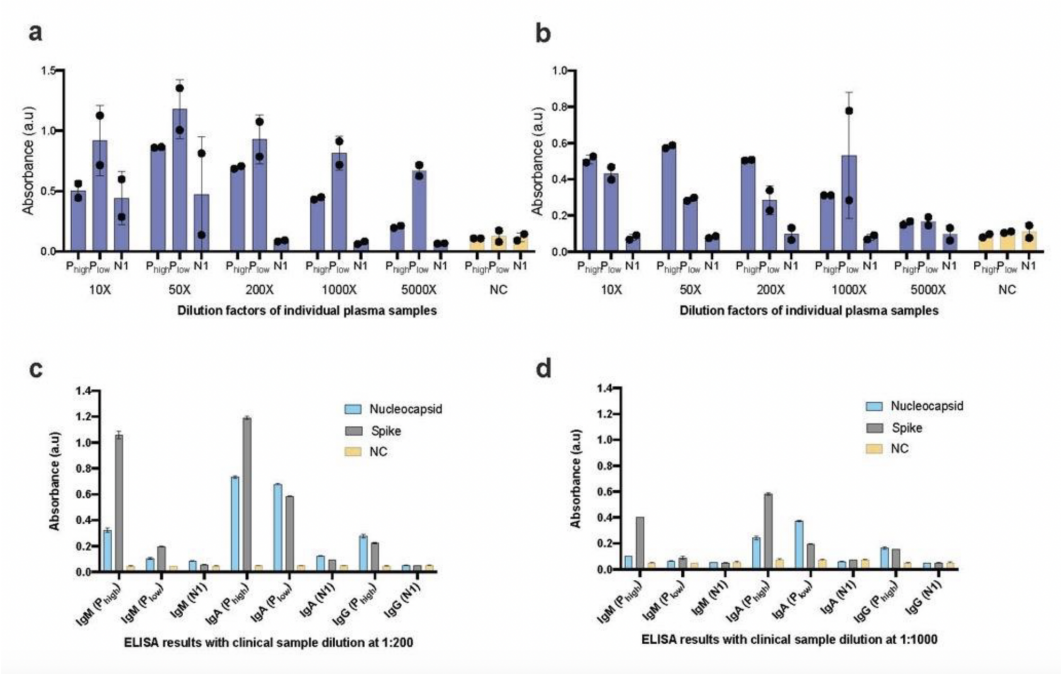



Figure B-7: **Optimization of the plasma dilutions using clinical plasma samples in a 96-well ELISA format.** A. Biotinylated goat anti-human IgG Biotin (A80-148B) and (A18821) had high signals for positive sample $P^{h\ i\ g\ h}$. All detection antibodies gave a similar low signal for negative clinical samples and PBS negative control background (Fig. B-4B). Overall, the anti-IgG Biotin (A80-148B) had a higher signal-to-noise ratio. B. Comparison of the performance of anti-IgG Biotin (A80-148B) with anti-IgG Biotin (109-006-170) using antigens S1, N, and S1-RBD. Anti-IgG Biotin (109-006-170) was selected for further experiments because it had higher signal-to-noise ratios for all antigens. The dilution factor of 200X (c) detected the less abundant antibody isotypes (IgM and IgA). Error bars represent the standard deviation of independent triplicate experiments, biological replicates.

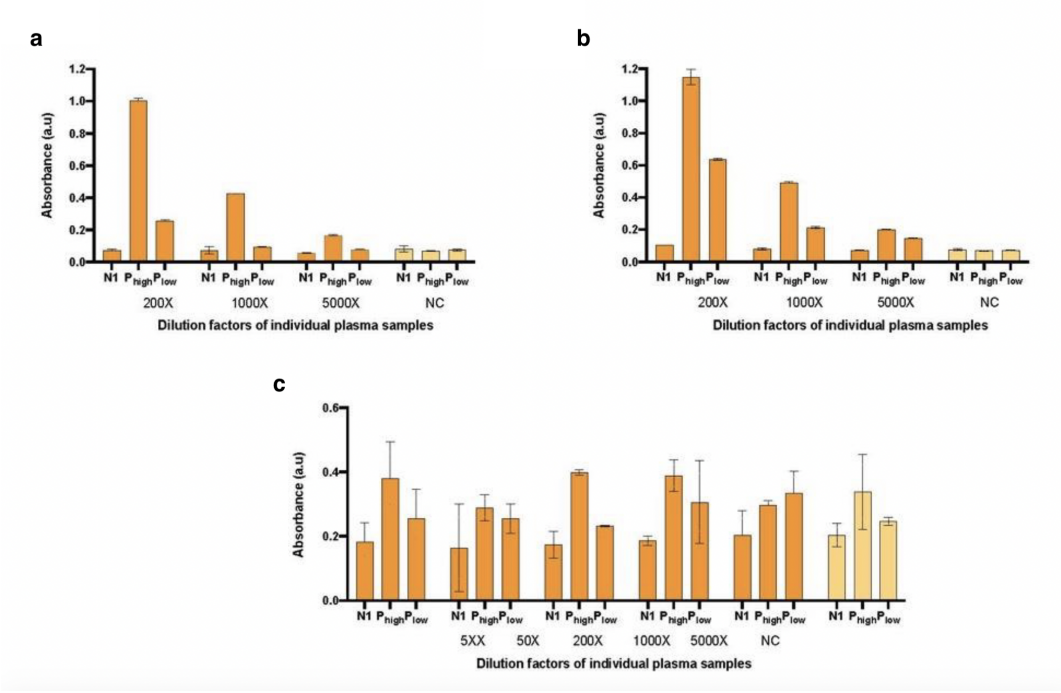



Figure B-8: **Optimization of the plasma dilutions using clinical plasma samples in a 96-well ELISA format.** Serial plasma dilutions were tested in an ELISA plate format with immobilized S1-RBD to detect human (A) IgM, (B) IgA and (C) IgG. We observed that dilutions between 200X-1000X show a high signal-to-noise ratio in ELISA when detecting IgG against both N and S1 protein antigens (a) and (b), respectively. The dilution factor of 200X detected the less abundant antibody isotypes (IgM and IgA), similar to results obtained in Fig. B-7C, indicating that a dilution factor of 200X was optimal for ELISA experiments. Error Error bars represent the standard deviation of independent triplicate experiments, biological replicates.

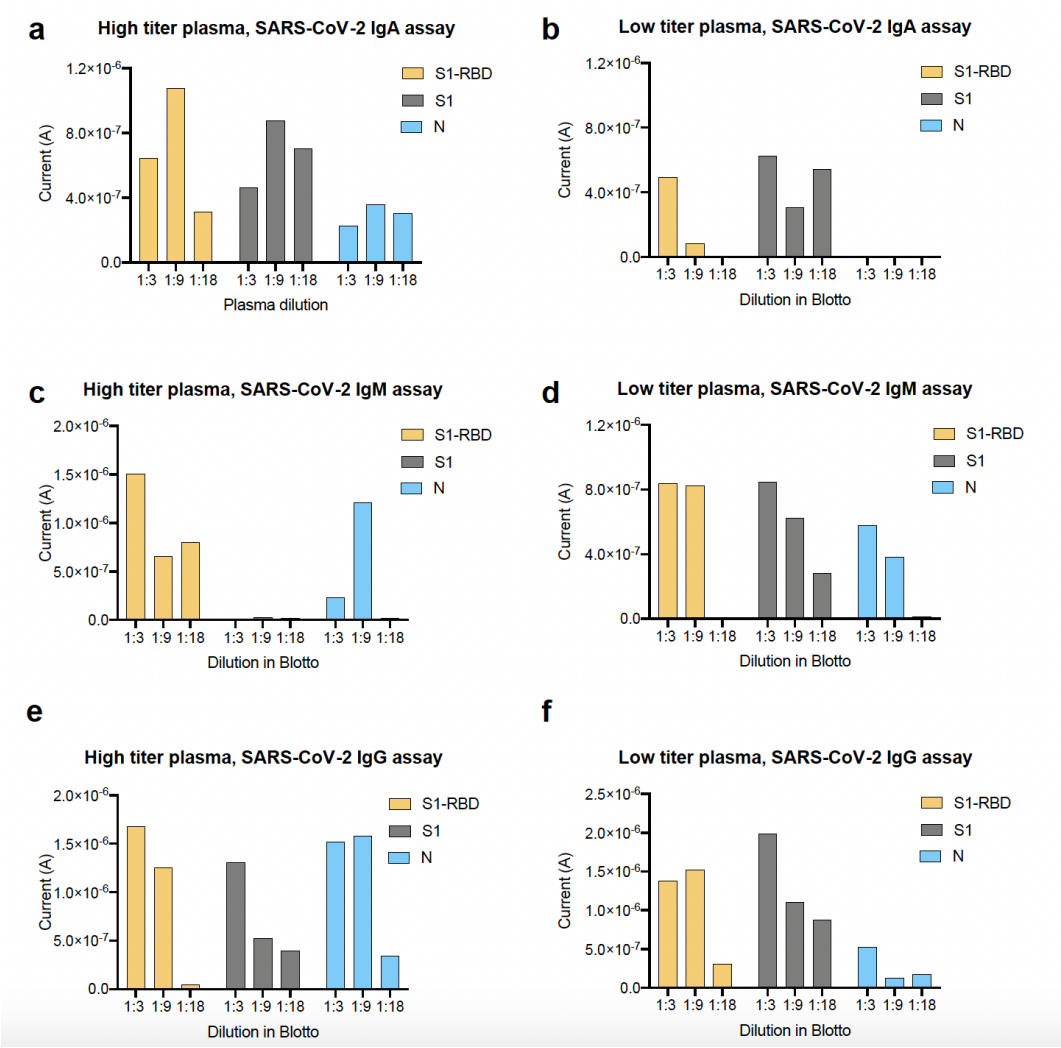



Figure B-9: **A sample dilution of 1:9 is optimal for IgG, IgM and IgA detection in the electrochemical platform.** High and low antibody titer plasma was serially diluted and tested on the electrochemical platform to detect anti-SARS-CoV-2 human (a, b) IgA, (c, d) IgM, and (e, f) IgG against S1, nucleocapsid, and S1-RBD antigens immobilized on the electrodes.

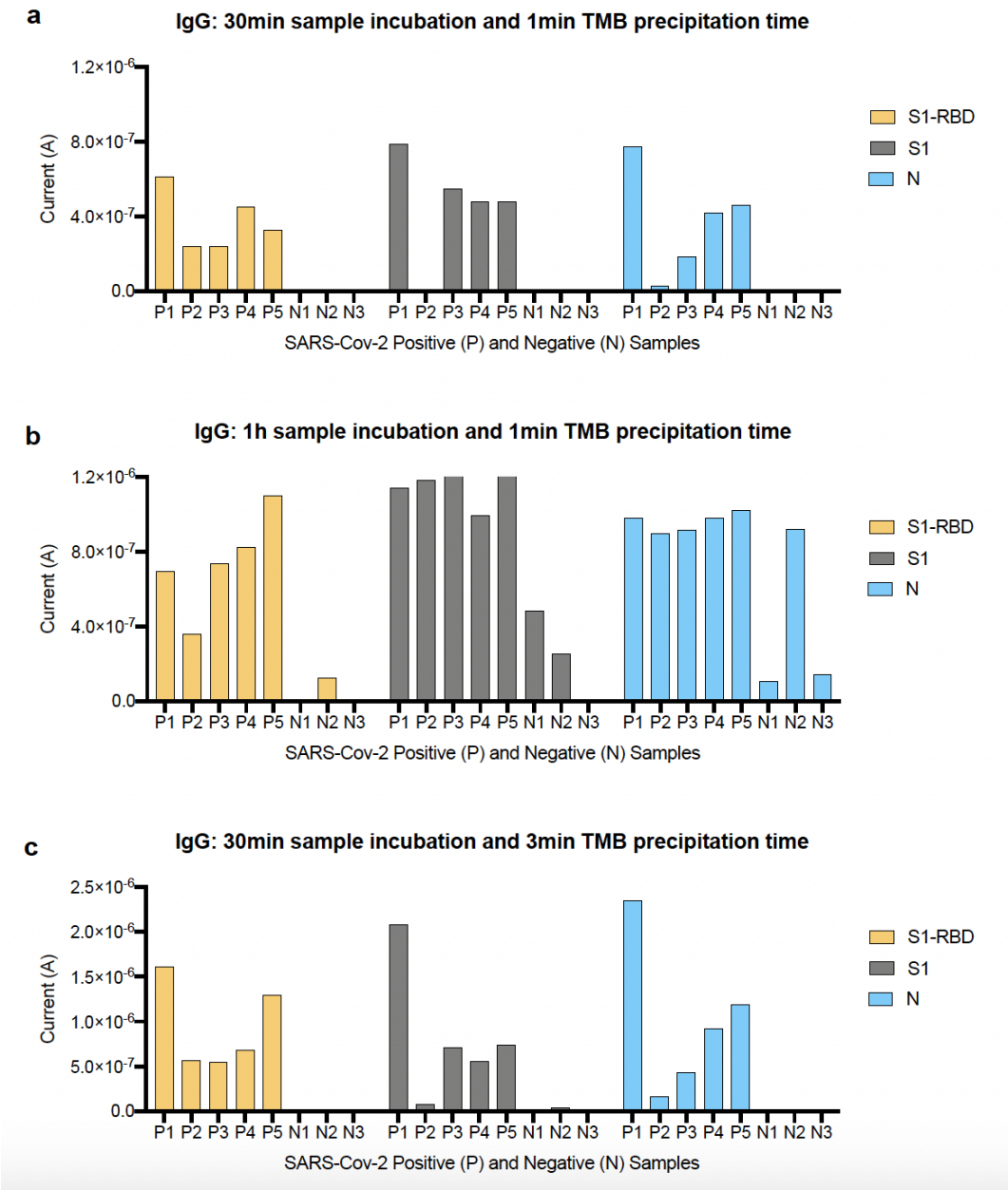



Figure B-10: **Optimization of the incubation and TMB precipitation times in 1:9 plasma dilutions.** Electrochemical biosensors were conjugated with the three antigens (S1-RBD, S1 and N) and used to detect IgG in a subset of positive and negative clinical plasma samples diluted at 1:9. We optimized the incubation and TMB precipitation times: (a) diluted plasma was incubated for 30 min and TMB was allowed to precipitate for 1 min (b) diluted plasma was incubated for 1h and TMB precipitation time was 1 min; (c) diluted plasma was incubated for 30 min and TMB reactions proceeded for 3 min. We observed that the optimal performance was obtained for (c) 3-min plasma incubation on the electrodes and 3-min TMB precipitation.

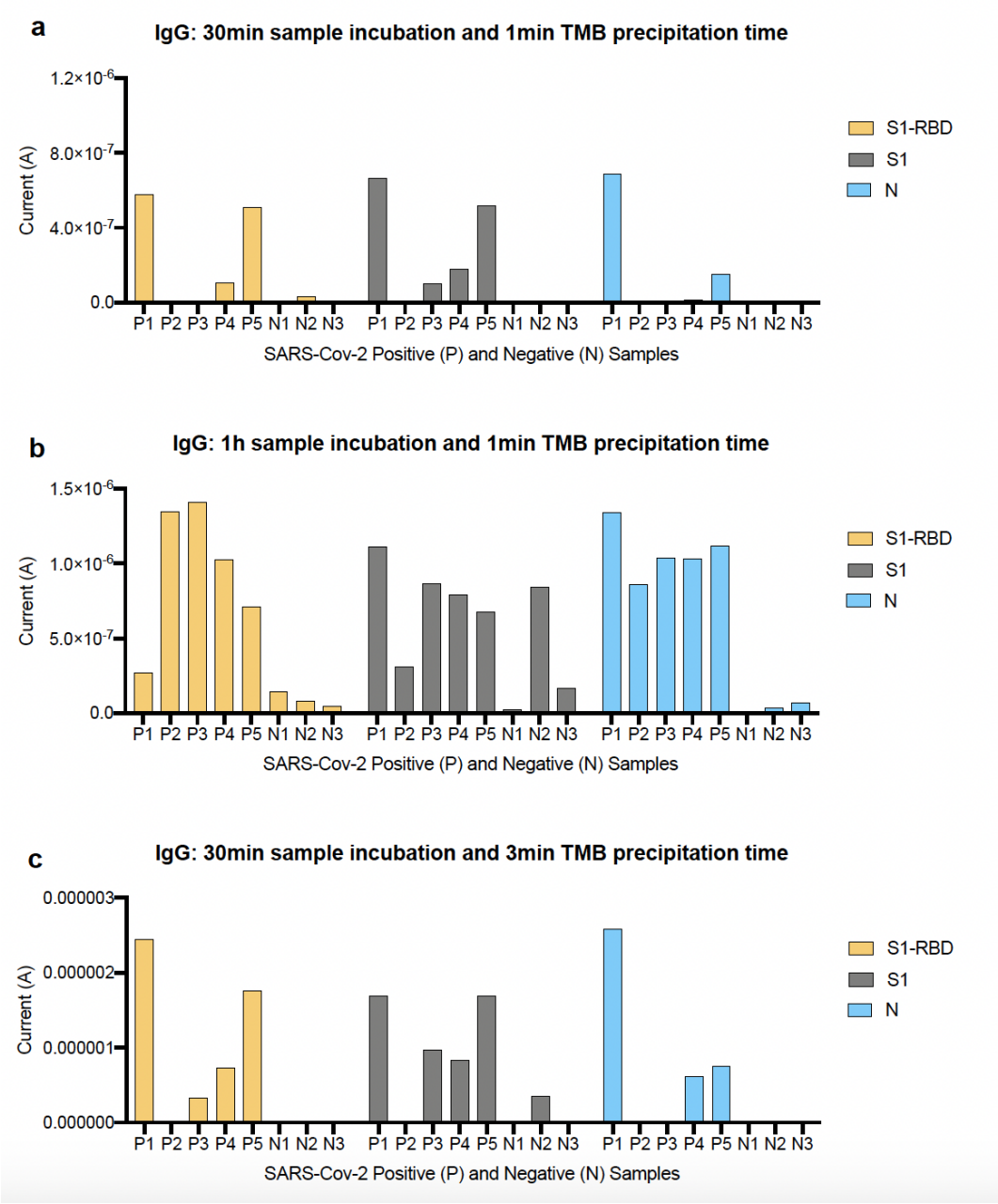



Figure B-11: **Optimization of incubation and TMB precipitation time at 1:18 plasma dilutions.** Electrochemical biosensors were conjugated with the three antigens (S1-RBD, S1 and N), and used to detect IgG in a subset of positive and negative clinical plasma samples diluted at 1:18. We optimized the incubation and TMB precipitation times: (a) diluted plasma samples were incubated for 30 min and TMB was allowed to precipitate for 1 min; (b) diluted plasma samples were incubated for 1h and TMB precipitation time was 1 min; (c) diluted plasma samples were incubated for 30 min and TMB reactions proceeded for 3 min. Optimal performance was observed for plasma diluted at 1:9 (Fig. 4-10A,c.)

| | | | | | | | | LOD (cp/μL) |
|---|---|---|---|---|---|---|---|---|
| Fluorescent assay | RNA (cp/μl) | 12.5 | 6.25 | 1 | 0.3 | 0.15 | 0 | 2.3 |
| | Replicates (+/total) | 5/5 | 5/5 | 3/5 | 2/5 | 2/5 | 0/5 | |
| EC assay | RNA (cp/μl) | 12.5 | 6.25 | 1 | 0.3 | 0.15 | 0 | 0.8 |
| | Replicates (+/total) | 5/5 | 5/5 | 3/5 | 3/5 | 2/5 | 0/5 | |

Table B.1: **Logit regression curve analysis data for device LOD.** Raw data used for the logit regression curve analysis to measure the limit of detection of fluorescent and electrochemical CRISPR-based SARS-CoV-2 RNA detection.

| | **CRISPR fluorescence** |
|---|---|
| AUC | 1.00 |
| 95% conf. int. | [1.00-1.00] |
| P value | <0.0001 |
| Cutoff | 8859 |
| Sens. | 1.00 [0.83-1.00] |
| Spec. | 1.00 [0.72-1.00] |
| N pos. | 19 |
| N neg. | 10 |

Table B.2: **CRISPR-based fluorescent assays accurately detect SARS-CoV-2 in clinical saliva samples.**

| **Primer** | **Concentration** | **Sequence** |
|---|---|---|
| FIP | 1.6 µM | TCAGCACACAAAGCCAAAAATTT ATTTTTCTGTGCAAAGGAAATTA AGGAG |
| BIP | 1.6 µM | TATTGGTGGAGCTAAACTTAAA GCCTTTTCTGTACAATCCCTTTG AGTG |
| F3 | 0.2 µM | CGGTGGACAAATTGTCAC |
| B3 | 0.2 µM | CTTCTCTGGATTTAACACACTT |
| LOOP F | 0.4 µM | TTACAAGCTTAAAGAATGTCTGA ACACT |
| LOOP B | 0.4 µM | TTGAATTTAGGTGAAACATTTGT CACG |

Table B.3: **Best performing LAMP primer sequences and their final concentrations in LAMP assays.** [134]

Table B.4: Cost analysis for LOC microfluidic device.

| Component | Unit Cost ($) | Quantity | Cost | Vendor | Product Number |
|---|---|---|---|---|---|
| **Microfluidic chip** | | | | | |
| 3D printed resin, 1L | 149.99 | 0.0058 | 0.87 | Formlabs | RS-F2-GPGR-04 |
| Power resistor | 1.43 | 2 | 2.85 | Digikey | A143815 TR |
| Microfluidic tape | 351 | 0.0009 | 0.31 | 3M | 70000126311 |
| PES membrane | 0.02 | 0.125 | 0.04 | Millipore | GPWP04700 |
| **Biochemical Assay** | | | | | |
| Proteinase K | 87 | 0.0023 | 0.20 | New England Biolabs | P8107S |
| LbCas12a | 250 | 0.0025 | 0.63 | New England Biolabs | M0653T |
| LAMP mix | 883 | 0.002 | 1.77 | New England Biolabs | E1700L |
| gRNA, reporter, and buffers | | | 0 | | |
| **Electrochemical Sensor chip** | | | | | |
| Glass and gold chip | 6 | 1 | 6 | Telic Company | |
| Reduced graphene oxide, tetraethylene pentamine functionalized | 379 | 0.00112 | 0.42 | Millipore Sigma | 806579 |
| Bovine Serum Albumin (IgG-Free, Protease-Free | 260 | 0.000007 | 0 | Jackson Immuno Research | 001-000-162 |
| Glutaralde hyde, 70% | 75.10 | 0.0002 | 0.02 | Millipore Sigma | G7776 |
| EDC (1-ethyl-3-(3-dimethylamino propyl) carbodiimide hydrochloride) | 132 | 0.001078 | 0.14 | Thermo Fisher Scientific | 22980 |
| N-hydroxysucc inimide (NHS) | 21.20 | 0.000322 | 0.01 | Millipore Sigma | 130672 |
| Poly-HRP Streptavidin | 266 | 0.00008 | 0.02 | Thermo Fisher Scientific | N200 |

Continued on next page

**Table B.4 – continued from previous page**

| Component | Unit Cost ($) | Quantity | Cost | Vendor | Product Number |
|---|---|---|---|---|---|
| TMB Enhanced One Component HRP Membrane Substrate | 103 | 0.0002 | 0.02 | Millipore Sigma | T9455 |
| **Total** | **$ 13.29** | | | | |

| **Voltages for Microfluidic Heating** | |
|---|---|
| Temperature (℃) | Voltage (V) |
| 55 | 4.7 |
| 95 | 7 |
| 65 | 4.85 |
| 37 | 2.55 |

Table B.5: **Voltages used for heating the LOC microfluidic device.**

# Appendix C

# MyRWA Ethnography

**Code book for Ethnography of MyRWA monitors**
**Environment:**

**Development**: having to do with construction and urbanization of the land within the watershed.

*"if there's one factor that drives the kind of environmental changes we track most closely, it's urbanization" (M1:37-39)*

**Bacteria**: relating to the bacteria most commonly found in water contamination related to run-off and sewage. For freshwater, this is primarily E.coli and for saltwater this is primarily Enterococcus.

*"And it's E.coli in freshwater, Enterococcus in salt water" (M1: 58)*

**Watershed**: hydrological events or activities that impact or take place within a watershed (here usually the Mystic River or Charles River watershed)

*"material builds up on these surfaces and it's washed very efficiently to rivers and streams" (M1: 108-109)*

**Dam**: Relating to the Amelia Earhart Dam

**Hydrology**: related to the movement of water over land, specifically within the watershed

*"and changes everything about the way water moves across the landscape and changes everything" (M1: 45-46)*

**Infrastructure**: Relating to city and urban infrastructure that impacts the local watershed *"So it's, when the pipe that's leading to Deer Island becomes, is sort of filled to capacity with this mix of rainwater and sewage" (M1: 72-73)*

**Urban**: Relating to activities that go on or are influenced by a densely populated area with mostly paved and non-absorbent surfaces.

*"phosphorus is driven, the excess phosphorus in urban areas is a function again of*

*impervious surface and material builds up on these surfaces" (M1: 107-108)*

**Rain**: relating to rain or rainfall
*"CSO's only trigger in big rainfall events" (M1: 70)*

**Climate change**: environmental changes due to climate change
*"the sort of root cause of cyanobacteria blooms which are becoming more frequent with climate change it turns out" (M1: 91-93)*

**Development**: related to urban development and the environmental impacts thereof
**"Uh and so there were saltmarsh there were, you know, exposed wetland areas that are now Upland developed areas." (M1: 30-31)**
**Nutrients**: Relating to the presence of nutrients (usually nitrogen, phosphorus, and/or calcium) in water
*"The root cause there is a nutrient, excess nutrient" (M1: 94)*

**Industrial**: Having to do with the presence of industry (non-residential)
*"It was, you know, big deserted, empty industrial spaces" (M1: 11-12)*

**Water Contamination**: Of or relating to the polluting of water bodies (chemical and/or biological) due to primarily to urban influences *"We also monitor for cyanobacteria, which as you probably know can produce these liver toxins and neurotoxins, things are again, a public health issue" (M1: 88-90)*

**Monitoring:**

**Convenience**: Related to the proximity/ease with which a monitoring activity can be done
*"I think proximity to where I live, I can just walk there pretty easy" (M2: 46)*

**Inconvenience**: Related to the difficulty with which a monitoring activity can be done
*"so we had to actually, like I put a stool in the river and I then the winter was particularly bad." (M2: 176-177)*

**Monitoring**: Related to the observation over time of water (and related environmental indicators)
*"So to say that, you know, if there's one factor that drives the kind of environmental changes we track most closely, it's urbanization" (M1: 38-39)*

**Quantitative**: any language that implies a measurement, technical instrument, or standardized protocol; to be used for water monitoring purposes
*"So in, the two big pollutants that we have our eye on in the water quality monitoring program we test for a variety of parameters, but the two main ones are bacteria, as an indicator of wastewater contamination of storm water" (M1:48-51)*

**Organizational**: Related to or in service an organized group (governmental or NGO) that performs a water monitoring-related services

**Sub-categories:**

*Organizational; MyRWA*: related to Mystic River Watershed Association activities

*Organizational; EPA*: related to Environmental Protection Agency activities

*Organizational; MWRA*: related to Massachusetts Water Resource Authority

*Organizational; USGS*: related to United States Geological Survey

*"And we, we we our annual water quality report card with EPA" (M1: 52-53)*

*"And there's a really interesting recent USGS study, did, they did a study in Madison, Wisconsin"*

**Water sampling**: Relating to the process of taking water for analysis

*"but the goal of the protocol for our samplers is to acquire a repeatable representative sample of the stream" (M1: 328-329)*

**Water sampling; complicated**: When a water sampling event is complicated

**Water sampling; automatic**: A continuous sampling (usually automated sampler)

**Copying**: Relating to a methodology or protocol that was copied from another organization

*"so that our, standard water quality sampling, parameters that the state uses and that the EPA uses to, to, measure water quality" (M1: 174-175)*

**Copying; CR**: Copying the methodology or protocols that the Charles River Watershed has performed

*"This same work was done in the Charles some years ago and that the order of reduction mandated by EPA is on that same scale more than 50%" (M1: 103-105)*

**Regulations**: Relating to existing or future regulations with regard to water quality

*"and there's a kind of suite of water quality, um, parameters that, that, that, that sort of carry weight, regulatory weight" (M1: 172-173)*

**Dataset**: Relating to the systematic time series collection of data that will be used together to assess the water. *"And we've been doing this now every month for 20 years at 15 sites. So we, I have that data." (M1: 180-181)*

**Accountability**: holding a person or organization accountable for their actions

*"we are both partners with EPA and kind of hold them accountable, you know"*

**CS vs CFT**: citizen scientist vs citizen field tech (a differentiation from being someone who does a sampling action vs someone who is can either choose the site of analysis or contributing to the analysis and interpretation of that sample)

*"they're not making, well, they're making some qualitative judgments about whether the sample has an odor or whether it has, what it's appearance is, but the goal of the protocol for our samplers is to acquire a repeatable representative sample of the stream" (M326-329)*

**Collaboration**: When two or more people or organization are working together or sharing information

*"But we're, we're often allies in helping the project defined by the Clean Water Act" (M1: 226-227)*

**Consistency**: The standardization of sample-to-sample collection and analysis

*"So that's the same lab doing both, their, their own testing, their own, their own crews go out and get water and they're tested in the same lab."*

**Adding Value**: An action that will increase the existing value of a monitoring program/dataset

*"And so we said, well, how can we add value? Well, we can monitor the tributaries and lakes that are attached to that are continuous with the Mystic River" (M1: 270-272)*

**Following protocol**: Using pre-determined standardized methods in order to accomplish a task (in this case, related to water quality measurements and monitoring)

*"observer following this protocol to go follow that procedure at that site and get the same sample." (M1: 305-306)*

**Monitoring Location**: Referencing a monitoring site

**Ad-hoc**: When a sampling event happens outside of a regulated schedule

*"they just sample wherever they can sample" (M2: 92-93)*

**Air quality**: Related to monitoring or sampling for air quality "And you can, you know, go on Amazon or anything and you, if you look for

*particulate portable, portable air quality monitor" (M2: 235-236)*

**Logistical difficulty**: There is a logistical difficulty involved with sampling or monitoring

*"I think the issue is, and it's maybe it's expensive or it's a lot of hassle to get this sample" (M2: 119)*

**Scheduled and Storm**: Related to a sampling procedure that consists of the scheduled samples as well as sampling after a storm event

**Temperature sensitive**: A sample that is temperature sensitive

*"you have to keep the samples no hotter than five degrees Celsius" (M2: 219)*

**M1**: Referencing M1 and his work with MyRWA

**Deer Island**: The Deer Island water testing facility

**Actor:**

**Surprise**: When something that seems counter-intuitive or unlikely, but is true, is mentioned

*"The, the kind of more surprising one to me though, and it's beginning now to be really quantified is organic material that that deposits itself on asphalt." (M1: 123-125)*

**Luck**: When an actor expresses that they got lucky

*"Sometimes we get lucky, if you count this as lucky, and wet weather events where if it's rained for at least a quarter or half an inch, within 12 hours of the sampling event," (M2: 65-67)*

**Hope**: When an actor is hoping for something

*"I was hoping we would have a more continuous monitoring potential, but apparently, you know, it's not possible" (M2: 102-103)*

**Monitoring Reason**: Reason that they are participating in a water monitoring program

**Qualification**: Qualifying a statement that was made

*"but I don't, I'm not engaged in any way in changing anything or rethinking, you know, when you have protocols that have been going on for a long time" (M3: 71-72)*

**Deferring**: When an actor defers responsibility or expertise to someone else

*"So I've been working, he, it's his station, I guess you could call it" (M3: 36)*

**Uncertainty**: When an actor is unsure about something

*"I don't know if salinity is important, maybe it's pH. I don't. I'm not sure. (M2: 148-149)*

**Sampling Story**: A story told by an actor about a sampling experience

**Other civic environmental engagement**: Activity that a water monitor participates in that relates to their local environment outside of MyRWA sampling

**Scientist**: When the actor indicates that they are involved in the sciences (through education and/or profession)

**Frustrated with labels**: When an actor is frustrated with being confined by a label

*"I guess, I get frustrated with the names business, like above the lakes It's not, it's called the Aberjona instead of the Mystic" (M3: 421-423)*

**Notes for myself:**

**Good Quote**: A quote I thought was important

**Interesting Thought**: Something that I am interested in exploring further or that was not intuitive

**For Memo**: Something that I would like to put in a memo

**Semi-structured interview questions for MyRWA Monitor**

Short interview protocol for MyRWA Monitors. These questions were only used for my interview with M1. Following that interview I updated the protocol.

Written: 4/13/20

Thank you so much for agreeing to participate in this study! Please feel free to answer as many questions as you are comfortable, and feel free to reach out if you have any questions or need clarification. Additionally, I am also happy to ask these questions over the phone and transcribe the responses for you to read if you do not have the time/resources to respond in this format. Your personal information will remain anonymous (and your sampling site as well if it is necessary for anonymity.

**Getting to know you:**

*When did you move to the Boston Area? What brought you here?*

*In what ways, if any, do you think the demographics of this area have changed over time?*

*Did you study anything related to the sciences or environment studies? Are you currently working within a field related to the sciences or environmental studies?*

**MyRWA program:**

*In what capacities are you involved in MyRWA? For how long?*

*What initially interested you in getting becoming a monitor?*

*Do you have a specific site or set of sites that you monitor regularly? Where are they and how did you choose or get assigned that location? What is the state of water quality of your site(s)?*

*How would you describe your role on a month-to-month basis for MyRWA? Do you see yourself as a citizen scientist, a citizem field tech, a monitor, or a name of your own assignment.? Why?*

*How do you think you incorporate the qualitative aspect of sampling (being in the physical space and using your senses to understand the environment) into the month-to-month measurement process, if at all? Do you think that matters?*

*Do you think your familiarity with the area allows you to participate in monitoring or community-directed activities in addition to the sampling you do?*

*Was there ever a time when you really wanted to sample a certain area or at a certain time that wasn't consistent with the monitoring protocols? What did you do?*

*Do you look at the water quality results from your sampling site(s)? Does that factor into your knowledge of the site and the subsequent measurements you take?*

*If you could choose a parameter or a site that is not the existing dataset to track (or pick one that already exists) to monitor, what and/or where would you choose and why?*

**Wrapping up**

We have about 5 minutes left. Is there anything you want to ask me or something you want to share as we finish up?

Thank you so much for your time looking through and answering these questions! I am hugely appreciative and look forward to sharing the results (maybe as a MyRWA meeting!) once I process the data.

**Updated Semi-structured interview questions for MyRWA Monitors**

Short interview protocol for MyRWA Monitors modofied based on information gathered from interviewing M1.

Written: 4/23/20

**Consent:**

I would like to audio record this interview so that I don't have to take notes while we speak. If you do not feel comfortable with this, I will be happy to take notes or not document it. Is this ok? I can send you the transcript from the interview to read over, correct, and approve. I will be keeping all of this information confidential, so

your name and identifying information will be hidden as much as possible. I would be happy to share the resulting report coming from this research.

**Getting to know you:**

*When did you move to the Boston Area? What brought you here?*

*In what ways do you think the demographics of this area have changed over time? In what ways do you think the area itself has changed over time?*

**MyRWA program:**

*In what capacities are you involved in MyRWA? For how long?*

*Do you have a specific site or set of sites that you monitor regularly? Where are they and how did you choose or get assigned that location?*

*How would you describe your role on a month-to-month basis for MyRWA? Do you see yourself as a citizen scientist or a field tech or monitor? Why?*

*How do you think you incorporate the qualitative aspect of sampling (being in the physical space and using your senses to understand the environment) into the measurement process, if at all?*

*Was there ever a time when you really wanted to sample a certain area or at a certain time that wasn't consistent with the monitoring protocols? What did you do?*

*Do you get a copy of the results for the water that you sample? Do you look at them? Does that factor into your knowledge of the site and the subsequent measurements you take?*

*If you could choose a parameter that is not the existing dataset to track (or pick one that already exists) to monitor, what would you choose?*

**Wrapping up**

*We have about 5 minutes left. Is there anything you want to ask me or something you want to share as we finish up?*

Thank them profusely for their time. Tell them how much you appreciated sitting how and how interesting this has been. Tell them the next time you think you will see them/follow up with them. Ask for information so you can follow up with the written interviews for analysis

# Bibliography


[1] Urban waters - the mystic river watershed, Jul 2022.

[2] Nicholas M Adams, Hali Bordelon, Kwo-Kwang A Wang, Laura E Albert, David W Wright, and Frederick R Haselton. Comparison of three magnetic bead surface functionalities for RNA extraction and detection. *ACS Appl. Mater. Interfaces*, 7(11):6062–6069, March 2015.

[3] Bethany Alender. Understanding volunteer motivations to participate in citizen science projects: a deeper look at water quality monitoring. *Journal of Science Communication*, 15(3):A04, 2016.

[4] M Monsur Ali, Feng Li, Zhiqing Zhang, Kaixiang Zhang, Dong-Ku Kang, James A Ankrum, X Chris Le, and Weian Zhao. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem. Soc. Rev.*, 43(10):3324–3341, May 2014.

[5] Nasir Ali, Rita de Cássia Pontello Rampazzo, Alexandre Dias Tavares Costa, and Marco Aurelio Krieger. Current nucleic acid extraction methods and their implications to Point-of-Care diagnostics. *Biomed Res. Int.*, 2017:9306564, July 2017.

[6] Zahir Ali, Rashid Aman, Ahmed Mahas, Gundra Sivakrishna Rao, Muhammad Tehseen, Tin Marsic, Rahul Salunke, Amit K Subudhi, Sharif M Hala, Samir M Hamdan, et al. iscan: An rt-lamp-coupled crispr-cas12 module for rapid, sensitive detection of sars-cov-2. *Virus research*, 288:198129, 2020.

[7] Kelly D Alley. *On the Banks of the Gaṅgā: When Wastewater Meets a Sacred River*. University of Michigan Press, 2002.

[8] Fatima Amanat, Daniel Stadlbauer, Shirin Strohmeier, Thi HO Nguyen, Veronika Chromikova, Meagan McMahon, Kaijun Jiang, Guha Asthagiri Arunkumar, Denise Jurczyszak, Jose Polanco, et al. A serological assay to detect sars-cov-2 seroconversion in humans. *Nature medicine*, 26(7):1033–1036, 2020.

[9] A Andino and I Hanning. Salmonella enterica: survival, colonization, and virulence differences among serovars. *The Scientific World Journal*, 2015, 2015.

[10] Ward Appeltans, Shane T Ahyong, Gary Anderson, Martin V Angel, Tom Artois, Nicolas Bailly, Roger Bamber, Anthony Barber, Ilse Bartsch, Annalisa Berta, et al. The magnitude of global marine species diversity. *Current biology*, 22(23):2189–2202, 2012.

[11] Jon Arizti-Sanz, Catherine A Freije, Alexandra C Stanton, Brittany A Petros, Chloe K Boehm, Sameed Siddiqui, Bennett M Shaw, Gordon Adams, Tinna-Solveig F Kosoko-Thoroddsen, Molly E Kemball, et al. Streamlined inactivation, amplification, and cas13-based detection of sars-cov-2. *Nature communications*, 11(1):1–9, 2020.

[12] James Ataria, (Rongomaiw, Melanie Mark-Shadbolt, (Ngāti Kah, Aroha Te Pareake Mead, (Ngāti Awa, Kevin Prime, (Ngāti Hin, Jim Doherty, (Ngāti Man, James Waiwai, (Ngāi Tūho, Tohe Ashby, (Ngāti Hin, Simon Lambert, (Ngāi Tūho, and Gary Owen Garner, (Ngāti Ran. Whakamanahia te mātauranga o te māori: empowering māori knowledge to support aotearoa's aquatic biological heritage. *N. Z. J. Mar. Freshwater Res.*, 52(4):467–486, October 2018.

[13] Hatice Ceren Ates, Anna Brunauer, Felix von Stetten, Gerald A Urban, Firat Güder, Arben Merkoçi, Susanna Maria Früh, and Can Dincer. Non-invasive diagnostics: Integrated devices for non-invasive diagnostics (adv. funct. mater. 15/2021). *Advanced Functional Materials*, 31(15):2170105, 2021.

[14] Monya Baker. Reproducibility crisis. *Nature*, 533(26):353–66, 2016.

[15] Karen J Bakker et al. *An uncooperative commodity: Privatizing water in England and Wales*. Oxford Geographical and Enviro, 2003.

[16] Manju Bansal. Dna structure: Revisiting the watson–crick double helix. *Current Science*, pages 1556–1563, 2003.

[17] Stephen R Barley and Beth A Bechky. In the backrooms of science: The work of technicians in science labs. eqw working papers wp11. 1993.

[18] Michael Batty. Big data, smart cities and city planning. *Dialogues Hum Geogr*, 3(3):274–279, November 2013.

[19] Genevieve Bell, Melissa Gregg, and Nick Seaver. *Data, now bigger and better!* Prickly Paradigm Press, 2015.

[20] Rebecca Bell, Rupert Sutherland, Daniel HN Barker, Stuart Henrys, Stephen Bannister, Laura Wallace, and John Beavan. Seismic reflection character of the hikurangi subduction interface, new zealand, in the region of repeated gisborne slow slip events. *Geophysical Journal International*, 180(1):34–48, 2010.

[21] Byron M Berenger, John M Conly, Kevin Fonseca, Jia Hu, Thomas Louie, Angela R Schneider, Takshveer Singh, William Stokes, Linda Ward, and Nathan

Zelyas. Saliva collected in universal transport media is an effective, simple and high-volume amenable method to detect sars-cov-2. *Clinical Microbiology and Infection*, 27(4):656–657, 2021.

[22] Sonja Berensmeier. Magnetic particles for the separation and purification of nucleic acids. *Appl. Microbiol. Biotechnol.*, 73(3):495–504, December 2006.

[23] Kelvin R Berryman. Distribution, age, and deformation of late pleistocene marine terraces at mahia peninsula, hikurangi subduction margin, new zealand. *Tectonics*, 12(6):1365–1379, 1993.

[24] Simon Elias Bibri. The anatomy of the data-driven smart sustainable city: instrumentation, datafication, computerization and related applications. *Journal of Big Data*, 6(1):59, July 2019.

[25] Karen H Bieluch, Theodore Willis, Jason Smith, and Karen A Wilson. The complexities of counting fish: engaging citizen scientists in fish monitoring. *Maine policy review*, 26(2), 2017.

[26] Geoffrey C Bowker. Biodiversity datadiversity. *Social studies of science*, 30(5):643–683, 2000.

[27] Sierra M Brooks and Hal S Alper. Applications, challenges, and needs for employing synthetic biology beyond the lab. *Nature Communications*, 12(1):1–16, 2021.

[28] James P Broughton, Xianding Deng, Guixia Yu, Clare L Fasching, Venice Servellita, Jasmeet Singh, Xin Miao, Jessica A Streithorst, Andrea Granados, Alicia Sotomayor-Gonzalez, et al. Crispr–cas12-based detection of sars-cov-2. *Nature biotechnology*, 38(7):870–874, 2020.

[29] Richard Bruch, Julia Baaske, Claire Chatelle, Mailin Meirich, Sibylle Madlener, Wilfried Weber, Can Dincer, and Gerald Anton Urban. Crispr/cas13a-powered electrochemical microfluidic biosensor for nucleic acid amplification-free mirna diagnostics. *Advanced Materials*, 31(51):1905311, 2019.

[30] Paolo Calistri, Laura Amato, Ilaria Puglia, Francesca Cito, Alessandra Di Giuseppe, Maria Luisa Danzetta, Daniela Morelli, Marco Di Domenico, Marialuigia Caporale, Silvia Scialabba, et al. Infection sustained by lineage b. 1.1. 7 of sars-cov-2 is characterised by longer persistence and higher viral rna loads in nasopharyngeal swabs. *International Journal of Infectious Diseases*, 105:753–755, 2021.

[31] Mark Cartwright, Martin Rottman, Nathan I Shapiro, Benjamin Seiler, Patrick Lombardo, Nazita Gamini, Julie Tomolonis, Alexander L Watters, Anna Waterhouse, Dan Leslie, Dana Bolgen, Amanda Graveline, Joo H Kang, Tohid Didar, Nikolaos Dimitrakakis, David Cartwright, Michael Super, and Donald E Ingber. A Broad-Spectrum infection diagnostic that detects Pathogen-Associated

molecular patterns (PAMPs) in whole blood. *EBioMedicine*, 9:217–227, July 2016.

[32] Ricardo Cavicchioli, William J Ripple, Kenneth N Timmis, Farooq Azam, Lars R Bakken, Matthew Baylis, Michael J Behrenfeld, Antje Boetius, Philip W Boyd, Aimée T Classen, et al. Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*, 17(9):569–586, 2019.

[33] Aleksandra Checinska, Alexander J Probst, Parag Vaishampayan, James R White, Deepika Kumar, Victor G Stepanov, George E Fox, Henrik R Nilsson, Duane L Pierson, Jay Perry, et al. Microbiomes of the dust particles collected from the international space station and spacecraft assembly facilities. *Microbiome*, 3(1):1–18, 2015.

[34] Aleksandra Checinska Sielaff, Camilla Urbaniak, Ganesh Babu Malli Mohan, Victor G Stepanov, Quyen Tran, Jason M Wood, Jeremiah Minich, Daniel McDonald, Teresa Mayer, Rob Knight, et al. Characterization of the total and viable bacterial and fungal communities associated with the international space station surfaces. *Microbiome*, 7(1):1–21, 2019.

[35] Janice S Chen, Enbo Ma, Lucas B Harrington, Maria Da Costa, Xinran Tian, Joel M Palefsky, and Jennifer A Doudna. Crispr-cas12a target binding unleashes indiscriminate single-stranded dnase activity. *Science*, 360(6387):436–439, 2018.

[36] Peng Chen, Jin Zhou, Yibin Wan, Huan Liu, Yongzheng Li, Zhaoxin Liu, Hongjian Wang, Jun Lei, Kai Zhao, Yiliang Zhang, et al. A cas12a ortholog with stringent pam recognition followed by low off-target editing rates for genome editing. *Genome biology*, 21(1):1–13, 2020.

[37] Yanju Chen, Yang Liu, Ya Shi, Jianfeng Ping, Jian Wu, and Huan Chen. Magnetic particles for integrated nucleic acid purification, amplification and detection without pipetting. *Trends Analyt. Chem.*, 127:115912, June 2020.

[38] Joanne Clapcott, Jamie Ataria, Chris Hepburn, Dan Hikuroa, Anne-Marie Jackson, Rauru Kirikiri, and Erica Williams. Mātauranga māori: shaping marine and freshwater futures. *N. Z. J. Mar. Freshwater Res.*, 52(4):457–466, October 2018.

[39] Peter JA Cock, Tiago Antao, Jeffrey T Chang, Brad A Chapman, Cymon J Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski, et al. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.

[40] Cathy C Conrad and Krista G Hilchey. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environ. Monit. Assess.*, 176(1-4):273–291, May 2011.

[41] Cathy C Conrad and Krista G Hilchey. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental monitoring and assessment*, 176(1):273–291, 2011.

[42] Zachary Crannell, Alejandro Castellanos-Gonzalez, Gayatri Nair, Rojelio Mejia, A Clinton White, and Rebecca Richards-Kortum. Multiplexed recombinase polymerase amplification assay to detect intestinal protozoa. *Analytical chemistry*, 88(3):1610–1616, 2016.

[43] Morgan Elisabeth Currie. *The Data-fication of Openness - The Practices and Policies of Open Government Data in Los Angeles.* PhD thesis, UCLA, 2016.

[44] Lauren M Czaplicki and Claudia K Gunsch. Reflection on molecular approaches influencing state-of-the-art bioremediation design: culturing to microbial community fingerprinting to omics. *Journal of environmental engineering (New York, NY)*, 142(10), 2016.

[45] Craig M Dalton, Linnet Taylor, and Jim Thatcher. Critical data studies: A dialog on data and space. *Big Data & Society*, 3(1):2053951716648346, 2016.

[46] Helena de Puig, Irene Bosch, James J Collins, and Lee Gehrke. Point-of-care devices to detect zika and other emerging viruses. *Annual review of biomedical engineering*, 22:371–386, 2020.

[47] Helena de Puig, Rose A Lee, Devora Najjar, Xiao Tan, Luis R Soenksen, Nicolaas M Angenent-Mari, Nina M Donghia, Nicole E Weckman, Audrey Ory, Carlos F Ng, et al. Minimally instrumented sherlock (misherlock) for crispr-based point-of-care diagnosis of sars-cov-2 and emerging variants. *Science Advances*, 7(32):eabh2944, 2021.

[48] Birgit Deiman, Pierre van Aarle, and Peter Sillekens. Characteristics and applications of nucleic acid sequence-based amplification (NASBA). *Mol. Biotechnol.*, 20(2):163–179, February 2002.

[49] Sarah Schaffer DeRoo, Natalie J Pudalov, and Linda Y Fu. Planning for a covid-19 vaccination program. *Jama*, 323(24):2458–2459, 2020.

[50] Xiong Ding, Kun Yin, Ziyue Li, Rajesh V Lalla, Enrique Ballesteros, Maroun M Sfeir, and Changchun Liu. Ultrasensitive and visual detection of sars-cov-2 using all-in-one dual crispr-cas12a assay. *Nature communications*, 11(1):1–10, 2020.

[51] Stefania Dispinseri, Massimiliano Secchi, Maria Franca Pirillo, Monica Tolazzi, Martina Borghi, Cristina Brigatti, Maria Laura De Angelis, Marco Baratella, Elena Bazzigaluppi, Giulietta Venturi, et al. Neutralizing antibody responses to sars-cov-2 in symptomatic covid-19 is persistent and critical for survival. *Nature communications*, 12(1):1–12, 2021.

[52] Antonia A Dominguez, Wendell A Lim, and Lei S Qi. Beyond editing: repurposing crispr–cas9 for precision genome regulation and interrogation. *Nature reviews Molecular cell biology*, 17(1):5–15, 2016.

[53] Anja Dörschug, Hagen Frickmann, Julian Schwanbeck, Elif Yilmaz, Kemal Mese, Andreas Hahn, Uwe Groß, and Andreas E Zautner. Comparative assessment of sera from individuals after s-gene rna-based sars-cov-2 vaccination with spike-protein-based and nucleocapsid-based serological assays. *Diagnostics*, 11(3):426, 2021.

[54] Paul Dourish and Edgar Gómez Cruz. Datafication and data fiction: Narrating data and narrating with data. *Big Data & Society*, 5(2):2053951718784083, 2018.

[55] Ramakrishnan Durairajan, Carol Barford, and Paul Barford. Lights out: Climate change risk to internet infrastructure. In *Proceedings of the Applied Networking Research Workshop*, pages 9–15, 2018.

[56] John L Durant and Kalsoum Abbasi. Water resources management in the mystic river, watershed i: Water quality history and challenges for the future. In *Watershed Management and Operations Management 2000*, pages 1–10. 2000.

[57] Stefan Elbe and Gemma Buckland-Merrett. Data, disease and diplomacy: Gisaid's innovative contribution to global health. *Global challenges*, 1(1):33–46, 2017.

[58] Max A English, Luis R Soenksen, Raphael V Gayet, Helena de Puig, Nicolaas M Angenent-Mari, Angelo S Mao, Peter Q Nguyen, and James J Collins. Programmable crispr-responsive smart materials. *Science*, 365(6455):780–785, 2019.

[59] Marisol Estrella and John Gaventa. Who counts reality? participatory monitoring and evaluation: A literature review. 1998.

[60] Deborah Estrin, Jason C Fisher, Robert Gilbert, Mark H Hansen, Thomas C Harmon, Jennifer Jay, William J Kaiser, Gaurav S Sukhatme, and Yu-Chong Tai. Distributed sensing systems for water quality assessment and management. 2007.

[61] Jesse Fajnzylber, James Regan, Kendyll Coxen, Heather Corry, Colline Wong, Alexandra Rosenthal, Daniel Worrall, Francoise Giguel, Alicja Piechocka-Trocha, Caroline Atyeo, et al. Sars-cov-2 viral load is associated with increased disease severity and mortality. *Nature communications*, 11(1):1–9, 2020.

[62] US FDA. Faqs on diagnostic testing for sars cov-2, 2020.

[63] Craig Fenwick, Antony Croxatto, Alix T Coste, Florence Pojer, Cyril André, Céline Pellaton, Alex Farina, Jérémy Campos, David Hacker, Kelvin Lau, et al.

Changes in sars-cov-2 spike versus nucleoprotein antibody responses impact the estimates of infections in population-based seroprevalence studies. *Journal of virology*, 95(3):e01828–20, 2021.

[64] Axel Fey, Stefan Eichler, Sébastien Flavier, Richard Christen, Manfred G Höfle, and Carlos A Guzmán. Establishment of a real-time pcr-based approach for accurate quantification of bacterial rna targets in water, using salmonella as a model organism. *Applied and environmental microbiology*, 70(6):3618–3623, 2004.

[65] Yaara Finkel, Orel Mizrahi, Aharon Nachshon, Shira Weingarten-Gabbay, David Morgenstern, Yfat Yahalom-Ronen, Hadas Tamir, Hagit Achdout, Dana Stein, Ofir Israeli, et al. The coding capacity of sars-cov-2. *Nature*, 589(7840):125–130, 2021.

[66] US Food, Drug Administration, Lucira COVID, et al. All-in-one test kit emergency use authorization letter of authorization, 19.

[67] Parinaz Fozouni, Sungmin Son, María Díaz de León Derby, Gavin J Knott, Carley N Gray, Michael V D'Ambrosio, Chunyu Zhao, Neil A Switz, G Renuka Kumar, Stephanie I Stephens, et al. Amplification-free detection of sars-cov-2 with crispr-cas13a and mobile phone microscopy. *Cell*, 184(2):323–333, 2021.

[68] Yu Fu, Xiaoming Zhou, and Da Xing. Integrated paper-based detection chip with nucleic acid extraction and amplification for automatic and sensitive pathogen detection. *Sens. Actuators B Chem.*, 261:288–296, May 2018.

[69] Jennifer Gabrys. Sensors and sensing practices: Reworking experience across entities, environments, and technologies, 2019.

[70] Jennifer Gabrys, Helen Pritchard, and Benjamin Barratt. Just good enough data: Figuring data citizenships through air pollution sensing and data stories. *Big Data & Society*, 3(2):2053951716679677, December 2016.

[71] Jennifer Gabrys, Helen Pritchard, and Benjamin Barratt. Just good enough data: Figuring data citizenships through air pollution sensing and data stories. *Big Data & Society*, 3(2):2053951716679677, 2016.

[72] Christian Gaebler, Zijun Wang, Julio CC Lorenzi, Frauke Muecksch, Shlomo Finkin, Minami Tokuyama, Alice Cho, Mila Jankovic, Dennis Schaefer-Babajew, Thiago Y Oliveira, et al. Evolution of antibody immunity to sars-cov-2. *Nature*, 591(7851):639–644, 2021.

[73] Eduardo Garcia-Rico, Ramon A Alvarez-Puebla, and Luca Guerrini. Direct surface-enhanced raman scattering (sers) spectroscopy of nucleic acids: From fundamental studies to real-life applications. *Chemical Society Reviews*, 47(13):4909–4923, 2018.

[74] Raphael V Gayet, Helena de Puig, Max A English, Luis R Soenksen, Peter Q Nguyen, Angelo S Mao, Nicolaas M Angenent-Mari, and James J Collins. Creating crispr-responsive smart materials for diagnostics and programmable cargo release. *Nature Protocols*, 15(9):3030–3063, 2020.

[75] Clifford Geertz. Local knowledge" and its limits: Some" obiter dicta. *The Yale Journal of Criticism*, 5(2):129, 1992.

[76] Ann Gerhardt. Bioindicator species and their use in biomonitoring. *Environmental monitoring*, 1:77–123, 2002.

[77] Anthony Giddens. *The constitution of society: Outline of the theory of structuration*. Univ of California Press, 1984.

[78] Peter B Gilbert, David C Montefiori, Adrian B McDermott, Youyi Fong, David Benkeser, Weiping Deng, Honghong Zhou, Christopher R Houchens, Karen Martins, Lakshmi Jayashankar, et al. Immune correlates analysis of the mrna-1273 covid-19 vaccine efficacy clinical trial. *Science*, 375(6576):43–50, 2022.

[79] Tanya Golubchik, Katrina A Lythgoe, Matthew Hall, Luca Ferretti, Helen R Fryer, George MacIntyre-Cockett, Mariateresa de Cesare, Amy Trebes, Paolo Piazza, David Buck, et al. Early analysis of a potential link between viral load and the n501y mutation in the sars-cov-2 spike protein. *MedRxiv*, 2021.

[80] Jonathan S Gootenberg, Omar O Abudayyeh, Max J Kellner, Julia Joung, James J Collins, and Feng Zhang. Multiplexed and portable nucleic acid detection platform with cas13, cas12a, and csm6. *Science*, 360(6387):439–444, 2018.

[81] Jonathan S Gootenberg, Omar O Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J Dy, Julia Joung, Vanessa Verdine, Nina Donghia, Nichole M Daringer, Catherine A Freije, et al. Nucleic acid detection with crispr-cas13a/c2c2. *Science*, 356(6336):438–442, 2017.

[82] Christelle Gramaglia and François Mélard. Looking for the cosmopolitical fish: monitoring marine pollution with anglers and congers in the gulf of fos, southern france. *Science, Technology, & Human Values*, 44(5):814–842, 2019.

[83] Allison J Greaney, Tyler N Starr, Pavlo Gilchuk, Seth J Zost, Elad Binshtein, Andrea N Loes, Sarah K Hilton, John Huddleston, Rachel Eguia, Katharine HD Crawford, et al. Complete mapping of mutations to the sars-cov-2 spike receptor-binding domain that escape antibody recognition. *Cell host & microbe*, 29(1):44–57, 2021.

[84] Alexander A Green, Pamela A Silver, James J Collins, and Peng Yin. Toehold switches: de-novo-designed regulators of gene expression. *Cell*, 159(4):925–939, 2014.

[85] Stephan Gromer, Heiko Merkle, and R Heiner Schirmer. Preparation and inhibitor studies. *Methods in enzymology*, 347(075.687902):535–00, 2002.

[86] Daniel F Gudbjartsson, Gudmundur L Norddahl, Pall Melsted, Kristbjorg Gunnarsdottir, Hilma Holm, Elias Eythorsson, Asgeir O Arnthorsson, Dadi Helgason, Kristbjorg Bjarnadottir, Ragnar F Ingvarsson, et al. Humoral immune response to sars-cov-2 in iceland. *New England Journal of Medicine*, 383(18):1724–1734, 2020.

[87] Stephen J Hadfield, Guy Robinson, Kristin Elwin, and Rachel M Chalmers. Detection and differentiation of cryptosporidium spp. in human clinical samples by use of real-time pcr. *Journal of clinical microbiology*, 49(3):918–924, 2011.

[88] Reza Hajian, Sarah Balderston, Thanhtra Tran, Tara DeBoer, Jessy Etienne, Mandeep Sandhu, Noreen A Wauford, Jing-Yi Chung, Jolie Nokes, Mitre Athaiya, et al. Detection of unamplified target genes via crispr–cas9 immobilized on a graphene field-effect transistor. *Nature biomedical engineering*, 3(6):427–437, 2019.

[89] Benjamin S Halpern, Kimberly A Selkoe, Fiorenza Micheli, and Carrie V Kappel. Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. *Conservation biology*, 21(5):1301–1315, 2007.

[90] Donna Jeanne Haraway. *The companion species manifesto: Dogs, people, and significant otherness*, volume 1. Prickly Paradigm Press Chicago, 2003.

[91] William T Harvey, Alessandro M Carabelli, Ben Jackson, Ravindra K Gupta, Emma C Thomson, Ewan M Harrison, Catherine Ludden, Richard Reeve, Andrew Rambaut, Sharon J Peacock, et al. Sars-cov-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*, 19(7):409–424, 2021.

[92] Richard Heeks and Satyarupa Shekhar. Datafication, development and marginalised urban communities: an applied data justice framework. *Inf. Commun. Soc.*, 22(7):992–1011, June 2019.

[93] Christian A Heid, Junko Stevens, Kenneth J Livak, and P Mickey Williams. Real time quantitative pcr. *Genome research*, 6(10):986–994, 1996.

[94] Benjamin Heidt, Williane F Siqueira, Kasper Eersels, Hanne Diliën, Bart van Grinsven, Ricardo T Fujiwara, and Thomas J Cleij. Point of care diagnostics in Resource-Limited settings: A review of the present and future of PoC in its most needed environment. *Biosensors*, 10(10), September 2020.

[95] Shichu Huang, Jaephil Do, Madhumita Mahalanabis, Andy Fan, Lei Zhao, Lisa Jepeal, Satish K Singh, and Catherine M Klapperich. Low cost extraction and isothermal amplification of DNA for infectious diarrhea diagnosis. *PLoS One*, 8(3):e60059, March 2013.

[96] Yuan Huang, Chan Yang, Xin-feng Xu, Wei Xu, and Shu-wen Liu. Structural and functional properties of sars-cov-2 spike protein: potential antivirus drug development for covid-19. *Acta Pharmacologica Sinica*, 41(9):1141–1149, 2020.

[97] Alan Irwin and Irwin a Staff. *Citizen Science: A Study of People, Expertise and Sustainable Development*. Psychology Press, 1995.

[98] Baweleta Isho, Kento T Abe, Michelle Zuo, Alainna J Jamal, Bhavisha Rathod, Jenny H Wang, Zhijie Li, Gary Chao, Olga L Rojas, Yeo Myong Bang, et al. Persistence of serum and saliva antibody responses to sars-cov-2 spike antigens in covid-19 patients. *Science immunology*, 5(52):eabe5511, 2020.

[99] Sumio Iwasaki, Shinichi Fujisawa, Sho Nakakubo, Keisuke Kamada, Yu Yamashita, Tatsuya Fukumoto, Kaori Sato, Satoshi Oguri, Keisuke Taki, Hajime Senjo, et al. Comparison of sars-cov-2 detection in nasopharyngeal swab and saliva. *Journal of Infection*, 81(2):e145–e147, 2020.

[100] Enedina Jiménez-Cardoso, Leticia Eligio-García, Adrián Cortés-Campos, Andrés Flores-Luna, Pedro Valencia-Mayoral, and Irma Lozada-Chávez. Changes in $\beta$-giardin sequence of giardia intestinalis sensitive and resistant to albendazole strains. *Parasitology research*, 105(1):25–33, 2009.

[101] Ashlee Jollymore, Morgan J Haines, Terre Satterfield, and Mark S Johnson. Citizen science for water quality monitoring: Data implications of citizen perspectives. *Journal of Environmental Management*, 200:456–467, 2017.

[102] Matthew R Jones and Helena Karsten. Giddens's structuration theory and information systems research. *MIS quarterly*, pages 127–157, 2008.

[103] M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. Galaxy zoo: Exploring the motivations of citizen science volunteers. September 2009.

[104] Julia Joung, Alim Ladha, Makoto Saito, Nam-Gyun Kim, Ann E Woolley, Michael Segel, Robert PJ Barretto, Amardeep Ranu, Rhiannon K Macrae, Guilhem Faure, et al. Detection of sars-cov-2 with sherlock one-pot testing. *New England Journal of Medicine*, 383(15):1492–1494, 2020.

[105] Michael M Kaminski, Omar O Abudayyeh, Jonathan S Gootenberg, Feng Zhang, and James J Collins. CRISPR-based diagnostics. *Nat Biomed Eng*, 5(7):643–656, July 2021.

[106] Michael M Kaminski, Miguel A Alcantar, Isadora T Lape, Robert Greensmith, Allison C Huske, Jacqueline A Valeri, Francisco M Marty, Verena Klämbt, Jamil Azzi, Enver Akalin, et al. A crispr-based assay for the detection of opportunistic infections post-transplantation and for the monitoring of transplant rejection. *Nature biomedical engineering*, 4(6):601–609, 2020.

[107] Martha Kaplan. Lonely drinking fountains and comforting coolers: paradoxes of water value and ironies of water use. *Cult. Anthropol.*, 26(4):514–541, 2011.

[108] Margot Karlikow, Severino Jefferson Ribeiro da Silva, Yuxiu Guo, Seray Cicek, Larissa Krokovsky, Paige Homme, Yilin Xiong, Talia Xu, Maria-Angelica Calderón-Peláez, Sigrid Camacho-Ortega, et al. Field validation of the performance of paper-based tests for the detection of the zika and chikungunya viruses in serum samples. *Nature biomedical engineering*, 6(3):246–256, 2022.

[109] Kazutaka Katoh and Daron M Standley. Mafft multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780, 2013.

[110] Navjot Kaur and Bhushan J Toley. Paper-based nucleic acid amplification tests for point-of-care diagnostics. *Analyst*, 143(10):2213–2234, May 2018.

[111] Michael Keller, David S Schimel, W William Hargrove, and Forrest M Hoffman. A continental strategy for the national ecological observatory network. *The Ecological Society of America: 282-284*, 2008.

[112] Max J Kellner, Jeremy G Koob, Jonathan S Gootenberg, Omar O Abudayyeh, and Feng Zhang. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.*, 14(10):2986–3012, October 2019.

[113] Robin Wall Kimmerer. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions, September 2013.

[114] Abby Kinchy. Citizen science and democracy: Participatory water monitoring in the marcellus shale fracking boom. *Science as Culture*, 26(1):88–110, 2017.

[115] Rob Kitchin and Gavin McArdle. What makes big data, big data? exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1):2053951716631130, 2016.

[116] Katarzyna M Koczula and Andrea Gallotta. Lateral flow assays. *Essays in biochemistry*, 60(1):111–120, 2016.

[117] Ti-Hsuan Ku, Tiantian Zhang, Hua Luo, Tony M Yen, Ping-Wei Chen, Yuanyuan Han, and Yu-Hwa Lo. Nucleic acid aptamers: an emerging tool for biotechnology and biomedical sensing. *Sensors*, 15(7):16281–16313, 2015.

[118] Cem Kuscu, Mahmut Parlak, Turan Tufan, Jiekun Yang, Karol Szlachta, Xiaolong Wei, Rashad Mammadov, and Mazhar Adli. Crispr-stop: gene silencing through base-editing-induced nonsense mutations. *Nature methods*, 14(7):710–712, 2017.

[119] Matthew A Lalli, Joshua S Langmade, Xuhua Chen, Catrina C Fronick, Christopher S Sawyer, Lauren C Burcea, Michael N Wilkinson, Robert S Fulton, Michael Heinz, William J Buchser, et al. Rapid and extraction-free detection

of sars-cov-2 from saliva by colorimetric reverse-transcription loop-mediated isothermal amplification. *Clinical chemistry*, 67(2):415–424, 2021.

[120] D Law. Virulence factors of escherichia coli o157 and other shiga toxin-producing e. coli. *Journal of Applied Microbiology*, 88(5):729–745, 2000.

[121] Rose A Lee, Helena De Puig, Peter Q Nguyen, Nicolaas M Angenent-Mari, Nina M Donghia, James P McGee, Jeffrey D Dvorin, Catherine M Klapperich, Nira R Pollock, and James J Collins. Ultrasensitive crispr-based diagnostic for field-applicable detection of plasmodium species in symptomatic and asymptomatic malaria. *Proceedings of the National Academy of Sciences*, 117(41):25722–25731, 2020.

[122] Rose A Lee, Joshua C Herigon, Andrea Benedetti, Nira R Pollock, and Claudia M Denkinger. Performance of saliva, oropharyngeal swabs, and nasal swabs for sars-cov-2 molecular detection: a systematic review and meta-analysis. *Journal of clinical microbiology*, 59(5):e02881–20, 2021.

[123] Sabina Leonelli. On the locality of data and claims about phenomena. *Philosophy of Science*, 76(5):737–749, 2009.

[124] Sabina Leonelli. What difference does quantity make? on the epistemology of big data in biology. *Big data & society*, 1(1):2053951714534395, 2014.

[125] Fan Li, Qinghua Ye, Moutong Chen, Baoqing Zhou, Jumei Zhang, Rui Pang, Liang Xue, Juan Wang, Haiyan Zeng, Shi Wu, et al. An ultrasensitive crispr/cas12a based electrochemical biosensor for listeria monocytogenes detection. *Biosensors and Bioelectronics*, 179:113073, 2021.

[126] Laam Li, Nilmini Mendis, Hana Trigui, James D Oliver, and Sebastien P Faucher. The importance of the viable but non-culturable state in human bacterial pathogens. *Frontiers in microbiology*, 5:258, 2014.

[127] Yi Li, Shiyuan Li, Jin Wang, and Guozhen Liu. Crispr/cas systems towards next-generation biosensing. *Trends in biotechnology*, 37(7):730–743, 2019.

[128] Romina Libster, Gonzalo Pérez Marc, Diego Wappner, Silvina Coviello, Alejandra Bianchi, Virginia Braem, Ignacio Esteban, Mauricio T Caballero, Cristian Wood, Mabel Berrueta, et al. Early high-titer plasma therapy to prevent severe covid-19 in older adults. *New England Journal of Medicine*, 384(7):610–618, 2021.

[129] Michelle Liebetreu. 50 years: Celebrating the past, present and future of the mystic river watershed, Mar 2022.

[130] JC Linnes, NM Rodriguez, L Liu, and CM Klapperich. Polyethersulfone improves isothermal nucleic acid amplification compared to current paper-based diagnostics. *Biomedical microdevices*, 18(2):1–12, 2016.

[131] Hai-bin Liu, Yu-Xuan Zang, Xin-jun Du, Ping Li, and Shuo Wang. Development of an isothermal amplification-based assay for the rapid visual detection of salmonella bacteria. *Journal of Dairy Science*, 100(9):7016–7025, 2017.

[132] Rui Liu, Xinghui Liu, Li Yuan, Huan Han, Muhammad Adnan Shereen, Jiesheng Zhen, Zhili Niu, Dong Li, Fang Liu, Kailang Wu, et al. Analysis of adjunctive serological detection to nucleic acid test for severe acute respiratory syndrome coronavirus 2 (sars-cov-2) infection diagnosis. *International immunopharmacology*, 86:106746, 2020.

[133] Xiaoqing Liu, Qiongying Yan, Jianfei Huang, Jing Chen, Zhengyang Guo, Zhongdong Liu, Lin Cai, Risheng Li, Yan Wang, Guowu Yang, et al. Influence of design probe and sequence mismatches on the efficiency of fluorescent rpa. *World Journal of Microbiology and Biotechnology*, 35(6):1–11, 2019.

[134] Zhuoming Liu, Laura A VanBlargan, Louis-Marie Bloyet, Paul W Rothlauf, Rita E Chen, Spencer Stumpf, Haiyan Zhao, John M Errico, Elitza S Theel, Mariel J Liebeskind, et al. Identification of sars-cov-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell host & microbe*, 29(3):477–488, 2021.

[135] Ivan Magriñá Lobato and Ciara K O'Sullivan. Recombinase polymerase amplification: Basics, applications and recent advances. *Trends Analyt. Chem.*, 98:19–35, January 2018.

[136] Yanni Alexander Loukissas. *All data are local: Thinking critically in a data-driven society*. MIT Press, 2019.

[137] Richard Louv and John W Fitzpatrick. *Citizen science: Public participation in environmental research*. Cornell University Press, 2012.

[138] Renfei Lu, Xiuming Wu, Zhenzhou Wan, Yingxue Li, Lulu Zuo, Jianru Qin, Xia Jin, and Chiyu Zhang. Development of a novel reverse transcription loop-mediated isothermal amplification method for rapid detection of sars-cov-2. *Virologica Sinica*, 35(3):344–347, 2020.

[139] John HT Luong, Keith B Male, and Jeremy D Glennon. Biosensor technology: technology push versus market pull. *Biotechnology advances*, 26(5):492–500, 2008.

[140] Sachit Mahajan, Jennifer Gabrys, and Joanne Armitage. Airkit: a citizen-sensing toolkit for monitoring air quality. *Sensors*, 21(12):4044, 2021.

[141] Gannon CK Mak, Peter KC Cheng, Stephen SY Lau, Kitty KY Wong, CS Lau, Edman TK Lam, Rickjason CW Chan, and Dominic NC Tsang. Evaluation of rapid antigen test for detection of sars-cov-2 virus. *Journal of Clinical Virology*, 129:104500, 2020.

[142] Charlotte Manisty, Thomas Alexander Treibel, Melanie Jensen, Amanda Semper, George Joy, Rishi K Gupta, Teresa Cutino-Moguel, Mervyn Andiapen, Jessica Jones, Stephen Taylor, et al. Time series analysis and mechanistic modelling of heterogeneity and sero-reversion in antibody responses to mild sars-cov-2 infection. *EBioMedicine*, 65:103259, 2021.

[143] Vivien Marx. Base editing a crispr way. *Nature Methods*, 15(10):767–770, 2018.

[144] Viktor Mayer-Schönberger and Kenneth Cukier. *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013.

[145] Brian Mayton, Gershon Dublon, Spencer Russell, Evan F Lynch, Don Derek Haddad, Vasant Ramasubramanian, Clement Duhart, Glorianna Davenport, and Joseph A Paradiso. The networked sensory landscape: Capturing and experiencing ecological change across scales. *Presence*, 26(2):182–209, 2017.

[146] Tara G McAllister, Jacqueline R Beggs, Shaun Ogilvie, Rauru Kirikiri, Amanda Black, and Priscilla M Wehi. Kua takoto te mānuka: mātauranga māori in new zealand ecology. *N. Z. J. Ecol.*, 43(3):1–7, 2019.

[147] Alaric McCarthy, Chris Hepburn, Nigel Scott, Katja Schweikert, Rachel Turner, and Henrik Moller. Local people see and care most? severe depletion of inshore fisheries and its consequences for māori communities in new zealand. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24(3):369–390, 2014.

[148] Duncan C McKinley, Abe J Miller-Rushing, Heidi L Ballard, Rick Bonney, Hutch Brown, Susan C Cook-Patton, Daniel M Evans, Rebecca A French, Julia K Parrish, Tina B Phillips, et al. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation*, 208:15–28, 2017.

[149] Christopher A Mebane, John P Sumpter, Anne Fairbrother, Thomas P Augspurger, Timothy J Canfield, William L Goodfellow, Patrick D Guiney, Anne LeHuray, Lorraine Maltby, David B Mayfield, et al. Scientific integrity issues in environmental toxicology and chemistry: Improving research reproducibility, credibility, and transparency. *Integrated environmental assessment and management*, 15(3):320–344, 2019.

[150] Benjamin Meyer, Christian Drosten, and Marcel A Müller. Serological assays for emerging coronaviruses: challenges and pitfalls. *Virus research*, 194:175–183, 2014.

[151] Mike Michael. "what are we busy doing?": Engaging the idiot. *Sci. Technol. Human Values*, 37(5):528–554, September 2012.

[152] L Scott Mills, Michael E Soulé, and Daniel F Doak. The keystone-species concept in ecology and conservation. *BioScience*, 43(4):219–224, 1993.

[153] Margaret S Morley, Bruce W Hayward, Jenny L Raven, Geoff A Foreman, and Hugh R Grenfell. Intertidal and shallow subtidal biota of mahia peninsula, hawkes bay. *Records of the Auckland Museum*, pages 11–41, 2006.

[154] Kary B Mullis. *The polymerase chain reaction*, volume 41. Springer science & business media, 1994.

[155] Cameron Myhrvold, Catherine A Freije, Jonathan S Gootenberg, Omar O Abudayyeh, Hayden C Metsky, Ann F Durbin, Max J Kellner, Amanda L Tan, Lauren M Paul, Leda A Parham, Kimberly F Garcia, Kayla G Barnes, Bridget Chak, Adriano Mondini, Mauricio L Nogueira, Sharon Isern, Scott F Michael, Ivette Lorenzana, Nathan L Yozwiak, Bronwyn L MacInnis, Irene Bosch, Lee Gehrke, Feng Zhang, and Pardis C Sabeti. Field-deployable viral diagnostics using CRISPR-Cas13. *Science*, 360(6387):444–448, April 2018.

[156] Kawin Nawattanapaiboon, Ekawat Pasomsub, Photchanathorn Prombun, Akanit Wongbunmak, Akarawit Jenjitwanich, Pantanat Mahasupachai, Purichaya Vetcho, Cholticha Chayrach, Natthapon Manatjaroenlap, Chonchanok Samphaongern, et al. Colorimetric reverse transcription loop-mediated isothermal amplification (rt-lamp) as a visual diagnostic platform for the detection of the emerging coronavirus sars-cov-2. *Analyst*, 146(2):471–477, 2021.

[157] Antoine Nkuba Ndaye, Ana Hoxha, Joule Madinga, Joachim Mariën, Martine Peeters, Fabian H Leendertz, Steve Ahuka Mundeke, Kevin K Ariën, Jean-Jacques Muyembe Tanfumu, Placide Mbala Kingebeni, et al. Challenges in interpreting sars-cov-2 serological results in african countries. *The Lancet Global Health*, 9(5):e588–e589, 2021.

[158] Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. The future of citizen science: emerging technologies and shifting paradigms. *Front. Ecol. Environ.*, 10(6):298–304, August 2012.

[159] Anna Christine Nilsson, Dorte Kinggaard Holm, Ulrik Stenz Justesen, Thøger Gorm-Jensen, Nanna Skaarup Andersen, Anne Øvrehus, Isik Somuncu Johansen, Jens Michelsen, Ulrik Sprogøe, and Søren Thue Lillevang. Comparison of six commercially available sars-cov-2 antibody assays—choice of assay depends on intended use. *International Journal of Infectious Diseases*, 103:381–388, 2021.

[160] Bo Ning, Tao Yu, Shengwei Zhang, Zhen Huang, Di Tian, Zhen Lin, Alex Niu, Nadia Golden, Krystle Hensley, Breanna Threeton, et al. A smartphone-read ultrasensitive and quantitative saliva test for covid-19. *Science advances*, 7(2):eabe3703, 2021.

[161] Eric Nost and Jenny Elaine Goldstein. A political ecology of data. *Environment and Planning E: Nature and Space*, 5(1):3–17, 2022.

[162] Tsugunori Notomi, Yasuyoshi Mori, Norihiro Tomita, and Hidetoshi Kanda. Loop-mediated isothermal amplification (LAMP): principle, features, and future prospects. *J. Microbiol.*, 53(1):1–5, January 2015.

[163] Reza Nouri, Zifan Tang, Ming Dong, Tianyi Liu, Aneesh Kshirsagar, and Weihua Guan. Crispr-based detection of sars-cov-2: A review from sample to result. *Biosensors and Bioelectronics*, 178:113012, 2021.

[164] Elinor Ostrom, Paul C Stern, and Thomas Dietz. Water rights in the commons. *Water Resources Impact*, 5(2):9–12, 2003.

[165] Isabel M Ott, Madison S Strine, Anne E Watkins, Maikel Boot, Chaney C Kalinich, Christina A Harden, Chantal BF Vogels, Arnau Casanovas-Massana, Adam J Moore, M Catherine Muenker, et al. Simply saliva: stability of sars-cov-2 detection negates the need for expensive collection devices. *MedRxiv*, 2020.

[166] Yang Pan, Daitao Zhang, Peng Yang, Leo LM Poon, and Quanyi Wang. Viral load of sars-cov-2 in clinical samples. *The Lancet infectious diseases*, 20(4):411–412, 2020.

[167] Maturada Patchsung, Krittapas Jantarug, Archiraya Pattama, Kanokpol Aphicho, Surased Suraritdechachai, Piyachat Meesawat, Khomkrit Sappakhaw, Nattawat Leelahakorn, Theerawat Ruenkam, Thanakrit Wongsatit, et al. Clinical validation of a cas13-based assay for the detection of sars-cov-2 rna. *Nature biomedical engineering*, 4(12):1140–1149, 2020.

[168] Rajesh Paul, Emily Ostermann, and Qingshan Wei. Advances in point-of-care nucleic acid extraction technologies for rapid diagnosis of human and plant diseases. *Biosensors and Bioelectronics*, 169:112592, 2020.

[169] Laura J Perovich, Sara Ann Wylie, and Roseann Bongiovanni. Chemicals in the creek: designing a situated data physicalization of open government data with the community. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):913–923, 2020.

[170] Duane L Pierson, Douglas J Botkin, Rebekah J Bruce, Victoria A Castro, Melanie J Smith, Cherie M Oubre, and C Mark Ott. Microbial monitoring of the international space station. In *8th International Workshop on Space Microbiology*, number JSC-CN-28760, 2013.

[171] Lakshmanane Premkumar, Bruno Segovia-Chumbez, Ramesh Jadi, David R Martinez, Rajendra Raut, Alena J Markmann, Caleb Cornaby, Luther Bartelt, Susan Weiss, Yara Park, et al. The receptor-binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in sars-cov-2 patients. *Science immunology*, 5(48):eabc8413, 2020.

[172] Acceptable Alternative Primer. Cdc 2019-ncov real-time rt-pcr diagnostic panel.

[173] Ye-Min Qu, En-Ming Kang, and Hai-Yan Cong. Positive result of sars-cov-2 in sputum from a cured patient with covid-19. *Travel medicine and infectious disease*, 34:101619, 2020.

[174] Ke Quan, Cuiping Yi, Xiaohai Yang, Xiaoxiao He, Jin Huang, and Kemin Wang. Fret-based nucleic acid probes: basic designs and applications in bioimaging. *TrAC Trends in Analytical Chemistry*, 124:115784, 2020.

[175] Brian A Rabe and Constance Cepko. Sars-cov-2 detection using isothermal amplification and a rapid, inexpensive protocol for sample inactivation and purification. *Proceedings of the National Academy of Sciences*, 117(39):24450–24458, 2020.

[176] M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. Galaxy zoo: Exploring the motivations of citizen science volunteers. *arXiv preprint arXiv:0909.2925*, 2009.

[177] Anne Rademacher. *Reigning the River*. Duke University Press, October 2011.

[178] Thandavarayan Ramamurthy, Amit Ghosh, Gururaja P Pazhani, and Sumio Shinoda. Current perspectives on viable but non-culturable (vbnc) pathogenic bacteria. *Frontiers in public health*, 2:103, 2014.

[179] M Mazhar Rathore, Awais Ahmad, Anand Paul, and Seungmin Rho. Urban planning and building smart cities based on the internet of things using big data analytics. *Computer Networks*, 101:63–80, June 2016.

[180] Jennifer N Rauch, Eric Valois, Sabrina C Solley, Friederike Braig, Ryan S Lach, Morgane Audouard, Jose Carlos Ponce-Rojas, Michael S Costello, Naomi J Baxter, Kenneth S Kosik, et al. A scalable, easy-to-deploy protocol for cas13-based detection of sars-cov-2 genetic material. *Journal of clinical microbiology*, 59(4):e02402–20, 2021.

[181] Pablo Rey-Mazón, Hagit Keysar, Shannon Dosemagen, Catherine D'Ignazio, and Don Blair. Public lab: Community-based approaches to urban and environmental health and justice. *Science and Engineering Ethics*, 24(3):971–997, 2018.

[182] Mere Roberts. Ways of seeing: Whakapapa. *Sites: a journal of social anthropology and cultural studies*, 10(1):93–120, 2013.

[183] Mere Roberts, Waerete Norman, Nganeko Minhinnick, Del Wihongi, and Carmen Kirkwood. Kaitiakitanga: Maori perspectives on conservation. *Pacific Conservation Biology*, 2(1):7–20, 1995.

[184] Lindsey R Robinson-McCarthy, Alexander J Mijalis, Gabriel T Filsinger, Helena De Puig, Nina M Donghia, Thomas E Schaus, Robert A Rasmussen, Raphael

Ferreira, Jeantine E Lunshof, George Chao, et al. Anomalous covid-19 tests hinder researchers. *Science*, 371(6526):244–245, 2021.

[185] Natalia M Rodriguez, Jacqueline C Linnes, Andy Fan, Courtney K Ellenson, Nira R Pollock, and Catherine M Klapperich. based rna extraction, in situ isothermal amplification, and lateral flow detection for low-cost, rapid diagnosis of influenza a (h1n1) from clinical specimens. *Analytical chemistry*, 87(15):7872–7879, 2015.

[186] Mohammed A Rohaim, Emily Clayton, Irem Sahin, Julianne Vilela, Manar E Khalifa, Mohammad Q Al-Natour, Mahmoud Bayoumi, Aurore C Poirier, Manoharanehru Branavan, Mukunthan Tharmakulasingam, et al. Artificial intelligence-assisted loop mediated isothermal amplification (ai-lamp) for rapid detection of sars-cov-2. *Viruses*, 12(9):972, 2020.

[187] Jonathan Sabaté Del Río, Olivier Y F Henry, Pawan Jolly, and Donald E Ingber. An antifouling coating that enables affinity-based electrochemical biosensing in complex biological fluids. *Nat. Nanotechnol.*, 14(12):1143–1149, December 2019.

[188] Sairaj Satarker and Madhavan Nampoothiri. Structural proteins in severe acute respiratory syndrome coronavirus-2. *Archives of medical research*, 51(6):482–491, 2020.

[189] Andrew Thomas Savage. Boston harbor: the anatomy of a court-run cleanup. *BC Envtl. Aff. L. Rev.*, 22:365, 1994.

[190] Caterina Scaramelli. Making sense of water quality: Multispecies encounters on the mystic river. *Worldviews: Global Religions, Culture, and Ecology*, 17(2):150–160, January 2013.

[191] Travis S Schlappi, Stephanie E McCalla, Nathan G Schoepp, and Rustem F Ismagilov. Flow-through capture and in situ amplification can enable rapid detection of a few single molecules of nucleic acids from several milliliters of solution. *Analytical chemistry*, 88(15):7647–7653, 2016.

[192] Krista L Schneider. *The Mystic River: Inventory, Analysis, and Opportunity : Issues Facing the Design and Development of a Comprehensive Master Plan*. Metropolitan District Commission, 1997.

[193] Cheryl B Schultz, Leone M Brown, Emma Pelton, and Elizabeth E Crone. Citizen science monitoring demonstrates dramatic declines of monarch butterflies in western north america. *Biological Conservation*, 214:343–346, 2017.

[194] Jan C Semenza. Cascading risks of waterborne diseases from climate change. *Nature Immunology*, 21(5):484–487, 2020.

[195] Nitin Kumar Singh, Jason M Wood, Fathi Karouia, and Kasthuri Venkateswaran. Succession and persistence of microbial communities and antimicrobial resistance genes associated with international space station environmental surfaces. *Microbiome*, 6(1):1–23, 2018.

[196] Barton E Slatko, Andrew F Gardner, and Frederick M Ausubel. Overview of next-generation sequencing technologies. *Current protocols in molecular biology*, 122(1):e59, 2018.

[197] Linda Tuhiwai Smith. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books Ltd., April 2021.

[198] Michael E Soule. What is environmental studies? *Bioscience*, 48(5):397–405, 1998.

[199] Sam C Staddon, Andrea Nightingale, and Shyam K Shrestha. The social nature of participatory ecological monitoring. *Soc. Nat. Resour.*, 27(9):899–914, September 2014.

[200] James T Staley, Richard W Castenholz, Rita R Colwell, John G Holt, Matthew D Kane, Norman R Pace, Abigail A Salyers, and James M Tiedje. The microbial world. *The Foundation of the Biosphere*, 32, 1997.

[201] Jessica C Stark, Ally Huang, Peter Q Nguyen, Rachel S Dubner, Karen J Hsu, Thomas C Ferrante, Mary Anderson, Ada Kanapskyte, Quinn Mucha, Jessica S Packett, Palak Patel, Richa Patel, Deema Qaq, Tyler Zondor, Julie Burke, Thomas Martinez, Ashlee Miller-Berry, Aparna Puppala, Kara Reichert, Miriam Schmid, Lance Brand, Lander R Hill, Jemima F Chellaswamy, Nuhie Faheem, Suzanne Fetherling, Elissa Gong, Eddie Marie Gonzalzles, Teresa Granito, Jenna Koritsaris, Binh Nguyen, Sujud Ottman, Christina Palffy, Angela Patel, Sheila Skweres, Adriane Slaton, Tarhonda Woods, Nina Donghia, Keith Pardee, James J Collins, and Michael C Jewett. BioBits™ bright: A fluorescent synthetic biology education kit. *Sci Adv*, 4(8):eaat5107, August 2018.

[202] Delphine Sterlin, Alexis Mathian, Makoto Miyara, Audrey Mohr, François Anna, Laetitia Claër, Paul Quentric, Jehane Fadlallah, Hervé Devilliers, Pascale Ghillani, et al. Iga dominates the early neutralizing antibody response to sars-cov-2. *Science translational medicine*, 13(577):eabd2223, 2021.

[203] Oliver W Stringer, Jennifer M Andrews, Hazel L Greetham, and Matthew S Forrest. TwistAmp® liquid: a versatile amplification method to replace PCR. *Nat. Methods*, 15:395, April 2018.

[204] Daan C Swarts, John van der Oost, and Martin Jinek. Structural basis for guide rna processing and seed-dependent dna targeting by crispr-cas12a. *Molecular cell*, 66(2):221–233, 2017.

[205] Xiao Tan, Justin H Letendre, James J Collins, and Wilson W Wong. Synthetic biology in the clinic: engineering vaccines, diagnostics, and therapeutics. *Cell*, 184(4):881–898, 2021.

[206] Rui Hua Tang, Hui Yang, Jane Ru Choi, Yan Gong, Shang Sheng Feng, Belinda Pingguan-Murphy, Qing Sheng Huang, Jun Ling Shi, Qi Bing Mei, and Feng Xu. Advances in paper-based sample pretreatment for point-of-care testing. *Crit. Rev. Biotechnol.*, 37(4):411–428, June 2017.

[207] Ruihua Tang, Min Li, Xueyan Yan, Mingyue Xie, Li Na Liu, Zedong Li, and Feng Xu. Comparison of paper-based nucleic acid extraction materials for point-of-care testing applications. *Cellulose*, 29(4):2479–2495, 2022.

[208] Rosina Taniwha. Wānanga–a distributive action. In *Māori and Pasifika Higher Education Horizons*. Emerald Group Publishing Limited, 2014.

[209] Kaiming Tao, Philip L Tzou, Janin Nouhin, Ravindra K Gupta, Tulio de Oliveira, Sergei L Kosakovsky Pond, Daniela Fera, and Robert W Shafer. The biological and clinical significance of emerging sars-cov-2 variants. *Nature Reviews Genetics*, 22(12):757–773, 2021.

[210] Dmitry Ter-Ovanesyan, Tal Gilboa, Roey Lazarovits, Alexandra Rosenthal, Xu Yu, Jonathan Z Li, George M Church, and David R Walt. Ultrasensitive measurement of both sars-cov-2 rna and antibodies from saliva. *Analytical Chemistry*, 93(13):5365–5370, 2021.

[211] Anne E Thessen, Benjamin Fertig, Jessie C Jarvis, and Adelaide C Rhodes. Data infrastructures for estuarine and coastal ecological syntheses. *Estuaries and coasts*, 39(2):295–310, 2016.

[212] Fang Tian, Bei Tong, Liang Sun, Shengchao Shi, Bin Zheng, Zibin Wang, Xianchi Dong, and Peng Zheng. Mutation n501y in rbd of spike protein strengthens the interaction between covid-19 and its receptor ace2. *BioRxiv*, 2021.

[213] Sanjay Sharma Timilsina, Nolan Durr, Mohamed Yafia, Hani Sallum, Pawan Jolly, and Donald E Ingber. Rapid antifouling nanocomposite coating enables highly sensitive multiplexed electrochemical detection of myocardial infarction and concussion markers. *medRxiv*, 2021.

[214] Gail Tipa and Laurel D Teirney. *A cultural health index for streams and waterways: a tool for nationwide use*. Ministry for the Environment Wellington, 2006.

[215] M Cristina Tirado, R Clarke, LA Jaykus, A McQuatters-Gollop, and JM Frank. Climate change and food safety: A review. *Food Research International*, 43(7):1745–1765, 2010.

[216] Norihiro Tomita, Yasuyoshi Mori, Hidetoshi Kanda, and Tsugunori Notomi. Loop-mediated isothermal amplification (lamp) of gene sequences and simple visual detection of products. *Nature protocols*, 3(5):877–882, 2008.

[217] Rebeca M Torrente-Rodríguez, Heather Lukas, Jiaobing Tu, Jihong Min, Yiran Yang, Changhao Xu, Harry B Rossiter, and Wei Gao. Sars-cov-2 rapidplex: a graphene-based multiplexed telemedicine platform for rapid and low-cost covid-19 diagnosis and monitoring. *Matter*, 3(6):1981–1998, 2020.

[218] Kenneth Tucker. Anthony giddens and modern social theory. *Anthony Giddens and Modern Social Theory*, pages 1–240, 1998.

[219] Buddhisha Udugama, Pranav Kadhiresan, Hannah N Kozlowski, Ayden Malekjahani, Matthew Osborne, Vanessa YC Li, Hongmin Chen, Samira Mubareka, Jonathan B Gubbay, and Warren CW Chan. Diagnosing covid-19: the disease and tools for detection. *ACS nano*, 14(4):3822–3835, 2020.

[220] Mickey Urdea, Laura A Penny, Stuart S Olmsted, Maria Y Giovanni, Peter Kaspar, Andrew Shepherd, Penny Wilson, Carol A Dahl, Steven Buchsbaum, Gerry Moeller, and Deborah C Hay Burgess. Requirements for high impact diagnostics in the developing world. *Nature*, 444 Suppl 1:73–79, November 2006.

[221] Mark J Valencia and David VanderZwaag. Maritime claims and management rights of indigenous peoples: Rising tides in the pacific and northern waters. *Ocean and Shoreline Management*, 12(2):125–167, 1989.

[222] Jan Van Elslande, Bram Decru, Stijn Jonckheere, Eric Van Wijngaerden, Els Houben, Patricia Vandecandelaere, Christophe Indevuyst, Melissa Depypere, Stefanie Desmet, Emmanuel André, et al. Antibody response against sars-cov-2 spike protein and nucleoprotein evaluated by four automated immunoassays and three elisas. *Clinical Microbiology and Infection*, 26(11):1557–e1, 2020.

[223] Jeroen JA van Kampen, David AMC van de Vijver, Pieter LA Fraaij, Bart L Haagmans, Mart M Lamers, Nisreen Okba, Johannes PC van den Akker, Henrik Endeman, Diederik AMPJ Gommers, Jan J Cornelissen, et al. Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (covid-19). *Nature communications*, 12(1):1–6, 2021.

[224] Jaco J Verweij, Janke Schinkel, Daphne Laeijendecker, Marianne AA van Rooyen, Lisette van Lieshout, and Anton M Polderman. Real-time pcr for the detection of giardia lamblia. *Molecular and cellular probes*, 17(5):223–225, 2003.

[225] Philip V'kovski, Annika Kratzel, Silvio Steiner, Hanspeter Stalder, and Volker Thiel. Coronavirus biology and replication: implications for sars-cov-2. *Nature Reviews Microbiology*, 19(3):155–170, 2021.

[226] Dean P Walker, James M Ataria, Kenneth F D Hughey, Patrick T Park, and John P Katene. Environmental and spatial planning with ngā atua kaitiaki: A mātauranga māori framework. *N. Z. Geog.*, 77(2):90–100, August 2021.

[227] G Terrance Walker, Melinda S Fraiser, James L Schram, Michael C Little, James G Nadeau, and Douglas P Malinowski. Strand displacement amplification—an isothermal, in vitro dna amplification technique. *Nucleic acids research*, 20(7):1691–1696, 1992.

[228] Liqian Wang, Xuejun Wang, Yungen Wu, Mingquan Guo, Chenjian Gu, Changhao Dai, Derong Kong, Yao Wang, Cong Zhang, Di Qu, et al. Rapid and ultrasensitive electromechanical detection of ions, biomolecules and sars-cov-2 rna in unamplified samples. *Nature Biomedical Engineering*, 6(3):276–285, 2022.

[229] Pengfei Wang, Manoj S Nair, Lihong Liu, Sho Iketani, Yang Luo, Yicheng Guo, Maple Wang, Jian Yu, Baoshan Zhang, Peter D Kwong, et al. Antibody resistance of sars-cov-2 variants b. 1.351 and b. 1.1. 7. *Nature*, 593(7857):130–135, 2021.

[230] Andrew M Waterhouse, James B Procter, David MA Martin, Michèle Clamp, and Geoffrey J Barton. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191, 2009.

[231] Joost Wesseling, Henri de Ruiter, Christa Blokhuis, Derko Drukker, Ernie Weijers, Hester Volten, Jan Vonk, Lou Gast, Marita Voogt, Peter Zandveld, et al. Development and implementation of a platform for public information on air quality, sensor measurements, and citizen science. *Atmosphere*, 10(8):445, 2019.

[232] Constantinos Kurt Wibmer, Frances Ayres, Tandile Hermanus, Mashudu Madzivhandila, Prudence Kgagudi, Brent Oosthuysen, Bronwen E Lambson, Tulio De Oliveira, Marion Vermeulen, Karin Van der Berg, et al. Sars-cov-2 501y. v2 escapes neutralization by south african covid-19 donor plasma. *Nature medicine*, 27(4):622–625, 2021.

[233] Roman Wölfel, Victor M Corman, Wolfgang Guggemos, Michael Seilmaier, Sabine Zange, Marcel A Müller, Daniela Niemeyer, Terry C Jones, Patrick Vollmar, Camilla Rothe, et al. Virological assessment of hospitalized patients with covid-2019. *Nature*, 581(7809):465–469, 2020.

[234] Anne L Wyllie, John Fournier, Arnau Casanovas-Massana, Melissa Campbell, Maria Tokuyama, Pavithra Vijayakumar, Joshua L Warren, Bertie Geng, M Catherine Muenker, Adam J Moore, et al. Saliva or nasopharyngeal swab specimens for detection of sars-cov-2. *New England Journal of Medicine*, 383(13):1283–1286, 2020.

[235] Wei Xu, Tian Jin, Yifan Dai, and Chung Chiun Liu. Surpassing the detection limit and accuracy of the electrochemical dna sensor through the application of crispr cas systems. *Biosensors and Bioelectronics*, 155:112100, 2020.

[236] Paul Yager, Gonzalo J Domingo, and John Gerdes. Point-of-care diagnostics for global health. *Annu. Rev. Biomed. Eng.*, 10:107–144, 2008.

[237] Chao Yan, Jinghua Cui, Lei Huang, Bing Du, Lu Chen, Guanhua Xue, Shaoli Li, Weiwei Zhang, Linqing Zhao, Yu Sun, et al. Rapid and visual detection of 2019 novel coronavirus (sars-cov-2) by a reverse transcription loop-mediated isothermal amplification assay. *Clinical Microbiology and Infection*, 26(6):773–779, 2020.

[238] Hsin-Chou Yang, Chun-houh Chen, Jen-Hung Wang, Hsiao-Chi Liao, Chih-Ting Yang, Chia-Wei Chen, Yin-Chun Lin, Chiun-How Kao, Mei-Yeh Jade Lu, and James C Liao. Analysis of genomic distributions of sars-cov-2 reveals a dominant strain type with strong allelic associations. *Proceedings of the National Academy of Sciences*, 117(48):30679–30686, 2020.

[239] Leonid Yurkovetskiy, Xue Wang, Kristen E Pascal, Christopher Tomkins-Tinch, Thomas P Nyalile, Yetao Wang, Alina Baum, William E Diehl, Ann Dauphin, Claudia Carbone, et al. Structural and functional analysis of the d614g sars-cov-2 spike protein variant. *Cell*, 183(3):739–751, 2020.

[240] Marjon Zamani, James M Robson, Andy Fan, Michael S Bono Jr, Ariel L Furst, and Catherine M Klapperich. Electrochemical strategy for low-cost viral detection. *ACS Central Science*, 7(6):963–972, 2021.

[241] Thomas G Zimmerman, Vito P Pastore, Sujoy K Biswas, and Simone Bianco. Embedded system to detect, track and classify plankton using a lensless video microscope. *arXiv preprint arXiv:2005.13064*, 2020.

[242] Tomer Zohar, Carolin Loos, Stephanie Fischinger, Caroline Atyeo, Chuangqi Wang, Matthew D Slein, John Burke, Jingyou Yu, Jared Feldman, Blake Marie Hauser, et al. Compromised humoral functional evolution tracks with sars-cov-2 mortality. *Cell*, 183(6):1508–1519, 2020.

[243] Lirong Zou, Feng Ruan, Mingxing Huang, Lijun Liang, Huitao Huang, Zhongsi Hong, Jianxiang Yu, Min Kang, Yingchao Song, Jinyu Xia, et al. Sars-cov-2 viral load in upper respiratory specimens of infected patients. *New England journal of medicine*, 382(12):1177–1179, 2020.

[244] Yiping Zou, Michael Glenn Mason, Yuling Wang, Eugene Wee, Conny Turni, Patrick J Blackall, Matt Trau, and Jose Ramon Botella. Nucleic acid purification from plants, animals and microbes in under 30 seconds. *PLoS Biol.*, 15(11):e2003916, November 2017.

[245] Uroš Zupančič, Pawan Jolly, Pedro Estrela, Despina Moschou, and Donald E Ingber. Graphene enabled low-noise surface chemistry for multiplexed sepsis biomarker detection in whole blood. *Advanced Functional Materials*, 31(16):2010638, 2021.