AFRL-AFOSR-VA-TR-2023-0419

Design of robust and accurate biosensing systems in living cells

Del Vecchio, Domitilla
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA, 02139
USA

08/08/2023
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
20230808	Final	START DATE 20200501	END DATE 20230430

4. TITLE AND SUBTITLE
Design of robust and accurate biosensing systems in living cells

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
	FA9550-20-1-0044	61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Domitilla Del Vecchio

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA 02139
USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	AFRL/AFOSR RTA2	AFRL-AFOSR-VA-TR-2023-0419

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Engineering biology promises revolutionary changes to the way we approach problems in areas ranging from energy and environment to health and medicine. In particular, engineering cells to concurrently sense multiple molecular species and compute a response based on these is going to be critical in a number of applications, including biosignature classification. In this project, we focus on the design of robust and accurate multi-input biosensors that compute the ratio between the levels of different molecular species. Despite tremendous progress in sensor design, capabilities for tracking the ratios of multiple biomarkers in a simple and deployable format have not been realized. Yet, ratiometric biomarker signatures carry key information about stress, fatigue, and cognitive overload in challenging environments. Additionally, although today we can, in principle, build complex genetic circuits comprising multiple genes, loads that genes apply to the cellular host couple independently regulated genes. This complicates design and makes the behavior of any genetic device, and of biosensors in particular, fragile and inaccurate. We thus propose to develop sensor design strategies that compensate for such effects on the output of interest.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	c. THIS PAGE		
U	U	U	UU	16

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
FREDERICK LEVE	696-9730

Standard Form 298 (Rev.5/2020)
Prescribed by ANSI Std. Z39.18

Design of robust and accurate biosensing systems in living cells

PI: Domitilla Del Vecchio, Mechanical Engineering, MIT
Collaborators: Jorge Chavez Benavides, Svetlana Harbaugh, and Kathryn Beabout, AFRL
AFOSR Award Number FA9550-20-1-0044

To the Attention of Dr. Frederick Leve

Abstract. Engineering biology promises revolutionary changes to the way we approach problems in areas ranging from energy and environment to health and medicine. In particular, engineering cells to concurrently sense multiple molecular species and compute a response based on these is going to be critical in a number of applications, including biosignature classification. In this project, we focus on the design of robust and accurate multi-input biosensors that compute the ratio between the levels of different molecular species. Despite tremendous progress in sensor design, capabilities for tracking the ratios of multiple biomarkers in a simple and deployable format have not been realized. Yet, ratiometric biomarker signatures carry key information about stress, fatigue, and cognitive overload in challenging environments. Additionally, although today we can, in principle, build complex genetic circuits comprising multiple genes, loads that genes apply to the cellular host couple independently regulated genes. This complicates design and makes the behavior of any genetic device, and of biosensors in particular, fragile and inaccurate. We thus propose to develop sensor design strategies that compensate for such effects on the output of interest. This will result in more accurate and predictable performance of sophisticated engineered biosensors and genetic circuits more generally. Finally, we propose a process for on-site detection of airborne virus, focusing on SARS-CoV2 as a test-bed, by leveraging the known RNA signatures of the target pathogen.

Part 1. Ratiometric sensing in living cells

Motivation of the Ratiometric sensing in living cells

The future operating environment of military missions is predicted to be agile, requiring deployment over long time, often in austere environments. Under these circumstances, easily deployable tools for monitoring warfighter biomarkers related to stress, fatigue and cognitive overload would allow for assessment of warfighter readiness and improvement of their performance in challenging environments. Varied stressors affect multiple interconnected cellular signaling pathways, resulting in complex systems confounded by wide variations between individuals. As a result, correlating absolute levels of biomarkers with behavior is often elusive; however, it has been shown that ratiometric signatures can be more informative [1, 2, 3, 4]. For example, the ratio of norepinephrine and cortisol was shown to discriminate PTSD patients from other psychiatric diagnosis subgroups [1] and the ratio of cortisol and DHEA-s could potentially identify stress resilient individuals for specific military needs [2, 3]. Therefore, the ability to monitor the ratio between biomarker levels is critical to assessing warfighter performance and to guiding, possibly automated, intervention strategies. In fact, these sensing capabilities could be eventually used to control different pathways to compensate for biomarker imbalances through the development of "smart" probiotics.

Despite tremendous progress in sensor design, capabilities for tracking the ratios of multiple biomarkers have not been realized. Recently, a "Marionette" strain of *E. coli* was engineered to harbor 12 distinct sensors with minimal crosstalk and high dynamic ranges [5]. While a major advancement in the detection of multiple inputs and regulation of gene expression, this strain has no capacity to detect the relationship between analytes. Furthermore, this strain was not optimized to reduce the coupling between multiple concurrently regulated genes, which arises due to intracellular resource competition [6, 7, 8]. Natural signaling pathways, such as BMP and Notch, have been shown to respond to multi-ligand inputs based on competitive receptor-ligand interactions and dynamic ligand signaling, respectively [9, 10]. However, these sensing mechanisms are highly specialized and cannot readily be coupled to human performance biomarker sensing elements.

In this project, we focus on the design of accurate and robust (to resource competition) multi-input biosensors that compute the ratio between the levels of different molecular species in a simple and deployable format.

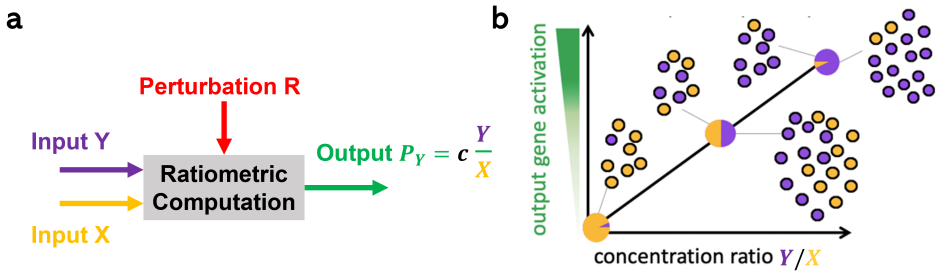


Figure 1: **Robust ratiometric response. (a)** A ratiometric response takes two signaling molecules as inputs (X and Y) and provides a protein P_Y as an output. The concentration of the output protein is proportional to the ratio between the concentrations of X and Y with proportionality constant c. R is a perturbation acting on the system that cannot be controlled and is unknown. If the output is independent of the perturbation R, the ratiometric response is robust to R. **(b)** Pictorial representation of the ratiometric input/output response. The output responds linearly to changes in the ratio Y/X.

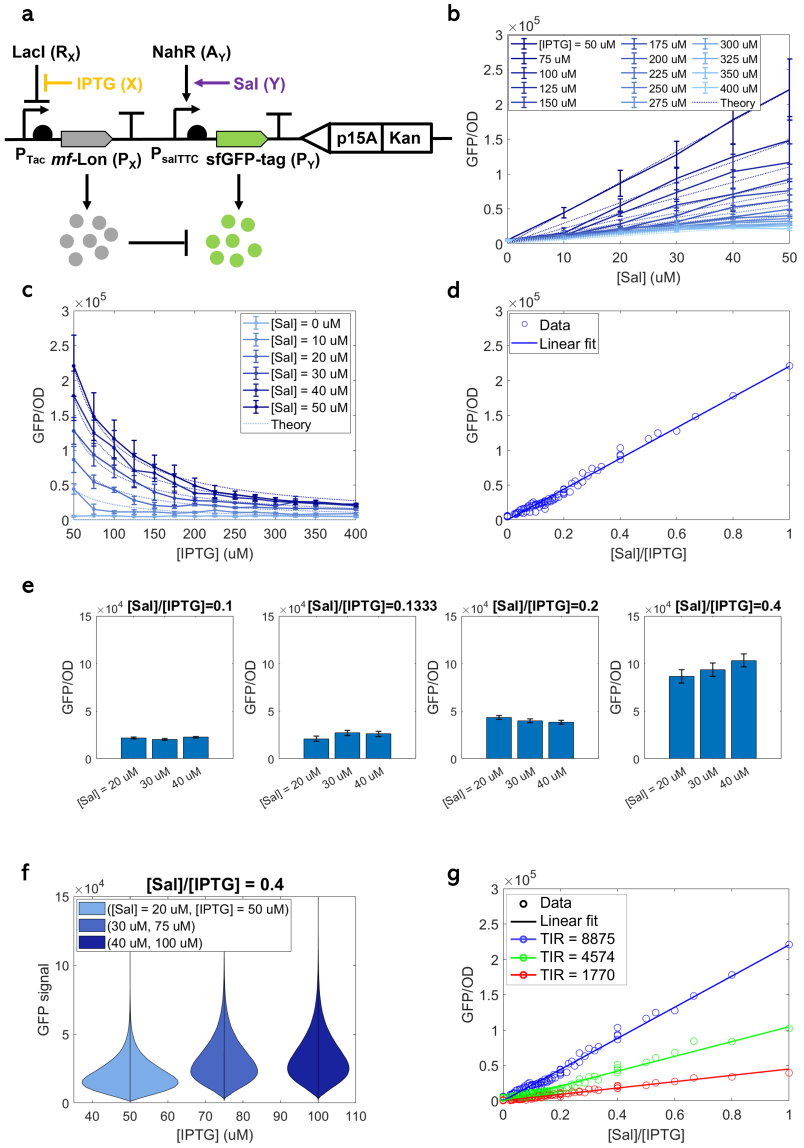


Figure 2: **Performance and tunability of the incoherent merger network genetic implementation.** **(a)** Genetic diagram of the incoherent merger network. This genetic circuit was created in three variants, depending on the choice of the RBS for sfGFP. Specifically, we chose TIRs as 1770, 4575, 8875. **(b)** Dose-response curves showing how GFP depends on Sal along with the reduced model. **(c)** Dose-response curves showing how GFP depends on IPTG along with the reduced model. **(d)** Response to the signal ratio Y/X. GFP/OD value of each circle is an average of 3 biologically independent replicates. **(e)** Response to the signal ratio Y/X for selected [Sal] and [IPTG] combinations in **(d)** that have the same ratio. **(f)** Distributions of output levels per cell for selected [Sal] and [IPTG] combinations that have the same ration (0.4). All data were measured by flow cytometry when OD value was close to 0.054. **(b)-(f)** were obtained with sfGFP's TIR = 8875. **(g)** Tunability of the ratiometric sensor. Blue, green and red colored plots represent TIR = 8875, 4575, 1770, respectively. Data in the line plots in panel (b)-(c) and data in the scatter plot in panel (d) and (g) represent mean values (± SD) of n = 3 biologically independent experiments.

Main results from the past year

During the past year, we focused on experimentally validating the ratiometric sensing ability and especially its quantitative tunability through the ribosome binding site strength as predicted from the model. As

a short summary of the model prediction, we had created an ODE model to describe the circuit's dynamics. Specifically, referring to Figure 2(a), the steady state value of the output (GFP) protein P_Y is given by

$$P_Y = \frac{\gamma k_Y K}{(k+\gamma)k_X} \cdot \frac{Y}{X} = c \cdot \left(\frac{Y}{X}\right), \quad c := \frac{\gamma k_Y K}{(k+\gamma)k_X}, \tag{1}$$

where K is the Michaelis-Menten constant of the protease reaction represented by the gray mf-Lon protein downregulating the GFP protein. Therefore, the steady state value of the output protein P_Y is proportional to the ratio Y/X of the inputs, and this ratio is independent of the resource level R (transcriptional or translational). In summary, the key design requirements to achieve robust ratiometric sensing are

(A0) The production rates of P_X (protease) and P_Y (output protein) are linear with the levels of the signaling molecules X and Y, respectively, and with the resource level R;

(A1) The protease enzymatic reaction is in the first order regime, that is, $P_Y \ll K$;

(A2) Dilution rate is negligible with respect to degradation rate, that is, $\gamma \ll P_{XT}(k+\gamma)/K$.

Regarding (A0), the linearity of the production rates of P_X and P_Y with the levels of X and Y, respectively, is needed for ratiometric computation, while the linearity of the production rates with the level of cellular resource R is needed to make the level of P_Y independent of R. Assumptions (A1) and (A2) can be both satisfied by taking a sufficiently strong protease (large catalytic constant k). Finally, we can tune the sensitivity c by varying the production rate constant k_Y of protein P_Y, which can be easily accomplished by varying the ribosome binding site (RBS) strength for P_Y.

Figure 2 shows that the output is a linear function of the input ratio and that it stays the same when the inputs change but keep the same ratio, thereby validating the key property of ratiometric sensing. Panel (g) of Figure 2 shows that the slope of the input/output line is quantitatively tunable by changing the RBS strength (measured by the TIR) of the appropriate factor, i.e., a two-fold increase in the TIR leads to a two-fold increase of the slope, in accordance with the mathematical model predictions (equation (1)).

As a second test, we demonstrated that indeed the input/output linear relationship does not depend on the cellular resources implicated in gene expression. To this end, we considered the ratiometric sensor functioning in an intra-cellular environment with other (competitor) genes that sequester resources away from the gene expression process (Figure 7). Indeed, in a practical operating environment, as sensor will work concurrently with other sensors and/or with other genetic circuit components, such as actuators to produce required drugs. In this situation, the cellular resources available to the sensor will decrease once other sensors or actuators become activated. It is therefore critical that the sensor input/output response remains unchanged despite the variability of cellular resources. To this end, we examined the input/output dose response curve with and without the protease tag when an RFP gene (competitor) was becoming activated. As expected from theory, the ratio computation was unaffected only when the protease tag was present on the RFP gene. When it was not, not only the ratiometric function was compromised but also the activation of RFP led the output of the sensor to decrease (Figure 7(b),(c)).

Summary. We have shown how ratiometric sensor design can incorporate an incoherent feedforward loop that assures robustness of the output to global changes in the cellular context that affect gene expression. This is the first ratiometric design engineered so far to the best of our knowledge. The input molecules can be adjusted based on the application by just swapping the promoters that control the output protein and the protease with promoters responsive to the molecules of interest. The same design would hold unchanged for translational sensors, as opposed to transcriptional sensors. Apart from the specific biomarker application, ratiometric sensing is generally required in applications where it is important the relative abundance of a molecule as opposed to the absolute abundance of it, such as in environmental sensing of pollutants in water or soil. Ratiometric sensing is also required when performing population control of multi-strain consortia, in

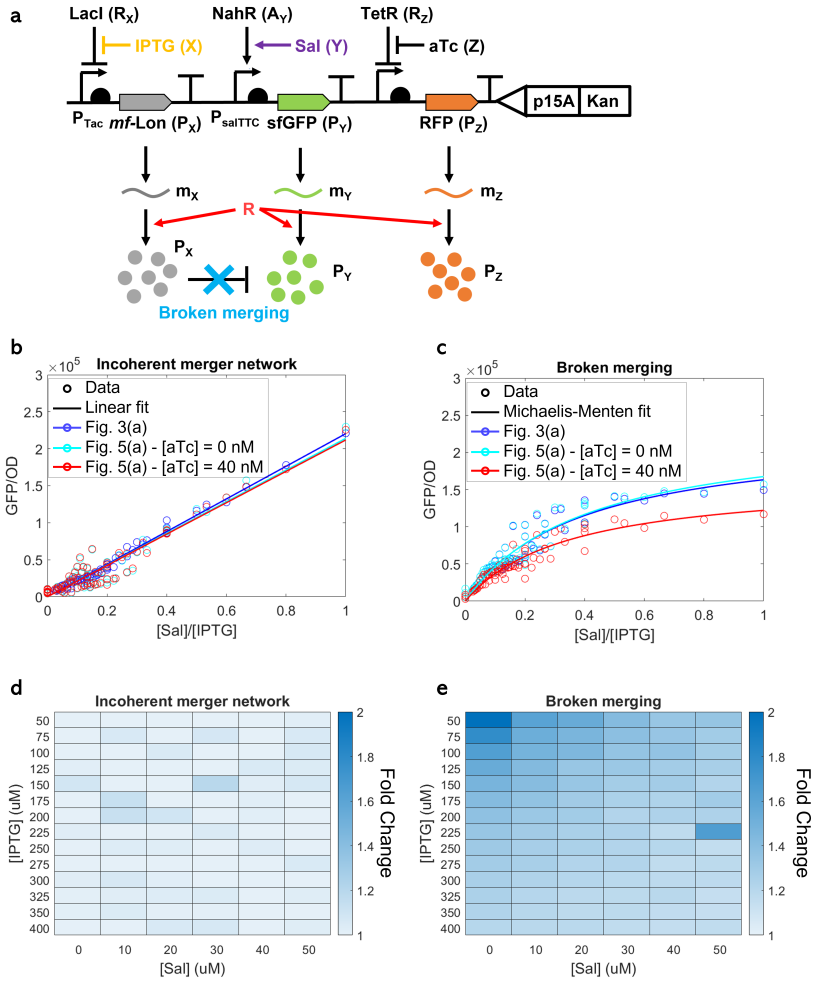


Figure 3: **The incoherent merger network is robust to perturbations in translational resources.** **(a)** Genetic diagram of the incoherent merger network with a resource competitor. Broken merging indicates that the GFP is devoid of the degradation tag and hence is not degraded by the protease. In the resource competitor, RFP expression is activated by addition of the aTc inducer. **(b)** Response of the incoherent merger network to the signal ratio Y/X. Blue colored plot shows the network without the resource competitor and cyan and red colored plots show the network with the resource competitor, with [aTc] = 0 nM and [aTc] = 40 nM, respectively. All solid lines represent linear fit. **(c)** Response of the broken merging to the signal ratio Y/X. Blue colored plot shows the network without the resource competitor and cyan and red colored plots show the network with the resource competitor, with [aTc] = 0 nM and [aTc] = 40 nM, respectively. All solid lines are obtained from the reduced model. **(b)-(c)** GFP/OD value of each circle is an average of 3 biologically independent replicates. **(d)** Heat map showing the fold change of the steady state value of the output protein P_Y of the incoherent merger network when aTc is added (40 nM) with varying IPTG and Sal concentrations. **(e)** Heat map showing the fold change of the steady state value of the output protein P_Y of the broken merging when aTc is added (40 nM) with varying IPTG and Sal concentrations. **(b)-(e)** were obtained with sfGFP's TIR = 8875. Data in the scatter plot in panel (b)-(c) represent mean values (± SD) of n = 3 biologically independent experiments.

which specific ratios among population sizes should be maintained. We therefore believe that this ratiometric sensing architecture will be broadly useful in a variety of applications in synthetic biology.

Products

- Ukjin Kwon, Hsin-Ho Huang, Jorge L Chávez, Kathryn Beabout, Svetlana Harbaugh, and Domitilla Del Vecchio. Incoherent merger network for robust ratiometric gene expression response. *Nucleic Acids Research*, 51(6): 2963–2973 , 2023.
- Ukjin Kwon, Hsin-Ho Huang and Domitilla Del Vecchio. Design of a long-term memory genetic toggle switch inspired by chromatin modification circuits. *Proc. of IEEE Conference on Decision and Control*, December 2023.
- Ukjin Kwon. Design of a robust ratiometric sensor and a long-term memory genetic toggle switch. *MIT PhD Thesis*, February 2023.

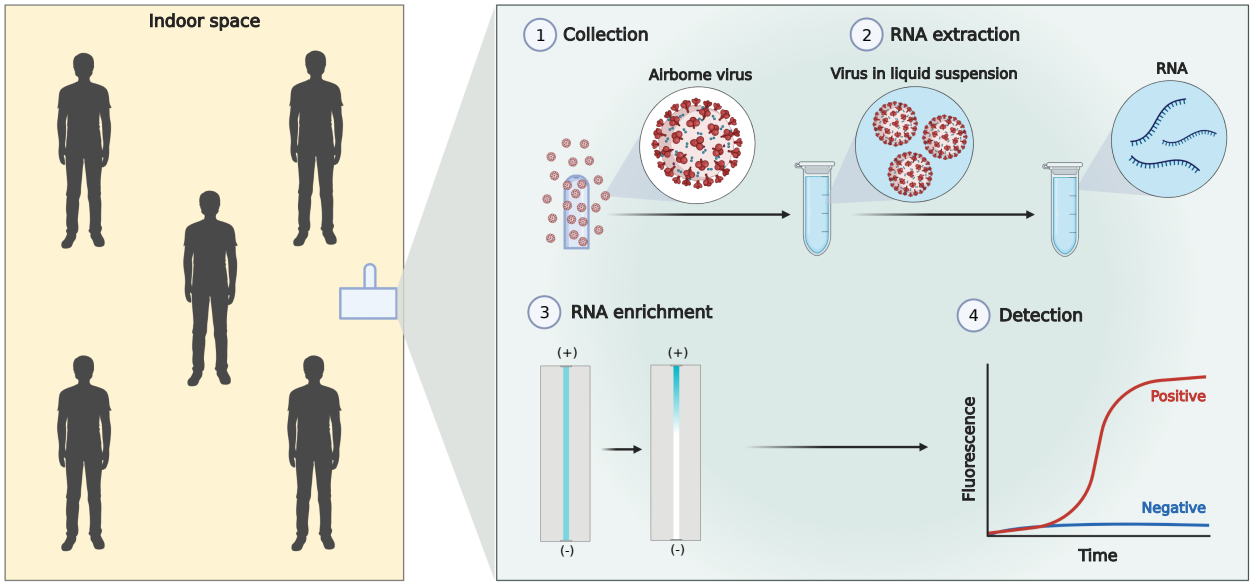


Figure 4: **Proposed concept for on-site detection of airborne pathogens.** A small platform collects the air through electrostatic precipitation (ESP), extracts and concentrates the genetic material, and then performs chemical detection through RT-qPCR. The platform outputs information within one hour of operation.

Part 2. Rapid detection of airborne viruses

Motivation for airborne pathogen detection

The COVID-19 pandemic has demonstrated how vulnerable we are to the spread of airborne infectious diseases. It is also not an isolated case, but rather it follows a series of other epidemics, although not as devastating, including SARS (2003), Swine Flu (2009), MERS (2012), Ebola (2014), and the 2018 United States adenovirus outbreak [11, 12, 13, 14, 15]. With a major epidemic every 2-6 years, we are at continuous risk for the spread of novel infectious diseases. At the same time, even common viruses, such as influenza, respiratory syncytial virus (RSV), and now also SARS-CoV-2, have major repercussions every year among vulnerable populations, leaving behind deaths and healthcare expenses [16, 17, 18].

Early detection is required for timely action and acting quickly has been proven one of the most effective means to contain an epidemic [19, 20, 21]. Currently, despite the numerous COVID-19 surveillance programs, there is an average delay of nearly 6 days between the onset of symptoms and diagnosis [22, 23]. Conventional detection methods depend on symptom-based viral testing, with an average viral shedding duration of 16 days observed among presymptomatic individuals [24, 25]. Environmental monitoring approaches that sample the air of indoor spaces and rapidly analyze it for the presence of pathogens is a more practical approach that could reduce costs, improve detection speed, and minimize disruption to schedules. Since timely actions can only be taken if detection is sufficiently fast and sensitive, the speed and sensitivity of detection are two critical features of an on-site airborne virus detection system.

Currently, there are no established processes for rapid detection of pathogens in indoor air. Common approaches include sampling the air with a pump and then processing the sample in a laboratory, which takes days (BioWatch program [26] and similar [27]). Technologies such as the Thermo Scientific's AeroSense Sampler and Bertin Technologies' Coriolis μ, with their separate PCR processes, can be used to detect common respiratory pathogens, including SARS-CoV-2, Flu A/B and/or RSV, and microbes on-site [28, 29]. However, the air samplers are expensive, the times of collection and analysis required to hit useful sensitivity

remain long, and the whole process requires technical expertise to run and interpret. These approaches are thus largely incompatible with rapid detection and with adoption in lower-resource locations where surveillance and containment are needed.

On-site bio-sensing technologies that have appeared could reportedly detect the presence of SARS-CoV-2, but only at concentrations as high as $31.8 \cdot 10^3$ genome copy/m^3 (gc/m^3) or 31.8 gc/L [28]. This value is among the highest concentrations of influenza virus during flu season and SARS-CoV-2 in COVID-19 hospital patient rooms, reported to be as low as $5 \cdot 10^3$ gc/m^3 or 5 gc/L [27, 30]. Although low, these concentrations can cause illness with high probability if inhaled for more than 1 hour [31]. Therefore, these methods are not sufficiently sensitive to inform actions that could make a difference. Insufficient sensitivity also plagues the Thermo Scientific's AeroSense Sampler and Bertin Technologies' Coriolis μ, with their separate PCR processes, which can detect common respiratory pathogens, including SARS-CoV-2, Flu A/B and/or RSV, and microbes on-site [28, 32, 29]. In summary, the ultra-low concentrations of virus in the air (as low as 1.84 gc/L in some cases [33, 34]) when compared to those in nasal swabs of sick patients ($10^4 - 10^7$gc/mL [35]) pose a unique challenge to the airborne virus detection problem.

Our main innovation is to include an RNA enrichment step that can concentrate the RNA extracted from the virus in a small volume so we can improve the sensitivity of the whole detection process (Figure 4). The enrichment principle uses electrophoresis in a small channel to concentrate the genetic material against a semipermeable membrane. Collection is performed through electrostatic precipitation (ESP), which is preferable to a pump since it is inexpensive, it is not noisy, and does not perturb the airflow of the room, which is important if using sensing to predict the distribution of viral particles in a room of interest. The added advantage of the ESP is that particles are collected on the surface of a small metal probe and, as such, we can extract the RNA in a small volume of water by covering the probe with a tiny layer of water and then by heating it up to perform heat-based RNA extraction. After hear extraction, we perform RNA enrichment, which concentrates the RNA in an even smaller volume, which is directly used for detection through RT-qPCR.

Main results from the past year

In the past year, we have has a number of significant breakthroughs. First of all, we optimized the efficiency of the air sampler that collects particles from the air through electrostatic precipitation (ESP). Second, we characterized the efficiency of the heat extraction method to release the RNA from a virus collected on the metal probe. Third, we optimized the enrichment-fold and speed of the enrichment device, and finally we performed an integration test to validate the whole limit of detection. For the collection process, we used a containment tent where we nebulized fluorescent beads in known amounts and measured the number of particles collected on the metal probe. For all the downstream steps (heat extraction, RNA enrichment, and chemical detection), we used inactivated SARS-CoV-2 particles and SARS-CoV-2 RNA. Although we used SARS-CoV-2 as a test bed, the process that we propose is applicable to any RNA or DNA virus whose genetic sequence is known.

We evaluate the performance of the whole process by its sensitivity C_{min}, that is, the lowest concentration of viral particles in the air (in genome copies per m^3 of air) that can be detected with a detection rate of at least 95%. To this end, we define the efficiency of the whole process η_{tot} as the ratio between the number of genome copies resulting in the 20 µL RT-qPCR reaction volume and the number of corresponding viral genome copies per m^3 of air. With η_{tot} defined this way, and with LoD the limit of detection of RT-qPCR (expressed in cp/µL), the sensitivity is given by

$$C_{min} = \frac{\text{LoD} \cdot V_{\text{PCR}}}{\eta_{tot}}, \tag{2}$$

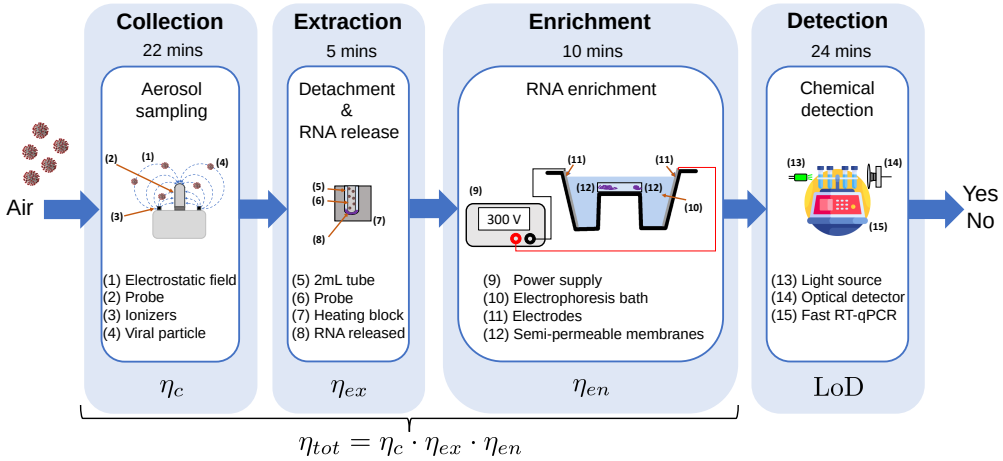


Figure 5: **Components of the airborne virus detection process.** Air particles are collected on the metal probe of an aerosol sampler, which uses electrostatic precipitation (ESP) and follows the design of [36]. After sampling for 22 minutes, the metal probe with collected particles is placed in a 2 mL tube with 220 µL of nuclease-free water and goes through heating at 98 °C for 5 minutes. The water with the sample is then entirely transferred to an enrichment device, which performs electrophoresis against a semipermeable membrane for 10 minutes. Upon completion, the 7.4 µL of sample next to the membrane is drawn and taken to chemical detection. This employs RT-qPCR with primers specific to the viral target and takes about 24 minutes when using a fast RT-qPCR machine. Each step performance is characterized in terms of its efficiency η. The variables η_c, η_{ex}, and η_{en} are the efficiencies of the collection, extraction, and enrichment steps, respectively, as defined in the text, LoD is the limit of detection of RT-qPCR, and η_{tot} represents the efficiency of the overall process.

where $V_{\text{PCR}} = 20\,\mu\text{L}$ is the final RT-qPCR reaction volume. We first estimate C_{min} by experimentally characterizing the efficiency of each of the process' steps in isolation, and by then calculating η_{tot} as the product of the efficiencies of each step (Figure 5). Specifically, we define the collection efficiency η_c as the ratio between the number of particles collected on the metal probe and the number of particles per m^3 of air. The extraction efficiency η_{ex} is defined as the ratio between the number of RNA particles extracted in 220 µL of water and the number of viral particles on the metal probe. The enrichment efficiency η_{en} is the ratio between the number of RNA molecules in the enriched fraction of the volume extracted (7.4 µL) and the number of RNA molecules in the enrichment device volume (250 µL).

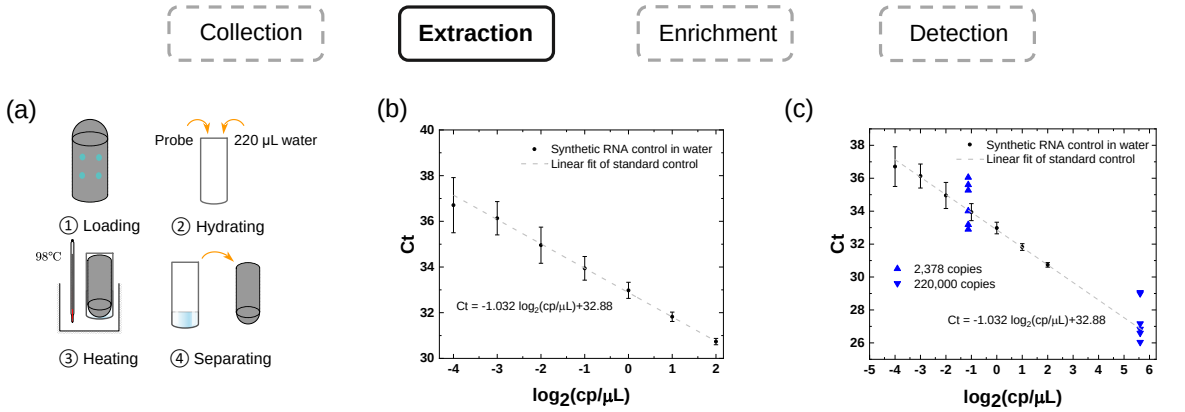
All the efficiencies evaluated are characterized by a mean value, $\bar{\eta}_c$, $\bar{\eta}_{ex}$, and $\bar{\eta}_{en}$, and a standard deviation σ_c, σ_{ex}, and σ_{en}, respectively. Specifically, given the evaluations of the efficiencies ($\eta_c = 0.10 \pm 0.001$, from the previous year), Figure 6 ($\eta_{ex} = 0.115 \pm 0.09$) and Figure 7 ($\eta_{en} = 0.14 \pm 0.02$), the mean value $\bar{\eta}_{tot}$ is given by

$$\bar{\eta}_{tot} = \bar{\eta}_c \cdot \bar{\eta}_{ex} \cdot \bar{\eta}_{en} = 0.00161.$$

The standard variation σ_{tot} of the whole process efficiency can then be evaluated by exploiting the formula for the uncertainty propagation of a product, that is,

$$\sigma_{tot} = \bar{\eta}_{tot} \cdot \sqrt{\left(\frac{\sigma_c}{\bar{\eta}_c}\right)^2 + \left(\frac{\sigma_{ex}}{\bar{\eta}_{ex}}\right)^2 + \left(\frac{\sigma_{en}}{\bar{\eta}_{en}}\right)^2} \approx 0.00127. \tag{3}$$

From this, the mean concentration in air that we can detect is $\bar{C}_{min} := \text{LoD} \cdot V_{\text{PCR}} / \bar{\eta}_{tot} = 0.25 \cdot 20 / 0.00161 \approx 3106$ particles/m^3. From this calculation, it is possible to conclude that the minimal concentrations $\bar{C}_{min}$ that our device is able to detect are between 649 and 5562 particles/m^3. These fall in the lower part of the concentration range of influenza virus detected in enclosed environments during flu season [37] or of SARS-CoV-2 virus detected in hospital rooms [27].



(d) Heat-extraction efficiency definition

$$\eta_{ex} = \frac{\text{copies detached}}{\text{copies loaded}} = \frac{\frac{20}{7.4} \cdot 2^{\frac{(Ct-32.88)}{-1.032}}}{\text{copies loaded}/220\ \mu L}$$

Heat-extraction efficiency calculation

	1	2	3	4	5	6	Statistics
2,378 copies	0.117	0.040	0.205	0.050	0.247	0.030	0.115±0.092
220,000 copies	0.176	0.152	0.258	0.036	0.033	0.122	0.129±0.086

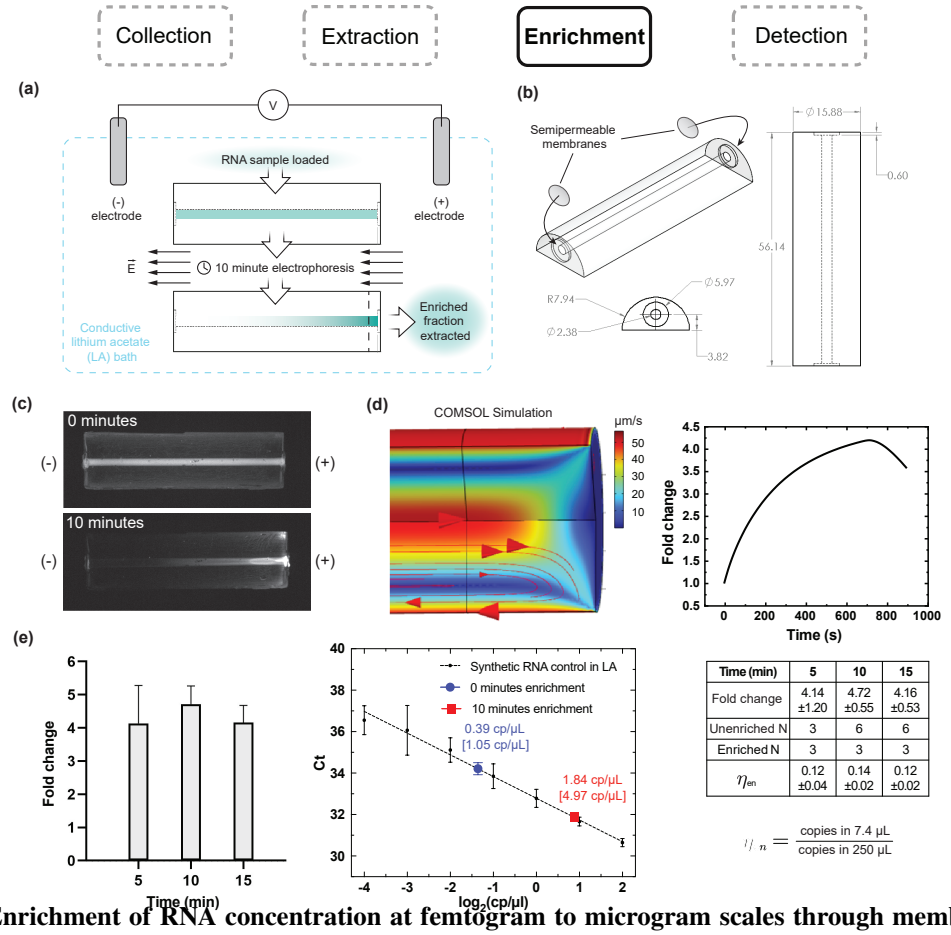
Figure 6: **Heat extraction characterization.** (a) A fixed number of genome copies of ATCC virus in a 10 µL volume is loaded on the metal probe in 1 µL droplets, dried in air for 5 minutes, and then collected in 2 mL microcentrifuge tubes. The tube is then heated at 98 °C for 5 minutes. The metal probe is removed after cool down. (b) RT-qPCR standard curves using synthetic SARS-CoV-2 RNA in water (N=24) (see "RT-qPCRs" in Materials and Methods for details). (c) Ct values after heat extraction and RT-qPCR corresponding to the two different numbers of genome copies plotted as blue triangles on the standard curve. For each copy number, N=6 biological replicates were obtained. (d) Definition and calculation of the heat-extraction efficiencies with two different genome copies.

In what follows, we describe the main advances for each step we made in the last year and then describe the integration test that we have done, which recapitulates the sensitivity calculated by the experimentally identified step-wise efficiencies.

Extraction. Figure 6 summarizes our characterization of the efficiency of the RNA extraction step. The metal probe is taken in a tube with a small volume of water to cover its entire surface, such that with heating, the RNA can be extracted from the particles attached to the metal probe. The efficiency that we estimated for this process is about 11 %.

Enrichment. Figure 7 shows the enrichment process, the devices, and the data obtained. We can enrich the concentration in a small initial volume of 250 µL, used for the heat extraction, by about 5 fold within only 10 minutes of operation in a micro volume of 7.4 µL, which we directly use in the downstream chemical detection reaction. As a comparison, spin-column based enrichment, vacuum concentrator, and magnetic bead approaches use industrial kits and require high-speed centrifugation [38], costly equipment [39], or hour-long laborious mixing and drying operations [40], which make them impractical and expensive for fielding. Compared to an on-chip isotachophoresis approach [41], our membrane-coupled enrichment process easily allows for concentration and detection of synthetic viral RNA samples with tenfold lower concentration while employing simpler buffers and equipment. Our enrichment process can, more generally, be used to concentrate samples in a number of other applications including nucleic acid fragment analysis and quality control, next-generation sequencing, and general biomolecular analysis. Here, we designed the enrichment device to be simple, inexpensive, and fieldable, but incorporation of parallelized microcapillaries, fluidic manifolds [42], electric field gradient focusing, or isotachophoresis [43, 41] could further improve the degree of enrichment at the expense of additional cost and complexity.

Integration test. To further validate the sensitivity and specificity of our process, we directly loaded the

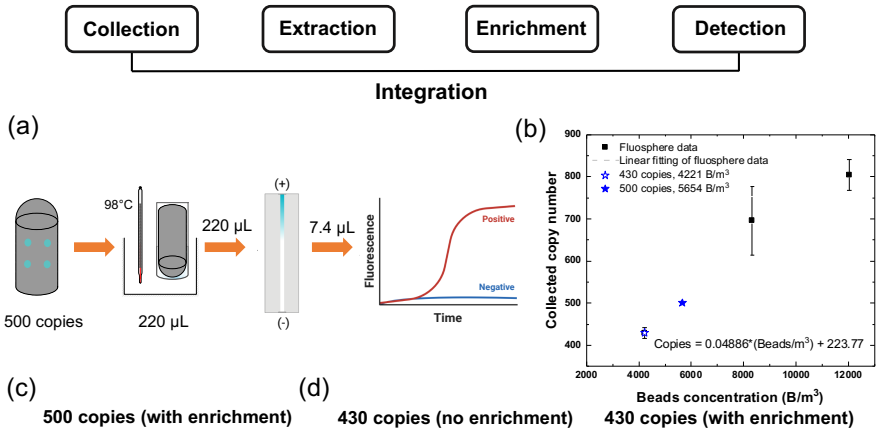


Time (min)	5	10	15
Fold change	4.14 ±1.20	4.72 ±0.55	4.16 ±0.53
Unenriched N	3	6	6
Enriched N	3	3	3
η_{en}	0.12 ±0.04	0.14 ±0.02	0.12 ±0.02

$$\eta_{en} = \frac{\text{copies in 7.4 μL}}{\text{copies in 250 μL}}$$

Figure 7: **Enrichment of RNA concentration at femtogram to microgram scales through membrane-coupled electrophoresis.** (a) Working principle of RNA enrichment. A 250 μL RNA sample buffered in 5 mM lithium acetate (LA) is loaded into a PDMS device consisting of a channel capped at both ends by semipermeable membranes. The device is immersed in an electrophoresis bath containing 5 mM LA and an applied voltage causes electrophoresis of RNA (teal) toward the positive terminal, concentrating the RNA against one of the membranes. An enriched fraction of RNA is then extracted adjacent to the membrane. (b) PDMS channel design for fast enrichment of nucleic acid. (Left) 3D view showing the location of semipermeable membranes in insets. (Center) Face-on view of channel and insets. (Right) Top view, rotated 90 degrees from that shown in (a). Dimensions are indicated in mm. (c) Images of microgram-scale enrichment of mammalian RNA stained with SYBR Green II in devices oriented as in (a), showing concentration profiles before enrichment (0 minutes) and after enrichment (10 minutes) at 2,000 V/m. Both images were obtained using identical exposure settings. (d) Multiphysics simulation of negatively charged particle migration in the enrichment device using COMSOL. (Left) Simulated velocity streamlines. (Right) Simulated time course of concentration fold change adjacent to the membrane (~3% of device volume). For details on the simulations, see "Materials and Methods: Simulation of nucleic acid enrichment in enrichment devices." (e) Femtogram-scale enrichment of ultra-low concentration synthetic SARS-CoV-2 RNA. (Left) Fold change enrichment of synthetic SARS-CoV-2 RNA quantified by RT-qPCR at 5, 10, and 15 minutes. (Center) 10 minutes of enrichment leads to a 4.7-fold enrichment of RNA. Concentrations before and after enrichment are shown on the RT-qPCR standard curve in blue and red respectively, with bracketed and unbracketed values representing the concentrations in the device and RT-qPCR reactions, respectively. (Right) Summary of data collected, including enrichment efficiency η_{en}.

metal probe with heat-inactivated SARS-CoV-2 viral particles and carried out extraction, enrichment, and detection (Figure 8(a)). Specifically, we loaded 500 copies of SARS-CoV-2 viral particles as this number corresponds to the amount of particles collected on the metal probe when the in-air particle concentration is about 5,654 particles/m^3 (Figure 8(b)) and this in-air concentration is at the high-end of the sensitivity $\bar{C}_{min}$



(c)

500 copies (with enrichment)

Repeat #	Ct value	Repeat #	Ct value
1	34.16	12	31.62
2	32.98	13	32.2
3	34.11	14	36.03
4	33.85	15	36.65
5	34.17	16	34.68
6	33.58	17	31.17
7	33.63	18	34.01
8	32.77	19	31.98
9	32.24	20	33.48
10	33.15	21	33.79
11	33.82	22	31.72
Detection rate:		22/22 (100%)	

(d)

430 copies (no enrichment)

Repeat #	Ct value	Repeat #	Ct value
1	ND	9	ND
2	36.74	10	ND
3	35.26	11	ND
4	35.28	12	35.33
5	ND	13	36.56
6	ND	14	ND
7	ND	15	36.62
8	34.78		
Detection rate:		7/15 (46.7%)	

430 copies (with enrichment)

Repeat #	Ct value	Repeat #	Ct value
1	36.58	9	36.23
2	36.45	10	33.53
3	35.67	11	31.04
4	33.78	12	36.74
5	33.78	13	36.85
6	ND	14	ND
7	33.58		
8	ND		
Detection rate:		11/14 (78.6%)	

Figure 8: **Experimental process for the integration test.** (a) **Loading:** the metal probe was loaded with either nuclease-free water (for NTCs) or SARS-CoV-2 ATCC-VP (for biological replicates) using 1 µL droplets with a total volume of 10 µL. **Extraction:** the metal probe was dried in a biosafety cabinet, collected in 2 mL tubes and 220 µL of water was added before heating at 98 °C for 5 minutes. **Enrichment:** the recovered liquid was supplemented with LA buffer and concentrated in the enrichment device for 10 minutes. **Detection:** 7.4 µL of enriched liquid was extracted from the positive end of the enrichment device and analyzed with RT-qPCR to determine the Ct value. (b) Linear regression analysis was performed over the three data points at the lowest concentration in air tested (from last year). The fitted line was used to convert the two copy numbers (500 and 430) back to in-air concentration. (c) Results of integration test starting with 500 copies. (d) Results of integration test when starting copies were reduced to 430 to further explore the LoD.

estimated above. For 500 copies, obtained a 100% detection rate for N=22. We then decreased the number to explore the limit of detection and loaded next 430 copies on the metal probe, corresponding to an in-air concentration of about 4,221 particles/m^3 (Figure 8(b)). For this number, we also followed two parallel routes after heat extraction to evaluate the benefit of incorporating the enrichment process for detection. Quantification of detection rate before and after enrichment at 430 copies shows that enrichment boosted the detection rate from 46.7% to 78.6% for N=15 and 14, respectively. This result also implies that the sensitivity of our overall process is between 4,221 particles/m^3 and 5,654 particles/m^3.

Summary. We characterized in a laboratory setting the performance of a rapid RT-qPCR-based detection process incorporating an electrophoretic step for nucleic acid enrichment. This process was integrated with pathogen collection using a compact, inexpensive, electrostatic air sampler, as a stand-alone system for early detection for airborne pathogens. To extract the viral RNA, we implemented a simple heat-extraction protocol to circumvent the need for commercial extraction kits, overall making the platform field-ready and relatively inexpensive. Our enrichment system uses small channel electrophoresis against a semipermeable membrane, and is able to concentrate RNA by nearly 5-fold in just 10 minutes. By virtue of our fast

enrichment system, we could reach a sensitivity of about 5,000 copies of virus per m^3 of air within 61 minutes of operation when using a fast RT-qPCR machine. This concentration, in turn, falls in the lower range of influenza virus and SARS-CoV-2 concentrations reported in the literature [44, 27]. By including our enrichment system in the detection process, we therefore make airborne virus detection sufficiently sensitive for practical application without the use of industrial kits [39, 45]. This also mitigates hardware reliance and brings the consumable cost down to below $1/test, much lower compared to the $6-$10/test when using commercially available extraction and concentration kits. While the majority of the equipment cost stems from the Fast RT-qPCR machine, this machine is capable of processing 16 samples in parallel within only 24 minutes, and could be implemented as a central processing station in facilities with distributed air samplers, which cost less than $250 each.

In the future, we seek to perform field tests within an hospital settings with positive and negative control groups, and to integrate and automate the whole process within a compact device for non-expert use.

Products

- Huifeng Du, Simone Bruno, Kalon J. Overholt, Sebastian Palacios, Hsin-Ho Huang, Carlos Barajas, Ben Gross, Cindy Lee, Haley K. Evile, Nuno Rufino de Sousa, Antonio Gigliotti Rothfuchs, and Domitilla Del Vecchio. Rapid and sensitive detection of airborne viruses using femtogram-scale RNA enrichment. *medRxiv: DOI: 10.1101/2023.07.03.23292170*, July 2023.

- Patent Application # 63149756. A SYSTEM FOR ON-SITE DETECTION OF VIRUS-LOADED AEROSOLS. Inventors: Domitilla Del Vecchio, Hsin-Ho Huang, Carlos Barajas, Kalon J. Overholt, Simone Bruno, Theodore W. Grunberg, 2021 (Pending).

References

[1] S. Wang, J. Mason, D. Charney, R. Yehuda, S. Riney, and S. Southwick, "Relationships between hormonal profile and novelty seeking in combat-related posttraumatic stress disorder," *Biological Psychiatry*, vol. 41, no. 2, pp. 145–151, 1997.

[2] R. M. Shia, L. K. McIntire, J. A. Hagen, C. D. Goodyear, L. N. Dykstra, and A. R. Myers, "Biomarker and biometric indices of physical exhaustion in the firefighting community," *Procedia Manufacturing*, vol. 3, pp. 5081–5087, 2015.

[3] R. M. Shia, J. A. Hagen, L. K. McIntire, C. D. Goodyear, L. N. Dykstra, and L. Narayanan, "Individual differences in biophysiological toughness: sustaining working memory during physical exhaustion," *Military medicine*, vol. 180, no. 2, pp. 230–236, 2015.

[4] D. Mauss, J. Li, B. Schmidt, P. Angerer, and M. N. Jarczok, "Measuring allostatic load in the workforce—a systematic review," *Industrial health*, 2014.

[5] A. J. Meyer, T. H. Segall-Shapiro, E. Glassey, J. Zhang, and C. A. Voigt, "Escherichia coli "marionette" strains with 12 highly optimized small-molecule sensors," *Nature chemical biology*, vol. 15, no. 2, pp. 196–204, 2019.

[6] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource competition shapes the response of genetic circuits," *ACS synthetic biology*, vol. 6, no. 7, pp. 1263–1272, 2017.

[7] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost lines describe the cellular economy of genetic circuits," *Biophysical journal*, vol. 109, no. 3, pp. 639–646, 2015.

[8] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis, "Quantifying cellular capacity identifies gene expression designs with reduced burden," *Nature methods*, vol. 12, no. 5, pp. 415–418, 2015.

[9] Y. E. Antebi, J. M. Linton, H. Klumpe, B. Bintu, M. Gong, C. Su, R. McCardell, and M. B. Elowitz, "Combinatorial signal perception in the bmp pathway," *Cell*, vol. 170, no. 6, pp. 1184–1196, 2017.

[10] N. Nandagopal, L. A. Santat, L. LeBon, D. Sprinzak, M. E. Bronner, and M. B. Elowitz, "Dynamic ligand discrimination in the notch signaling pathway," *Cell*, vol. 172, no. 4, pp. 869–880, 2018.

[11] Centers for Disease Control and Prevention, "Revised U.S. Surveillance Case Definition for Severe Acute Respiratory Syndrome (SARS) and Update on SARS Cases — United States and Worldwide, December 2003," 2003.

[12] Centers for Disease Control and Prevention, "The 2009 H1N1 Pandemic: Summary Highlights, April 2009-April 2010," 2009.

[13] Centers for Disease Control and Prevention, "Severe Respiratory Illness Associated with Middle East Respiratory Syndrome Coronavirus (MERS-CoV) — Worldwide, 2012–2013," 2013.

[14] Centers for Disease Control and Prevention, "CDC Releases Detailed History of the 2014-2016 Ebola Response in MMWR," 2016.

[15] H. M. Biggs, X. Lu, L. Dettinger, S. Sakthivel, J. T. Watson, and S. W. Boktor, "Adenovirus-associated influenza-like illness among college students, pennsylvania, usa," *Emerging infectious diseases*, vol. 24, no. 11, p. 2117, 2018.

[16] C. L. Hansen, S. S. Chaves, C. Demont, and C. Viboud, "Mortality associated with influenza and respiratory syncytial virus in the us, 1999-2018," *JAMA network open*, vol. 5, no. 2, pp. e220527–e220527, 2022.

[17] S. Yan, D. Weycker, and S. Sokolowski, "Us healthcare costs attributable to type a and type b influenza," *Human vaccines & immunotherapeutics*, vol. 13, no. 9, pp. 2041–2047, 2017.

[18] Y. Choi, A. Hill-Ricciuti, A. R. Branche, W. D. Sieling, L. Saiman, E. E. Walsh, M. Phillips, A. R. Falsey, and L. Finelli, "Cost determinants among adults hospitalized with respiratory syncytial virus in the united states, 2017–2019," *Influenza and Other Respiratory Viruses*, vol. 16, no. 1, pp. 151–158, 2022.

[19] I. M. Blake, P. Chenoweth, H. Okayasu, C. A. Donnelly, R. B. Aylward, and N. C. Grassly, "Faster detection of poliomyelitis outbreaks to support polio eradication," *Emerging infectious diseases*, vol. 22, no. 3, p. 449, 2016.

[20] B. Impouma, M. Roelens, G. S. Williams, A. Flahault, C. T. Codeço, F. Moussana, B. Farham, E. L. Hamblion, F. Mboussou, and O. Keiser, "Measuring timeliness of outbreak response in the world health organization african region, 2017–2019," *Emerging infectious diseases*, vol. 26, no. 11, p. 2555, 2020.

[21] Resolve to Save Lives, "Timeliness metrics: an opportunity for epidemic response tracking and improvement," 2021.

[22] The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) — china, 2020," *China CDC Weekly*, vol. 2, p. 113, 2020.

[23] C. Faes, S. Abrams, D. Van Beckhoven, G. Meyfroidt, E. Vlieghe, N. Hens, and B. C. G. on COVID-19 Hospital Surveillance, "Time between symptom onset, hospitalisation and recovery or death: statistical analysis of belgian covid-19 patients," *International journal of environmental research and public health*, vol. 17, no. 20, p. 7560, 2020.

[24] Centers for Disease Control and Prevention, "Overview of Testing for SARS-CoV-2, the virus that causes COVID-19," 2023.

[25] C. Chen, C. Zhu, D. Yan, H. Liu, D. Li, Y. Zhou, X. Fu, J. Wu, C. Ding, G. Tian, *et al.*, "The epidemiological and radiographical characteristics of asymptomatic infections with the novel coronavirus (covid-19): a systematic review and meta-analysis," *International Journal of Infectious Diseases*, vol. 104, pp. 458–464, 2021.

[26] National Research Council, *BioWatch and public health surveillance: Evaluating systems for the early detection of biological threats: Abbreviated version*. National Academies Press, 2011.

[27] J. A. Lednicky, M. Lauzardo, Z. H. Fan, A. Jutla, T. B. Tilly, M. Gangwar, M. Usmani, S. N. Shankar, K. Mohamed, A. Eiguren-Fernandez, *et al.*, "Viable sars-cov-2 in the air of a hospital room with covid-19 patients," *International Journal of Infectious Diseases*, vol. 100, pp. 476–482, 2020.

[28] P. F. Horve, L. Dietz, D. Northcutt, J. Stenson, and K. Van Den Wymelenberg, "Evaluation of a bioaerosol sampler for indoor environmental surveillance of severe acute respiratory syndrome coronavirus 2," *PLoS One*, vol. 16, no. 11, p. e0257689, 2021.

[29] A. H. Abeykoon, M. Poon, S. M. Firestone, M. A. Stevenson, A. K. Wiethoelter, and G. A. Vincent, "Performance evaluation and validation of air samplers to detect aerosolized coxiella burnetii," *Microbiology Spectrum*, vol. 10, no. 5, pp. e00655–22, 2022.

[30] W. Yang, S. Elankumaran, and L. C. Marr, "Concentrations and size distributions of airborne influenza a viruses measured indoors at a health centre, a day-care centre and on aeroplanes," *Journal of the Royal Society Interface*, vol. 8, no. 61, pp. 1176–1184, 2011.

[31] M. Prentiss, A. Chu, and K. K. Berggren, "Finding the infectious dose for covid-19 by applying an airborne-transmission model to superspreader events," *PLoS One*, vol. 17, no. 6, p. e0265816, 2022.

[32] N. Rufino de Sousa, L. Shen, D. Silcott, C. J. Call, and A. G. Rothfuchs, "Operative and technical modifications to the coriolis® μ air sampler that improve sample recovery and biosafety during microbiological air sampling," *Annals of work exposures and health*, vol. 64, no. 8, pp. 852–865, 2020.

[33] S. L. Miller, W. W. Nazaroff, J. L. Jimenez, A. Boerstra, G. Buonanno, S. J. Dancer, J. Kurnitski, L. C. Marr, L. Morawska, and C. Noakes, "Transmission of sars-cov-2 by inhalation of respiratory aerosol in the skagit valley chorale superspreading event," *Indoor air*, vol. 31, no. 2, pp. 314–323, 2021.

[34] P. Y. Chia, K. K. Coleman, Y. K. Tan, S. W. X. Ong, M. Gum, S. K. Lau, X. F. Lim, A. S. Lim, S. Sutjipto, P. H. Lee, *et al.*, "Detection of air and surface contamination by sars-cov-2 in hospital rooms of infected patients," *Nature communications*, vol. 11, no. 1, p. 2800, 2020.

[35] A. Mawaddah, H. S. Genden, S. G. Lum, and M. B. Marina, "Upper respiratory tract sampling in covid-19," *The Malaysian journal of pathology*, vol. 42, no. 1, pp. 23–35, 2020.

[36] N. Rufino de Sousa, N. Sandström, L. Shen, K. Håkansson, R. Vezozzo, K. Udekwu, J. Croda, and A. Rothfuchs, "A fieldable electrostatic air sampler enabling tuberculosis detection in bioaerosols," *Tuberculosis*, vol. 120, p. 101896, 12 2019.

[37] Y. Zhang, G. Ren, J. Buss, A. J. Barry, G. C. Patton, and N. A. Tanner, "Enhancing colorimetric loop-mediated isothermal amplification speed and sensitivity with guanidine chloride," *BioTechniques*, vol. 69, pp. 178–185, Sept. 2020.

[38] N. Shukla, I. K. Yan, and T. Patel, "Multiplexed detection and quantitation of extracellular vesicle rna expression using nanostring," *Extracellular RNA: Methods and Protocols*, pp. 177–185, 2018.

[39] I. Sánchez, F. Betsou, and W. Mathieson, "Does vacuum centrifugal concentration reduce yield or quality of nucleic acids extracted from ffpe biospecimens?," *Analytical biochemistry*, vol. 566, pp. 16–19, 2019.

[40] S. Chea, S. Lay, M. Oum, G. Tang, C. Hou, M. Vanaerschot, C. Yek, C. Tato, J. Manning, and V. Ahyong, "Integra magbead dna and rna extraction for isolated colonies," *protocols.io*, 2022.

[41] A. Ramachandran, D. Huyke, E. Sharma, M. Sahoo, C. Huang, N. Banaei, B. Pinsky, and J. Santiago, "Electric field-driven microfluidics for rapid crispr-based diagnostics and its application to detection of sars-cov-2," *PNAS*, no. 47, pp. 29518–29525, 2020.

[42] W. Qiao, C. Wang, Z. Ding, and et al, "A two-stage electrophoretic microfluidic device for nucleic acid collection and enrichment," *Microfluidics and Nanofluidics*, no. 77, 2016.

[43] G. Garcia-Schwarz, A. Rogacs, S. Bahga, and J. Santiago, "On-chip isotachophoresis for separation of ions and purification of nucleic acids.," *J Vis Exp.*, p. e3890, 2012.

[44] S. E. W. Yang and L. C. Marr, "Concentrations and size distributions of airborne influenza a viruses measured indoors at a health centre, a day-care centre and on aeroplanes," *J. R. Soc. Interface*, vol. 8, no. 61, pp. 1176–1184, 2011.

[45] G. Buonanno, A. Robotto, E. Brizio, L. Morawska, A. Civra, F. Corino, D. Lembo, G. Ficco, and L. Stabile, "Link between sars-cov-2 emissions and airborne concentrations: Closing the gap in understanding," *Journal of Hazardous Materials*, vol. 428, p. 128279, 2022.

Design of robust and accurate biosensing systems in living cells

PI: Domitilla Del Vecchio, Mechanical Engineering, MIT
Collaborator: Jorge Chavez Benavides, AFRL
AFOSR Award Number FA9550-20-1-0044

To the Attention of Dr. Frederick Leve

Abstract. Engineering biology promises revolutionary changes to the way we approach problems in areas ranging from energy and environment to health and medicine. In particular, engineering cells to concurrently sense multiple molecular species and compute a response based on these is going to be critical in a number of applications, including biosignature classification. In this project, we focus on the design of robust and accurate multi-input biosensors that compute the ratio between the levels of different molecular species. Despite tremendous progress in sensor design, capabilities for tracking the ratios of multiple biomarkers in a simple and deployable format have not been realized. Yet, ratiometric biomarker signatures carry key information about stress, fatigue, and cognitive overload in challenging environments. Additionally, although today we can, in principle, build complex genetic circuits comprising multiple genes, loads that genes apply to the cellular host couple independently regulated genes. This complicates design and makes the behavior of any genetic device, and of biosensors in particular, fragile and inaccurate. We thus propose to develop design strategies to compensate for such a coupling on the output of interest. This should result in more accurate and predictable performance of engineered biosensors and genetic circuits more generally. Finally, we propose to design a point-of-need airborne sensor of SARS-CoV2 by leveraging the known RNA signatures of the virus.

Part 1. Ratiometric sensing in living cells

Motivation of the Ratiometric sensing in living cells

The future operating environment of military missions is predicted to be agile, requiring deployment over long time, often in austere environments. Under these circumstances, easily deployable tools for monitoring Warfighter biomarkers related to stress, fatigue and cognitive overload would allow for assessment of Warfighter readiness and improvement of their performance in challenging environments. Varied stressors affect multiple interconnected cellular signaling pathways, resulting in complex systems confounded by wide variations between individuals. As a result, correlating absolute levels of biomarkers with behavior is often elusive; however, it has been shown that ratiometric signatures can be more informative [1, 2, 3, 4]. For example, the ratio of norepinephrine and cortisol was shown to discriminate PTSD patients from other psychiatric diagnosis subgroups [1] and the ratio of cortisol and DHEA-s could potentially identify stress resilient individuals for specific military needs [2, 3]. Therefore, the ability to monitor the ratio between biomarker levels is critical to assessing Warfighter performance and to guiding, possibly automated, intervention strategies. In fact, these sensing capabilities could be eventually used to control different pathways to compensate for biomarker imbalances through the development of "smart" probiotics.

Despite tremendous progress in sensor design, capabilities for tracking the ratios of multiple biomarkers have not been realized. Recently, a "Marionette" strain of *E. coli* was engineered to harbor 12 distinct sensors with minimal crosstalk and high dynamic ranges [5]. While a major advancement in the detection of multiple inputs and regulation of gene expression, this strain has no capacity to detect the relationship between analytes. Furthermore, this strain was not optimized to reduce the coupling between multiple concurrently regulated genes, which arises due to intracellular resource competition [6, 7, 8]. Natural signaling pathways, such as BMP and Notch, have been shown to respond to multi-ligand inputs based on competitive receptor-ligand interactions and dynamic ligand signaling, respectively [9, 10]. However, these sensing mechanisms are highly specialized and cannot readily be coupled to human performance biomarker sensing elements.

In this project, we focus on the design of accurate and robust (to resource competition) multi-input biosensors that compute the ratio between the levels of different molecular species in a simple and deployable format.

Main Results

In Figure 1a, we propose transcriptional (TX) sensors based ratiometric sensor. Input X and Y can be a negative inducer of a repressor protein or a positive inducer of an activator protein. Repressor (activator) protein is a transcriptional factor that binds to the promoter site and turns off (enhances) the transcription of the gene [11]. Therefore, more X (Y) produces more P_1 (P_2). In this design, P_1 is a protease that targets P_2 for degradation [12], and the steady state output protein level P_2 is, under suitable parameter conditions, proportional to the ration $\frac{Y}{X}$. First, when we do not consider the resource competition, $P_2 \propto \frac{Y}{X}$ can be proved by using a simple mathematical model of gene expression [11] and of protease activity [12]. Specifically, we assume that production of P_1 and P_2 are proportional to input signals X and Y, respectively, which can be modeled as $\emptyset \xrightarrow{k_1 X} P_1$ and $\emptyset \xrightarrow{k_2 Y} P_2$. We model the degradation of P_2 by P_1 as an enzymatic reaction $P_1 + P_2 \underset{b}{\overset{a}{\rightleftharpoons}} C \xrightarrow{k} P_1$, in which C is a complex. We also consider dilution, $P_1, P_2, C \xrightarrow{\gamma} \emptyset$, where γ is the dilution rate constant. By letting $P_{1T} = P_1 + C$ be the total concentration of P_1, we obtain the following

reduced (through time-scale separation) ordinary differential equation (ODE) model:

$$\begin{aligned}\frac{d}{dt}P_{1T} &= k_1 X - \gamma P_{1T},\\ \frac{d}{dt}P_2 &= k_2 Y - (\gamma + k)\frac{P_2 P_{1T}}{K + P_2} - \gamma P_2,\end{aligned} \tag{1}$$

in which $K = \frac{b+k+\gamma}{a}$ is the Michaelis-Menten constant of the protease enzymatic reaction. By setting the time derivatives to zero and solving for P_2 and P_{1T}, we obtain the steady state value of P_2. Specifically, assuming that (A1) $P_2 \ll K$ (protease enzymatic reaction is in the first order regime), (A2) $\gamma \ll \frac{P_{1T}(k+\gamma)}{K}$ (dilution is negligible compared to degradation of P_2), and (A3) production of P_1 and P_2 are proportional to input signals X and Y, respectively (binding of the transcription factor to the promoter site is in the linear regime), then, we obtain that the steady state level of P_2 is given by

$$P_2 = \frac{k_2 K \gamma}{k k_1} \cdot (\frac{Y}{X}) = f \cdot (\frac{Y}{X}), f := \frac{k_2 K \gamma}{k k_1} \tag{2}$$

which gives the ration between X and Y. Figure 1b is obtained from simulating model (1) and plotting the steady state value of P_2 as a function of X and Y plotted using logarithmic scale. If indeed $P_2 = f(X/Y)$, in log scale we obtain $\log P_2 = \log f + \log X - \log Y$, and hence we should expect a plane when plotting the steady state of the simulation, which is the case as seen in Figure 1b. Additionally, Figure 1c shows that the slope estimate f obtained from the approximation (2) holds well under the stated assumptions. Also, Figure 1d shows that the theoretically predicted P_2 value from (2) matches well with Gillespie simulations (Y is fixed, X is changed). Therefore, our system is robust to instrinsic noise. We conclude that this circuit topology can implement ratiometric sensing.

Now we consider the resource competition, especially ribosome competition, and reflect it in our mathematical model. Both P_1 and P_2 require ribosome to be translated from their mRNA. When there are more (less) resource, more (less) P_1 and P_2 are produced, and P_1 degrades P_2. Figure 1d shows this inherent IFFL structure in our ratiometric circuit. Therefore, the steady state output protein level P_2 is, under suitable parameter conditions, proportional to the ration $\frac{Y}{X}$, even with the resource competition. Without loss of generality, let X be a negative inducer of a repressor protein, R_1. Then production of P_1 from X can be modeled as $X + R_1 \underset{b_1}{\overset{a_1}{\rightleftharpoons}} C_1$, $R_1 + D_1 \underset{b_2}{\overset{a_2}{\rightleftharpoons}} C_2$, $D_1 \xrightarrow{k_1} m_1 + D_1$, $m_1 + R \underset{b_3}{\overset{a_3}{\rightleftharpoons}} C_3 \xrightarrow{k_3} m_1 + P_1 + R$, where C_1, C_2 and C_3 are complexes, D_1 and m_1 are DNA and mRNA that produce P_1, respectively, and R is free ribosome. Let Y be a positive inducer of an activator protein, A_1. Then production of P_2 from Y can be modeled as $Y + A_1 \underset{b_4}{\overset{a_4}{\rightleftharpoons}} C_4$, $D_2 + C_4 \underset{b_5}{\overset{a_5}{\rightleftharpoons}} C_5 \xrightarrow{k_5} m_2 + D_2 + C_4$, $m_2 + R \underset{b_6}{\overset{a_6}{\rightleftharpoons}} C_6 \xrightarrow{k_6} m_2 + P_2 + R$, where C_4, C_5 and C_6 are complexes, D_2 and m_2 are DNA and mRNA that produce P_2, respectively. We also model the degradation of P_2 by P_1 as an enzymatic reaction $P_1 + P_2 \underset{d}{\overset{a}{\rightleftharpoons}} C \xrightarrow{k} P_1$, in which C is a complex, and consider dilution as well, $C_i, C, P_i, m_i \xrightarrow{\gamma} \emptyset$, where γ is the dilution rate constant. By letting $R_{1T} = R_1 + C_1$ be the total repression protein concentration, $A_{1T} = A_1 + C_4$ be the total activator protein concentration, $D_{1T} = D_1 + C_2$ and $D_{2T} = D_2 + C_5$ be the total DNA concentrations, $R_T = R + C_3 + C_6$ be the total ribosome concentration, $P_{1T} = P_1 + C$ be the total concentration of P_1, we obtain the following reduced (through time-scale separation) ordinary differential equation (ODE) model:

$$\begin{aligned}\frac{d}{dt}P_{1T} &= k_3 \frac{m_1/K_3}{1 + m_1/K_3 + m_2/K_6} R_T - \gamma P_{1T},\\ \frac{d}{dt}P_2 &= -(k+\gamma)\frac{P_2 P_{1T}}{K + P_2} + k_6 \frac{m_2/K_6}{1 + m_1/K_3 + m_2/K_6} R_T - \gamma P_2,\end{aligned} \tag{3}$$

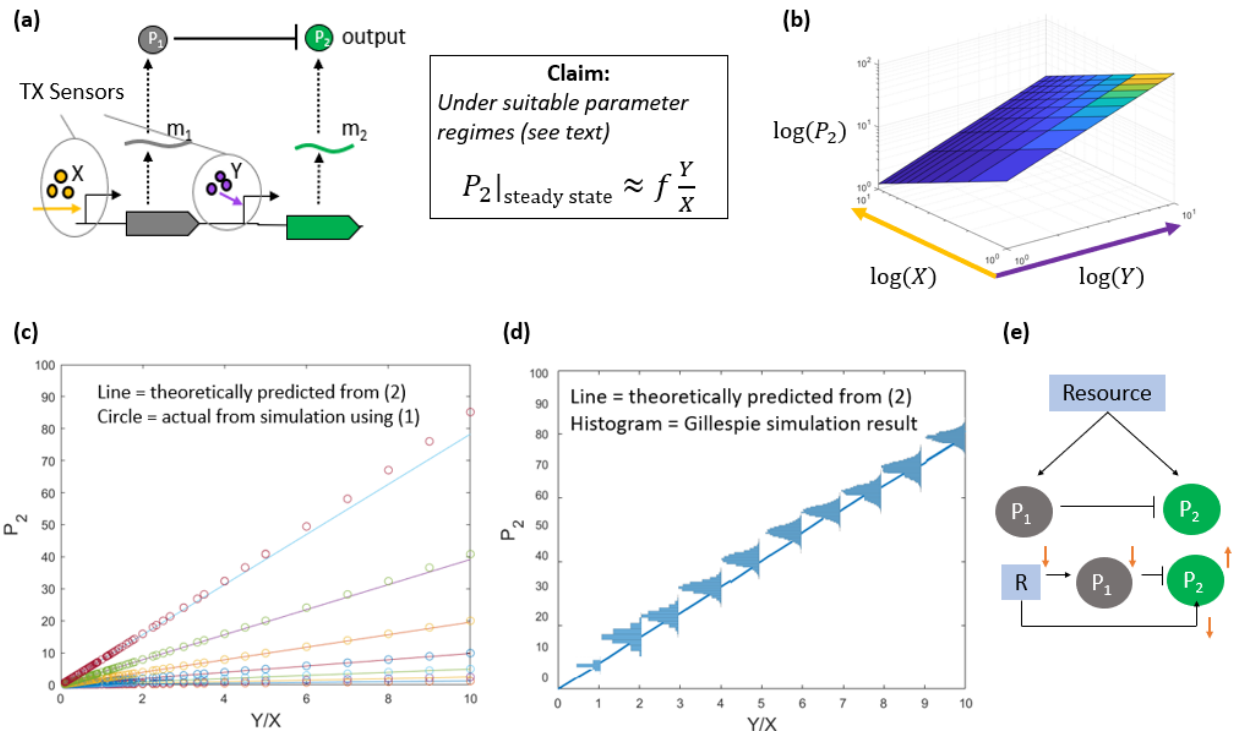


Figure 1: **Circuit design for transcriptional sensors based ratiometric sensor, which is robust to resource competition.** (a) Input X and Y can be a negative inducer of a repressor protein or a positive inducer of an activator protein. Therefore, more X (Y) produces more P_1 (P_2). In this design, P_1 is a protease that targets P_2 for degradation. Under suitable parameter regimes (see text), the steady state output protein level P_2 is proportional to $\frac{Y}{X}$. (b) Graph is obtained from simulating model (1) and plotting the steady state of P_2 as a function of X and Y plotted in logarithmic scale. (c) Theoretically predicted P_2 value from (2) matches well with the simulation result using (1) under the stated assumptions. (k_2 is changed) (d) Theoretically predicted P_2 value from (2) matches well with Gillespie simulations (Y is fixed, X is changed and k_2 is fixed). Therefore, our system is robust to instrinsic noise. (e) More resource (i.e. ribosome) produces more P_1 and P_2, and P_1 degrades P_2, so this circuit has inherent IFFL structure. Therefore, with proper parameter regimes, this circuit is robust to resource competition.

in which $K_1 = \frac{b_1}{a_1}$, $K_2 = \frac{b_2}{a_2}$, $K_4 = \frac{b_4}{a_4}$ are dissociation constants, $K_3 = \frac{b_3+k_3+\gamma}{a_3}$, $K_5 = \frac{b_5+k_5+\gamma}{a_5}$, $K_6 = \frac{b_6+k_6+\gamma}{a_6}$ and $K = \frac{b+k+\gamma}{a}$ are Michaelis-Menten constants. Here, we should pay attention that ODEs of both P_{1T} and P_2 contain $\frac{1}{1+m_1/K_3+m_2/K_6}$ term, which arise from the ribosome competition. By setting the time derivatives to zero and solving for P_2 and P_{1T}, we obtain the steady state value of P_2. Specifically, assuming that (A1) $P_2 \ll K$ (protease enzymatic reaction is in the first order regime), (A2) $\gamma \ll \frac{P_{1T}(k+\gamma)}{K}$ (dilution is negligible compared to degradation of P_2), (A3) $C_3 \ll K_5$, $K_2 \ll R_1$, $K_1 \ll X$, $Y \ll K_4$ (binding of the transcription factor to the promoter site is in the linear regime), $\frac{1}{1+m_1/K_3+m_2/K_6}$ term is canceled out and we obtain the steady state level of P_2 as

$$P_2 = \frac{\gamma k_5 k_6 K K_1 K_3 D_{2T} A_{1T} R_{1T}}{k_1 k_3 (k+\gamma) K_2 K_4 K_5 K_6 D_{1T}} \cdot (\frac{Y}{X}), \tag{4}$$

which is proportional to $\frac{Y}{X}$, even with the resource competition. For the future work, a similar work of Figure 1b and c should be done with (3) and (4).

To implement the ratiometric sensor (Figure 2), we choose orthogonal Lon protease (*mf*-Lon) from *Mesoplasma florum* [13] and superfolder GFP (sfGFP) protein [14] as the P_1 and P_2 proteins, respectively. By adding a cognate degradation tag of *mf*-Lon protease to the C-terminal of sfGFP protein, sfGFP protein can be specifically degraded by *mf*-Lon protease. We specifically choose pdt#3 tag from [13], which has the largest catalytic rate constant value, k ($P_1 + P_2 \underset{b}{\overset{a}{\rightleftharpoons}} C \xrightarrow{k} P_1$). Larger k is essential for the (A1) ($P_2 \ll K = \frac{b+k+\gamma}{a}$) to be satisfied. We select transcriptional repressor LacI and its inducer IPTG (the input X) to control *mf*-Lon protease expression which is driven by P_{Tac} promoter. Similarly, transcription activator NahR and its inducer Sal (the input Y) regulates sfGFP expression that is transcribed by P_{salTTC} promoter. The LacI and NahR transcriptional regulators are encoded in the genome and constitutively expressed in the host bacterium *E. coli* Marionette strain [5]. The reason that we choose IPTG and Sal is because they have the widest linear range of input/output relationship (Figure 3 of [5]), and this is crucial for the (A3) to be satisfied. We build the ratiometric sensor in the plasmid that uses p15A as the replication origin and kanamycin as the selection marker.

Figure 3 shows the experiment data. For the experiment, overnight culture was prepared by inoculating a $-80°C$ glycerol stock in 1000 μL M9 (+kanamycin) media in a 1.5 ml microcentrifuge tube and grew at $30°C$, 220 rpm in a horizontal orbiting shaker for 12 h. Overnight culture was first diluted to an initial optical density at 600 nm (OD_{600nm}) of 0.001 in 200 µL growth medium per well in a 96-well plate (Falcon, 351172) and grew for 1.5 h to ensure exponential growth before induction. The 96-well plate was incubated at $30°C$ in Tecan infinite M Nano+ microplate reader in static condition and was shaken at a fast speed for 3 s right before OD and fluorescence measurements and sampling interval was 5 min. Excitation and emission wavelengths to monitor GFP fluorescence were 488 and 548 nm, respectively. To ensure enough time to reach a steady state GFP/OD signal while the cells are in exponential growth, the cell culture was diluted with fresh growth medium to OD_{600nm} of 0.005 when OD_{600nm} approached 0.14 at the end of each batch. 2 batches were conducted for a total experiment time of up to 10 hours until GFP/OD reaches steady state. The steady state GFP/OD value is computed from the last batch of each experiment. For the inducers, we choose [IPTG] from 0 to 400 [μM] and for [Sal] from 0 to 50 [μM], which are in the linear range of input/output relationship. As we expected, the output P_2 decreases as the input X (IPTG) increases (Figure 3(a)) and it increases as the input Y (Sal) increases (Figure 3(b)). Also, the output P_2 is proportional to the ratio $\frac{Y}{X}$ of the two inputs, with $R^2 = 0.892$ (Figure 3(c)).

Summary and future work

In this part of the project, we focused on the design of robust and accurate multi-input biosensors that compute the ratio between the levels of different molecular species. The experiment results (Figure 3)

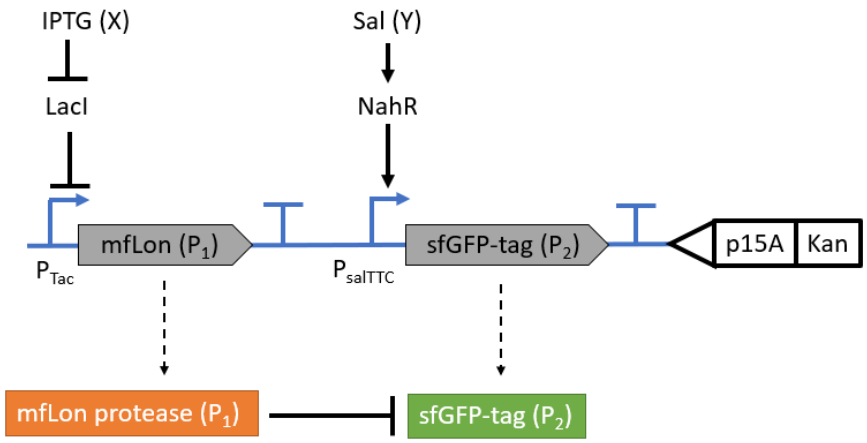


Figure 2: **Genetic construct of the ratiometric sensor.** Transcriptional repressor LacI and its inducer IPTG (the input X) control *mf*-Lon protease expression and transcriptional activator NahR and its inducer Sal (the input Y) regulates sfGFP expression. *Mf*-Lon protease degrades sfGFP.

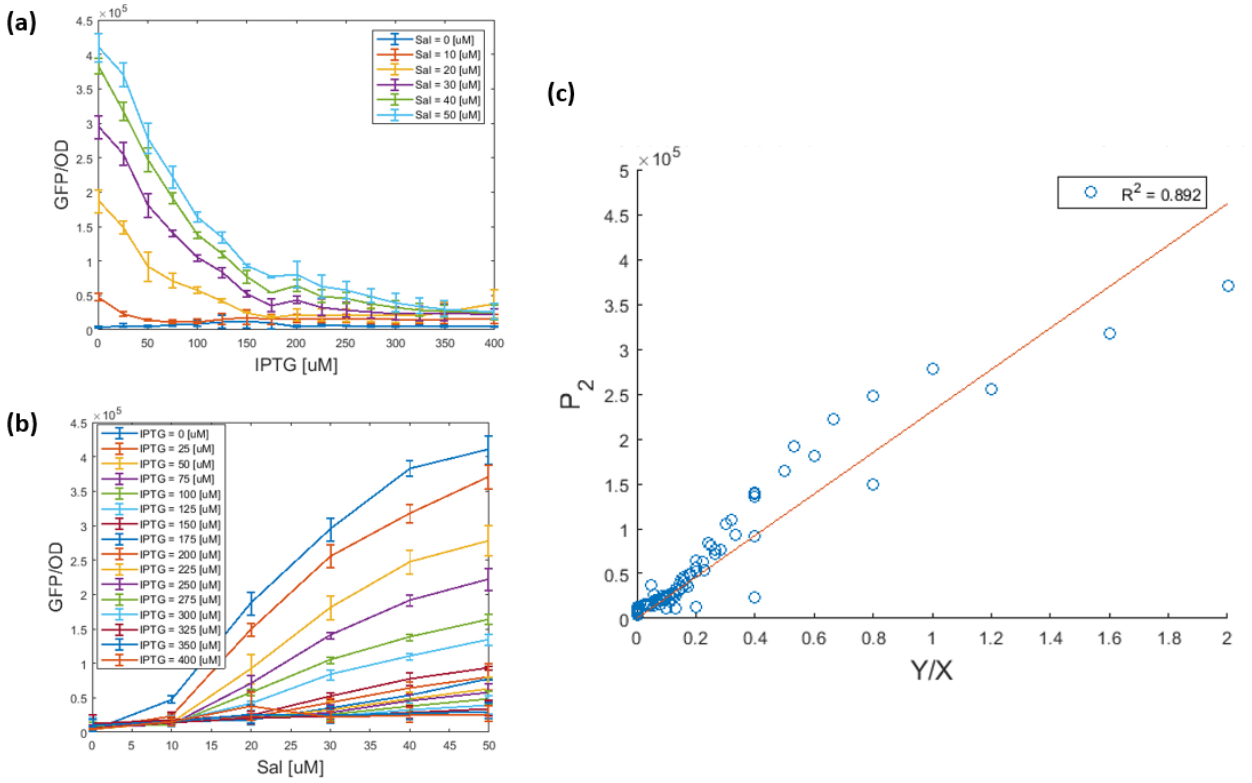


Figure 3: **Performance of the ratiometric sensor.** (a) The input X (IPTG inducer) can decrease the output P_2 (sfGFP protein). (b) The input Y (Sal inducer) can increase the output P_2 (sfGFP protein). (c) The output P_2 is proportional to the ratio $\frac{Y}{X}$ of the two inputs. The error bar indicates standard deviation of three biological replicates.

showed that the output P_2 is proportional to the ratio of two inputs, $\frac{Y}{X}$, which implied that our sensor functioned well as a ratiometric sensor. However, there are some remaining problems to be solved. Figure 4

indicates that more *mf*-Lon protease ([IPTG] more than 200 [μM]) causes larger growth rate reduction. To solve this issue, instead of using the external protease like *mf*-Lon protease, we plan to use potyvirus protease (i.e. TEV) and the endogenous protease (ClpSAP) machinery for the target protein degradation [15]. Potyvirus protease cleavages N-terminal of the labeled target protein and expose the amino acid sequence signal called the 'degron'. Then the endogenous protease machinery of E coli will carry out the degradation of the protein. Since it utilizes the endogenous protease machinery of *E. coli* instead of the external protease, we should be able to avoid the growth rate issue (Figure 5b). However, more GFP sequesters (for [IPTG] from 100 [μM] to 200 [μM], for [Sal] from 20 [μM] to 50 [μM]) *mf*-lon protease, so growth rate is not reduced in this case. This leads us to design Figure 5a. We hypothesize that constitutively expressed RFP with the same pdt#3 tag can sequester *mf*-Lon protease as well, so it may solve the growth rate issue. Furthermore, theory predicts that we can tune the coefficient of the ration by tuning the ribosome binding site (RBS) strengths, up to the extent where the assumptions for ratiometric computation are not violated. We therefore propose to construct libraries with varying RBS strengths to demonstrate the tunability of the sensor.

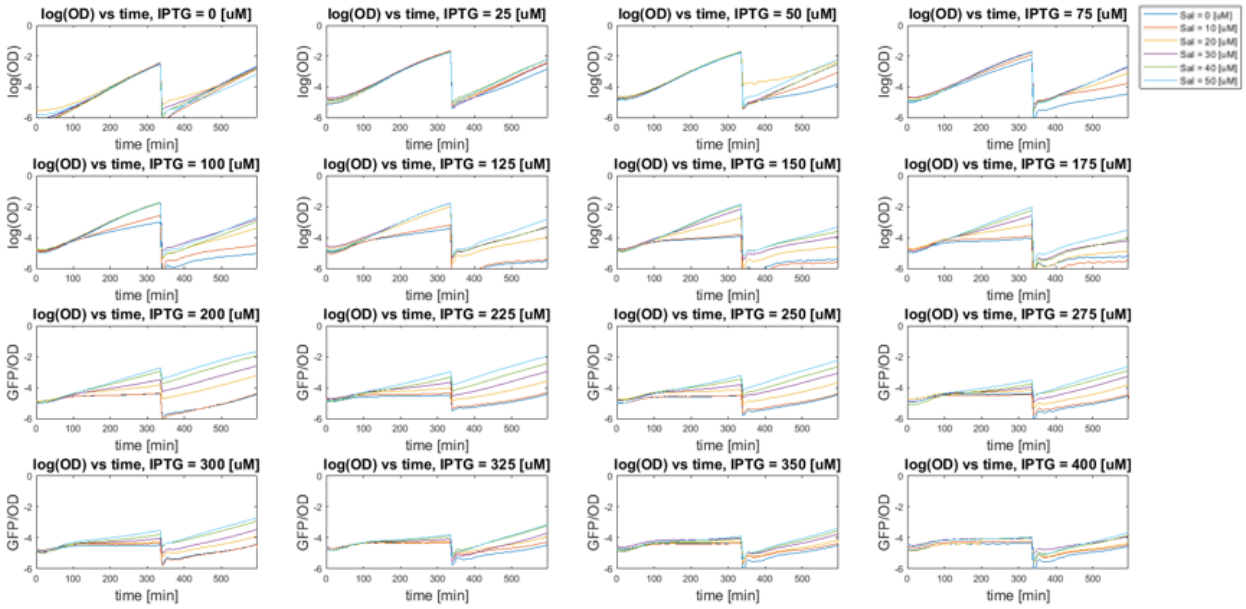


Figure 4: **Growth rate issue of the ratiometric sensor.** $\log(\mathrm{OD} - \mathrm{OD}_{\mathrm{blank}})$ vs time graphs with different [IPTG] and [Sal]. Growth rate is significantly decreased by overexpressed *mf*-Lon protease ([IPTG] more than 200 [μM]). More GFP sequesters (for [IPTG] from 100 [μM] to 200 [μM], for [Sal] from 20 [μM] to 50 [μM]) *mf*-lon protease, so growth rate is not reduced in this case.

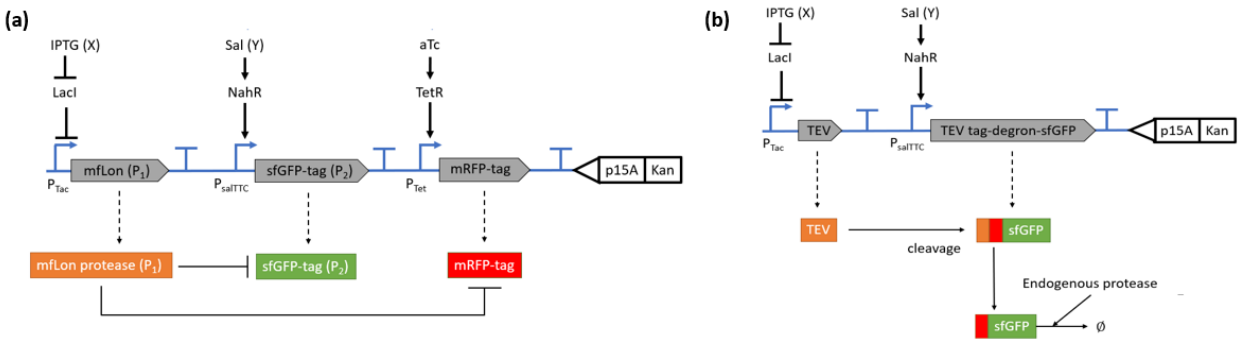


Figure 5: **Potential solutions to solve the growth rate decreasing issue** (a) Constitutively expressed mRFP with the same pdt#3 tag can be added to Figure 2. (b) Potyvirus protease (i.e. TEV) and the endogenous protease (ClpSAP) machinery based ratiometric sensor.

Part 2. airborne COVID biosensor

Motivation of the airborne COVID biosensor

The ability to perform on-site detection of infectious virus particles in aerosol in indoor environments is critical to prevent further spread of the disease. Coronavirus disease 2019 (COVID-19) pandemic so far has caused more than 174 million people with confirmed infection and more than 3.7 million people died worldwide. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of this disease. Airborne transmission has been identified as the dominant route for the spread of COVID-19 [16, 17, 18]. In particular, viable SARS-CoV-2 has been recovered from aerosols in hospital rooms of COVID-19 patients, indicating potential aerosol-induced infection [19]. As a consequence of the risk of infection through virus-loaded aerosol, the World Health Organization (WHO) updated the guidelines for infection prevention [20]. Also, the Center for Disease Control and Prevention (CDC) has officially recognized aerosols as a potential transmission mechanism [21]. A commentary signed by 239 researchers from the scientific and medical community around the globe are advocating preventive measures to mitigate this route of airborne transmission [22]. While these measures are critical during the current pandemic, many will remain in place to different degrees even after the pandemic is gone, just like many measures for homeland security have remained in place after 9/11. In fact, the SARS-CoV2 pandemic has been preceded by a number of other epidemics [SARS (2003), Swine Flu (2009), MERS (2012), Ebola (2014), the 2018 United States adenovirus outbreak], with about one major epidemic every 2-6 years, indicating continuous risk for the spread of infectious diseases. Example measures include improved air ventilation by increased fresh air in-flow in HVAC systems, space decontamination, and quarantine. One major gap is that these measures are often taken "open loop", that is, without knowing whether the virus is present or not in the air. This often leads to unnecessary costs for HVAC systems operation, services for decontamination, and workforce reduction due to quarantine. To fill this gap, this motivates us to devise a biosensor that collects aerosols and measures the viral load in the air on-site. This device will inform whether HVAC systems can safely recirculate air to save energy or need to exchange 100% of the air for occupants' safety, evacuation is necessary, or the space needs decontamination. It will further serve as a tool for early detection of high-risk biological threats that can spread from other countries when used in airplanes, airports, and cruise ships. In the long run, the system can become a platform for detection of any biological pathogen with a known genetic signature.

Main Results

We develop the process for on-site detection of virus in aerosol (Figure 6). At a high-level, the system first collects the aerosol from room(s) of interest on the metal probe of the sampling device. Then, the collected particles are detached from the probe and concentrated in water using a vortex mixer. Once in water, viral RNA is extracted by a 5 min heating step at 90 °C. After RNA extraction, the system performs chemical detection by a one-step RT-LAMP reaction at 65 °C. A portable fluorescence reader monitors fluorescence level produced by RT-LAMP reaction for 25 min. If within the 25 min the fluorescence level is above a pre-set threshold, then the system outputs "YES", otherwise it outputs "NO".

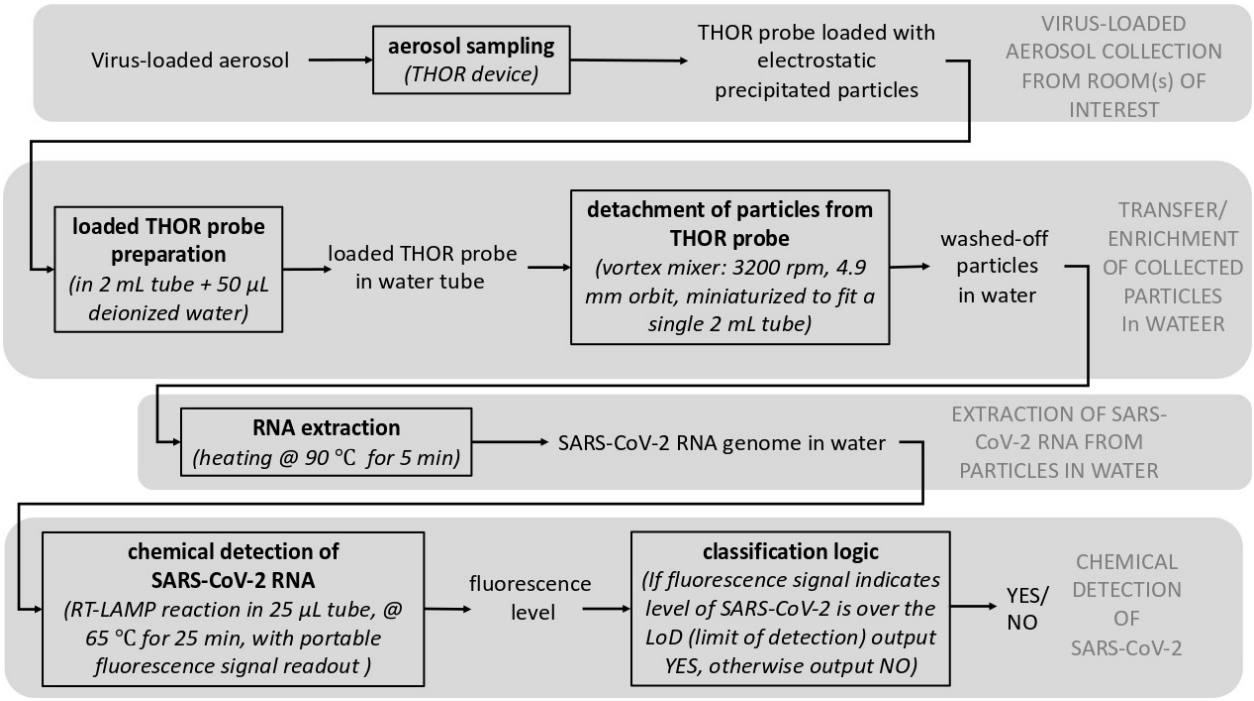


Figure 6: **Process for on-site detection of virus in aerosol.** The system specifically focuses on SARS-CoV-2, but the same process can be used for any other virus by replacing the RT-LAMP primers.

We design the stand-alone biosensor (Figure 7) that comprises a mobile air sampler device based on electrostatic precipitation (ESP) and a chemical detection kit. The detection kit allows detaching, enriching, and detecting the presence of low amounts of virions collected on the metal probe of the sampler with least human intervention. The key challenge here is to transfer the particles from the metal probe into water at sufficiently high concentrations to be detected by the RT-LAMP, even though the concentration in air may be low.

We explore the limit of the detection (LoD) of the RT-LAMP. Specifically, the RT-LAMP reaction uses the WarmStart LAMP 2x Master Mix kit (New England Biolabs, catalog E1700), 1 µM SYTO9 fluorescent dye, and a 10x stock of the combined primer sets to set up a 25 µL reaction with nuclease-free water (Zymo Research, catalog W1001) in a well of the 96-well plate (Axygen, catalog number PCR-96-LC480-W). The primer set N2 [24] and the primer set Orf1a-HMSe [25] target the N gene and a non-conserved region of the SARS-CoV-2 Orf1a gene, respectively. By using two primer sets concurrently, it improved the LoD by 2-fold consistent with other work [24]. The RT-LAMP method can be tailored to any other virus of interest by simply changing the primer set. In order to determine the LoD and the rate of false positive and false negatives, we run the RT-LAMP reaction using synthesized SARS-CoV-2 RNA target, incubated with

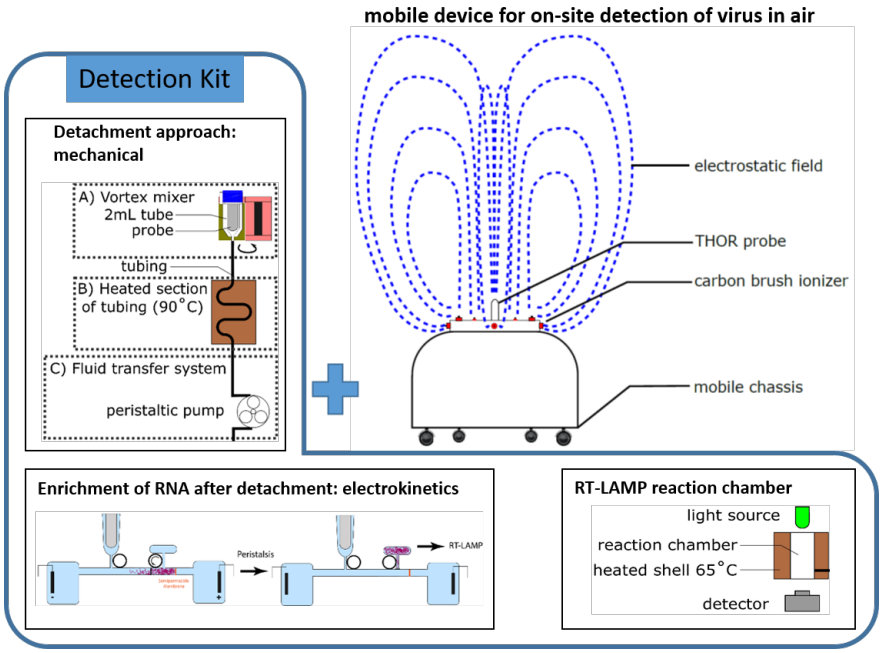


Figure 7: **Mobile platform for on-site detection of virus in air.** It comprises (i) a published air sampler based on electrostatic precipitation (ESP), called THOR [23], (ii) a system for maximizing detachment of viral material and concentration into aqueous solution, (iii) a system for RNA enrichment, (iv) and an RT-LAMP reaction chamber with fluorescence detector.

a constant temperature at 65 °C In order to assess false positive rate, we run the reaction for 1.5 hours. In order to monitor in real-time fluorescence of SYTO9, we run the one-step reaction in the Roche LightCycler 480. We conducted N = 50 biologically independent experiments to determine the LoD of the optimized RT- LAMP condition described above. We defined the LoD as the number of copies of viral RNA per 25 µL that can be detected with at least 95% positive rate (Figure 8). Therefore, the LoD is 50 copies per 25 µL.

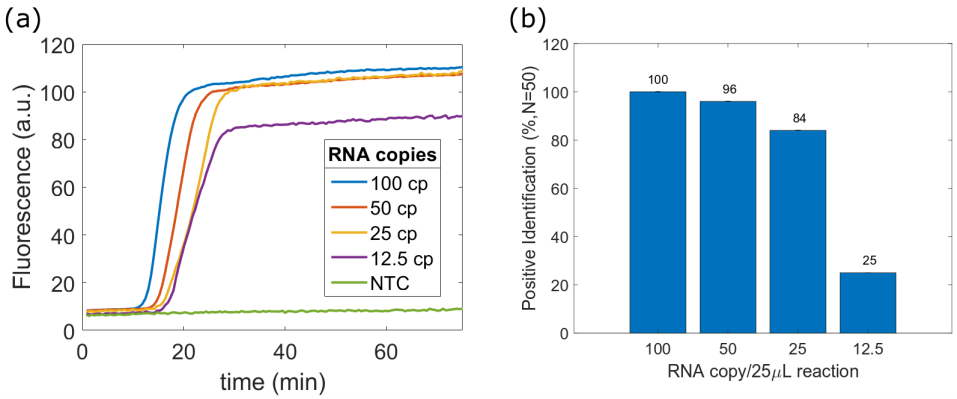


Figure 8: **RT-LAMP performance.** (a) Typical amplification curves from different RNA copy (cp) and the non-template control (NTC). (b) Limit of detection (LoD) is determined as 50 cp/25 µL reaction with 96% positive identification rate and the amplification time is between 13.8 and 15 min with 95% confidence interval (N=50 independent experiments). We use sterile de-ionized water as the non-template control (NTC). In term of the amplification-time threshold as 38.8 min, the false-positive rate of NTC is zero in 31 NTC samples.

To gauge the overall collection efficiency of the air sampler called THOR [23], we conduct the sampling experiments inside a containment cube (Abetement Technologies, AG3000MCCK) as shown in Figure 9a. Solutions of 1 µm diameter FluoSphere (Thermo Fisher, F13081), emulating viral particles, were properly diluted in filtered water to the desired concentrations and they were introduced in the tent through four nebulizers (Omron, NEC801) located at the four top corners of the tent. The THOR device was located in the tent in a central position, 20 cm from the ground and 1.5 meters from the nebulizers. At the end of the sampling phase, the collector was transferred in a 2 mL microcentrifuge tube and filtered water was added to the tube with the collector inside. Then, after vortexing the tube, the collector was removed thorough a magnet and a defined amount of CountBright Absolute Counting Beads (Invitrogen, C36950) was added to the sample left in the tube and analyzed using a BD Accuri C6 flow cytometer (BD Biosciences, San Jose, CA).

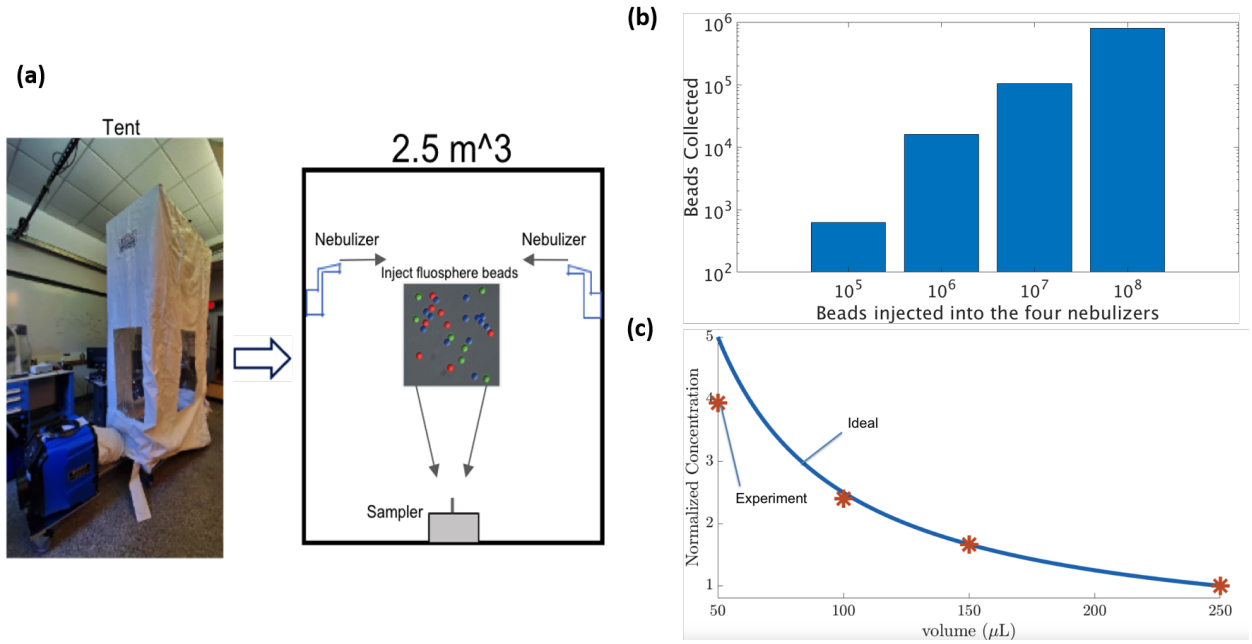


Figure 9: **(a) Experimental test-bed.** Containment tent for air sampler testing. **(b) Titration experiment.** We vary the number of FluoSpheres introduced into the nebulizers and measure the number of beads collected in water using the flow cytometer. The ratio is roughly equal to 0.01. **(c) The concentration of particles collected can be increased by reducing the extraction volume of de-ionized water.** In the graph the solid line represents the ideal behaviour (i.e., reduction of volume does not affect the mechanics of the removal process) of the collected particles concentration with respect to the extraction volume. The red dots represent the four experiments conducted, each one with a different extraction volume (i.e., 50 µL, 100 µL, 150 µL, 250 µL).

We first performed a titration experiment (Figure 9b) to determine the ratio between the total number of particles placed in the nebulizers and the particles collected in water using a de-ionized water volume of 250 µL, the amount of water that completely covers the probe in the 2 mL tube. The results indicate that the ratio of beads collected by the THOR sampler in water over beads introduced into the nebulizers is roughly equal to

$$r = \frac{\text{number of beads collected in water}}{\text{number of beads introduced in nebulizer}} \approx 0.01. \tag{5}$$

We note that the number of beads introduced in the nebulizer is larger than the number of beads injected

in the tent [data not shown]. Thus, the reported value of r is an under-estimation of the ratio between the number of particles collected in water and those present in the air.

We further optimize the detachment procedure that transfers collected particles from the metal probe to de-ionized water. The procedure uses a vortex mixer to generate mechanical agitation on 250 µL of water in a 2 mL tube. In order to optimize the process to enable detecting a lower concentration of particles in the air, we decreased the volume of de-ionized water added to the 2 mL tube, reasoning that this would increase the in-water concentration of the sample. This would be the case under the assumption that this reduction would not adversely affect the mechanics of the removal process. In particular, we conducted one experiment (beads injected $\approx 2 \times 10^5$ per m^3) simultaneous nebulization and collection with $t_{nebulization} = t_{collection} = 25$ minutes) four times, reducing each time the volume of water added in the tube. We considered 50 µL, 100 µL, 150 µL, and 250 µL as volumes of de-ionized water to add to the 2 mL tube. The results are shown in Figure 9c. Indeed, by reducing the volume we were able to increase the concentration of collected particles by up to four times. In particular, this plot indicates that the number of beads collected in water (at least for the concentration in air tested) is practically not affected by decreasing the water volume in the tube to 100 µL.

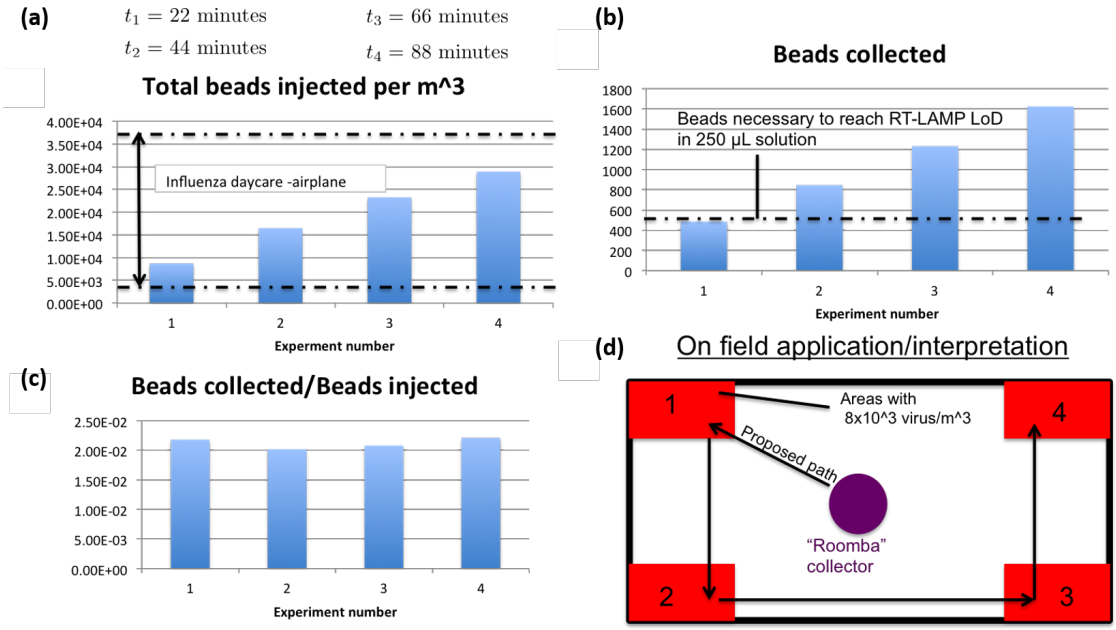


Figure 10: **On-field conditions experiment.** (a) Estimated concentration of beads injected in the tent and nebulization time (ti) for each of the experiments (i = 1, 2, 3, 4). Dashed lines delimit the typical range of influenza virus concentration found in airplanes and day care centers during flu season [26]. (b) The number of beads collected in water using the process described in the text and a de-ionized water volume of 250 µL. The dashed line denotes the number of bead required (in a 250 µL solution) to reach the calculated RT-LAMP LoD. (c) The ration between the number of beads collected in water and the estimated number of beads injected in the tent (constant). (d) Application scenarios emulated by the four experiments.

The second experiment that we performed had the objective of testing the sampling and collection efficiency in an environment that resembles on-field conditions, such as having the device roaming in a closed room with one or more sick people (see Figure 10d). To this end, we considered a continuous but small rate of nebulization (simulating someone speaking or breathing) and a simultaneous collection (device roaming and collecting around the room). In particular, we conducted four experiments that can be described as

follows:

- Experiment 1: turn on the first nebulizer for 22 minutes;
- Experiment 2: turn on the first nebulizer for 22 minutes, then turn off and turn on the second nebulizer for 22 minutes;
- Experiment 3: turn on the first nebulizer for 22 minutes, then turn off and turn on the second nebulizer for 22 minutes, then turn off and turn on the third nebulizer for 22 minutes;
- Experiment 4: turn on the first nebulizer for 22 minutes, then turn off and turn on the second nebulizer for 22 minutes, then turn off and turn on the third nebulizer for 22 minutes, then turn off and turn on the fourth nebulizer for 22 minutes.

The THOR device was kept on for the entire experiment (i.e. $t_{nebulization} = t_{sampling}$). For experiments 1-4 the total number of FluoSpheres placed inside the nebulizers approximately corresponds to a concentration inside the containment cube as indicated by the values shown in Figure 10a. These concentrations are within typical values found in airplanes/day care centers during influenza season [26], as indicated by the dashed lines. The results of these experiments are shown in Figure 10b,c. Based on these results we conclude that for an expected on-field viral concentration, increasing the sampling time will increase the amount of particles collected. This implies that, if there is one or more sick people in a room, the THOR will be able to collect an amount of particles equal or higher than the RT-LAMP LoD after a sufficiently long time. Given that the RT-LAMP LoD with our primers is given by 50 copies per 25 μL reaction (i.e. 2 copies per μL), we have that the minimum number of beads collected in 250 μL water that can be detected by the RT-LAMP is given by 500 (See the horizontal dash line in Figure 10b). Therefore, the smallest concentration of particles in the air that we could detect is given by 8740 beads per m^3, after collecting for t_{min} = 22 minutes. It is important to point out that we injected the particles constantly for the entire experimental time (22 minutes), meaning that the instantaneous concentration in the tent is lower than 8740 beads per m^3 (Figure 10a, the experiment 1).

Summary and future work

In this part of the project, we have characterized the overall collection efficiency of the air sampler, and optimized the detachment procedure with fluosphere beads that emulates virus particles in air. The results suggested that this sampling procedure can approximately exceed the determined LoD as 50 copies/25 μL of RT-LAMP detection (Figure 11) and thus, suggested the feasibility of the developed methodology from collection to detection of virus particles in aerosol. In the future, we will conduct optimizations of the process (Figure 6) and build a prototype (Figure 7) that integrates the optimized individual steps from collection to detection of virus particles. This prototype then could serve as the base for future integration with the HVAC system.

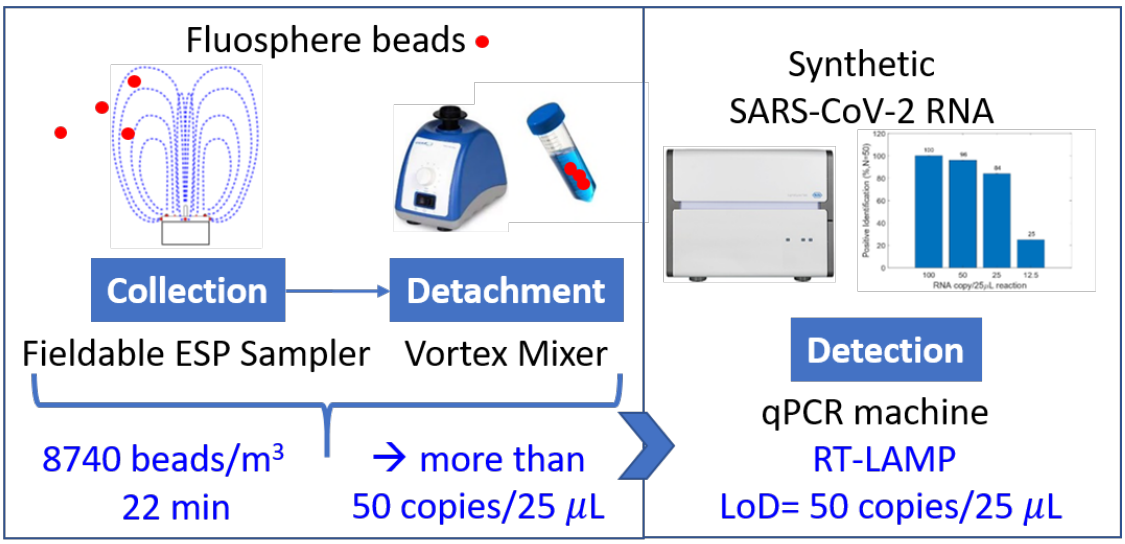


Figure 11: **Graphical summary of feasibility assessment.** With about 8740 beads/m3 of fluosphere beads in air, by sampling for more than 22 minutes and by optimized mechanical detachment, we can obtain more than 50 copies of beads per 25 uL. This is above the LoD of RT-LAMP characterized using a qPCR machine.

References

[1] S. Wang, J. Mason, D. Charney, R. Yehuda, S. Riney, and S. Southwick, "Relationships between hormonal profile and novelty seeking in combat-related posttraumatic stress disorder," *Biological Psychiatry*, vol. 41, no. 2, pp. 145–151, 1997.

[2] R. M. Shia, L. K. McIntire, J. A. Hagen, C. D. Goodyear, L. N. Dykstra, and A. R. Myers, "Biomarker and biometric indices of physical exhaustion in the firefighting community," *Procedia Manufacturing*, vol. 3, pp. 5081–5087, 2015.

[3] R. M. Shia, J. A. Hagen, L. K. McIntire, C. D. Goodyear, L. N. Dykstra, and L. Narayanan, "Individual differences in biophysiological toughness: sustaining working memory during physical exhaustion," *Military medicine*, vol. 180, no. 2, pp. 230–236, 2015.

[4] D. Mauss, J. Li, B. Schmidt, P. Angerer, and M. N. Jarczok, "Measuring allostatic load in the workforce—a systematic review," *Industrial health*, 2014.

[5] A. J. Meyer, T. H. Segall-Shapiro, E. Glassey, J. Zhang, and C. A. Voigt, "Escherichia coli "marionette" strains with 12 highly optimized small-molecule sensors," *Nature chemical biology*, vol. 15, no. 2, pp. 196–204, 2019.

[6] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource competition shapes the response of genetic circuits," *ACS synthetic biology*, vol. 6, no. 7, pp. 1263–1272, 2017.

[7] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost lines describe the cellular economy of genetic circuits," *Biophysical journal*, vol. 109, no. 3, pp. 639–646, 2015.

[8] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis, "Quantifying cellular capacity identifies gene expression designs with reduced burden," *Nature methods*, vol. 12, no. 5, pp. 415–418, 2015.

[9] Y. E. Antebi, J. M. Linton, H. Klumpe, B. Bintu, M. Gong, C. Su, R. McCardell, and M. B. Elowitz, "Combinatorial signal perception in the bmp pathway," *Cell*, vol. 170, no. 6, pp. 1184–1196, 2017.

[10] N. Nandagopal, L. A. Santat, L. LeBon, D. Sprinzak, M. E. Bronner, and M. B. Elowitz, "Dynamic ligand discrimination in the notch signaling pathway," *Cell*, vol. 172, no. 4, pp. 869–880, 2018.

[11] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press, 2014.

[12] D. E. Cameron and J. J. Collins, "Tunable protein degradation in bacteria," *Nature biotechnology*, vol. 32, no. 12, pp. 1276–1281, 2014.

[13] D. E. Cameron and J. J. Collins, "Tunable protein degradation in bacteria," *Nature Biotechnology*, vol. 32, no. 12, pp. 1276–1281, 2014.

[14] J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, and G. S. Waldo, "Engineering and characterization of a superfolder green fluorescent protein," *Nature Biotechnology*, vol. 24, no. 1, pp. 79–88, 2005.

[15] J. Fernandez-Rodriguez and C. A. Voigt, "Post-translational control of genetic circuits using potyvirus proteases," *Nucleic acids research*, vol. 44, no. 13, pp. 6493–6502, 2016.

[16] R. Zhang, Y. Li, A. L. Zhang, Y. Wang, and M. J. Molina, "Identifying airborne transmission as the dominant route for the spread of COVID-19," *Proceedings of the National Academy of Sciences*, vol. 117, no. 26, pp. 14857–14863, 2020.

[17] E. A. Meyerowitz, A. Richterman, R. T. Gandhi, and P. E. Sax, "Transmission of SARS-CoV-2: A review of viral, host, and environmental factors," *Annals of Internal Medicine*, vol. 174, no. 1, pp. 69–79, 2021.

[18] E. L. Anderson, P. Turnham, J. R. Griffin, and C. C. Clarke, "Consideration of the aerosol transmission for COVID-19 and public health," *Risk Analysis*, vol. 40, no. 5, pp. 902–907, 2020.

[19] J. A. Lednicky, M. Lauzardo, Z. H. Fan, A. Jutla, T. B. Tilly, M. Gangwar, M. Usmani, S. N. Shankar, K. Mohamed, A. Eiguren-Fernandez, C. J. Stephenson, M. M. Alam, M. A. Elbadry, J. C. Loeb, K. Subramaniam, T. B. Waltzek, K. Cherabuddi, J. G. Morris, and C.-Y. Wu, "Viable SARS-CoV-2 in the air of a hospital room with COVID-19 patients," *International Journal of Infectious Diseases*, vol. 100, pp. 476–482, 2020.

[20] World Health Organization, "Transmission of SARS-CoV-2: implications for infection prevention precautions," 2020.

[21] Centers for Disease Control and Prevention, "Scientific brief: SARS-CoV-2 Transmission," 2021.

[22] L. Morawska and D. K. Milton, "It is time to address airborne transmission of coronavirus disease 2019 (COVID-19)," *Clinical Infectious Diseases*, 2020.

[23] N. R. de Sousa, N. Sandström, L. Shen, K. Håkansson, R. Vezozzo, K. I. Udekwu, J. Croda, and A. G. Rothfuchs, "A fieldable electrostatic air sampler enabling tuberculosis detection in bioaerosols," *Tuberculosis*, vol. 120, p. 101896, 2020.

[24] Y. Zhang, G. Ren, J. Buss, A. J. Barry, G. C. Patton, and N. A. Tanner, "Enhancing colorimetric loop-mediated isothermal amplification speed and sensitivity with guanidine chloride," *BioTechniques*, vol. 69, no. 3, pp. 178–185, 2020.

[25] B. A. Rabe and C. 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*, vol. 117, no. 39, pp. 24450–24458, 2020.

[26] W. Yang, S. Elankumaran, and L. C. Marr, "Concentrations and size distributions of airborne influenza a viruses measured indoors at a health centre, a day-care centre and on aeroplanes," *Journal of The Royal Society Interface*, vol. 8, no. 61, pp. 1176–1184, 2011.