NANOPORE SENSING OF PCR-AMPLIFIED PATHOGENIC DNA

Simon King

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Master's Degree in Physics

Department of Physics
Faculty of Science
University of Ottawa

Table of Contents

List of Figures

Main

Appendix A

List of Abbreviations

Abbreviation	Definition
Ag/AgCl	Silver/Silver Chloride
bp / kbp	base pair / kilobase pair
BSA	Bovine Serum Albumin
CBD	Controlled Breakdown
cDNA	complimentary DNA
CRC	Colorectal Cancer
C_q	Cycle of quantification
CUSUM	Cumulative Sum algorithm
ddPCR	Droplet Digital PCR
DMSO	Dimethyl Sulphoxide
dNTP	deoxynucleoside triphosphate
dsDNA / ssDNA	double stranded DNA / single stranded DNA
GAS	Group A Streptococcus
I-V curve	Current vs Voltage curve
KCl	Potassium Chloride
LAMP	Loop-mediated isothermal amplification
LFA	Lateral Flow Assay
LiCl	Lithium Chloride
LPF	Low-Pass Filter
miRNA	microRNA
mM / nM	millimolar (10^{-6} moles/L) / nanomolar (10^{-9} moles/L)
mRNA	messenger RNA
NA	Nucleic Acid
nA / pA	nanoamp (10^{-9} amperes) / picoamp (10^{-12} amperes)
NEAR	Nicking Enzyme Amplification Reaction
ng	nanogram (10^{-9} grams)
nm / µm	nanometer (10^{-9} meters) / micrometer (10^{-6} meters)
nt	nucleotide
PCR	Polymerase Chain Reaction
PEG	Polyethelene glycol
PNA	Peptide Nucleic Acid
PSD	Power Spectral Density
qPCR	real-time PCR
RMS	Root Mean Square noise
RNA	Ribonucleic Acid
RPA	Recombinant Polymerase Amplification
RT	Reverse Transcription
SiN_x	Silicon Nitride
SiO_2	Silicon Dioxide
SNR	Signal-to-Noise Ratio
SSNP	Solid-State Nanopore
TEM	Transmission Electron Microscope
µs	microsecond (10^{-6} seconds)

Summary

Solid-state nanopore sensors are an emerging platform for performing single-molecule characterization of biomolecules such as DNA. With the advent of Controlled Breakdown (CBD), creating a simple, tunable, ultra-low concentration sensing device *in situ* has enabled their direct integration with a host of platforms. The simplicity and sensitivity in performing measurements allows nanopore-based technologies to find uses which enhance existing methods.

One such promising avenue for nanopore-based sensing is in the detection of infectious diseases, where early and accurate identification of the causative pathogen is essential for successful patient outcomes. Conventional assays, while effective, often have limitations in speed, cost, or target sensitivity, and may benefit from nanopore sensing approaches. However, solid-state nanopores currently lack the ability to discriminate between biomarkers sharing identical size and charge densities, such as sequentially-heterogeneous strands of DNA.

Addressing the weakness of both conventional assays and nanopores could come from combining nanopore sensing with the polymerase chain reaction (PCR), a well-established and highly-selective nucleic acid amplification scheme. PCR is designed to produce large quantities of identical fragments of DNA, known as amplicons, if a user-defined parent copy is present. After the PCR process has finished, the signals produced by this population of amplicons on a nanopore sensor should therefore be indicative to the presence of a DNA biomarker in the starting sample.

As PCR reactions can use a mix of different proteins, detergents, and other molecules, the challenge lies in ensuring the compatibility of these reagents with a nanopore, and determining whether the background signal they produce can be discriminated from an amplicon signal. To this end, this thesis investigates PCR-nanopore compatibility, and experimentally demonstrates a nanopore signal-classification technique to successfully identify the presence of chromosomal DNA from Group A Streptococcus (GAS), an infectious bacterium responsible for strep throat, in samples derived from clinical extracts.

Sommaire

Les capteurs nanopores à l'état solide sont une plateforme émergente pour la caractérisation unimoléculaire de biomolécules comme l'ADN. Avec l'avènement du claquage contrôlé (CBD) pour la fabrication de nanopore, la création *in situ* d'un dispositif de détection simple, électronique et sensible, a permis leur intégration directe dans une multitude de plateformes. La simplicité et la sensibilité des mesures effectuées permettent des technologies basées sur les nanopores de trouver des utilisations qui améliorent les méthodes existantes.

L'une de ces voies prometteuses pour la détection basée sur les nanopores est le diagnostic des maladies infectieuses, où l'identification précoce et précise de l'agent pathogène responsable est essentielle pour le bien-être des patients. Les tests conventionnels, bien qu'efficaces, présentent souvent des limites en termes de vitesse, de coût ou de sensibilité de la cible, et peuvent donc bénéficier des approches de détection par nanopore. Cependant, les nanopores à l'état solide n'ont actuellement pas la capacité de distinguer les biomarqueurs partageant une taille et une densité de charge identiques, comme les brins d'ADN de séquences hétérogènes.

Pour remédier aux faiblesses des tests conventionnels et des nanopores, il est possible de combiner la détection par nanopores avec la réaction en chaîne par polymérase (PCR), un schéma d'amplification des acides nucléiques bien établi et hautement sélectif. La PCR est conçue pour produire de grandes quantités de fragments identiques d'ADN, appelés amplicons, si une copie mère définie par l'utilisateur est présente. Une fois le processus PCR terminé, les signaux produits par cette population d'amplicons sur un capteur nanopore devraient donc indiquer et quantifier la présence d'un biomarqueur d'ADN dans l'échantillon de départ.

Comme les réactions PCR peuvent utiliser un mélange de différentes protéines, détergents et autres molécules, le défi consiste à assurer la compatibilité de ces réactifs avec un nanopore, et à déterminer si le signal de fond qu'ils produisent peut être discriminé d'un signal d'amplicon. À cette fin, cette thèse étudie la compatibilité PCR-nanopore, et démontre expérimentalement une technique de classification du signal nanopore pour identifier avec succès la présence d'ADN chromosomique du streptocoque du groupe A (GAS), une bactérie infectieuse responsable de l'angine streptococcique, dans des échantillons provenant d'extraits cliniques.

Acknowledgements

I would like to thank all my colleagues in the Tabard-Cossa lab. Conducting research in such a collegial environment has been a real privilege, and working with all of you has made my time spent here much more enjoyable than it could have been anywhere else.

I'd like to acknowledge the effort that Kyle has put into developing the software I used for this thesis research. I'm proud to have helped him improve his software so much by breaking it repeatedly. I also appreciate his guidance in both experimental design and scientific writing, which markedly improved the quality of my research. Plus he runs a great D&D campaign.

Matt was my supervisor during much of my time as a co-op student with the lab, and shares a large part of the blame for my decision to pursue graduate studies. He has always been available to answer a question or give advice, and his excellent day-to-day management of the lab gave me space to focus on my research. Most important though, has been his willingness to share a laugh and a beer with me after work. I'm privileged to call him my friend.

I'd also like to thank Vincent for giving me the opportunity to pursue my studies with his lab, and his patience and mentorship during the worst of a worldwide pandemic-induced lockdown. As my graduate supervisor and as a person, he has been truly exceptional.

My wonderful girlfriend Ashley has been my steadfast partner throughout the entire scholastic process from undergrad co-op to master's thesis. I wouldn't have made it this far without her love and support.

Finally, I'd like to thank my parents, Brock and Helen, for all the gifts they've given me, and the loving environment in which I could grow. My mother, Helen, instilled in me a deep-seated curiosity of the world around me, while my father, Brock, provided me all the philosophical tools to help make sense of it. For that I feel truly grateful.

Statement of Originality

Research was conducted during my work as an MSc student under the supervision of Dr. Vincent Tabard-Cossa at the Department of Physics of the University of Ottawa. All text, data and figures are of my own creation unless otherwise indicated, with sources acknowledged.

Chapter 1 contains an introductory overview of nanopore and PCR concepts, as well as a review of the relevant background literature.

Chapter 2 and Appendix A contain research from a manuscript that has been published as an article in ACS Sensors:

> King, S.; Briggs, K.; Slinger, R.; Tabard-Cossa, V. *"Screening for Group A Streptococcal disease via Solid-State Nanopore Detection of PCR Amplicons."* ACS Sensors 7, 1, 207–214 (2022) - https://pubs.acs.org/doi/10.1021/acssensors.1c01972

I performed all experiments and analysis for the data presented in this chapter and wrote the first draft of the manuscript. Dr. Robert Slinger provided clinical samples. Drs. Kyle Briggs and Vincent Tabard-Cossa edited the text.

Chapter 3 contains an outlook and details some preliminary research I have performed since conclusion of the PCR work presented in Chapter 2, in collaboration with Dr. Erin McConnell of the T-Cossa Lab.

Chapter 1: Introduction to Nanopore Sensing and Infectious Disease Assays

1.1 Introduction

In this section I will present an overview of the major aspects related to my thesis work, in order to provide a degree of background on these different areas of research. In the first section, I will review the basics of nanopore sensing, including the instrumentation used to record the data of my thesis and a description of the analysis performed to extract molecular information. In the second section, I will provide a high-level review of some of the current techniques used in infectious disease testing, focusing on PCR-based tests, and explain the background of this workhorse amplification strategy for many detection tools. I will then discuss attempts to pair nanopore DNA sensing with techniques for discriminating biomarkers by genetic sequence.

A. NANOPORE SENSORS

1.2 Nanopore Basics

At its most fundamental, a nanopore is just a tiny hole in a thin membrane separating two volumes of an electrolyte. Any ion or molecule that passes from one reservoir to the other must travel through this aperture. Due to nanometer size scales of pore length and diameter, larger molecules must pass through one at a time. When coupled with a means of driving and detecting these translocations, a nanopore becomes single-molecule sensor with a myriad of applications, including DNA sequencing, dilute analyte detection, protein characterization and even data storage.[1]

Broadly speaking, there are two types of nanopores: biological and solid-state (Figure 1). Biological nanopores were initially designed to emulate those found in a cell: a transmembrane protein channel embedded in a phospholipid bilayer.[2] The advantages of biopores are their

atomic precision in shape, which enables their use in sequencing applications. Furthermore, their properties can be precisely tuned with site-directed mutagenesis for better adaptation to various targets. They are not without limitations however: both the lipid membrane and the protein channel are sensitive to their environment, meaning biopores require carefully controlled conditions to perform properly.[3]

This has led more recently to the development of solid-state nanopores (SSNPs), in which a pore is formed in an inorganic material. SSNPs are made in mechanically robust membranes which can tolerate a wide range of experimental conditions including pH, temperature, and voltage. They also have tunable pore sizes and depths. Solid-state nanopores have been formed in glass, quartz, and dielectric materials such as silicon nitride (SiN). These insulator-material pores can be formed by drilling with a high-powered TEM beam,[4] however this requires expensive and cumbersome equipment. A simpler, quicker method of nanopore fabrication has more recently been developed by the Tabard-Cossa lab, using a technique known as controlled dielectric breakdown (CBD),[5] which is explained in Section 1.5, below.

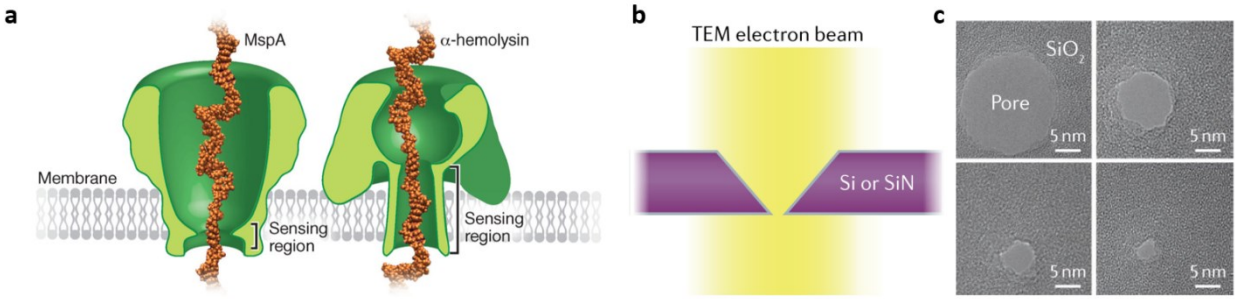

Figure 1: Depictions of biological and solid state nanopores. a) Biological nanopores embedded in a lipid bilayer [Image taken from *Three Decades of Nanopore Sequencing* (Nat. Biotechnol.) with permission] b) Formation of a solid-state nanopore in Si/SiO_2 with a high-energy TEM beam c) TEM image of a solid state nanopore. [Images taken from *Solid-state nanopore sensors* (Nat. Reviews) with permission]

1.3 Nanopore Sensing

The application of a low, sub-1V, electric potential across the membrane pulls charged ions through a nanopore. This flow of charges produces measurable ionic current. The conductance (σ), the ohmic inverse of the resistance, provides a voltage-independent means of characterizing the ionic throughput of a pore. This is modeled as a combination of a cylindrical pore's access resistance[6] and its geometry,[7] giving

$$G = \frac{I}{V} = \sigma\left(\frac{4l}{\pi d^2} + \frac{1}{d}\right)^{-1} \quad (1)$$

where *I* is the measured current, *V* is the applied potential, *σ* is the bulk solution conductivity, *l* is the length and *d* is the diameter of a cylindrical pore. With measurement of the pore conductance, the diameter of the pore may be extracted provided the pore length is known.

If a molecule is pulled along the electric field and into the pore, it effectively blocks a portion of this ionic current. It therefore follows that

$$\Delta G = G_{open\ pore} - G_{blocked\ pore} = G_{molecule} \quad (2)$$

where $G_{open\ pore}$ is the conductance of an unoccluded pore and $G_{blocked\ pore}$ is the conductance of the pore when blocked by said molecule. The measurement of this decrease in conductance can be used to infer properties of the molecule. Extracted characteristics can include molecular length, cross-sectional area, shape, and conformation. For example, the conductance change from a DNA translocation, when related to a simple model for cylindrical dsDNA blockage,

$$\Delta G = G_{DNA} = \frac{\sigma \pi d_{DNA}^2}{4l} \quad (3)$$

can be used to determine either the length of a nanopore channel or the expected blockage depth of DNA translocation, using the established diameter of d_{DNA}=2.2nm. This allows us to establish a characteristic blockage depth for identifying DNA translocations in a complex mixture (Figure 2).

The number of molecules translocating for a given amount of time, or the capture rate, is directly related to the concentration, though capture rates vary based on pore geometry and so require some internal standard to precisely quantify. Alternative ways of calculating capture rate can be found in recent work done by the Tabard-Cossa lab.[8]

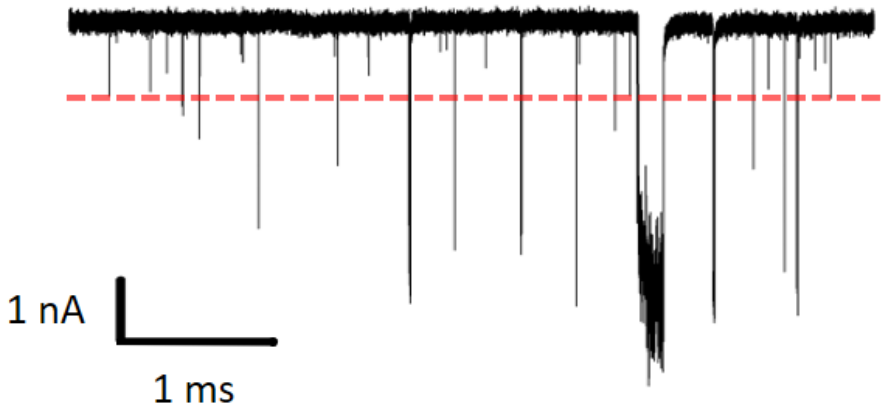

Figure 2: Current over time through a nanopore at an applied potential. As a complex mixture of DNA, proteins and other individual molecules translocate through a pore, varying amounts of current are blocked. Unfolded DNA translocations can be distinguished from others around a characteristic blockage depth (dashed line) based on equation 3, above.

1.4 Sensing Instrumentation

The blockage magnitude and passage time of single-molecule translocations through a pore can be small. Often on the scale of nanoamps and microseconds, the resolution and characterization of these translocations requires sensitive instrumentation. A typical nanopore

measurement setup is presented in Figure 3. A nanopore device is housed in a millifluidic cell and connected to a low-noise current amplifier – either an Axopatch 200B or a VC100 Chimera – via a headstage. The nanopore-fluidic cell and a connected headstage, which applies potentials and measures currents, are housed in a faraday cage to minimize external electromagnetic noise. During the sensing process, headstage-amplified analog signals from the nanopore are digitized and recorded to a computer via a data acquisition (DAQ) card, which interfaces between the amplifier and the computer running custom software.

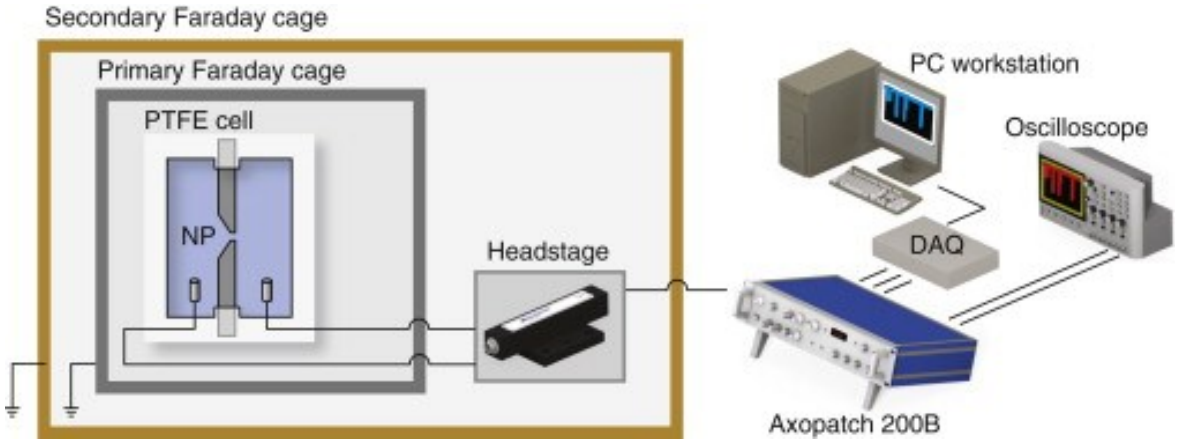

Figure 3: Schematic of a typical nanopore sensing setup. The nanopore is housed in a flowcell, which is in turn protected from external EM noise in a faraday cage. The headstage interfaces with an Axopatch or Chimera (not depicted) current amplifier. Analog amplifier readings are converted to digital via the Data Acquisition (DAQ) card and recorded to a computer for analysis.[9]

1.5 Forming a nanopore by Controlled Breakdown

The controlled breakdown (CBD) technique for fabricating nanopores consists of the application of 5-20V potentials across a thin (10-30nm) dielectric membrane immersed in an electrolyte such as KCl or LiCl.

The resulting electric field produces a measurable, highly localized, leakage current across the chip. A portion of this leakage current is believed to be produced by electron tunneling through imperfections in the membrane known as traps.[10] Over time, these traps begin to accumulate at

a site of higher conductance from this leakage current, amplifying its strength in a positive-feedback mechanism. Once a critical number of traps have accumulated, the membrane material breaks down, leaving behind a pore. Further enlarging can be achieved through the application of small pulses of alternating potentials, until the desired size is reached.[11] This method allows for nanopore fabrication in a short period of time and with inexpensive equipment, and in the same apparatus as that used for single-molecule sensing (Figure 4).

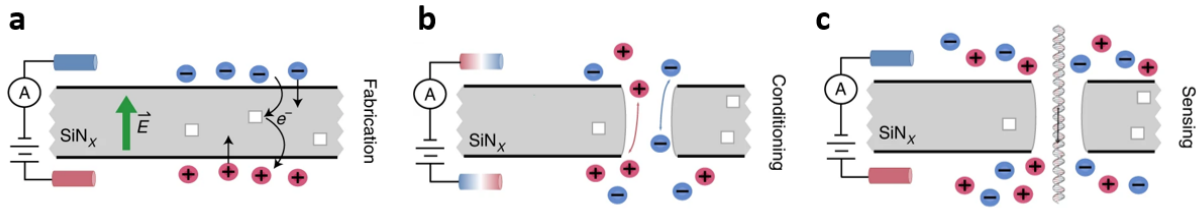

Figure 4: Schematic of in-situ nanopore processes. a) During fabrication, high potential across the SiN_x membrane leads to the creation and accumulation of traps (white boxes), which localize and amplify tunneling until dielectric breakdown occurs. b) Once a pore has formed, it may be enlarged, or conditioned, through the application of short, lower-potential pulses. c) Single-molecule sensing such as for DNA can occur in-situ with fabrication. [image taken from *Solid-state nanopore fabrication by automated controlled breakdown* (Nat. Protocols) with permission]

1.6 Fabrication Instrumentation

The nanopore fabrication setup used in this thesis makes use of custom instrumentation and software developed in part by the author during coop placements and summer research internships.[12] Nanopore chips possess a free-standing 40x40 µm window of exposed SiN_x, approximately 10 nm thick in their center. This is supported by a 5x5 mm square, 200 µm thick piece of Si substrate (Figure 5). Some chip types have additional layers of SiO_2, which has been shown to improve signal-to-noise characteristics during sensing. These chips are fabricated at the wafer scale and were purchased commercially for the work perform in this thesis.

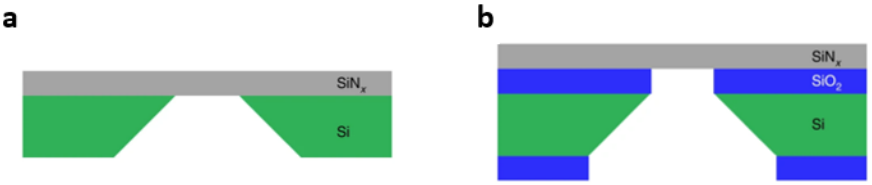

Figure 5: Silicon chip types used in the fabrication of nanopores. Schematic depiction (not to scale) of a) Si/SiNx and b) Si/SiO2/SiNx chips, demonstrating the exposed SiNx windows. [image taken from *Solid-state nanopore fabrication by automated controlled breakdown* (Nat. Protocols) with permission]

The chip is sandwiched between two elastomer gaskets with central apertures and housed in two 3D-printed millifluidic half-cells (Figure 6). When pressed together the clamping clips hold the cells together, applying sufficient pressure to the gaskets to maintain a fluidic seal between the chip surface and fluidic channels in the cell assembly.

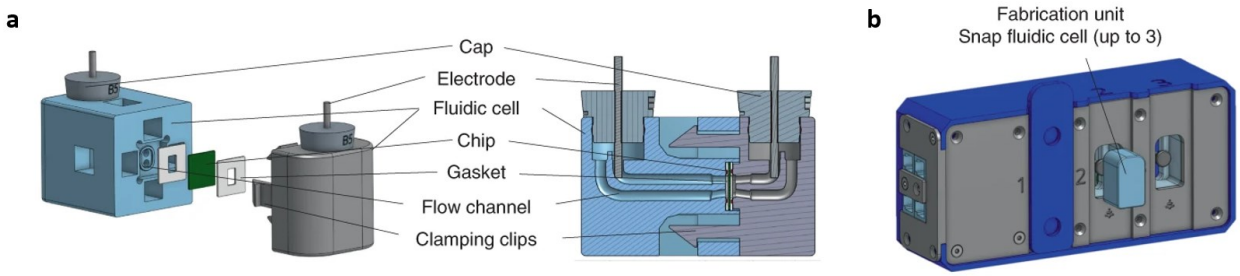

Figure 6: Schematic of millifluidic flowcell and fabrication unit. a) Labelled cross-sectional renders of both halves of a fluidic cell with a mounted Si/SiN_x chip. b) Render of a fluidic cell inserted into a nanopore fabrication unit, which can accommodate and fabricate multiple fluidic cell assemblies simultaneously. [image taken from *Solid-state nanopore fabrication by automated controlled breakdown* (Nat. Protocols) with permission]

Once sealed, electrolyte can be flushed through each millifluidic channel to create a conductive pathway on either side the SiN_x membrane. Ag/AgCl electrodes are inserted into this electrolyte at each end of the channels, and then capped to minimize evaporative losses to their surroundings. The fluidic cell assembly is placed into a nanopore fabrication unit, which is connected to a computer running a custom CBD program (CBDsoft).

This program instructs the fabrication unit to apply a gradually increasing potential through the flowcell assembly and monitors the leakage current across the SiN_x membrane (Figure 7). When a CBD event occurs, the sudden increase in current passes an offset threshold and signals to the software that a pore has formed, quickly halting the fabrication process before the nascent pore grows beyond a size suitable for sensing.

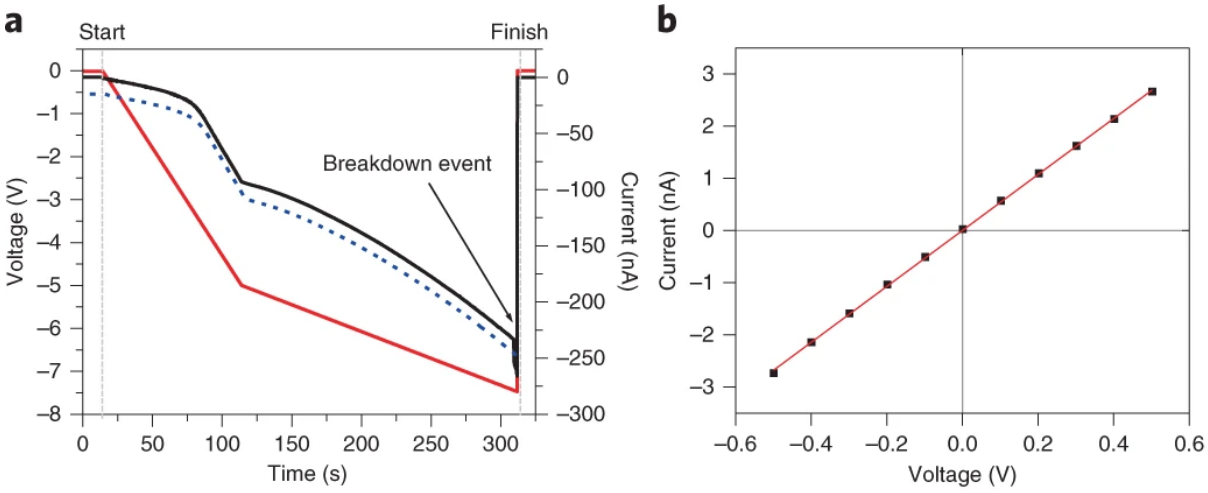

Figure 7: CBDsoft outputs during nanopore fabrication. a) Fabrication curve. The voltage (red, left axis) decreases over time, producing a gradual, proportional change in the leakage current (black, right axis) until a breakdown event leads to a sudden spike in current, crossing the offset threshold (blue dashed line) which halts fabrication. b) I-V curve. A series of potentials from -500 to 500 mV are applied. The resultant open-pore current is measured, and the pore size can then be determined from equation 2. [image taken from *Solid-state nanopore fabrication by automated controlled breakdown* (Nat. Protocols) with permission]

The pore is then characterized with a current-voltage (I-V) response curve to establish its size. The ohmic relationship between a series of applied potentials and the resultant current, when plotted and fitted, provides the nanopore's conductance. As nanopores formed by CBD have been determined to be approximately cylindrical in shape,[13] equation 1 can be used to obtain a first-order approximation of the pore diameter from this conductance value.

Pores fabricated by CBD are often initially 2-3 nm in diameter, smaller than many biomolecules of interest. After fabrication and size measurement, if below a pre-set size, the pore

is gradually enlarged through an alternating series of short, smaller (<4 V) voltage pulses. The nanopore is intermittently measured until it has achieved its desired size, whereupon it may be used for sensing.

1.7 Nanopore Data Analysis

The analysis of nanopore translocation signals involves the extraction of parameters from individual events to construct a statistically significant body of data. The process of identifying, extracting, and analyzing these data is presented in the following sections.

1.7.1. CUSUM+ event identification and fitting

The Cumulative Sums (CUSUM) algorithm has been adapted from a process control tool as a means of automating identification and fitting of translocations from nanopore sensing experiments.[14] This algorithm works by assuming that baseline, open-pore current is piecewise constant with Gaussian-distributed noise. Once identified, these abrupt changes in the mean baseline current are segmented and fitted into blockage levels by the algorithm. This thesis used a custom implementation of the CUSUM+ algorithm developed by Dr Kyle Briggs (https://github.com/shadowk29/CUSUM), which has since been adapted into an enhanced user interface (Figure 8).

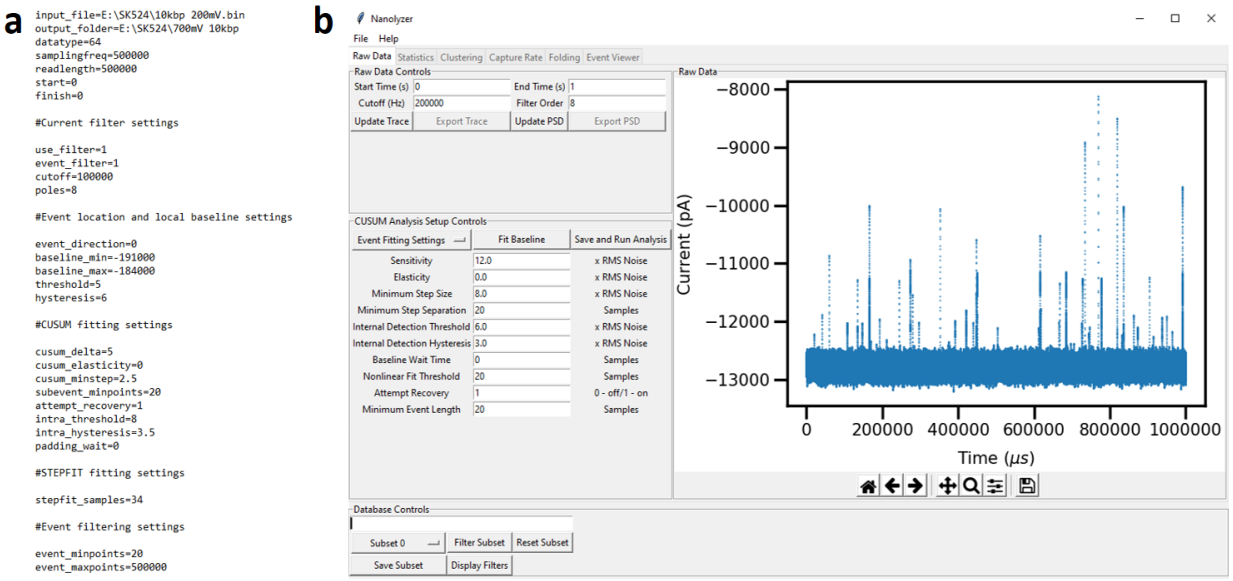

Figure 8: CUSUM software interface for event fitting an analysis. a) Initial inputs for CUSUM algorithm entered as a text document. b) Subsequent development of CUSUM inputs into a GUI incorporating visual feedback for enhanced analysis.

There are two sets of user inputs in the CUSUM event fitting process. The first set instructs the program on how to separate event signals from baseline noise. Passing the data first through a low-pass filter (LPF) minimizes higher-frequency noise contributions. Selecting the appropriate LPF cutoff frequency is important for maximizing the signal-to-noise ratio (SNR): shallow blockages can be hidden under excessive high-frequency noise, while a LPF cutoff frequency set too low will attenuate translocation signals.

In order to estimate the local baseline, the program requires the range of currents to be considered be bounded by the user. In this way, if the baseline current changes in the absence of a translocation event (for example, if the pore clogs), the program will recognize the data as unsuitable and disregard it.

The second set of inputs instructs CUSUM on how to 'fit,' or construct the time-averaged current sublevels internal to the translocation event data identified above. What follows is an overview of some of the more relevant of these parameters.

Detection Threshold establishes a blockage depth, based on the multiple of the open-pore RMS current, at which an event is detected. This threshold is required for distinguishing translocations from smaller, transient spikes in noise. A threshold set too high will cause CUSUM to ignore many otherwise valid translocations, while care must be taken not to set one so low as to count noise spikes as events.

For event fitting, the Sensitivity determines how closely CUSUM attempts to match its fit with the raw translocation current, in order to identify sublevel segments within an event. These sublevels, indicative of changes in structural elements passing through the pore, are of particular use when characterizing non-uniform molecules such as proteins or labelled DNA.

CUSUM assumes instantaneous step changes in current for fitted events. However, bandwidth limitations cause measured translocations to produce a more continuous change in current from fully open pore to fully molecule-occluded. The Minimum Step Separation fitting parameter instructs the algorithm how many samples make up this change so that the non-constant portions of the current are not overfitted in a 'staircase'-type pattern.

Finally, the investigator may be only interested in shorter or longer translocations, and so can instruct CUSUM to ignore fitting events that fall outside the Minimum and Maximum Event Length inputs.

1.7.2. Extracted Event Characteristics

Each fitted event has dozens of parameters. Among the most important (Figure 9) are:

- Number of Levels: The number of fitted segments in the event. By default, the program counts the baseline before and after the translocation as a level, so a simple single-file blockage has three levels as counted by CUSUM.
- Maximum Blockage: The largest difference between the local baseline and the fitted blockage levels within an event.

- Maximum Deviation: The highest amount of current blocked at any point during the translocation. This is always higher than the maximum blockage, as it is a single point that is sensitive to noise and not an average over the level.
- Dwell Time: The sum of durations of all the non-baseline fitted levels in the event.
- Residuals: The root mean square of the differences between all points of the current and the local fit value. This is useful as an indicator of CUSUM fit accuracy.
- Event time: The elapsed recording time at which the fitted event begins. This can be used to construct a simple time-averaged capture rate.

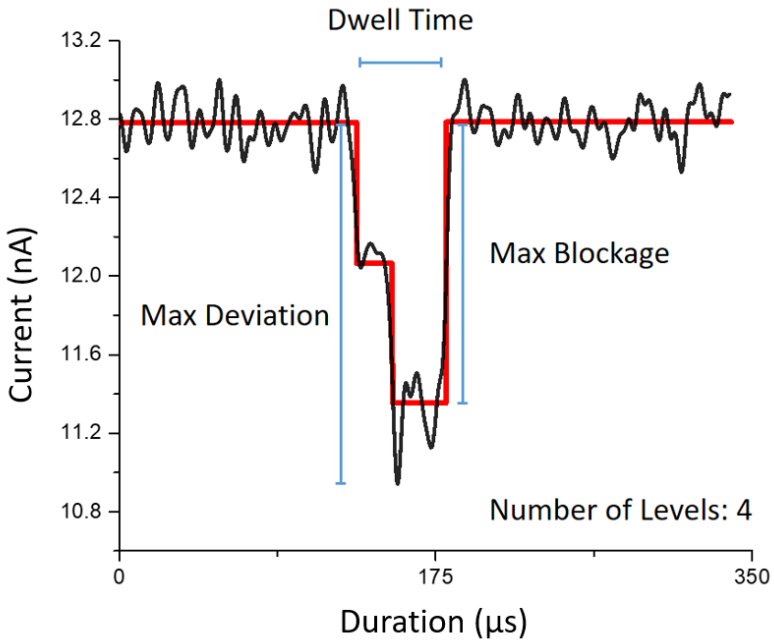

Figure 9: Current trace with CUSUM fit overlayed. Raw current data (black) is fitted by CUSUM into an event consisting of discrete levels of current (red). Extracted event characteristics such as Dwell Time, Maximum Deviation and Maximum Blockage are visually depicted. The pre- and post-translocation open-pore current levels are added to the two fitted blockage levels for a total of 4 fitted levels.

Once the CUSUM algorithm has finished identifying and fitting event parameters, the fitted metadata can be plotted (Figure 10) and filtered in a variety of ways to facilitate understanding of the underlying experiment.

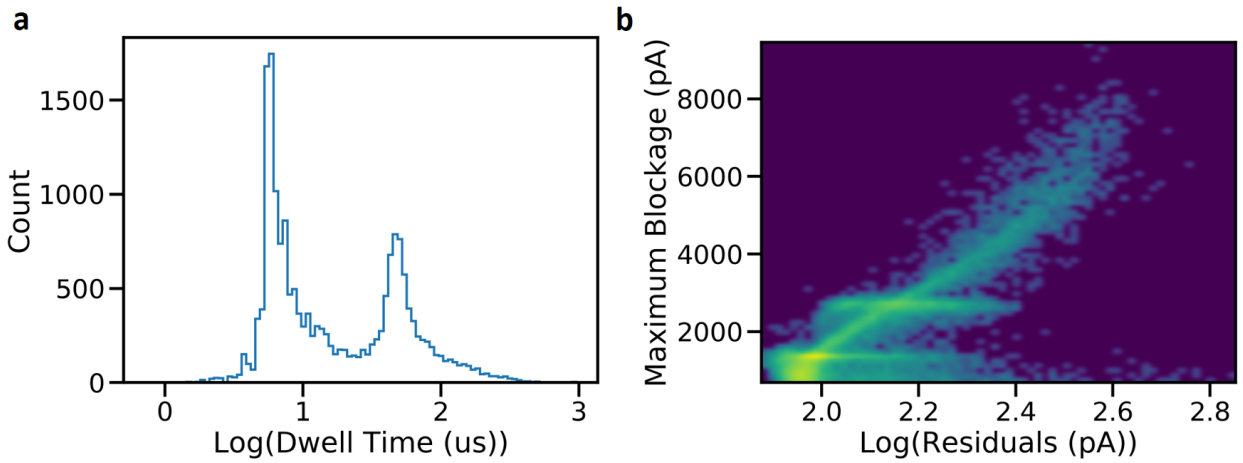

Figure 10: Plotted CUSUM-fitted event parameters. a) A histogram allows for the visualization of event dwell time distribution. Two peaks are indicative of different molecular lengths. b) A heat-map comparison of maximum blockage and residuals indicates whether the CUSUM algorithm has had difficulty fitting deeper/shallower events.

For example, to find information about molecule geometry a histogram of event dwell times can highlight differences in length – shorter molecules pass more quickly – and resolvable peaks can indicate the number of species present. A 2-dimensional histogram, with bin counts based on colour intensity, can be used to compare the frequency of paired parameters such as event residual with maximum blockage, to determine how the quality of fit follows the fitted depth.

1.8 Noise

Using sensitive equipment to detect pico- or nano-amp changes in ionic current brings with it challenges – chief among them noise. Noise is the measurement uncertainty of the sensing equipment. It has multiple sources – both mechanical and electrical – depending on the

frequency at which they dominate. For these experiments, noise is measured by its power spectral density (PSD) – the amount of power each constituent frequency brings to the open-pore current at a fixed potential. The total current PSD noise can be represented with the following polynomial[9]

$$S_I(f) = \frac{a_0}{f} + a_1 + a_2 f + a_3 f^2 \left(\frac{A^2}{Hz}\right) \tag{4}$$

where f is the frequency and a_0, a_1, a_2 and a_3 represent different noise contributors, which scale and vary based on different parameters. The aggregate noise of the PSD is represented by the root-mean-square (RMS) noise, or the square-root of the integral of the PSD

$$RMS = \sqrt{\int_1^{f_{high}} S_I(f) df} \ (A) \tag{5}$$

where f_{high} is the high-frequency cutoff of a filter or the bandwidth of the acquisition software.

For the work done for this thesis, the low-frequency noise had to be below a certain threshold (<250 pA_{rms} at a bandwith of 900 kHz) so as to have a well-behaved pore to capture and detect single molecules. Also, because the molecules measured are for the most part short dsDNA amplicons, high bandwidth is required to detect these fast (µs-scale) passing molecules. Therefore, to maintain a decent SNR of 7, high-frequency noise had to be minimized using SiO_2-coated chips, and the data digitally low-pass filtered at 500 kHz during analysis.

B: INFECTOUS DISEASE DIAGNOSIS

1.9 Molecular Basis of PCR

The response of an organism to its environment that allows it to survive or thrive is mediated by a complex biochemical interplay, which translates genetic information into appropriate cellular activity. One such activity, DNA replication, allows the organism to create copies of its genome for future descendants. The process varies between organisms but, for eukaryotes, follows the same general cellular process where each step is regulated by an enzyme: helical dsDNA is unwound and separated into individual single strands, which serve as templates for the new DNA to be synthesized, by an enzyme called helicase. Short fragments of complimentary RNA, known as primers, are assembled on regions of the template ssDNA strands by the enzyme primase. The DNA/RNA bound primer region is then recognized by a DNA polymerase. This enzyme travels along the template ssDNA and incorporates complimentary nucleotide bases into the nascent copy DNA strand, until both ssDNA templates become identical dsDNA copies.

One way to emulate DNA replication *in vitro* is with PCR, a well-established[15] molecular biology technique, which employs select elements from organismal DNA replication. The result is a large quantity of specific, identical dsDNA fragments from a small number of initial copies. These have many further applications in research and clinical work, including genome sequencing, DNA fingerprinting, and detection of bacteria and viruses.[16]

A PCR mixture (Figure 11) at its most fundamental [17] consists of the following:

- A template strand of DNA containing a sequence of interest to be amplified
- Short (16-30 nt) single-stranded primers that are complimentary to regions on the template DNA which bookend the sequence of interest
- A mix of all four main dNTPs (dATP, dGTP, dCTP and dTTP) that serve as the building blocks of the synthesis of new DNA

- A heat-stable DNA polymerase (such as *Taq* polymerase) that catalyzes the replication of DNA

Structurally, DNA polymers have an orientation. A single strand of DNA has two unique ends, termed 5' and 3', and DNA polymerases work in one direction only: from the 5' end (upstream) towards the 3' end (downstream). Furthermore, complimentary strands of ssDNA hybridize in opposite-facing directions. As a result, for conventional PCR two different primers must be employed: one upstream of the sequence of interest on each template ssDNA. These are typically referred to as forward and reverse primers.

With these components assembled in the same vessel along with a suitable buffer, the reaction is regulated by a thermocycler, which controls the PCR process through timed, temperature-controlled steps:

1. The temperature is raised to 95 °C, at which point the hydrogen bonds keeping dsDNA together are disrupted, causing it to **denature** into separate ssDNA
2. The temperature is then lowered so that primer **annealing** to complimentary ssDNA can occur. The exact temperature depends on the primer length and sequence but is typically in the 55-65 °C range. DNA polymerases require this primer-template hybridization structure for binding.
3. The temperature is raised to 72 °C, the optimal temperature at which polymerases associate with the primer-template structure and begin **extension** of the copy strand in the 5'→3' direction. This happens by sequential incorporation of free dNTPs which are complimentary with bases in the single-stranded parent.

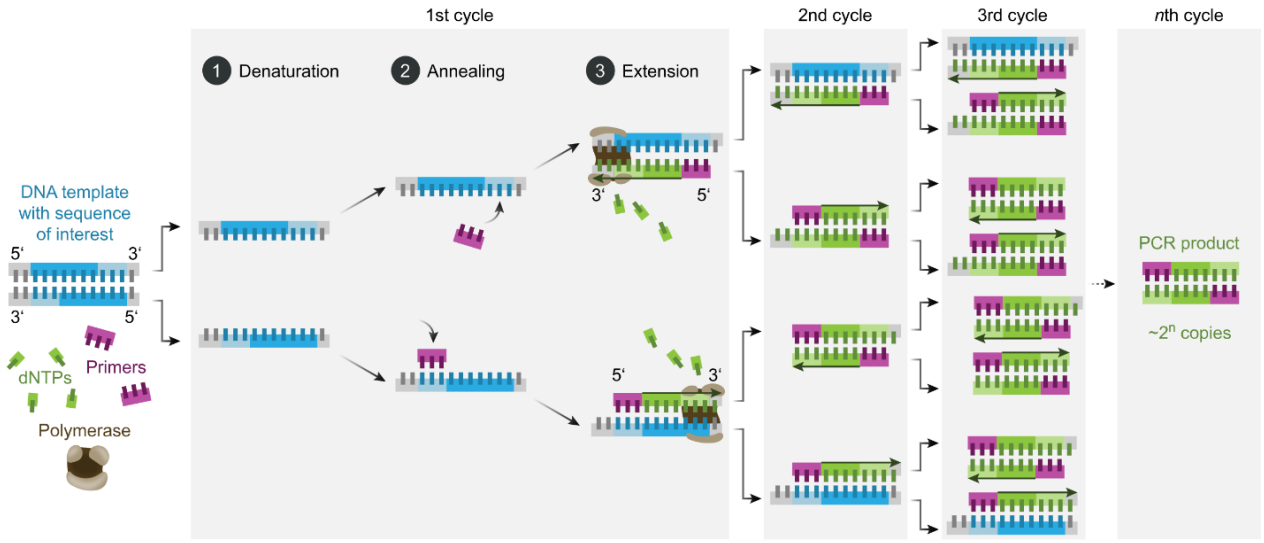

Figure 11: Main reagents and steps of PCR. The original sequence of interest is in both blues, with the primer binding sites in light blue. The forward and reverse primers each act on one strand of the parent DNA. Each three-step cycle doubles the number of amplicons. Temperature control of each cycle step is typically achieved with a pre-programmed thermocycler. [image taken from Enzoklop under CC BY-SA 4.0]

Each cycle of the preceding three steps (denaturation, annealing and extension) approximately doubles the number of dsDNA molecules containing the sequence of interest, also known as amplicons. This increase in amplicon copy-number means that cycling the temperature, typically 25-40 times, can generate more than 10^9 identical copies of the original sequence in <1 hour (note that the number of copies generated dependent on the starting concentrations of template, primer, dNTPs and other operating conditions).

Though PCR theoretically doubles the amount of DNA with every cycle, the true amplification efficiency of the reaction can vary based on several factors.[18] Primer pair design is an important consideration; primers with more than a single site-specific sequence will bind to different regions of the template DNA and produce off-target amplification, while self-complimentary primers will dimerize and can present a competing amplification target for DNA polymerase to act upon. Amplicon production is also highly sensitive to Mg^{2+} ion concentration, which complexes with dNTPs during replication and enhances DNA polymerase activity. Various other "enhancer" compounds have also been found to speed specificity or efficiency, including

detergents, glycerol, and Dimethyl sulphoxide (DMSO). For PCR performed on whole-cell extracts compounds such as BSA, a blood protein, and gp32[19], a single-stranded binding protein, have been found to enhance reaction rates by removing endogenous inhibitors.

1.10 qPCR

The amplicon products of a conventional PCR reaction are only analyzed once all the reaction cycles have completed, as they often require separation and staining techniques. This is known as end-point analysis,[20] and is of limited use for determining quantitative results such as initial target copy number. Real-time PCR, also known as qPCR, circumvents this limitation by indirectly measuring amplicon production following each reaction cycle. This is accomplished with fluorescent-labeled primers or dyes that bind to double-stranded DNA. Instrumentation consists of a conventional PCR thermocycler equipped with optical sensors to measure the excitation emission of the reaction mixture after each PCR cycle (Figure 12).

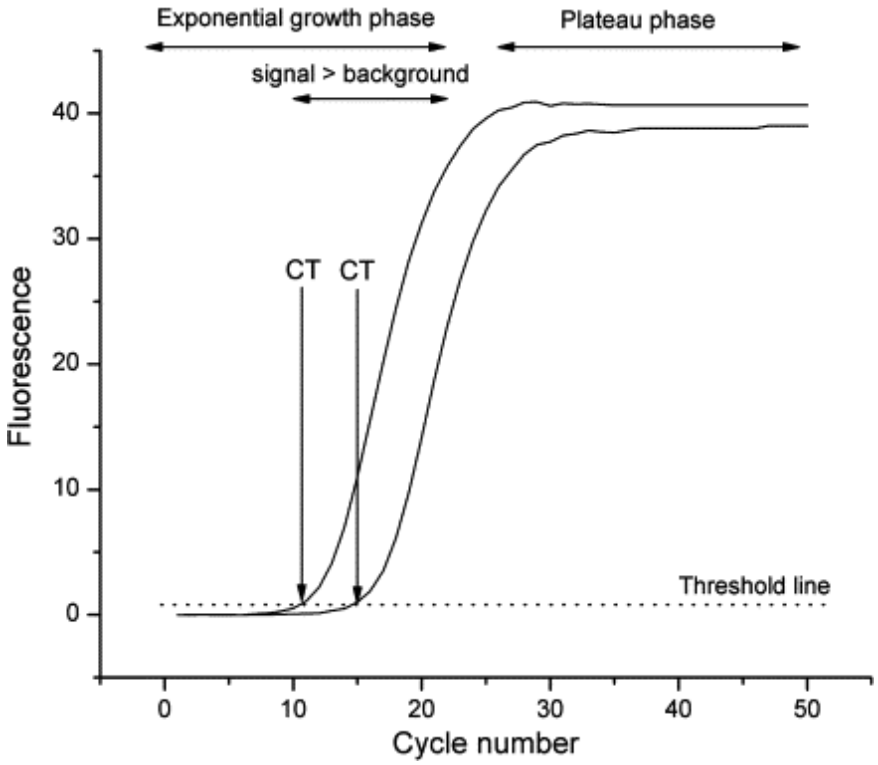

Figure 12: qPCR fluorescence curves. Two curves highlighting different C_q numbers (listed as C_t in this reproduced image) at 10 and 15 cycles, at which point the signal fluorescence exceeds background levels. The C_q of a series of standard dilutions can be used to accurately estimate initial copy number of the sample. [image taken from The Real-Time Polymerase Chain Reaction (Mol. Aspects of Medicine) with permission][21]

The fluorescence readings of a real-time PCR lag the cycle count as the fluorescence is initially below detectible levels by the optical instrumentation. The number of cycles required to cross this threshold, referred to as the quantification cycle, or C_q (occasionally C_t),[22] is proportionally related to the starting number of template DNA molecules. At later cycles, the reaction moves from exponential amplicon growth to a plateau as reagents or indicator molecules - primers, dNTPs or fluorescent markers - are depleted. A quantitative initial copy-number can be obtained by comparing the C_q value of the sample with a calibration curve constructed of C_q values from serial dilutions of a known copy-number standard with the same amplification conditions.

1.11 Clinical Diagnostic techniques

1.11.1 Sensitivity and Specificity

In the field of infectious disease diagnosis, identifying the cause of illness is essential to providing the appropriate treatment. For example, a patient being prescribed an antibiotic treatment will only respond positively if suffering from a bacterial infection, and improper prescription can contribute to deleterious public health outcomes.[23] Conventional diagnosis by physicians relates the presence of an infectious pathogen with a patient's symptoms. However, symptoms are often nonspecific: a fever or a sore throat may bring a patient to seek treatment, but each has many different causes. Successful treatment depends on the proper identification of the pathogenic origin of an infectious disease, so health-care providers often turn to clinical approaches.

Techniques for investigating infections are often defined by their sensitivity and specificity. Sensitivity refers to a test's ability to identify a particular disease, while specificity is the test's ability to rule out other possible diseases. As frequency-based measures, sensitivity is defined as

$$Sensitivity = \frac{\#\ tested\ positive}{\#\ true\ positive} \tag{6}$$

while specificity is defined as

$$Specificity = \frac{\#\ tested\ negative}{\#\ true\ negative} \tag{7}$$

A good clinical test is therefore both sensitive (avoids false negatives), and specific (will avoid false positives) in its identification. As a highly specific molecular technique, and one sensitive to low copy number, PCR and qPCR are useful clinical tools for the diagnosis of infectious disease.[24] Furthermore, with the ability to determine initial template DNA levels, qPCR techniques have found success when used to determine infectious viral loads for appropriate treatment levels.[25]

1.11.2 Light Microscopy

One of the oldest forms of diagnosis, microscopy consists of visual inspection of prepared patient samples under a light microscope.[26] The clinician searches for tissue irregularities or the presence of pathogenic organisms. While comparatively simple with respect to equipment requirements, there are two major limitations to light microscopy. The first is the diffraction limit: a light microscope cannot resolve objects smaller than 0.2 µm, the upper end of the viral size scale. HIV, Ebola, Zika and other smaller viruses go undetected without the use of supplementary techniques and equipment.[27] Secondly, quality assurance varies,[28] as the detection of infectious diseases is largely dependent on the expertise of the technician. Sample selection, fixation and staining are all delicate, multi-choice processes whose outcomes vary depending on the pathogen under investigation as well as the technician's skill level. Furthermore, even with appropriate slide preparation, sensitivity is hampered by visual identification at low-pathogen numbers.

1.11.3 Cell Culture

Culture techniques are seen as a gold standard in modern infectious disease diagnosis. A collected sample is grown in a favourable artificial environment for isolation and identification of pathogenic organisms. This technique is inexpensive and unspecific: unlike slide preparation for microscopy, the examining clinician does not need *a priori* knowledge, as a universally favourable growth medium will theoretically allow most pathogens to proliferate. In practice the largest

disadvantages stem from the fact that this technique relies upon living processes. First, not every organism responds favourably to an artificial culturing environment. The technique finds most success with bacterial and fungal species, but struggles with viral targets because they require appropriate cell lines as hosts for replication.[29] Second, the diagnostic speed is reliant on cell growth, making the organism's multiplication-time the primary time-limiting process. While awaiting culture diagnostic results, which can take days, many medical professionals opt to begin treating the likeliest cause of an infection. Often this entails use of antibiotics, the overuse of which contributing to the development of resistant strains such as *C. difficile,* as well as other adverse effects. [30]

1.11.4 Serology

Serological techniques employ labelled detection antibodies, either directly against microbial antigens or indirectly against immune-response antibodies. Typically used against pathogens that are difficult to culture or identify with other techniques, serological assays also lend themselves well to point-of-care tests as they are easy to perform and can produce results in less than 15 minutes.

One common example of a serological test is the lateral flow assay (LFA), in which a sample is pulled along a nitrocellulose strip via capillary action (Figure 13). During the capillary flow the sample passes first through a conjugation pad, which is loaded with detection antibodies conjugated to indicator molecules. In the presence of targets in the migrating sample, the conjugated particles bind and are pulled along as the sample reaches a test line. This band contains capture antibodies fixed to the membrane which are specific to the target molecules, pulling them out of the passing sample. The visual appearance of the test band on the strip thus confirms the presence of target molecules in the sample.[31] A control line, containing fixed capture antibodies specific to the detection antibodies themselves, is located downstream and helps indicate if proper flow has occurred.

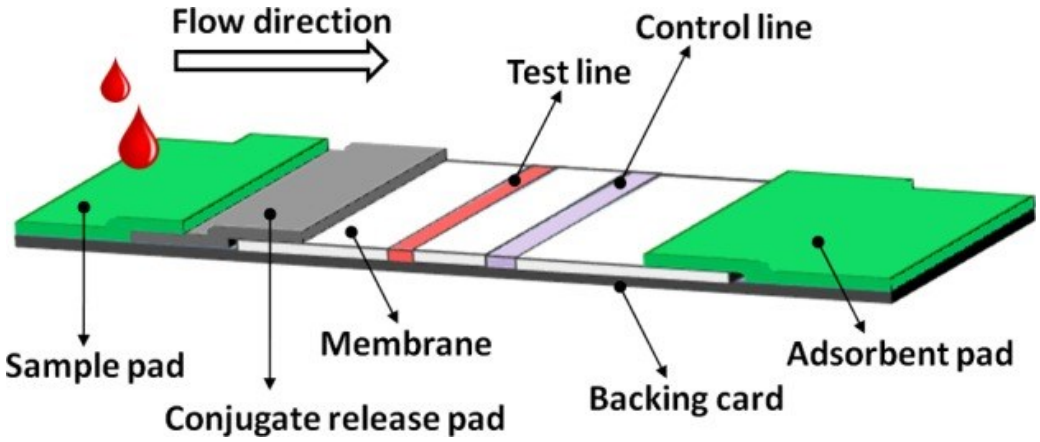

Figure 13: Lateral flow test strip. Schematic representation of a typical test strip. The capillary pull of the adsorbent pad causes the loaded sample to pass through each element in succession. Antibodies in the test line and control line are bound to the membrane, through which the sample flows. The backing card is typically made of an inert material.[32] [image taken from *Lateral Flow Assays* (Essays Biochem) with permission]

Serological tests in general tend to suffer from low specificity: antibodies may bind to nonunique binding sites, and similarities between target antigens and off-target molecules can produce false positives. Furthermore, serological tests which measure antibody production can have poor sensitivity at varying stages of an infection (Figure 14), or to infections which do not trigger a strong immune response, such as in immunocompromised patients.[26].

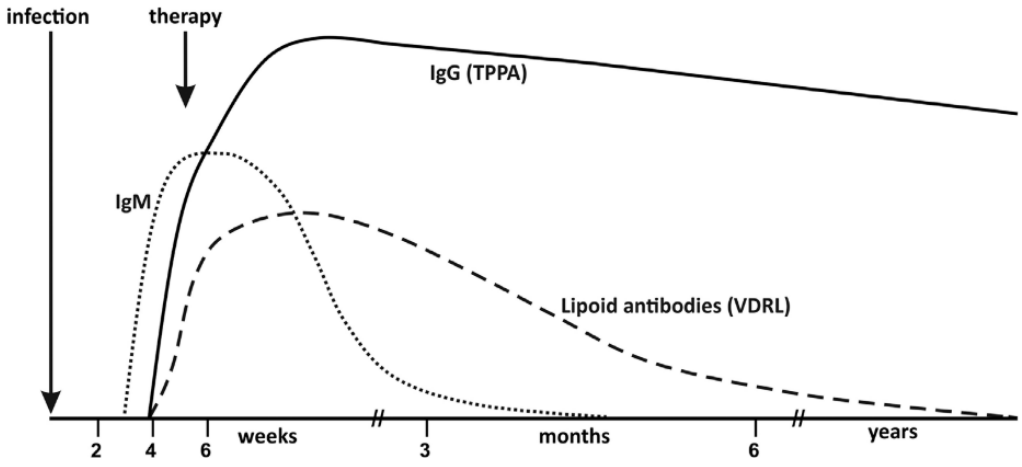

Figure 14: Serology of Treponema Pallidum (Syphilis). Different immune response antibody levels (IgM, IgG and Lipoid antibodies) vary depending on the infection stage and degree of treatment, impacting serology diagnostic sensitivity. [image taken from *On detection thresholds–a review on diagnostic approaches in the infectious disease laboratory and the interpretation of their results* (Acta Topica) with permission]

1.12 Other Nucleic Acid Techniques

Other PCR-derived amplification techniques have been developed in recent years. Many are isothermal, removing the need for a lengthy thermocycler-controlled process and thus produce DNA amplicons faster. Techniques such as loop-mediated isothermal amplification (LAMP), nicking-enzyme amplification reaction (NEAR), and recombinant polymerase assay (RPA) are well suited for rapid point-of-care diagnostics. Some of these techniques have even found their way into marketable products, such as in Abbott's ID NOW or Oxford Nanopore's GridION/lamPORE assay, demonstrating their ease-of-use and potential.

1.12.1 Loop-mediated Isothermal Amplification

The LAMP method, first discovered in 2000, produces nucleic acid products of varying size and structure, using a polymerase enzyme with a high strand-displacing activity at a constant 60 – 65 °C in place of heat-based cycling. While the amplicons are not exact copies of the template DNA the main advantage of the LAMP process, beyond isothermal operation, lies in their stem-loop shape. The large, bulky conformations adopted by LAMP amplicons makes them easier to detect by immunoassay antibodies and fluorescence compared to standard dsDNA[33].

LAMP relies on the careful identification of four or six primer sites on the template DNA (Figure 15). These sites are selected such that, upon initial replication, the single-stranded copy is displaced from the template and folds on itself as a self-hybridizing loop structure. The dumbbell shape presents a single-stranded loop at each end, which serve as initiation sites for further rounds of primer annealing, DNA synthesis/displacement and loop formation. The process continues indefinitely, with amplicons consisting of increasingly numerous concatemers, until the reaction is halted by reducing the temperature.

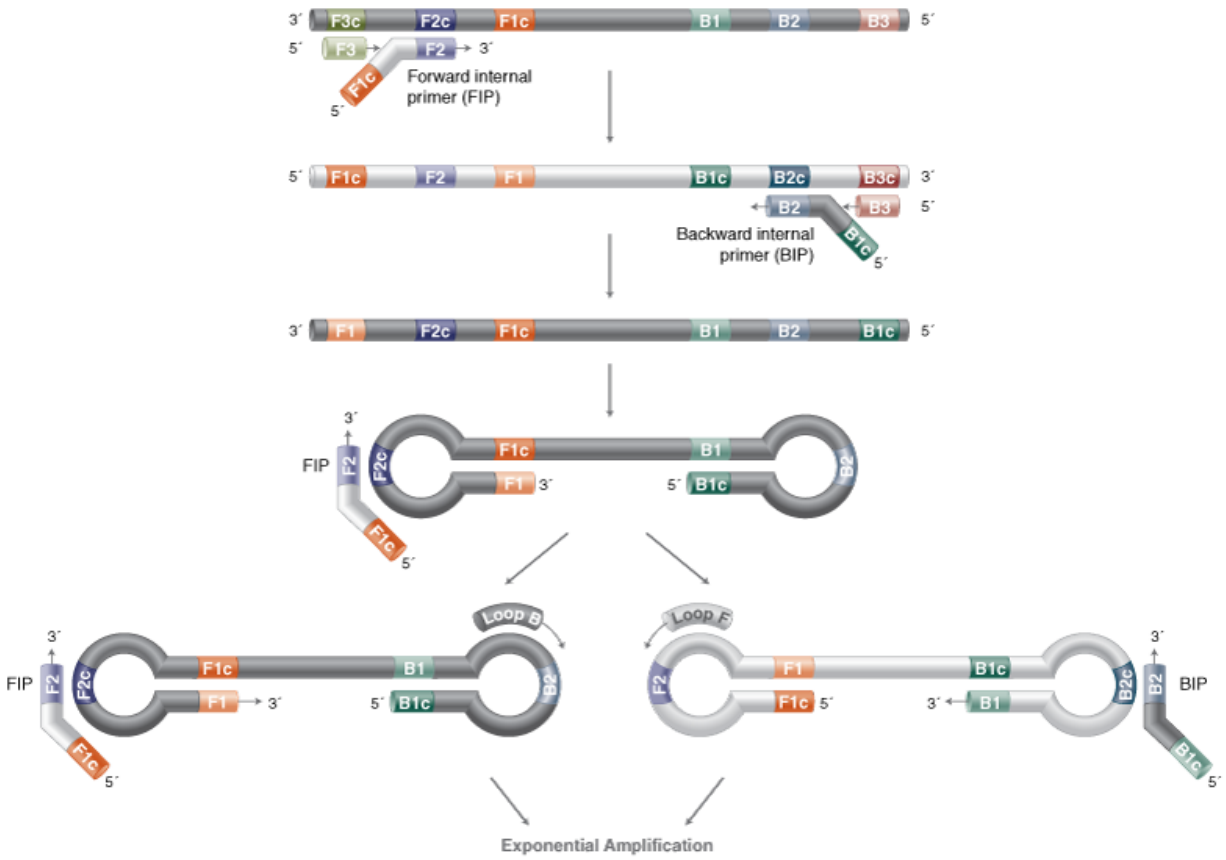

Figure 15: Initial LAMP cycle. Carefully designed 'internal' primers introduce a region of self-complementarity into the amplicon, causing it to form a single-stranded loop structure. A primer site in each loop structure then be used for further extension and loop formation. The number of loops and therefore overall amplification rate increases exponentially. [figure taken from New England Biolabs product information]

While LAMP allows investigators to rapidly produce a large quantity of high signal-producing DNA, downstream applications for amplicons are limited. Furthermore, varying concatemer lengths and structures make true quantitative measurements difficult to assess.

1.12.2 Recombinase Polymerase Amplification

Developed in 2006,[34] RPA, like LAMP, is an isothermal amplification technique that uses a strand-displacing polymerase in place of heat-cycling. Unlike LAMP however, RPA amplicons are faithful dsDNA copies of the template DNA. A further strength is that RPA reactions can occur at low temperatures: optimally at 37 – 42 °C, but amplicon production has been observed as low as room temperature. This reduces heating requirements, and thus potential cost and complexity, considerably – researchers have conducted RPA simply by holding the reaction flask in hand.

RPA differs from other nucleic acid amplification schemes by use of both a recombinase, capable of complexing with an oligonucleotide primer, as well as single-stranded binding proteins, which cooperatively bind and stabilize ssDNA. During the reaction (Figure 16), recombinase-primer complexes facilitate strand exchange at double-stranded primer sites. The displaced ssDNA is kept from rehybridizing and ejecting the primer by single-stranded binding proteins, while the polymerase binds the primer site and synthesizes/displaces along the template strand. If both a forward and reverse primer are present, the reaction produces amplicons at a geometrically increasing rate.

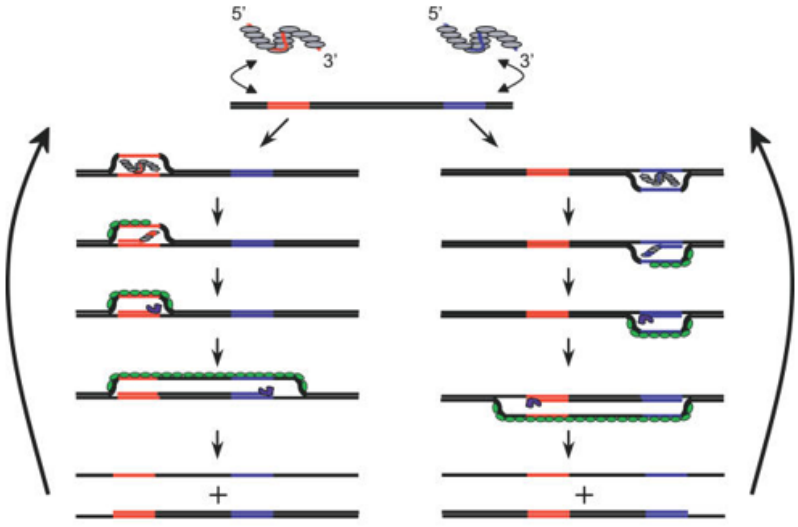

Figure 16: Single cycle of RPA. Recombinase-primer complexes are in grey and red/blue. Forward (red) and reverse (blue) primer sites are highlighted on two dsDNA templates (black). Single-stranded binding proteins (green) prevent strand ejection of primer and nascent copy. [Image adapted from *DNA Detection Using Recombination Proteins* (PLoS Biology) under CC BY-SA 4.0]

1.13 Nanopore-based approaches to DNA identification

As biosensors, nanopores offer the utmost sensitivity in analytical sensing by virtue of single-molecule detection. In the context of diagnostics, the only true constraint on sensitivity is the rate of mass transfer from sample to pore. With a typical capture rate of 1 Hz/nM for standard operating conditions, nanopores generally are able to detect hundreds to thousands of molecules in the nano- or high picomolar concentration range in a reasonable amount of time (minutes). Nucleic acid biomarkers in clinical samples can range in concentration from the high nanomolar to the attomolar range, where just a handful of copies exist for microlitres of sample. Depending on target concentration of interest, amplification may be needed to access low copy-number samples rapidly.

Currently however, solid-state nanopores are still unable to sequence DNA, as opposed to their biological counterparts[35] and thus cannot yet independently discriminate between different target sequences. To obtain specificity, researchers often make use of labelled hybridization assays in which sequences are recognized by complimentary nucleic acids ligated to indicator molecules. These labels produce unique ionic blockage profiles when translocating through a pore. In this way the specificity of a nanopore sensor becomes determined by the specificity of the amplification reaction and not the ability of the pore to distinguish unique sequences.

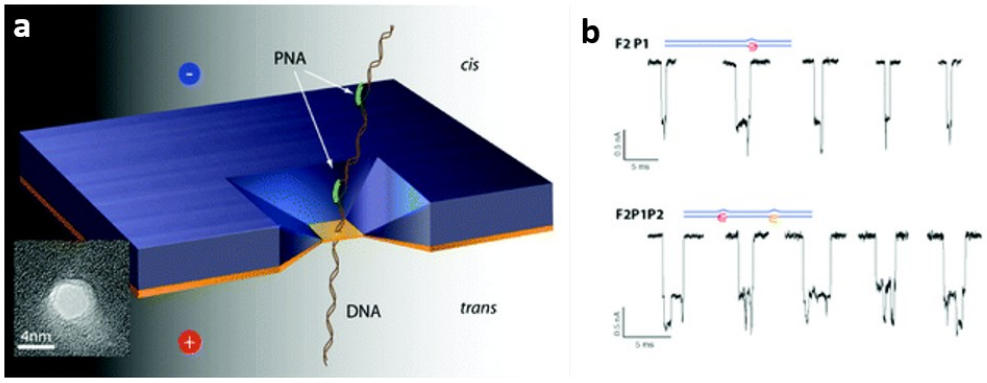

Figure 17: PNA-tagged DNA translocating through a nanopore. a) Schematic of a PNA-DNA complex during translocation, and b) Current blockage profiles produced by single- and double-tagged PNA-DNA strands. The two profiles can be distinguished by the number of sharp, deep blockages which correspond to PNA hybridization sites. [image adapted from *Nanopore Based Sequence Specific Detection of Duplex DNA for Genomic Profiling* (Nano Letters) with permission]

The investigation of different labeling schemes began with the Meller group exploring the use of peptide nucleic acids (PNAs) as target sequence detection markers. These were used to competitively bind complementary portions of double-stranded genomic DNA fragments at multiple sites, as PNA-DNA interactions have a higher binding energy than solely DNA-DNA ones. These PNA-tagged portions of dsDNA molecules were detected while translocating through a nanopore (Figure 17) as they produced unique, deep, sharp current blockages compared with dsDNA alone, demonstrating the potential for nanopore-based DNA barcoding.[36]

This work was supplemented with efforts from the Dunbar group, in collaboration with the Tabard-Cossa lab, in aiding detection of PNA-tagged DNA. They employed PEG-linked PNA probes complimentary to the CFTRΔF508 cystic-fibrosis mutation codon, using a range of nanopore sizes and geometries. Upon translocation, 324bp dsDNA fragments containing this mutation were identified by the deeper, longer blockages produced by the bulk-adding 10 kDa PEG-PNA complexes when compared to DNA/PNA alone, PEG alone, and wild-type fragment sequences in which no complexes formed.[37]

Following a similar target sequence probe approach, Hall et al. used a nanopore-based assay to directly detect the presence of a conserved sequence of HIV-1B's RNA genome at nanomolar concentrations.[38] This assay consisted of short (<250bp) biotinylated and monovalent streptavidin-linked ssDNA probes (Figure 18).

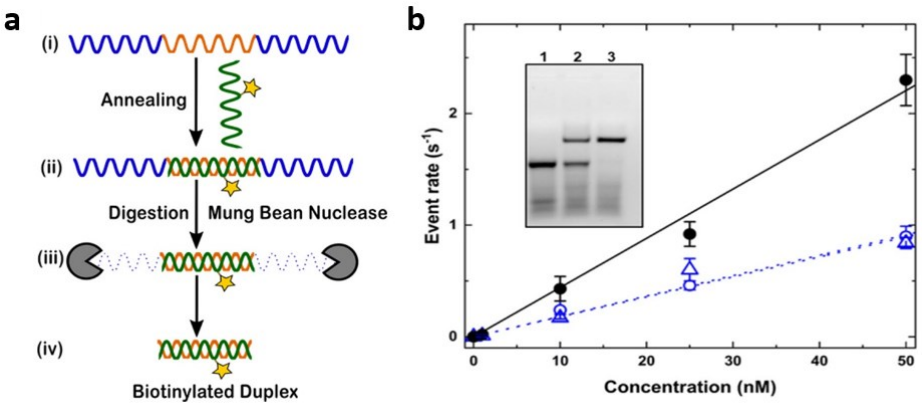

Figure 18: Monovalent streptavidin-linked ssDNA probe assay. a) Schematic of the hybridization assay for nanopore sensing. Complimentary portions of viral RNA and the biotinylated ssDNA probe hybridize, with all remaining single-stranded fragments digested by Mung Bean Nuclease, leaving behind a biotinylated duplex structure. b) Capture rate of the monovalent streptavidin-linked duplexes at varying probe concentrations, using two different probes (black circles) compared with only one (blue circle or triangle) [image taken from *Direct Detection of Conserved Viral Sequences and Other Nucleic Acid Motifs with Solid-State Nanopores* (ACS Nano) with permission]

While on their own the probes translocated through the pore too quickly to be detected by the low bandwidth of their recording system (<100kHz), in the presence of complimentary portions of the single-stranded RNA viral sequence, the probes hybridized and slowed nanopore translocation speeds to within the realm of their instrument measurement. Furthermore, employing two probes complimentary to different regions of the viral genome roughly doubled the event signal, suggesting multi-probe signal amplification is a practical alternative to PCR for low-concentration target detection.

Dr. Eric Beamish, during his doctoral studies in the Godin lab at the University of Ottawa, explored labeling schemes using the principles of DNA origami, in which conventionally linear strands fold into different typologies based on regions of discontinuous complementarity.[39] The structure consisted of a linear single strand of M13mp18 DNA hybridized with 186 'tile' complimentary single-stranded sequences of 39 bases in length (Figure 19). Two of these tiles, each 1/3 of the way along either end of the M13mp18 strand, had unhybridized overhangs complimentary to a portion of the Zika NS5 gene. In the presence of this gene fragment, the two

overhangs pulled the M13mp18 into a looped 'nanoswitch', which could be readily resolved from its linear counterpart in a nanopore by increased current blockage. The fraction of looped translocations was also demonstrated to be sensitive to Zika gene levels.

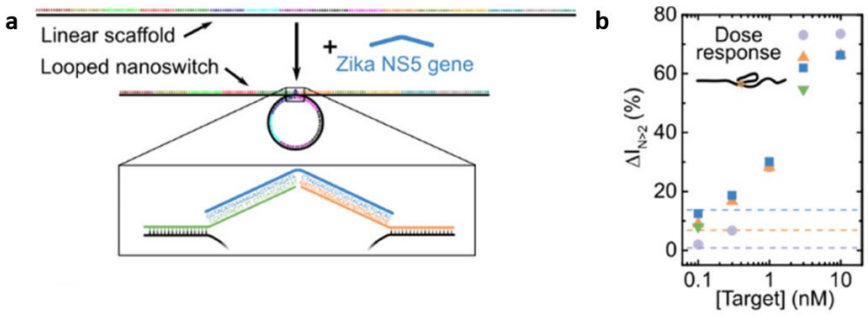

Figure 19: Zika Nanoswitch Sensing. a) Depiction of the linear scaffold (top), which hybridizes with the Zika NS5 gene to form a looped nanoswitch structural conformation (bottom). b) Results of nanopore sensing of nanoswitch structures after incubation with Zika NS5 target. Looped nanoswitch structures, with a greater blockage state than the linear scaffold, appear in greater fractions at higher target concentrations, demonstrating a degree of dose response. [image adapted from *Programmable DNA Nanoswitch Sensing with Solid-State Nanopores* (ACS Sensors) with permission]

A similar assay was developed for the detection of multiple microRNA (miRNA) diagnostic biomarkers for human cancers. This approach used several unique DNA nanostructures: 98bp probe sets with short ssDNA binding sites.[40] These lessened the complexity of the analysis, as well as the risk of false positives due to the folded passages of long DNA nanoswitches. The probe set overhangs, complementary to different miRNA targets, linked the probes into assembled complexes which translocated at deeper blockage levels than the probes alone. Moreover, the specificity of this assay was investigated by introducing probe sets for other miRNA targets with unique, resolvable translocation characteristics during multiplexed sensing experiments (Figure 20).

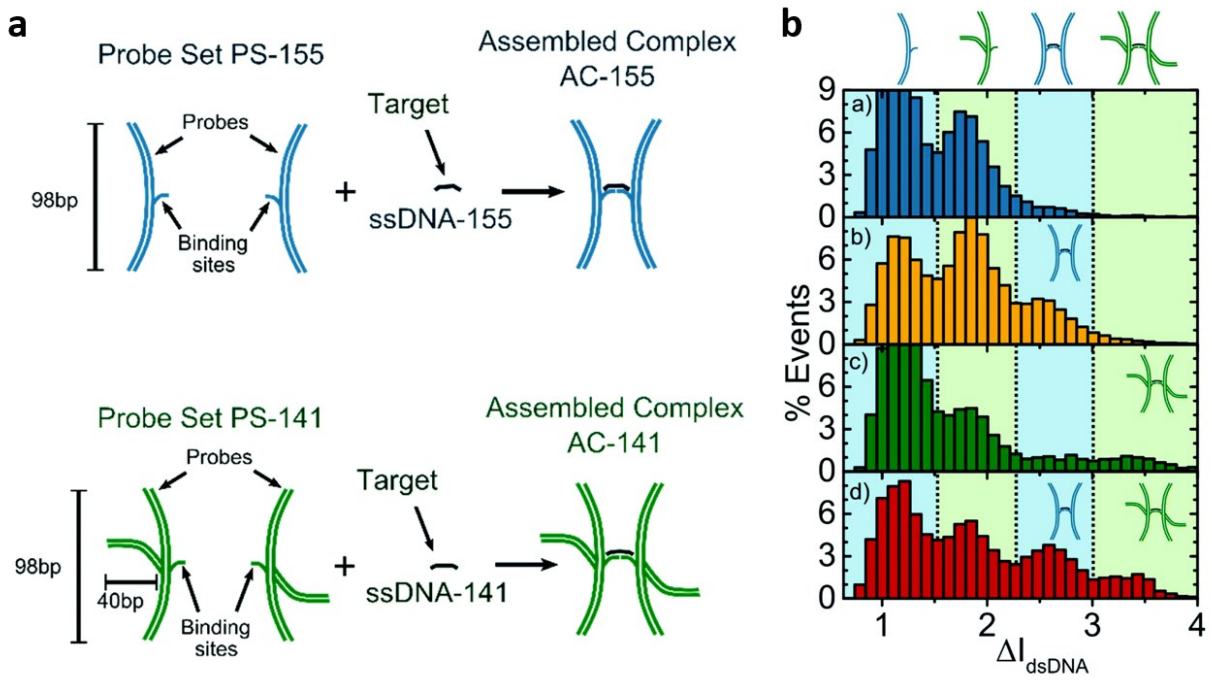

Figure 20: DNA Probe Set sensing. a) Depiction of two different probe sets, with different DNA structures, which formed assembled complexes in the presence of their respective biomarker targets. b) Blockage depth histograms of simultaneous nanopore sensing with both probe sets and no ssDNA targets (blue), only ssDNA-155 (yellow), only ssDNA-141 (green) and both targets (red). The probe sets and assembled complexes can be resolved by different blockage depths. [image adapted from *Digital counting of nucleic acid targets using solid-state nanopores* (Nanoscale) with permission]

Nanopores have also been used to detect polymorphic insertion or deletion mutations, which can determine the virulence or drug-resistance of two different bacterial organisms. The Meller group were able to detect large insertion or deletions directly via differences in translocation times through a pore. Single nucleotide variations or short insertions and deletions could be detected after mutation site restriction-enzyme digestion which produced two smaller, nanopore-resolvable fragments if the mutation was present.[41]

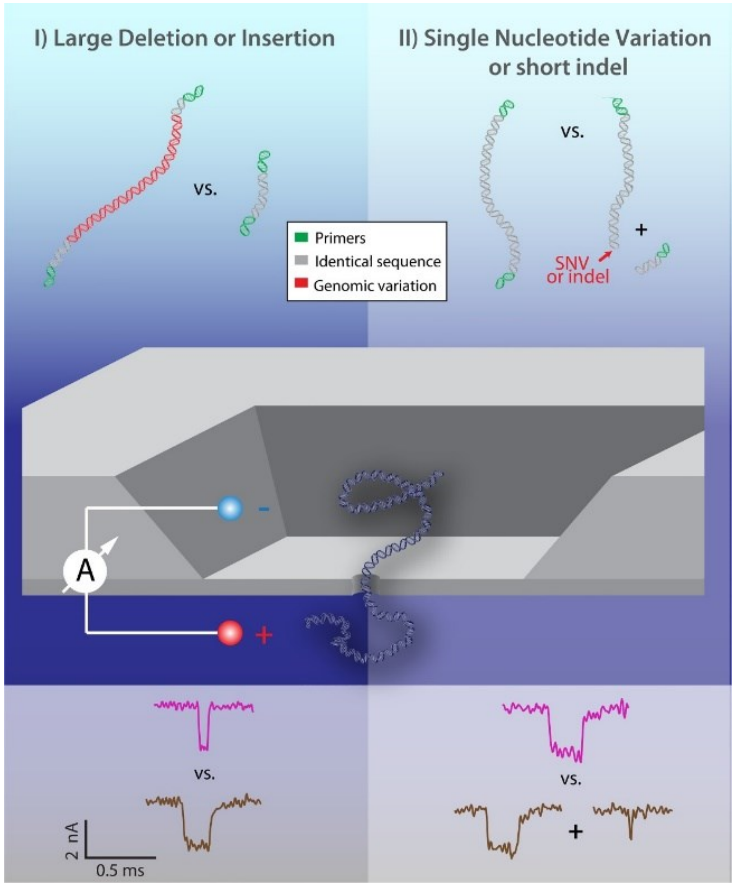

Figure 21: Nanopore detection of polymorphic mutations. Depiction of i) large (left) and ii) short (right) insertion/deletion mutations and their respective signal profiles after cleavage with mutation site-recognition enzyme. [image taken *from Genomic pathogen typing using solid-state nanopores* (PLoS One) with permission]

Research into nanopore-based point-of-care devices were successful employing disposable nanopore test strips for the detection and quantification of target proteins. Dunbar et al. incubated a saliva swab with an engineered PNA-peptide detection molecule containing a portion of the HIV-1 envelope glycoprotein. When this DNA/PNA-peptide complex was exposed to HIV antibodies (Figure 22), it produced deeper and longer blockage profiles upon nanopore translocation, easily resolvable from unbound DNA/PNA or uncomplexed molecules.[42]

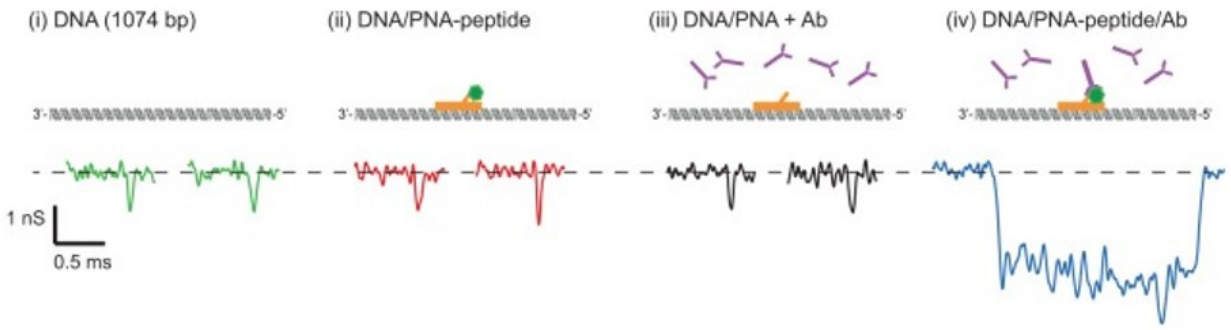

Figure 22: PNA-peptide signal enhancement. Schematic of labelled DNA and their representative blockage profiles. The PNA/DNA-peptide/Ab complex produced deep, long blockages easily resolvible over the DNA/PNA-peptide or the DNA/PNA with unbound Ab. [figure taken from *Nanopore-based target sequence detection* (PLoS One) permission]

Most recently, the Meller group adapted reverse transcription (RT) qPCR paired with nanopore sensing (coined RT-qNP) to quantify relative gene-specific cDNA expression levels in a label-free manner.[43] RNA was first extracted from cells, and sequences of interest were converted to double-stranded cDNA using RT. Samples would then undergo a limited amount (1-8 PCR cycles) of amplification. Following this, all remaining RNA and proteins were digested with sequential addition of first RNase I, then Proteinase K, and the sample was sensed directly by a nanopore for the capture rate of the cDNA. This technique was used to analyze the expression levels of metastasis-associated mRNA targets MACC1 (360bp) and S100A4 (123bp) in human colorectal cancer (CRC) cells, relative to a reference gene. Furthermore, RT-qNP was found to be sensitive to amplicon production from 50ng of whole-cell CRC RNA after only a handful of amplification cycles; a much lower C_q value when compared to RT-qPCR (Figure 23).

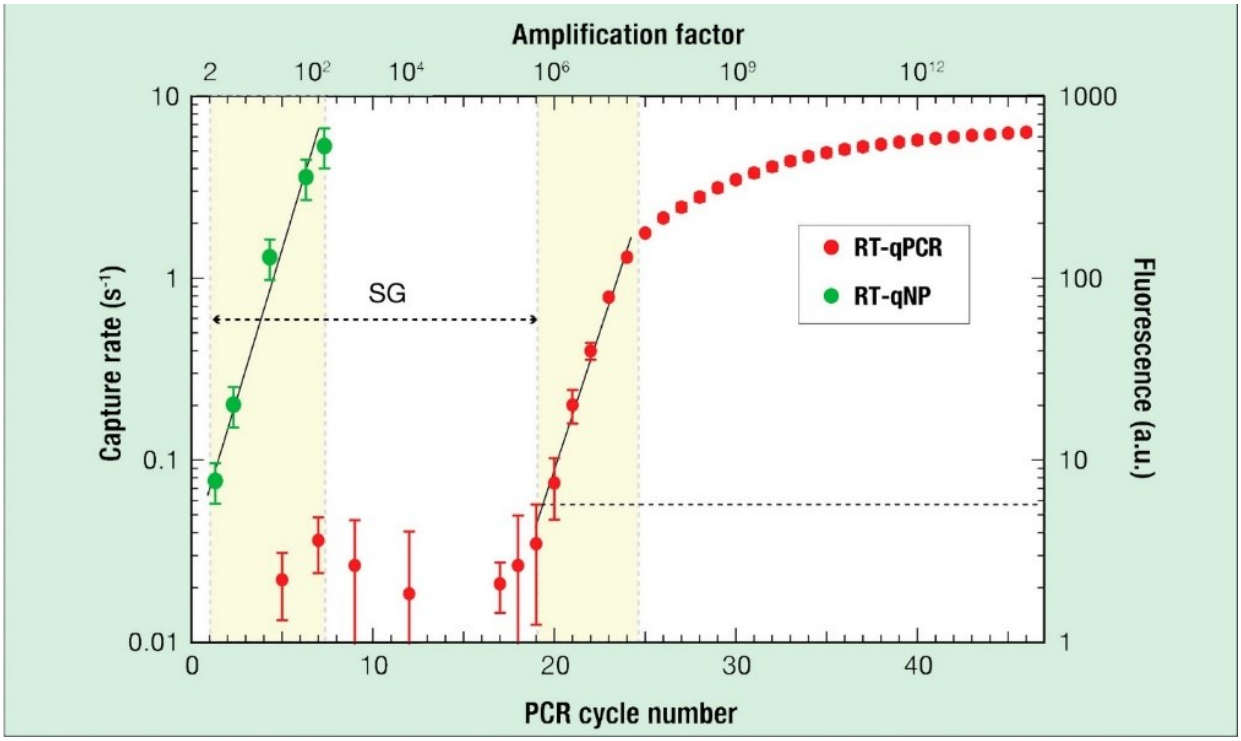

Figure 23: Comparison of RT-qNP with RT-qPCR at selected PCR cycle counts. The RT-qNP results were recorded by capture rate (left axis) while the RT-qPCR results by total fluorescence (right axis). The dashed line represents the limit of detection for the qPCR instrumentation, and SG refers to the sensitivity gap between the two methods' C_q values. [image taken from *Quantification of mRNA Expression Using Single-Molecule Nanopore Sensing* (PLoS One) with permission]

These earlier studies for the detection of nucleic acid biomarkers using solid-state nanopores helped showcase some of the potential of the nanopore method for infectious disease diagnostics. While prior work showcases some interesting schemes to specifically recognize DNA sequences related to a pathogen, they often worked from synthetic or purified samples and often did not investigate those obtained from clinical contexts. In the next chapter I will present the development of a solid-state nanopore-based method for the detection of amplicons from a PCR amplification reaction and demonstrate the use of these nanosensors for the screening of Group A streptococcus (GAS) in throat swabs from children with pharyngitis.

Chapter 2: Screening for Group A Streptococcal disease via Solid-State Nanopore Detection of PCR Amplicons

This chapter is adapted from a manuscript published article:

Simon King, Kyle Briggs, Robert Slinger, and Vincent Tabard-Cossa, "*Screening for Group A Streptococcal Disease via Solid-State Nanopore Detection of PCR Amplicons*" ACS Sensors 7, 1, 207–214 (2022)

https://doi.org/10.1021/acssensors.1c01972

The study was conceived by Dr. Vincent Tabard-Cossa of the University of Ottawa and Dr Robert Slinger of the Children's Hospital of Eastern Ontario. Dr. Slinger acquired the clinical samples and performed the DNA extraction. I performed all the experimental work, analyzed the resulting data, and wrote the first draft of the manuscript. Drs. Kyle Briggs and Tabard-Cossa helped with interpretation and editing of the manuscript.

2.1 Materials and Methods

Sample Collection and Preparation

Samples were collected with BD ESwabs (BD, 220245) on children being tested for GAS at the Emergency Department of the Children's Hospital of Eastern Ontario (CHEO). DNA was extracted from throat swabs using the Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo D6005) following kit instructions.

Gel Benchmarking

2% agarose (w/v) gel was made in 1X TE buffer with Sybr Safe DNA Gel Stain (Thermofisher, S33102). 5 µL sample aliquots were diluted in 6X DNA Gel Loading Dye (Thermofisher, R0611) with water.

Commercial PCR

Commercial PCR was performed using Thermofisher's PowerUP Sybr Green Master Mix (A25741) according to manufacturer-provided directions for "Standard Cycling Mode (primer T_m < 60°C)" over 25 cycles instead of the instructed 40 cycles.

Nanopore Optimized PCR

Nanopore optimized PCR was performed using a customized master mix. A 50µL preparation of nanopore-optimized PCR master mix contained 5 µL 10X Taq Buffer (NEB, B9014S), 0.25 µL Taq Polymerase (NEB, M0273S), 2.5 µL 10 µM DNAse B Forward Primer (0.5 µM – in final volume) [IDT 5'-AGT TGA TTC CAA GAG CTG TCG TG-3'], 2.5 µL 10 µM DNAse B Reverse Primer (0.5 µM) [IDT 5'-TTA ATG GTG TAG CCA TTA GCT GTG TT-3'], 1 µL dNTP mix (NEB, N0447S) 5 µL 10 mM $MgCl_2$ (2.5 mM), and 31.75 µL Nuclease-free ultrapure water (Thermofisher, 10977015). 2 µL target DNA extract was incubated in this mixture and thermocycled. Starting from 60 s at 95°C, the mixture was subjected to 25 cycles (15 s at 95°C, 15 s at 55°C, then 30 s at 68°C), followed by a cooldown to 60°C for 420 s then reduction to 4°C to stabilize until use.

Nanopore Fabrication

Nanopores are fabricated using controlled breakdown,[5] employing the equipment and protocols outlined in detail in Waugh et al.[44] Briefly, pores were fabricated in nominally 12 nm thick SiN_x membranes purchased from Norcada Inc. (NBPX5004Z-60O-Hi RES) in 1 M KCl pH 8 using a linear voltage ramp, followed by conditioning and enlargement in 3.6 M LiCl pH 8 until a low-noise pore in the size range of 5-8 nm was reached.

Data Acquisition and Analysis

Prior to sensing with the nanopore, 28 µL of the sample following the PCR amplification step was diluted with 72 µL of 3.6 M LiCl 10 mM HEPES pH 8. From that sensing mixture, 40 µL is used to perform a nanopore experiment. Amplicon signal was verified using NoLimits 100 bp DNA fragments (ThermoFisher, SM1441)

PCR amplicons are sensed in 3.6 M LiCl pH 8 using a 200 mV voltage bias. Data are acquired using a Chimera VC100 amplifier sampled at 4.17 MHz with a bandwidth of 1 MHz and are software low-pass filtered at 500 kHz prior to analysis. Events are fitted using a custom implementation of the CUSUM+ algorithm.[45]

2.2 Introduction

Molecular methods for infectious disease diagnostic are gradually being introduced into clinical use, and this trend will only accelerate because of the Covid-19 pandemic.[46–49] Many of these molecular tests typically employ a DNA amplification method to generate many copies of the target sequence and then rely on downstream optical sensors to detect a fluorescent signal generated by the presence of amplified nucleic acid products (i.e., amplicons) in the sample. However, the optical components and the use of dyes as labels, required for fluorescence-based detection, can impose some limitations on the cost, size, and robustness of the instrument or reagents. These factors may limit the use of molecular diagnostics at the point of care and in rural

or remote settings. Alternative detection methodologies that can improve on speed, sensitivity and accuracy are being actively explored.[50,51]

The use of solid-state nanopores for electrical detection of PCR products is one such strategy that can offer less costly, more miniaturizable, and more rugged alternative to optical sensors. Solid-state nanopores are molecular-sized openings in thin membranes that allow the passage of individual biomolecules immersed in a conductive solution. Under an applied voltage, a steady stream of ions transits the nanopore, creating a measurable electrical current that is disrupted by the passage of a single biomolecule. The characteristic electrical signature generated by a target molecule can enable target identification. [52–54]

To demonstrate this potential, a few solid-state nanopore sensing methods have been developed to detect specific nucleic acid (NA) targets. PNA probes have been used to bind to specific DNA sequences,[36,55] and employed to detect the cystic-fibrosis mutation codon.[37] Solid-state nanopores have also been used to detect polymorphic insertion or deletion mutations in DNA, which determine the virulence or drug-resistance of two different bacterial organisms, after cleavage with mutation site recognition enzymes. Conserved sequences of the HIV RNA genome have been detected using biotinylated ssDNA probes and monovalent streptavidin.[38] DNA nanostructures were used to report the presence of a Zika virus gene by folding in its presence, as well as a microRNA biomarker for human lung cancer.[39–41] A loop-mediated isothermal amplification scheme was used to generate many amplicon copies that can be counted to develop qualitative and quantitative nucleic acid tests.[56] RNA targets, including SARS-CoV-2, were also quantified by reverse transcription to produce gene specific cDNA that could be counted without necessarily requiring amplification, though low-level of amplification steps improved the detection speed in cases of rare transcripts.[43] Multiplexed detection of prostate cancer biomarkers (circulating tumor DNA or microRNA) has also been developed with electro-optical nanopore platforms.[57,58] Using solid-state nanopores for direct detection of NA targets from clinical specimens is challenging, partly due to the complexity of most biofluids, which can clog the pore or severely complicate downstream data analysis by generating a variety of translocation signals, but also to the requirement for a high volume of statistics and a high degree of reproducibility needed for confidently obtaining useful clinical results. This requires employing

many individual solid-state nanopore sensors, often of similar geometries, which is presently achieved by performing and repeating many measurements serially but would be accelerated if done in parallel with arrays of nanopores.

Here, we performed a proof-of-concept study to determine the feasibility of pathogen detection from clinical samples with solid-state nanopores. We targeted group A streptococcus (GAS) in throat swabs from children with pharyngitis. GAS is a very important pediatric bacterial pathogen, accounting for 20-30% of pharyngitis episodes in children and leading to many pediatric emergency department visits.[59,60] The complications of untreated GAS pharyngitis include acute rheumatic fever, scarlet fever, tonsillopharyngeal abscess, and streptococcal toxic shock syndrome.[61] Antibiotic therapy is used to prevent these severe complications and to reduce the duration of pharyngitis symptoms. Clinicians have traditionally depended on bacterial culture results from throat swabs to identify those infected with GAS. More recently, molecular tests for GAS that amplify organism DNA have been developed,[62] but, as noted, such tests are often costly and devices used for testing may be less adapted for handheld operations at the point of need, limiting the widespread use of these tests. Nanopore-based diagnostics may be able to overcome these barriers, while potentially offering simple workflows and high throughput, and maintaining high sensitivity.

In this work, we demonstrate the utility of solid-state nanopores as the sensing element in a screening test for the presence of Group A streptococcus (GAS) in throat swabs. We report the optimizations and workflow adaptations required to make PCR amplification compatible with downstream solid-state nanopore detection, present an analysis scheme based on statistical differences in capture rates to make a call on a clinical sample, and demonstrate that our homebrew PCR nanopore assay agrees well with gold standard cell-culture and qPCR-based tests results.

2.3 Results and Discussion

In this study, screening for an infectious agent in a clinical sample is based on electrical identification of individual PCR amplicons using a solid-state nanopore sensor. To enable detection with an acceptably low level of false positives (specificity) and false negatives (sensitivity), the nanopore must distinguish target amplicons from all other molecules present in the sample or introduced during the amplification reaction, including small molecules, genomic DNA, proteins, and enzymes. The electrophoretic and electroosmotic forces in and around the pore can capture these unwanted molecules, and their collision and interaction with the pore as well as their full translocation can lead to measurable blockages in the ionic current signal that must be differentiated from the amplicon signal. The degree with which the nanopore sensor can report the detection of these molecules is a function of the pore properties (pore size, surface chemistry), operating condition (salt concentration, pH, voltage), the properties of the off-target molecules in the sample (size, charge, and propensity to stick to and clog pores in a given set of operating conditions), and the response time of the amplifier-pore system.[63–71] Consequently, we first performed control experiments to investigate the nanopore response to different PCR mixtures to identify adequate operating conditions for sensing amplicons.

2.3.1 Nanopore Detection from Commercial PCR mixtures

Polymerase Chain Reaction (PCR) is the gold standard technique for amplifying nucleic acids,[15,17,72] and numerous kits are available commercially. These kits contain several reagents to optimize the amplification reaction following their protocol. Commercial kits are also often used with a downstream fluorescent optical readout and cannot *a priori* be assumed to be compatible with nanopore sensing. These off-the-shelf PCR master mix kits usually include components such as bovine serum albumin (BSA) and detergents (e.g., Tween20) used to stabilize enzymes during storage and minimize losses to walls of plastic tubes as well as additional proprietary or undisclosed ingredients to increase PCR yield and product lifetime. These molecules, which can

be in relatively high concentrations, are known to produce signals when translocating a nanopore, but also to stick to the pore walls, increasing the low-frequency noise of the nanopore signal and often leading to clogs (reversible or permanent).[73–76] Characterization of the signals caused by any PCR master-mix is therefore a critical control for any analysis of the amplification reaction with a nanopore sensor.

Figure 24a-c shows the results of a PCR mixture from a commercial kit sensed by a 6 ±2 nm nanopore in 3.6M LiCl pH 8 at 200 mV (note the uncertainty in the pore size is due to the uncertainty in the effective membrane thickness, likely smaller than the nominal membrane thickness of 12 nm).[77,78] To perform this measurement, a commercial master mix (see Methods section) was added to a sensing salt to mimic the background signals expected of a PCR amplification reaction during nanopore sensing.

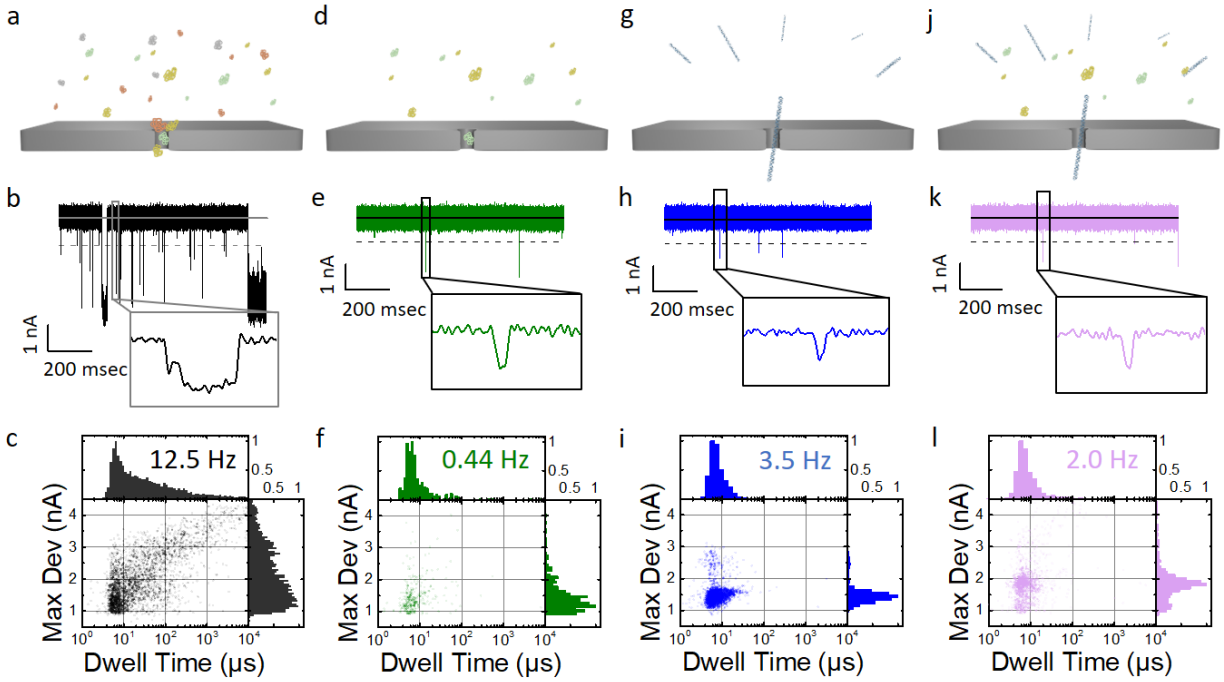

Figure 24: Nanopore sensing of different PCR mixtures. a) Schematic representation of a nanopore sensing a commercial PCR reaction mixture; b) 1-s current trace (black) of the commercial master mix; c) Maximum blockage depth of an event versus its passage time for the entire ~10 minutes dataset, comprised of 8,078 single-molecules events (black dots); d) Schematic of a nanopore sensing of a homebrew PCR mixture. e) 1-s current trace (green) of nanopore-optimized homebrew PCR master mix; f) Maximum blockage depth of an event versus its passage time for all 297 events identified during ~10 minutes of recording with nanopore-optimized homebrew PCR mixture (green dots); g) Schematic representation of a nanopore sensing DNA amplicons; h) 1-s current trace (blue) of 100bp dsDNA at 75 nM; i) Maximum blockage depth of an event versus its passage time for all 2,779 events identified during ~10 minutes with 100bp dsDNA (blue dots); j) Schematic of a nanopore sensing of a homebrew PCR-amplified reaction mixture; k) 1-s current trace of homebrew PCR-amplified reaction mixture containing homebrew mixture components and generated 100bp target amplicons from a GAS positive clinical sample; l) Maximum blockage depth of an event versus its passage time for all 1,444 events identified during ~10 minutes of recording with of PCR-amplified reaction mixture using homebrew PCR mixture containing 100bp target amplicons from a GAS positive clinical sample. Data for b) and c) acquired on the results of a PCR mixture from a commercial kit sensed by a 6 ± 2 nm pore, while for i), k), l) on a 5 ± 1 nm diameter pore (with 7 ± 2 nm effective thickness) and e), f), h) on the same pore that grew to 7 ± 1 nm diameter pore (with 8 ± 2 nm effective thickness) many hours later, all at 200 mV in 3.6M LiCl pH8. Current traces displayed with a low-pass Bessel filter at 500 kHz

Figure 24b shows a representative current trace of this experiment, sampled at 4.17 MS/s and low pass filtered at 500 kHz. The trace shows many different event shapes, including temporary clogs, and ends with the nanopore irreversibly clogging after about 10 minutes. Figure 24c provides additional detail on the characteristics of the signals generated, showing a range of blockage depths from 1.0-3.5 nA and a wide distribution of passage times from 3-1000 μs.

We attribute the large number of events and the complexity of the signals to the variety of molecules translocating the nanopore. The events observed can result from a complex combination of signals generated by the presence of detergents, fluorescent dyes, uracil-DNA glycosylase (UDG), and proteins included to improve stability and amplicon yield. Furthermore, under the high salt concentration condition used here, it is also possible for proteins to agglomerate and change structure. The data of Figure 24c revealed 8,078 single-molecule events in a 645 s of recording, exhibiting a high capture rate of molecules of 12.5 Hz. A representative sampling of signal shapes attributed to the master mix and 100bp dsDNA can be found in Appendix A1.

The passage of linear unfolded dsDNA molecules is expected to produce blockages of approximately 1.5 ± 0.2 nA, depending on the value of the effective pore length. The presence of the amplicon population would therefore be difficult to discern from this commercial mixture. While the pore properties and operating conditions could be adjusted to potentially permit a better identification of amplicons from this commercial PCR mixture, the sensing conditions are far from optimized. Analysis is also complicated by the large number of events detected and the propensity of pore to clog that severely limits the nanopore lifetime and lowers experimental yield. To overcome these limitations, a homebrew PCR mixture was developed to be better suited to solid-state nanopore operation.

2.3.2 Nanopore Detection from Homebrew PCR mixture

To develop a PCR mixture compatible with nanopore detection (i.e., that minimizes the events from background and avoids frequent clogging), we optimized a reaction mixture starting from a minimal set of components. Details on this process can be found in Appendix A2. Characterization of this homebrew PCR mixture under optimized nanopore sensing conditions are shown in Figure 24d-f. In this negative control experiment, the solution contained the homebrew PCR reagents without any template (or genomic DNA), primer or amplicon present mixed in 3.6 M LiCl pH8 following 25 cycles of PCR amplification (see Methods section). The PCR reagents include Taq Polymerase at 25 units/mL, Tris-HCl at 10 mM, KCl at 50 mM, $MgCl_2$ at 4 mM, and dNTPs at 200 µM.

The current trace (Figure 24e) using this optimized mixture reveals far fewer events compared to Figure 24b for a similar sized pore (~5nm) at the same applied voltage of 200 mV.) using this optimized mixture reveals far fewer events compared to Figure 1.b for a similar sized pore (~5nm) at the same applied voltage of 200 mV. Additional current traces of the homebrew PCR mixture on additional pore sizes 6 ± 2 nm are shown in Appendix A3. The solid black line represents the mean current, while the dashed line is the current blockage threshold at 7 standard deviations (SD) from the mean used to identify the beginning of a translocation event during our analysis. With these analysis settings, 297 events were detected in 679 s, for a capture rate of 0.44 Hz. The plot of the maximum blockage versus dwell time of an event (Figure 24f) shows a mean blockage depth of 1.2 ± 0.3 nA and passage time of 5 ± 3 µs. While some events in this negative control experiments generate blockages that overlap with the estimated blockage depth (~1.5 nA) and mean passage time (< 10 µs) for single-file linear dsDNA in these conditions, this homebrew PCR mixture produces fast and relatively very few background signals, and most importantly is far less prone to clogging the nanopore when compared with the commercial kit. These features should greatly facilitate data collection and analysis when characterizing clinical samples.

Next, we characterized the signal from a positive control, wherein a 100 bp dsDNA strand was added as a means by which to highlight the expected result of amplification of a target sequence from an infectious agent. The results of this positive control are shown in Figure 24g-i. Similar to the negative control, the presence of 75 nM of 100 bp, rod-like, dsDNA produces events with a mean blockage of 1.4 ± 0.2 nA. While these events are also translocating quite rapidly, they are more tightly distributed in passage time and blockage depth as compared to the events from the homebrew PCR mixture. Distinguishing an amplicon from a molecule in the homebrew PCR mixture on an individual event basis can be challenging, as the two populations show some overlap, although the frequency of events is much greater for this concentration of DNA, with 2,779 single-molecule events in a 779 s of recording. We note that the fast passage times of the short DNA molecules approaches the response time of our nanopore-amplifier system and we carefully set the analysis parameters (low-pass filter settings and detection threshold value) to provide appropriate signal-to-noise ratio to detect the blockage level of dsDNA while minimizing attenuation. For the average residence time of ~2-5 microseconds, a filter of 500 kHz was chosen to minimize attenuation of signals <2 μs.

Finally, to explore if the presence of amplicons in a homebrew PCR mixture can be discerned, we performed the full amplification reaction of a clinical GAS-positive sample using the homebrew PCR protocol and examined the result with a nanopore. Here, in addition to the homebrew PCR reagents and the amplicons, the sample contains genomic DNA and potentially other biomolecules from the clinical sample leftover from the assay workflow (see section below for assay workflow details). The results of this experiment performed on the same pore are shown in Figure 24j-l. As expected, the populations of events observed resembles a superposition of both previously tested controls, with some overlap (Figure 24f and 22i). However, with the reduced background signal from the homebrew mixture, the presence of the population of amplicons is now more easily identifiable. Equipped with this homebrew PCR mixture, and an understanding on how to discern the presence of an amplicon population, we next turned to testing of clinical samples.

2.3.3 Assay Workflow and Testing of Clinical Samples

For this study we received clinical samples from residual throat swab specimens submitted to the Eastern Ontario Regional Laboratory Association (EORLA) from de-identified patients at the Children's Hospital of Eastern Ontario (CHEO). These swabs were previously used to diagnose the presence of Group A streptococcus (GAS) bacterial infections in children using cell culture techniques. The results from these clinical tests were used for comparison with the nanopore PCR assay workflow.

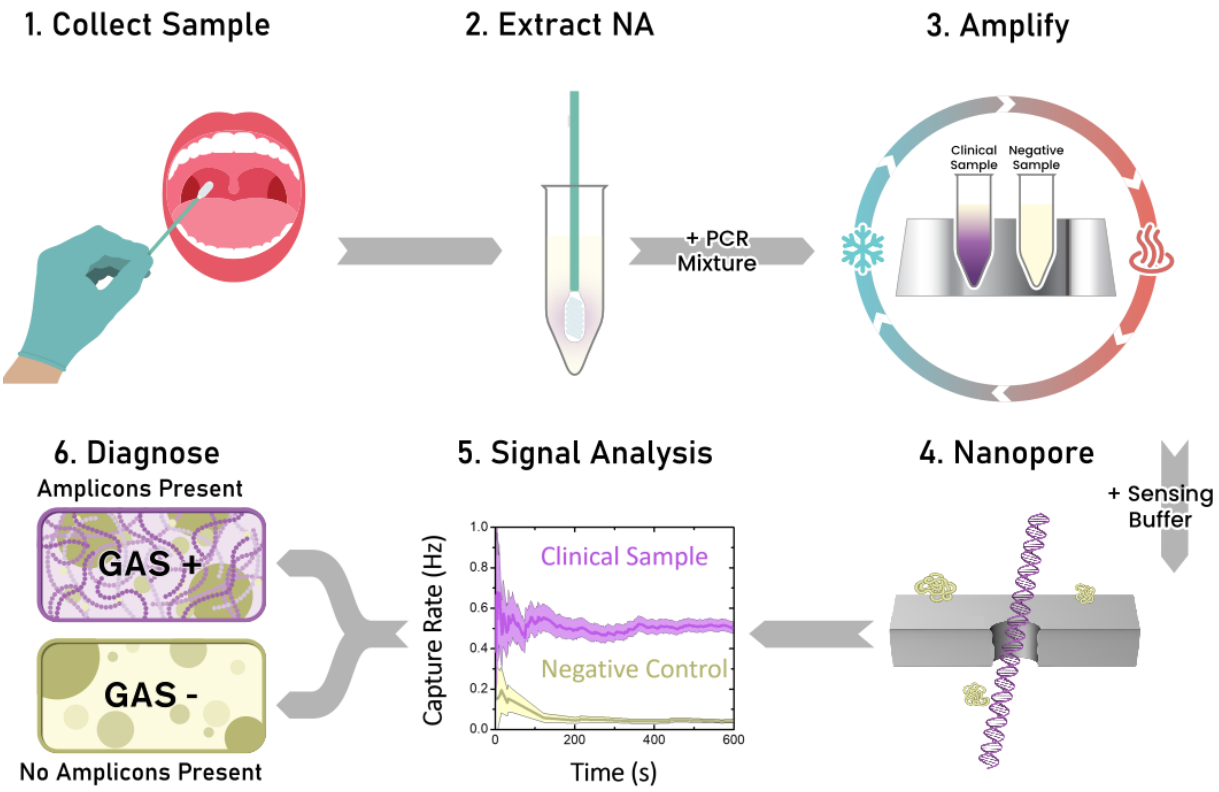

Figure 25: Schematic representation of the assay workflow. Step 1: a throat swab is collected from patients with BD ESwabs. Step 2: DNA is extracted from the throat swab using the Quick-DNA Fungal/Bacterial Miniprep Kit and purified by centrifugation. Step 3: PCR is performed (25 cycles) with a homebrew reaction mixture. Step 4: The sample is mixed with sensing solution and injected in the fluidic flow cell to sense with the nanopore for a few minutes. Step 5: Capture rate of single molecules from sample and blank are compared on the same pore. Step 6: The sample is called GAS positive if a significant amplicon population is present compared to background.

Figure 25 is a schematic of the workflow used to amplify a region of the genome of GAS using specially designed primers previously clinically validated by qPCR.[79] Briefly, throat swabs were collected from patients expressing strep-like symptoms when visiting the Emergency Department at CHEO. The sample was obtained by DNA extraction and purified by centrifugation, before undergoing PCR with the optimized components (see Methods). After PCR, the reaction mixture was diluted, combined with salts for nanopore sensing and run on a nanopore for approximately ten minutes. Each nanopore used for the assay also saw a blank negative control consisting of a template-free homebrew PCR reaction mixture. Following nanopore sensing each set of events were analyzed based on the signal profile produced by the blank to classify the background (blank) population. This classified signal is used to identify a subpopulation located around the expected amplicon translocation profile of a sample. The classified amplicon subpopulation is then compared with the classified signal of the blank with a *t*-test.

2.3.4 Nanopore Capture Rate-Based Screening

Here, we analyzed data sets with events from multiple populations of molecules, such as in Figure 24l, by a method of population classification based on blockage depth and dwell time. We discuss an alternative approach by which such data can be analyzed, and single-molecule events from different population separated, in Appendix A4. From the analysis of multiple positive and negative controls, we identified the expected range of blockage depths and dwell times for the amplicon population by Gaussian fitting of the two distributions. This knowledge was then applied the PCR-amplified clinical samples to construct a filter to select all events within an ellipse drawn on a log-passage time versus blockage depth plot with each major axis spanning 2 SD about the expected amplicon peak on an individual pore basis, and to calculate the capture rate of events so selected (the “signal rate”) as shown in Appendix A4 (Figure A8). Of note, this elliptical classification provides a conservative estimate on the number of amplicons; some true

amplicon signals are invariably excluded mostly due to intra- and inter-pore variabilities in passage times. Appendix A4 provides examples of some of this variability.

To demonstrate the utility of solid-state nanopores as the sensing element in a screening test for the presence of Group A streptococcus (GAS) in throat swabs, we used the assay workflow of Figure 25 and compared on the same nanopore the results of a clinical sample with a negative control containing only the PCR mixture. Both sample and control underwent PCR amplification and nanopore sensing under identical conditions. Because even the homebrew PCR mixture produces some events that overlap with the amplicon population, the single-molecule event detection rate from the subpopulation classified as amplicon was used to determine whether the PCR amplification process had generated enough amplicons in the clinical sample to exceed the rate at which signals were generated in the negative control. Differences attributable to matrix composition by sample type were found to be negligible. A comparison between pre- and post- amplification positive and negative clinical samples, and the blank can be found in Appendix A5 and A6. If a statistically significant increase in the amplicon-classified signal rate over the background was identified from the same population clusters, a clinical sample was then identified as positive. This significance was measured at a two-tailed confidence interval of 95%.

If more precise quantitative analysis is desired, we have previously demonstrated an applicable technique (controlled counting) that makes use of a known-concentration standard to account for variations between nanopore and over time in the measured capture rate of a target molecule.[8] This method together with the use of an amplicon that produces events that do not overlap with the signals from the PCR mixture would allow precise quantitative measurement of the amplicon concentration. Nevertheless, for the purposes of our proof-of-concept study, which focuses simply on screening for the presence or absence of pathogenic DNA, the accuracy and precision of the capture rate data is not as important as qualitatively identifying the GAS-positive samples that produced a clear signal over the negative samples. Click or tap here to enter text.

To assess the performance of our assay and analysis workflow on clinical samples, four throat swab samples (two GAS-positive, two GAS-negative) were assessed. The

positivity/negativity of the samples for GAS was determined by conventional plating technique (i.e. bacterial culture), and verified by ddPCR. Overall, the data reported in the main text comprises >10,000 single-molecule events from 11 experiments acquired on 3 nanopores. The results of these GAS screening tests are presented in Figure 26.

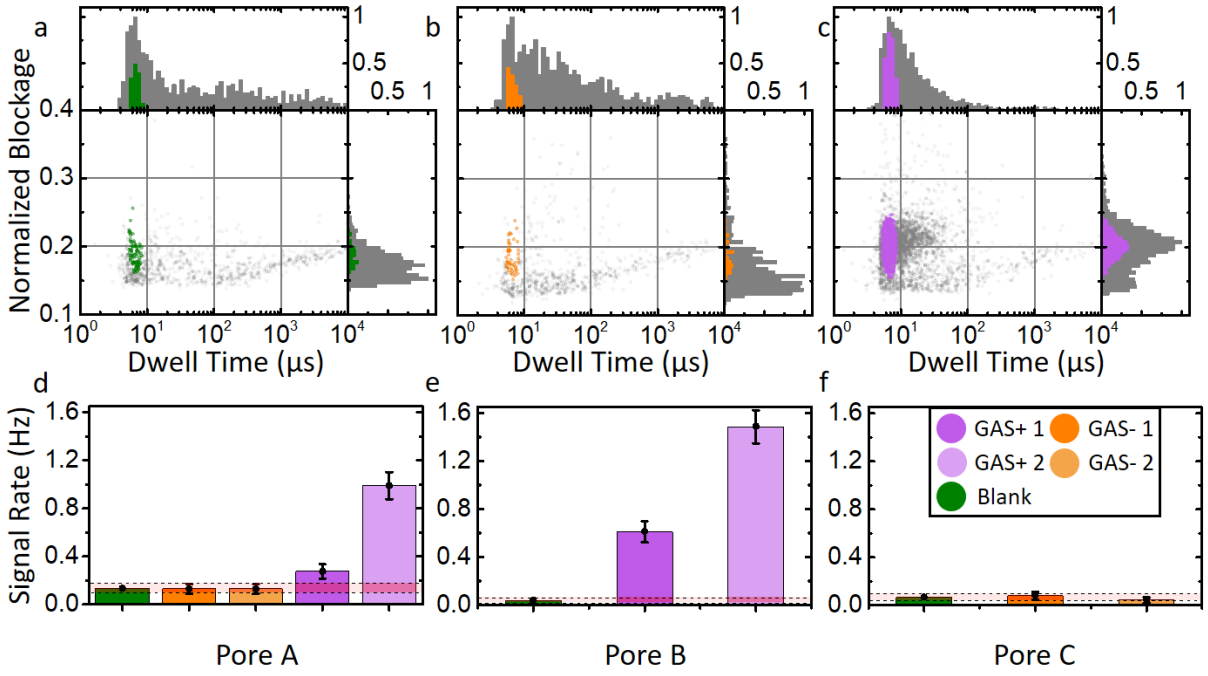

Figure 26: Nanopore analysis of clinical samples. a-c) Blockage fraction as a function of open-pore current vs passage time of all (gray) and amplicon-classified (green, orange, purple) events collected over ten minutes on the same pore with a) blank control (n = 816 total, n = 72 classified events), b) GAS-negative (n = 662 total, n = 47 classified events), and c) GAS-positive (n = 2,598 total, n = 591 classified events) samples amplified by PCR. d-f) Signal rate of amplicon-classified events after 300 s of recording for each sample. Red area bounded by dashed lines represents the 95% confidence interval of the Blank signal rate. All GAS-neg sample signal rates fall within the blank CI, while GAS-pos sample signal rates exceed the blank CI. Data acquired on pore A (6 ±1 nm dia.), pore B (5 ±1 nm dia.) and pore C (7 ±1 nm dia.) at 200 mV in 3.6M LiCl pH8.

Figure 26a-c depicts the normalized blockage depth and dwell time for all events as well as the classified signals for different samples collected for ~10 minutes on the same pore. The same 2 SD criteria, to classify the amplicon population, is applied on all samples on the same

pore, though the centroid of the ellipsoid and the actual value of the SD varies between pores. These classified events are shown in color (green, orange, purple) over the entire data set (grey). Analysis of additional clinical samples can be found in Appendix A7. The classified signal rate is presented in Figure 26d-f, which shows the capture rates after 300 s of recording is enough to correctly identify the sample with a 95% confidence interval. The measurements on each sample were repeated on at least two nanopores to verify reproducibility and consistency of results of our classification approach. The results are summarized in Table 1.

Table 1: Nanopore Signal Rate and Initial Copy Number. Outcomes of all samples tested with the nanopore assay along with initial copy numbers as determined by ddPCR. The signal rate scales with the initial copy number for the same given pore.

Pore ID	Sample type	Initial Copy Number (#/µL)	Signal Rate (Hz)
Pore A	Blank	n/a	0.14 ± 0.04
Pore A	GAS- 1	0.7 ± 0.9	0.13 ± 0.04
Pore A	GAS- 2	undetectable	0.13 ± 0.04
Pore A	GAS+ 1	180 ± 20	0.28 ± 0.06
Pore A	GAS+ 2	1400 ± 100	1.0 ± 0.1
Pore B	Blank	n/a	0.05 ± 0.02
Pore B	GAS+ 1	180 ± 20	0.56 ± 0.09
Pore B	GAS+ 2	1400 ± 100	1.4 ± 0.1
Pore C	Blank	n/a	0.07 ± 0.03
Pore C	GAS- 1	0.7 ± 0.9	0.08 ± 0.03
Pore C	GAS- 2	undetectable	0.04 ± 0.02

Variations in the capture rate between the negative samples and blank are reduced to statistical noise when taking the classified signal detection rate, while the positive samples maintain their statistically higher signal rate. Demonstration of the validity of simple capture rate averaging over time can be found in Appendix A8. Note that the diagnosis results from all three

pores are consistent, even if the capture rates for the three pores presented in Table 1 show some variation from pore to pore, which is expected despites being closely related in size, ranging from 5 ± 1 nm to 7 ± 1 nm, and being used under otherwise identical operating conditions.[8]

The low signal rate for the GAS-negative clinical samples shows that very few amplicons were generated as compared to blank control samples. The low level of amplification from these GAS-negative clinical samples could be explained by the presence of the extremely low target concentration (here estimated to be <1 copy per μL) not picked up by bacterial culture (not surprisingly, PCR is more sensitive than conventional plating techniques) or from nonspecific binding of the primers to other genomic DNA present in the throat swab sample. From this limit of blank, we can estimate a limit of detection for this homebrew PCR assay with a few minutes of downstream detection with a single nanopore down to seconds for a small array of tens of pores. The time-evolution of the capture rate is shown in Appendix Figure A16

The nanopore screening results agree well with cell cultures and ddPCR techniques. Due to the difference in components and technique between PCR and ddPCR, a direct quantitative comparison between amplicon-classified capture rate and starting copy number is not possible, though the same trend of higher capture rates for higher starting copy numbers was observed. Using the technique presented here, we can make a confident call on the presence or absence of strep-A in a clinical throat swab sample after 25 cycles of PCR with our homebrew reaction mixture, using a purely electrical nanopore readout that obviates the need for optical labelling of the resulting amplicons.

2.4 Conclusion

We have demonstrated that the reagent mixture for PCR can easily be adapted for nanopore sensing, to avoid clogging and spurious events that can prevent measurements or complicate data analysis. We have also provided convincing proof of concept results that nanopores are suitable as the readout mechanism for nucleic-acid amplification-based assay.

Going forward, it will also be interesting to investigate isothermal methods such as NEAR, RPA or LAMP,[80–82] and to build real-time detection and analysis into the assay to reduce the time-to-answer. While thermocycling comes with requirements for temperature control that most likely precludes real-time sensing of the amplification reaction because of the constantly changing pore conductance, changes in capture, higher, etc., isothermal techniques will likely allow for nanopore measurements in real time, permitting the time to answer for amplification tests assayed with nanopores to match or potentially event exceed fluorescence-based readouts, given that the PCR cycle threshold C_p can be lower when paired with a nanopore single-molecule detection than it is with optical instrumentation.[43]

While our results show that solid-state nanopores can find utility in point-of-care diagnostics, they also demonstrate clearly that bringing this test into the clinical space requires practical technology development to scale up and parallelize the sensing modality beyond what is currently possible. We hope that these results, and the inspiring proof of concept demonstrations being performed by others in the diagnostics space with solid-state nanopores, will motivate this essential next step in realizing the full potential of solid-state nanopore sensors.

Chapter 3: Discussion and Outlook

3.1 Integrating nanopore sensing into nucleic acid amplification

The research objective of this thesis was the proof-of-concept demonstration of a nanopore-compatible PCR-based assay for the detection of infectious organisms in a clinical context. For this, I developed a simplified, sample-optimized PCR mixture that is compatible with nanopore sensing, producing fewer background translocation events and leading to less signal conflation with 100 bp DNA than off-the-shelf commercial kits. This allowed me to employ PCR amplification on samples extracted from clinical swabs, and to measure the end-point production of 100 bp amplicons by combining the reaction mixture with an electrolytic solution and directly sensing with a nanopore. I also developed a technique for assessing the presence of amplicon-classified signals based on contrasting the expected signal profile of DNA translocations with the signal profile generated by the template-free reaction mixture. This signal-classification approach was shown to be effective in determining the presence of GAS-infectious samples, as validated with more conventional diagnostic techniques, including ddPCR and cell culturing.

Investigating the compatibility of PCR reagents with nanopores, developing an experimental scheme for sample-PCR-nanopore integration, and successfully identifying the production of GAS-indicative amplicons has been shown to be viable. Future work will investigate refinements to the nucleic acid amplification technique, with a hopeful eye towards eventual integration with a nanopore sensor for point-of-care use. However, there are several hurdles to overcome before a robust, PCR-nanopore assay can be realized as a clinical diagnostic technique.

The degree of amplification proved to be challenging and required appropriate PCR cycle selection in which the appropriate number of cycles had to be chosen for the set of GAS samples. Too few cycles, and the GAS+ sample did not produce sufficient amplicons and made the detection slow (>hours), while too many led to amplification in GAS- samples. This was possibly due to the asymptomatic presence of GAS bacteria in conventionally-diagnosed negative samples by plate culture, which is not as sensitive as molecular tests and prone to interpretation differences. This issue is exacerbated by the end-point limitation of conventional PCR, which only

allows for a binary test of the presence of template DNA. Incorporating real-time analysis, such as qPCR, would alleviate this by quantifying pathogen load, allowing clinicians to differentiate benign from active, or early from peak infection.[83]

A qPCR-nanopore assay would require some form of integration between amplification and nanopore sensing, as the product evolution over time requires constant reporting. This may necessitate the development of a PCR mixture capable of performing in the high-salt environment required for electrical nanopore sensing. The polymerase used in the homebrew PCR scheme for this thesis comes from *Thermus aquaticus*, a thermophilic archeal microorganism. Though extremely heat-tolerant, the polymerase loses enzymatic activity at salt concentrations >100 mM,[84] while low-salt nanopore sensing remains challenging for the resolution of smaller charged molecules such as DNA amplicons.[85] Modern PCR employs a variety of polymerases for different conditions, however identifying or designing one both for high heat- and salt-tolerance would be a challenge. Avoiding direct integration of PCR reagents with sensing solution may be a better solution, though it could require a more complex fluidic device setup.

Cycled-temperature nucleic acid amplification schemes like PCR also require the use of equipment, such as thermoelectric heat pumps, to rapidly heat and cool a sample. These devices emit heat and electrical noise into their surrounding environments, both of which interfere with nanopore sensing. The time required to ramp up and down in temperature every cycle is also time-consuming and hinders the applicability of PCR as a rapid test. An isothermal amplification reaction which does away with the need for a thermocycler would allow closer integration of reaction mixture with nanopore sensor and may save time in running results.

Automating future analysis to provide a diagnostic call in real-time removes a bottleneck in which the prospective researcher or technician must sift through multiple experimental data sets to generate appropriate amplicon-classification parameters. Additionally, leaving signal-classification to the discretion of the end user introduces the potential for unconscious biasing of results. To that end, developing software that automates this process is essential for speeding any future time-to-results.

Dr. Kyle Briggs of the T.-Cossa lab has recently been investigating a solution to this problem by use of unsupervised cluster analysis, a machine-learning implementation in which data is sorted into groups by pattern recognition algorithms. Ideally, this solution would allow a clustering-based program to identify the correct classification parameters without direct human input.

Surmounting all of these challenges, though nontrivial, is worthwhile. It would finally unlock the potential of abiotic nanopore-based devices as cheap, low-power, and robust diagnostic tools. Pairing nanopore sensors with molecular methods allows for high-quality reporting on virtually any organismal-based infection, anywhere in the world. Proximity to specialist technicians and expensive, bulky laboratory equipment would no longer be the limiting factor for a great number of healthcare outcomes. A doctor treating patients in a remote rural setting, an epidemiologist rushing to identify a contagious disease outbreak, or the citizen of a country with unequal access to healthcare could all hugely benefit from a decentralized device that speedily diagnoses pathogenic DNA.

The potential impact on world health is enormous, but also within the realm of the possible. An example approach could employ very similar elements to those investigated for this thesis. A handheld, battery-powered device would contain a cartridge receptacle, with each disposable cartridge housing a fluidic cell with a blank chip. For diagnosis, the user would simply load a clinical specimen such as blood or saliva into the device, select an organismal genome to screen against from an extensive primer library, and insert the cartridge. The device would then fabricate and condition a nanopore to the appropriate noise and size characteristics, while in parallel controlling a nucleic acid amplification reaction with the appropriate primers. The resulting product could be sensed immediately following, with the results analyzed in real-time. The diagnosis could be given in minutes, at relatively low-cost, and with high reproducibility.

3.2 Recombinase Polymerase Assay and Nanopore Sensing

Research in this direction has started in the T.-Cossa lab, and in collaboration with Dr. Erin McConnell, I have explored the use of the Recombinase Polymerase Amplification (RPA) method of nucleic acid amplification, following the completion of the PCR-nanopore results shown in this thesis. Using an RPA-based nanopore sensing approach allows for the cycle-free amplification of targets, without the special primer requirements of other isothermal nucleic acid methods such as Loop-mediated isothermal amplification (LAMP). Initial compatibility tests have found the RPA reaction mixture to be compatible with nanopore sensors, which appear able to resolve 1 kbp dsDNA translocations as shown in Figure 27.

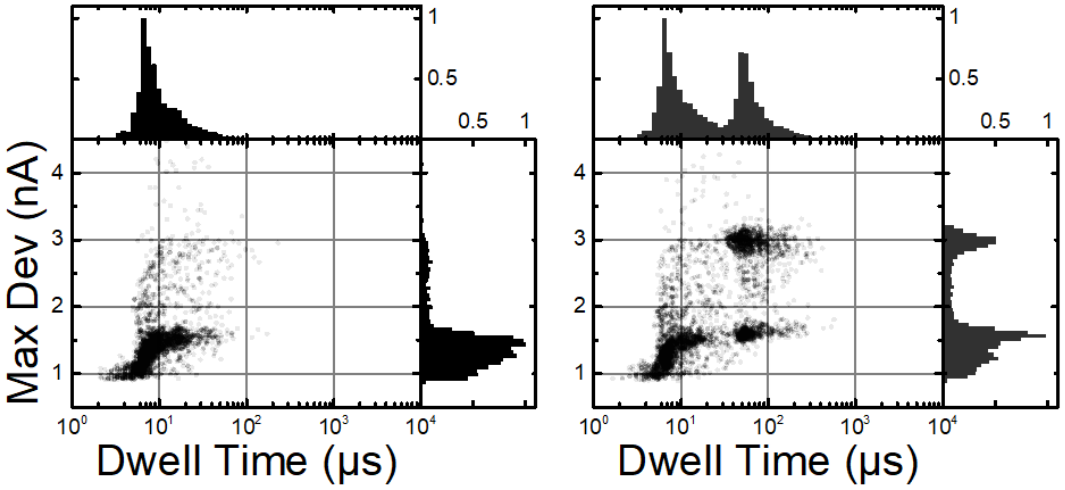

Figure 27: Validation of RPA master mix on a Nanopore. a) scatter plots (with projected histograms) of the maximum blockage vs dwell time of the single-molecule events in the RPA reaction mixture alone (N = 4434 events) after 922 s of sensing b) scatter plots (with projected histograms) of the maximum blockage vs dwell time of the RPA reaction mixture spiked with 15 nM of 1 kbp dsDNA (N = 8859 events) after 884 s of sensing

These results show that the RPA mixture produces a distribution of short, shallow blockages much like the homebrew PCR mixture. DNA translocations of long fragments are clearly visible at ~70 µs at the 1.6 nA (unfolded) and 3 nA (folded) levels. However, as the master mix appears to block current levels very similar to that of unfolded DNA, selecting for a longer

amplicon length (>500 bp) may prove essential for resolution between the two components by dwell time.

Appendix A

A1: Nanopore signals from PowerUp Sybr Green PCR Master Mix

To demonstrate the complexity of the commercial master mix kit, we have selected approximately 200 overlayed time- and current-normalized current traces of Sybr Green translocation events (Figure A1, left). The ensemble of events follows a wide range of blockage depths with no discernable blockage pattern. This can be shown in contrast with Figure A1 (right), where most of the ~200 overlayed traces of 100 bp dsDNA translocation events appear to reach a blockage depth of -1200 +/- 100 pA.

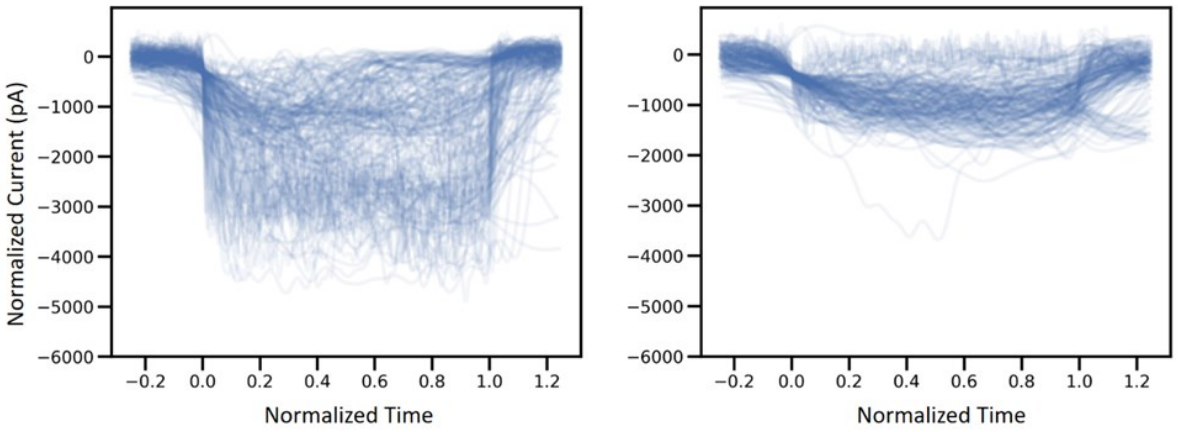

Figure A1: Translocation signal patterns. Normalized translocation patterns of PowerUp Sybr Green PCR Master Mix (left) and 100 bp dsDNA (right). The master mix signals range in blockage depth over the course of translocation while the 100 bp signals follow a more well-defined pattern.

A2: Homebrew PCR Optimization process

Following our initial testing, we set out to optimize the PCR master mix for nanopore sensing. The goal of this effort was to reduce the capture rate of off-target master mix components, both to simplify analysis and reduce the likelihood of pore clogging before completion of an assay.

The first step was to move to a bare-bone implementation of a PCR mixture – containing only nucleotides, primers, Taq polymerase and a detergent-free buffer. This does not contain any of the proprietary components found in a commercial kit and should reduce the number and complexity of signals from the master mix. However, this created two inter-related problems that needed to be overcome:

- First, using a stripped-down PCR mixture comes at the cost of reduced amplicon production for a given number of PCR cycles. This reduces nanopore effectiveness by increasing the amount of time needed to collect enough events to make a call.
- Second, increasing the number of PCR cycles to increase amplicon yield is impracticable past a certain point, as even a target-free sample will begin to produce amplicons (due to non-specific interactions) and produce false positives.

To solve these optimization problems, we explored the influence of $MgCl_2$ and cycle number on amplicon yield using our basic PCR mixture. Gel images showing the yield for various samples as a function of PCR cycle number are shown in Figure A2. Selecting the $MgCl_2$ concentration which provided the largest amplification enhancement, and the cycle number which minimized negative sample amplification, gave us an optimized level of 4.0 mM $MgCl_2$ and 25 cycles, respectively.

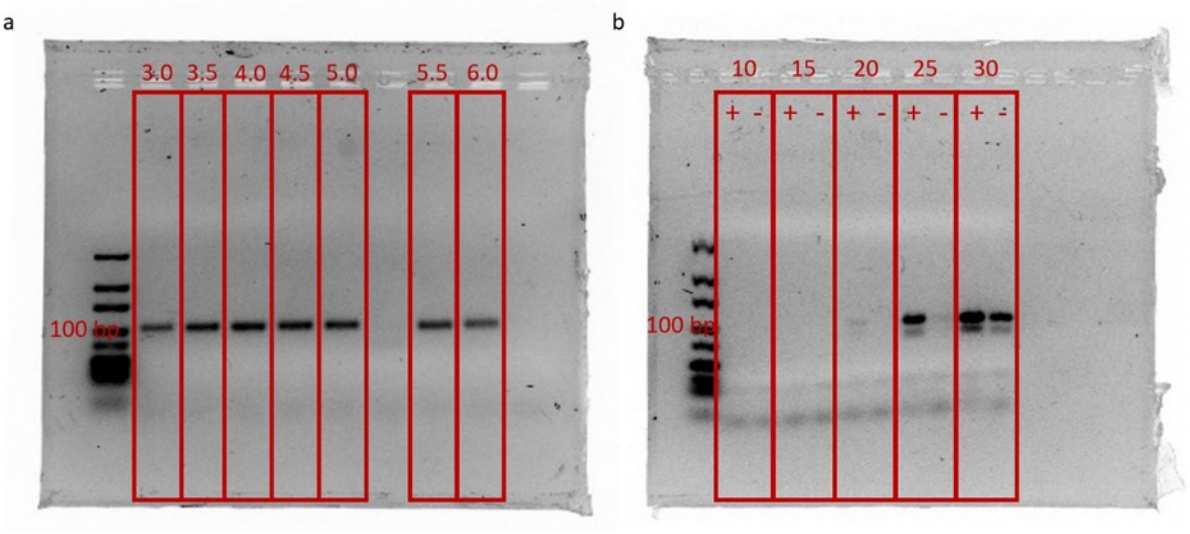

Figure A2: Gel screening for optimal PCR amplification conditions. a) PCR amplification of a positive Group A Streptococcus sample with homebrew master mix and varying amounts of $MgCl_2$. The total amount of $MgCl_2$ added to the reaction was varied from 3.0-6.0 mM in 0.5 mM increments from left to right. All bands migrated a similar distance to the 100bp band on the DNA ladder (far left). **b)** PCR amplification of a positive and a negative Group A Streptococcus samples with homebrew master mix. The number of cycles in the reaction was varied from 10-30 in 5-cycle increments (from left to right). While the positive sample produces the darkest band at 30 cycles, the negative sample also produces a band at 30 cycles. There is a strong positive signal at 25 cycles with only a very faint negative signal.

We used the results of Figure A2 in conjunction with nanopore sensing results to guide cycle number optimization for the homebrew assay, using the GAS-negative sample (GAS- #1) that shows the strongest PCR signal in the commercial test and the GAS-positive sample (GAS+ #1) with the weakest signal to determine the threshold needed to confidently make a call. Using these results, a cycle value of 25 was selected as providing an optimal balance between sensitivity and specificity.

Of note is the fact that all clinical samples we have tested, regardless of their established GAS status, will eventually show a positive result on the gel imaging as a function of PCR cycle number. It is therefore important that the cycle number be optimized to balance between sensitivity (low false negative rate) and specificity (low false positive rate), since it appears that given enough PCR cycles all samples will report GAS-positive.

A3: Homebrew PCR mixture event signatures across different pores

Scatter plots of maximum current deviation (the largest instantaneous deviation of the current from the open pore baseline) versus passage time for each single molecule event detected during a ~10 min recording of the nanopore signal for the homebrew PCR across 3 different pores. The data from Figure 24f looks similar to the left and middle plots of Figure A3 and to Figure A14 (left), while the right most plot in Figure A3 is more simialar to the homebrew signals shown in Figure 26 and Figure A13 (left). While there is some variation in how the homebrew PCR mixture behaves from pore to pore, what is important for our analysis framework is the result of the *t*-test comparing the signal rate for the blank with that for the clinical samples within the ellipse of the classified amplicon population (see Appendix A4), rather than the precise event signatures of the blank control.

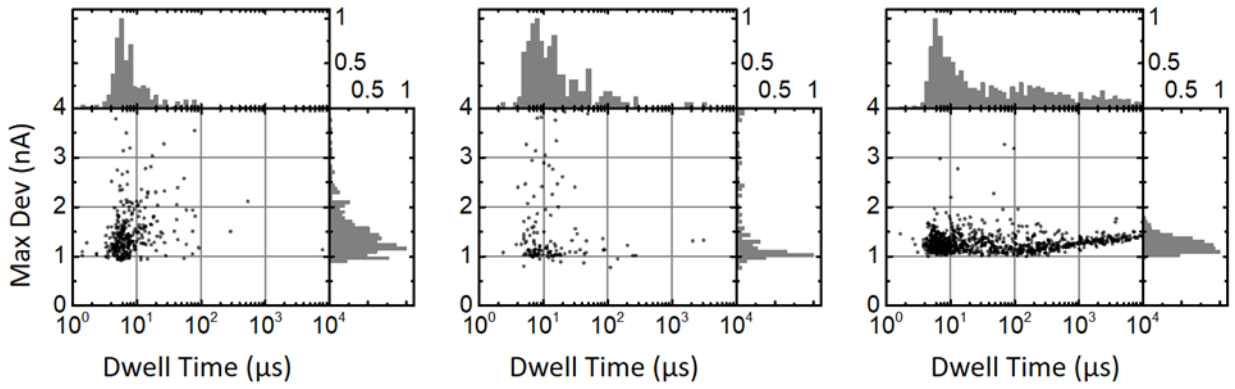

Figure A3: HMM event signatures of different pores. Scatter plots of maximum current deviation versus passage time for each single molecule events detected during a ~10 min recording of 3 nanopores (4.9 nm (left, ~0.5 Hz), 5.8 nm (middle, ~0.25 Hz) and 7.4 nm (right, ~1.3 Hz) for the homebrew master mix (HMM) components in 3.6 M LiCl at 200 mV.

A4: Method of classifying events

Filtering the events to classify those predominantly produced by amplicon translocation requires comparison of short DNA fragment behaviour with signals produced by the homebrew master mix. From Figures Figure ***A4*** and Figure ***A5***, we can determine that 100 bp dsDNA translocations have mean blockages of 1.5 ±0.3 nA and dwell times of 7 ±4 µs for the operating conditions used in this study. We thus locate the expected amplicon peak and background peaks of event histograms.

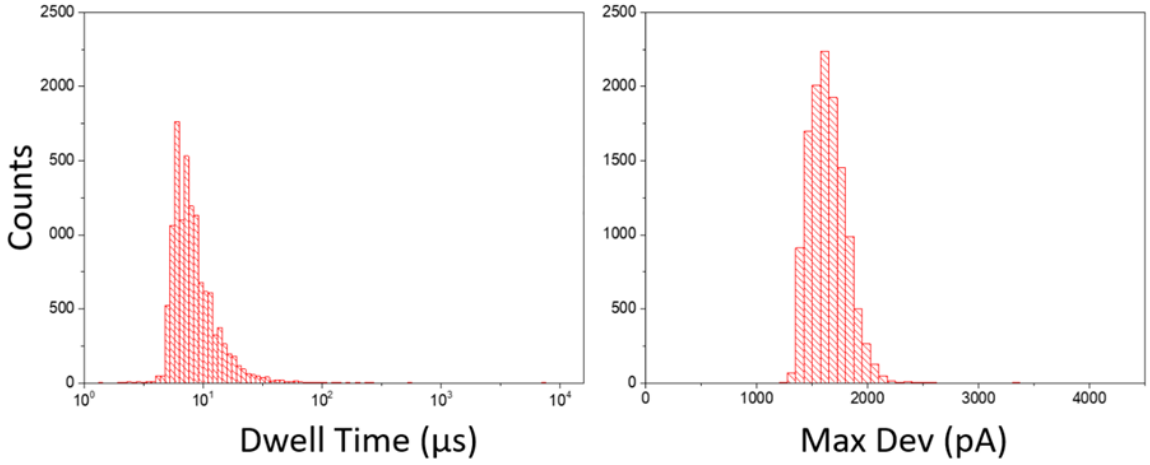

Figure A4. 100bp translocation histograms. Histograms of dwell time (left) and maximum current deviation (right) of translocation events with 100 bp dsDNA on a 4.9 nm pore.

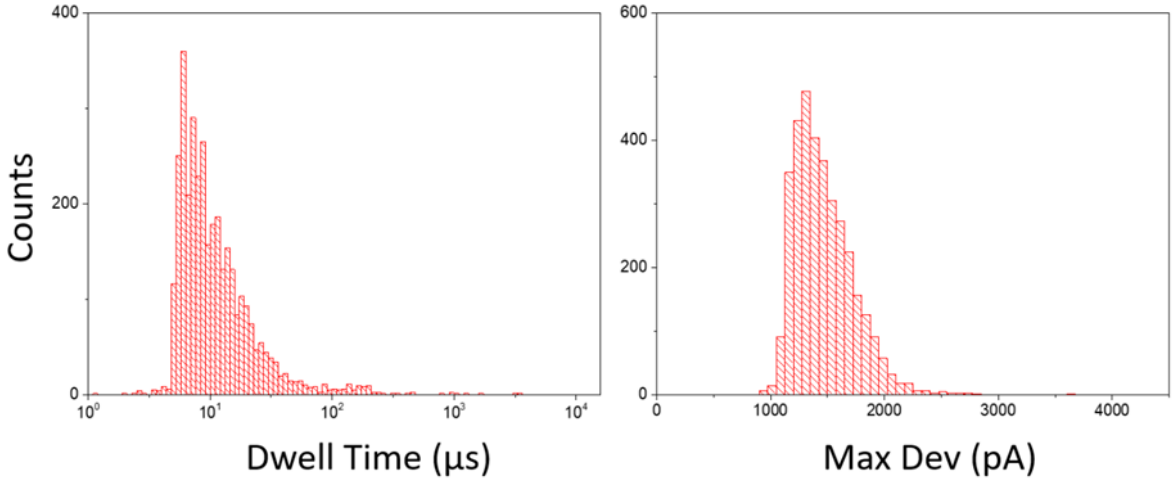

Figure A5: 100bp translocation histograms. Histograms of dwell time (left) and maximum current deviation (right) of translocation events with 100 bp dsDNA on a 6.6 nm pore.

Next, we considered the homebrew master mix signals of a given pore. Classifying the subpopulation signal for each pore requires analysis of its blank sample for proper peak selection (Figure A6). The background events, while convoluted with the 100 bp events temporally, had a characteristic “tail” of longer dwell times and on average events with shorter blockages than those of the 100 bp.

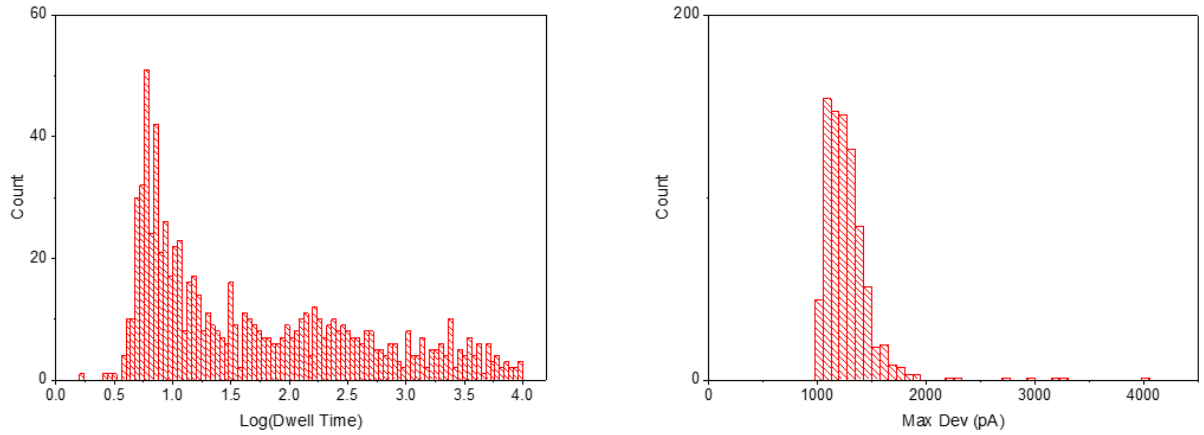

Figure A6: HMM translocation histograms. Histograms of dwell time (left) and maximum current deviation (right) of translocation events with homebrew master mix on a 5.8 nm pore.

Using these, we compared with the signals generated by a positive sample on the same pore to locate the appearance of an expected peak due to the presence of amplicons. We then fitted a Gaussian peak to this expected amplicon distribution (Figure A7). From this fit we extract the mean and standard deviation we use in our signal classification.

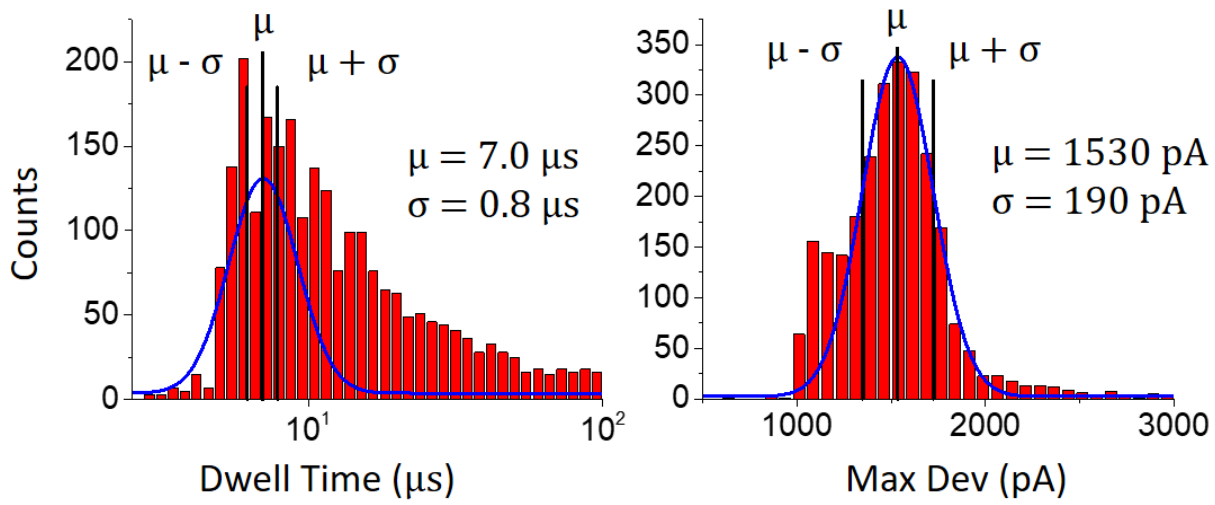

Figure A7: **Histograms and fitted Gaussian curves for signal parameter selection.** Dwell time (left) and maximum current deviation (right) of translocation events with PCR-amplified GAS-positive sample on a 5.8 nm pore.

The signal classification employs a elliptical filter centered on the mean values of the dwell time and maximum blockage extracted from Figure A7, with major axes spanning two standard deviations about the mean, as shown in Figure A8.

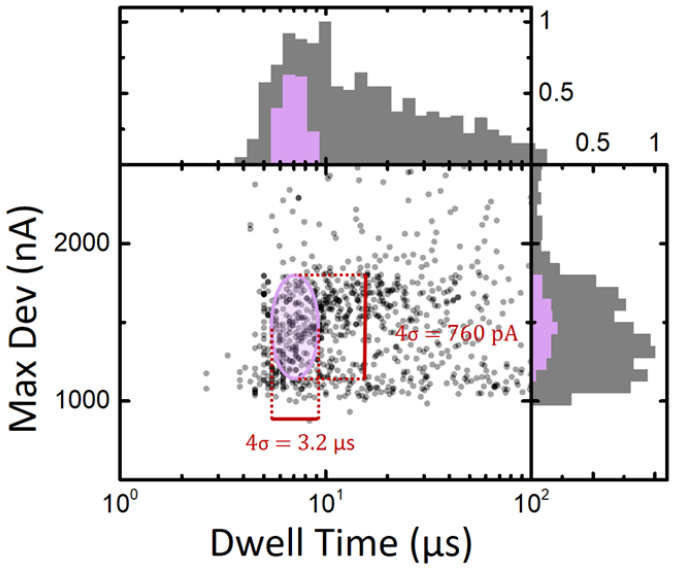

Figure A8: Signal classification filter applied to a GAS-positive sample. The elliptical filter is centered at the mean obtained above (7.0 μs, 1530 pA), with major axes spanning ±2SD (3.2 μs and 760 pA) on either side.

A5: Matrix Effect of Clinical Sample

To explore whether background molecules in the clinical samples (i.e. the matrix effects) would produce a different signal rate compared to the pure Homebrew PCR mixture (blank), we tested GAS positive and GAS negative clinical samples on a nanopore before performing any amplification. For those control experiments, the samples were combined with the homebrew PCR reagents and were sensed for a few minutes without amplification. For this, we selected the highest-amplicon producing samples of each sample type, i.e. GAS+ sample #2 and GAS- sample #1, which have the greatest GAS chromosome copy-number of the two samples (see Table 1 from main text) and should therefore contribute the most background signals. Figure A9 shows the scatter plot of the current blockage versus dwell time for each event (in grey), and those events falling in the amplicon classification of both samples (in colored). Figure A10 shows the signal rates of each of these samples after 600 s of sensing. Both samples, pre-amplification, show very low and almost identical signal rates of <0.05 Hz. The difference in signal rate between the two samples is well within the 95% error interval, which demonstrates the similarity in matrix contributions to the signal rate between positive and negative samples.

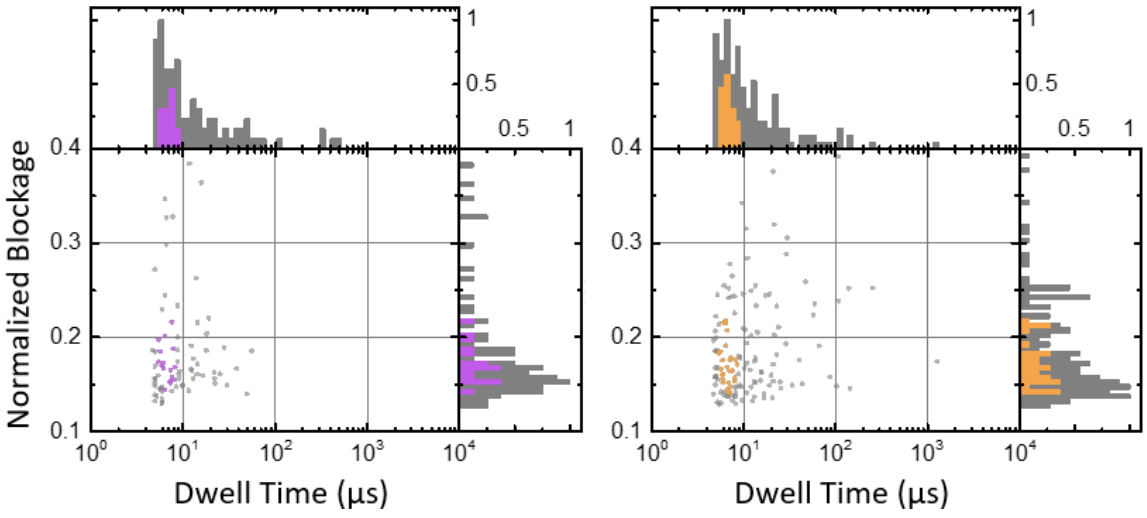

Figure A9: Pre-PCR signal classification of GAS+ and GAS- samples. Classification using method described in Section A4 above, on Pore D (7 ±1 nm dia.), which saw unamplified GAS-pos2 (N=91 events, N=16 classified events, purple, during 605 s of recording), unamplified GAS-neg1 (N=142 events, N=30 classified events, orange, during 632 s of recording). The elliptical signal classification is centered at 7.0 µs and 1535 pA, with major axes of 3.4 µs and 750 pA.

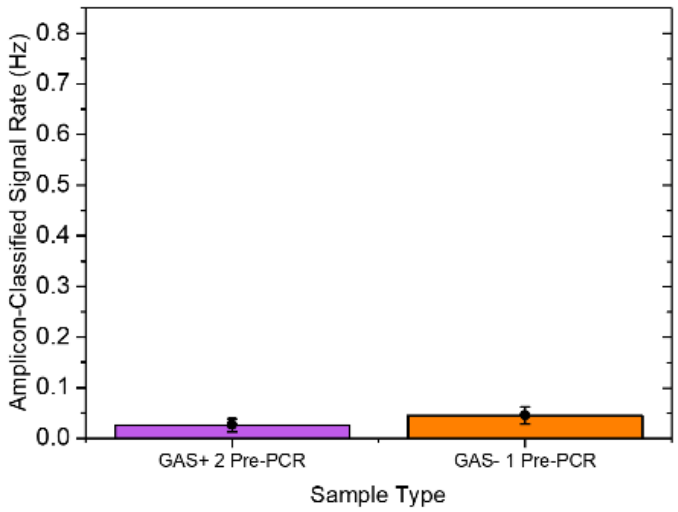

Figure A10: Pre-PCR signal rates of GAS+ and GAS- samples. Signal rate of amplicon-classified events after 600 s of recording for each sample. The low GAS+ 2 pre-PCR sample signal rate (0.026 Hz) and the GAS- 1 pre-PCR signal rate (0.045 Hz) overlap, or are close to overlap, within their error bars. Data acquired on pore D (7 ±1 nm dia.), at 200 mV in 3.6M LiCl pH8.

A6: Pre- and Post-PCR GAS+ Sample #2 Signal Rates

Here we compare the nanopore translocation and capture characteristics pre- and post-amplification on the GAS+ #2 clinical sample, along with a template-free blank (i.e., the homebrew PCR mixture). The results are presented below in Figure A11, which displays the signal classification of each sample, and Figure Figure A12, displaying their signal rates. While the pre-PCR clinical sample exhibits a slightly higher signal rate than the blank, it falls within the blank sample 95% confidence interval. The post-PCR clinical sample has a significantly higher signal rate and demonstrates that major changes in signal rate are directly attributable to the amplification process as opposed to matrix effects.

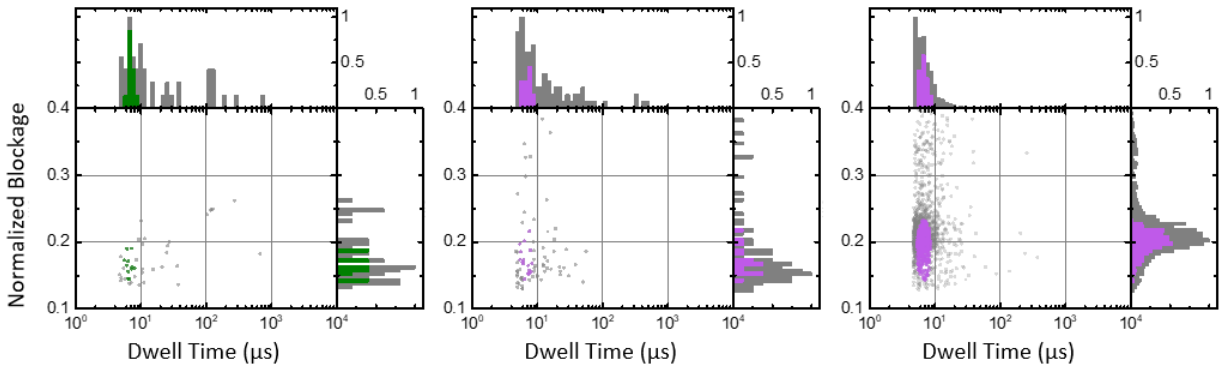

Figure A11: Signal Classification of Pore D. Classification using method described in A4 above, on Pore D (7 ±1 nm dia.), which saw Blank (N=47 events, N=11 classified events, left green), unamplified GAS-pos2 (N=91 events, N=16 classified events, middle purple), amplified GAS-pos2 (N=1456 events, N=508 classified events, right purple). The elliptical signal classification is centered at 7.0 µs and 1535 pA, with major axes of 3.4 µs and 750 pA.

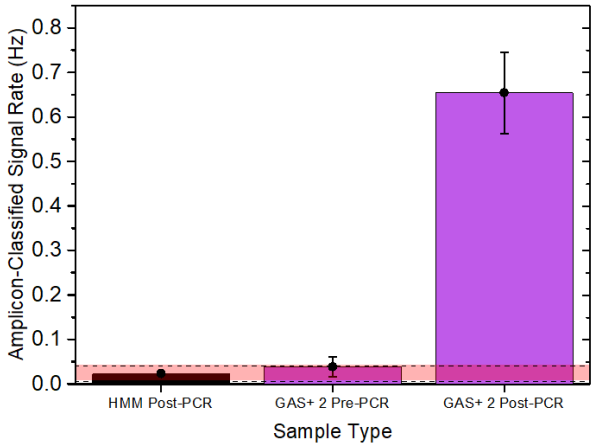

Figure A12: Pre- and post-PCR signal rates of GAS+ 2. Signal rate of amplicon-classified events after 300 s of recording for each sample. Red area bounded by dashed lines represents the 95% confidence interval of the PCR-amplified template-free (blank) Homebrew Master Mix signal rate (0.023 Hz). The GAS+ 2 pre-PCR sample signal rate (0.038 Hz) falls within the blank CI, while the GAS+ 2 post-PCR sample signal rate (0.65 Hz) exceeds the blank CI. Data acquired on pore D (7 ±1 nm dia.), at 200 mV in 3.6M LiCl pH8.

A7: Signal Classifications

Here we present the entire set of classified events against the background of signals. For ease of comparison we have also reproduced in Figure A13 the same as shown in Figure 26 of Chapter 2.

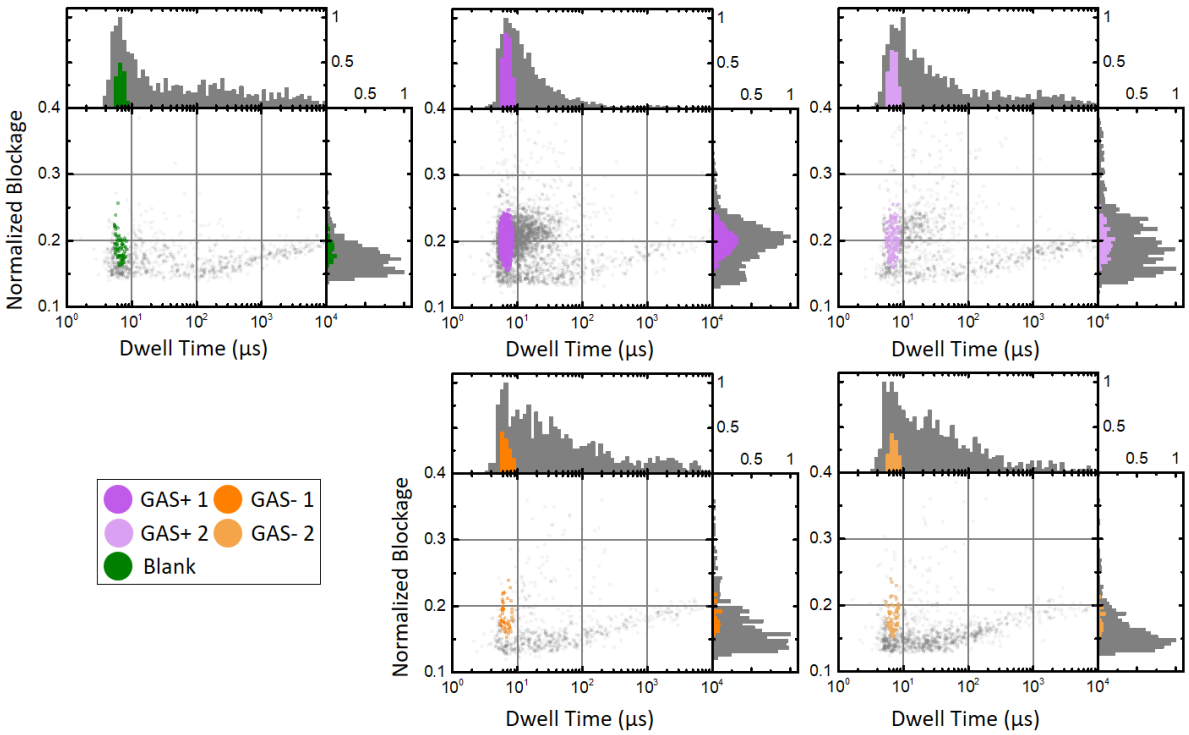

Figure A13: Signal Classification of Pore A. Classification using method described in Section A4 above, on Pore A (6 ±1 nm dia.), which saw Blank (N=816 events, N=72 classified events, green), GAS-pos1 (N=2598 events, N=591 classified events, purple), GAS-pos2 (N=1130 events, N=137 classified events, light purple), GAS-neg1 (N=662 events, N=47 classified events, orange), GAS-neg2 (N=935 events, N=63 classified events, light orange). The elliptical signal classification is centered at 7.0 µs and 1530 pA, with major axes of 3.2 µs and 780 pA

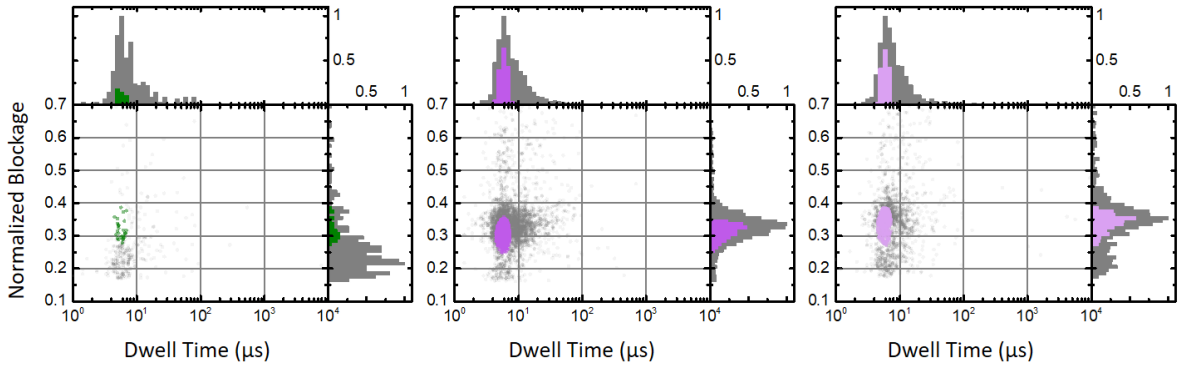

Figure A14: Signal Classification of Pore B. Classification using method described in Section A4 above, on Pore B (5 ±1 nm dia.), which saw Blank (N=297 events, N=25 classified events, green), GAS-pos1 (N=2935 events, N=784 classified events, purple) and GAS-pos2 (N=1443 events, N=360 classified events, light purple). The elliptical signal classification is centered at 6.0 µs and 1700 pA, with major axes of 3.2 µs and 620 pA.

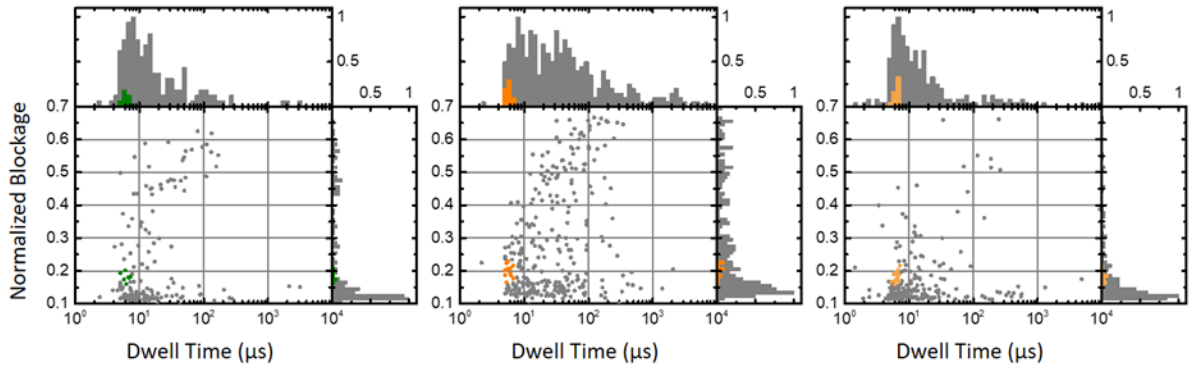

Figure A15: Signal Classification of Pore C. Classification using method described in A4 above, on Pore C (7 ±1 nm dia.), which saw Blank (N=160 events, N=6 classified events, green), GAS-neg1 (N=334 events, N=12 classified events, orange), GAS-neg2 (N=217 events, N=13 classified events, light orange). The elliptical signal classification is centered at 6.0 µs and 1400 pA, with major axes of 3.2 µs and 700 pA.

A8: Capture Rate time evolution

To demonstrate the validity of averaging the number of events with time as a simple method for calculating the capture rate, we have reproduced here the recorded event classification over the entire nanopore sensing experiment. Figure A16 shows that the number of single molecules counted linearly increase with time so that the slope can be used to extract the capture rate, and that the capture rate remains constant over the course of the sensing experiment.

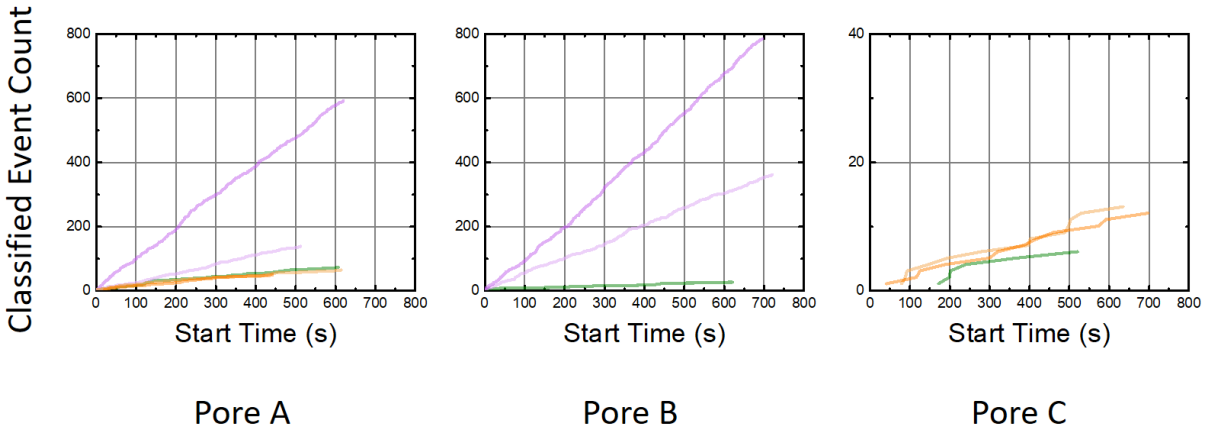

Figure A16: Cumulative number of classified events as a function of time for all samples on Pores A, B and C. Pore A saw Blank (610 s, green), Gas-pos1 (622 s, purple), GAS-pos2 (532 s, light purple), GAS-neg1 (447 s, orange), and GAS-neg2 (619 s, light orange). Pore B saw Blank (678 s, green), Gas-pos1 (697 s, purple) and GAS-pos2 (728 s, light purple). Pore C saw Blank (664 s, green), GAS-neg1 (773 s, orange), and GAS-neg2 (724 s, light orange).

Figure A17 shows the evolution of the capture rate versus time (proportional to the single molecule count), while Figure A18 shows the evolution of the capture rate versus classified event count (proportional to the time). These plots can help determine how long a nanopore needs to sense to make a call. Clearly, for the samples and the pores under the operating conditions used here, the GAS+ capture rates are readily distinguishable from the background in <200 s, while the GAS- signals always remain within the blank capture rate confidence interval.

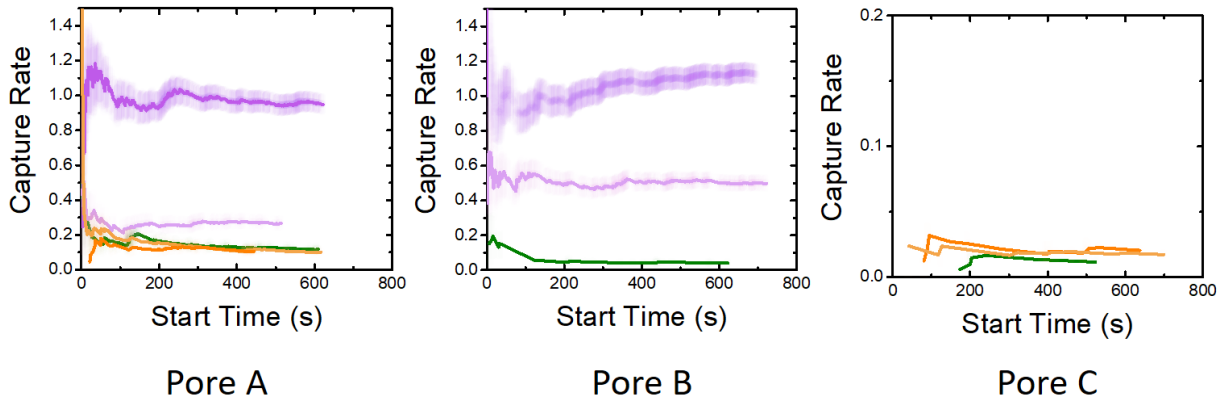

Figure A17: Capture rate of classified events as a function of time for all samples on Pores A, B and C. Pore A saw Blank (610 s, green), Gas-pos1 (622 s, purple), GAS-pos2 (532 s, light purple), GAS-neg1 (447 s, orange), and GAS-neg2 (619 s, light orange). Pore B saw Blank (678 s, green), Gas-pos1 (697 s, purple) and GAS-pos2 (728 s, light purple). Pore C saw Blank (664 s, green), GAS-neg1 (773 s, orange), and GAS-neg2 (724 s, light orange).

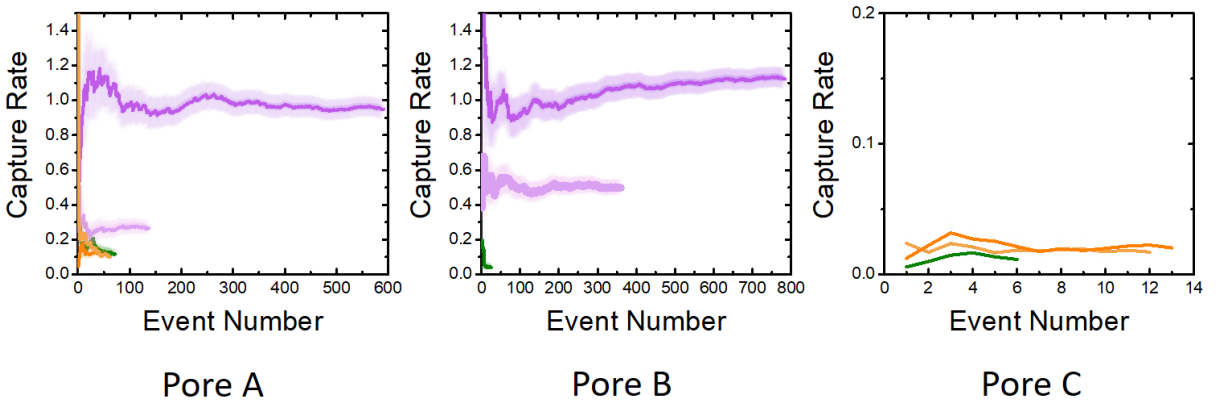

Figure A18: Capture rate of classified events as a function of classified event count for all samples on Pores A, B and C. Pore A saw Blank (610 s, green), Gas-pos1 (622 s, purple), GAS-pos2 (532 s, light purple), GAS-neg1 (447 s, orange), and GAS-neg2 (619 s, light orange). Pore B saw Blank (678 s, green), Gas-pos1 (697 s, purple) and GAS-pos2 (728 s, light purple). Pore C saw Blank (664 s, green), GAS-neg1 (773 s, orange), and GAS-neg2 (724 s, light orange)

References

1. Fried, J. P. *et al.* In situ solid-state nanopore fabrication. *Chemical Society Reviews* vol. 50 4974–4992 (2021).

2. Bayley, H. & Cremer, P. S. Stochastic sensors inspired by biology. *Nature* vol. 413 226–230 (2001).

3. Shi, W., Friedman, A. K. & Baker, L. A. Nanopore Sensing. doi:10.1021/acs.analchem.6b04260.

4. Storm, A. J., Chen, J. H., Ling, X. S., Zandbergen, H. W. & Dekker, C. Fabrication of solid-state nanopores with single-nanometre precision. *Nature Materials* vol. 2 537–540 (2003).

5. Kwok, H., Briggs, K. & Tabard-Cossa, V. Nanopore fabrication by controlled dielectric breakdown. *PLoS ONE* **9**, e92880 (2014).

6. Vodyanoy, I. & Bezrukov, S. M. Sizing of an ion pore by access resistance measurements. *Biophysical Journal* **62**, 10–11 (1992).

7. Kowalczyk, S. W., Grosberg, A. Y., Rabin, Y. & Dekker, C. Modeling the conductance and DNA blockade of solid-state nanopores. *Nanotechnology* **22**, 315101 (2011).

8. Charron, M., Briggs, K., King, S., Waugh, M. & Tabard-Cossa, V. Precise DNA Concentration Measurements with Nanopores by Controlled Counting. *Analytical Chemistry* **91**, 12228–12237 (2019).

9. Tabard-Cossa, V. Instrumentation for Low-Noise High-Bandwidth Nanopore Recording. in *Engineered Nanopores for Bioanalytical Applications: A Volume in Micro and Nano Technologies* 59–93 (Elsevier Inc., 2013). doi:10.1016/B978-1-4377-3473-7.00003-0.

10. Polishchuk, I., King, T. J. & Hu, C. Physical origin of SILC and noisy breakdown in very thin silicon nitride gate dielectric. in *Annual Device Research Conference Digest* 20–21 (2001). doi:10.1109/drc.2001.937855.

11. Beamish, E., Kwok, H., Tabard-Cossa, V. & Godin, M. Precise control of the size and noise of solid-state nanopores using high electric fields. *Nanotechnology* **23**, 405301–405308 (2012).

12. Waugh, M. *et al.* Solid-state nanopore fabrication by automated controlled breakdown. *Nature Protocols* **15**, (2020).

13. Briggs, K., Kwok, H. & Tabard-Cossa, V. Automated fabrication of 2-nm solid-state nanopores for nucleic acid analysis. *Small* **10**, 2077–2086 (2014).

14. Raillon, C. *et al.* Fast and automatic processing of multi-level events in nanopore translocation experiments. *Nanoscale, Royal Society of Chemistry* **14**, (2012).

15. Saiki, R. K. *et al.* Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354 (1985).

16. Yalçinkaya, T. & Koçoglu, T. Polymerase chain reaction. *Mikrobiyoloji bülteni* vol. 26 373–378 (1992).

17. Saiki, R. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491 (1988).

18. Moller, M. J. M. and S. G. PCR: Second Edition. 1–305 (2006) doi:10.4324/9780203002674.

19. Kreader, C. A. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology* **62**, 1102–1106 (1996).

20. Pryor, R. J. & Wittwer, C. T. Real-time polymerase chain reaction and melting curve analysis. *Methods in molecular biology (Clifton, N.J.)* **336**, 19–32 (2006).

21. Kubista, M. *et al.* The real-time polymerase chain reaction. *Molecular Aspects of Medicine* vol. 27 95–125 (2006).

22. Kralik, P. & Ricchi, M. A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Frontiers in Microbiology* vol. 8 108 (2017).

23. Solomon, S. L. & Oliver, K. B. Antibiotic resistance threats in the United States: Stepping back from the brink. *American Family Physician* vol. 89 938 (2014).

24. Huggett, J. Polymerase chain reaction and infectious diseases. in *The PCR Revolution: Basic Technologies and Applications* (ed. Bustin, S. A. E.) 173–188 (Cambridge University Press, 2009). doi:10.1017/CBO9780511818974.013.

25. Watzinger, F., Ebner, K. & Lion, T. Detection and monitoring of virus infections by real-time PCR. *Molecular Aspects of Medicine* vol. 27 254–298 (2006).

26. Hahn, A. *et al.* On detection thresholds–a review on diagnostic approaches in the infectious disease laboratory and the interpretation of their results. *Acta Tropica* vol. 205 105377 (2020).

27. Lehmann, M. *et al.* Quantitative multicolor super-resolution microscopy reveals tetherin HIV-1 interaction. *PLoS Pathogens* **7**, 1002456 (2011).

28. Mekonen, A. *et al.* Factors which contributed for low quality sputum smears for the detection of acid fast bacilli (AFB) at selected health centers in Ethiopia: A quality control perspective. *PLoS ONE* vol. 13 (2018).

29. Leland, D. S. & Ginocchio, C. C. Role of cell culture for virus detection in the age of technology. *Clinical Microbiology Reviews* **20**, 49–78 (2007).

30. Ghazi, I. M. *et al.* Antibiotic utilization and opportunities for stewardship among hospitalized patients with influenza respiratory tract infection. in *Infection Control and Hospital Epidemiology* vol. 37 583–589 (2016).

31. Hu, J. *et al.* Advances in paper-based point-of-care diagnostics. *Biosensors and Bioelectronics* vol. 54 585–597 (2014).

32. Koczula, K. M. & Gallotta, A. Lateral flow assays. *Essays in Biochemistry* **60**, 111 (2016).

33. Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research* **28**, e63 (2000).

34. Piepenburg, O., Williams, C. H., Stemple, D. L. & Armes, N. A. DNA Detection Using Recombination Proteins. *PLoS Biology* **4**, 1115–1121 (2006).

35. Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nature Biotechnology* vol. 34 518–524 (2016).

36. Singer, A. *et al.* Nanopore based sequence specific detection of duplex DNA for genomic profiling. *Nano Letters* **10**, 738–742 (2010).

37. Morin, T. J. *et al.* Nanopore-Based Target Sequence Detection. *PLOS ONE* **11**, e0154426 (2016).

38. Sethi, K. *et al.* Direct Detection of Conserved Viral Sequences and Other Nucleic Acid Motifs with Solid-State Nanopores. *ACS Nano* **15**, 8474–8483 (2021).

39. Beamish, E., Tabard-Cossa, V. & Godin, M. Programmable DNA Nanoswitch Sensing with Solid-State Nanopores. *ACS Sensors* **4**, 2458–2464 (2019).

40. Beamish, E., Tabard-Cossa, V. & Godin, M. Digital counting of nucleic acid targets using solid-state nanopores. *Nanoscale* **12**, 17833–17840 (2020).

41. Squires, A. H., Atas, E. & Meller, A. Genomic pathogen typing using solid-state nanopores. *PLoS ONE* **10**, e0142944 (2015).

42. Morin, T. J. *et al.* A handheld platform for target protein detection and quantification using disposable nanopore strips. *Scientific Reports* **8**, (2018).

43. Rozevsky, Y. *et al.* Quantification of mRNA Expression Using Single-Molecule Nanopore Sensing. *ACS Nano* **14**, 57 (2020).

44. Waugh, M. *et al.* Solid-state nanopore fabrication by automated controlled breakdown. *Nature Protocols* **15**, 122–143 (2019).

45. Briggs, K. CUSUM v3.3.3 Nanopore Analysis Software Suite. (2020) doi:10.5281/zenodo.595675.

46. Weile, J. & Knabbe, C. Current applications and future trends of molecular diagnostics in clinical bacteriology. *Analytical and Bioanalytical Chemistry* **394**, 731–742 (2009).

47. Bissonnette, L. & Bergeron, M. G. Diagnosing infections—current and anticipated technologies for point-of-care diagnostics and home-based testing. *Clinical Microbiology and Infection* **16**, 1044–1053 (2010).

48. Gonzalez de Castro, D., Clarke, P. A., Al-Lazikani, B. & Workman, P. Personalized Cancer Medicine: Molecular Diagnostics, Predictive biomarkers, and Drug Resistance. *Clinical Pharmacology & Therapeutics* **93**, 252–259 (2013).

49. Hu, J. *et al.* Advances in paper-based point-of-care diagnostics. *Biosensors and Bioelectronics* **54**, 585–597 (2014).

50. Xue, L. *et al.* Solid-state nanopore sensors. *Nature Reviews Materials* **5**, 931–951 (2020).

51. Kelley, S. O. *et al.* Advancing the speed, sensitivity and accuracy of biomolecular detection using multi-length-scale engineering. *Nature Nanotechnology* **9**, 969–980 (2014).

52. Bell, N. A. W. & Keyser, U. F. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. *Nature Nanotechnology 2016 11:7* **11**, 645–651 (2016).

53. Walt, D. R. Optical Methods for Single Molecule Detection and Analysis. *Analytical Chemistry* **85**, 1258–1263 (2012).

54. He, L. *et al.* Digital Immunoassay for Biomarker Concentration Quantification using Solid-State Nanopores. *Nature Communications* **Accepted**, (2021).

55. Singer, A., Rapireddy, S., Ly, D. H. & Meller, A. Electronic barcoding of a viral gene at the single-molecule level. *Nano Letters* **12**, 1722–1728 (2012).

56. Tang, Z., Choi, G., Nouri, R. & Guan, W. Loop-Mediated Isothermal Amplification-Coupled Glass Nanopore Counting Toward Sensitive and Specific Nucleic Acid Testing. *Nano Letters* **19**, 7927–7934 (2019).

57. Cai, S. *et al.* Single-molecule amplification-free multiplexed detection of circulating microRNA cancer biomarkers from serum. *Nature Communications 2021 12:1* **12**, 1–12 (2021).

58. Burck, N. *et al.* Nanopore Identification of Single Nucleotide Mutations in Circulating Tumor DNA by Multiplexed Ligation. *Clinical Chemistry* **67**, 753–762 (2021).

59. Schappert SM & Burt CW. Ambulatory Care Visits to Physician Offices, Hospital Outpatient Departments, and Emergency Departments: United States, 2001-02. *Vital and Health Statistics* **13**, 1–66 (2006).

60. Diseases, A. A. P. C. on I. Red Book (2018): Report of the Committee on Infectious Diseases, 31st Edition. (2018) doi:10.1542/9781610021470.

61. Langlois, D. M. & Andreae, M. Group A Streptococcal Infections. *Pediatrics In Review* **32**, 423–430 (2011).

62. Church, D. L., Lloyd, T., Larios, O. & Gregson, D. B. Evaluation of Simplexa Group A Strep Direct Kit compared to hologic group a streptococcal direct assay for detection of group a streptococcus in throat swabs. *Journal of Clinical Microbiology* **56**, (2018).

63. Davenport, M. *et al.* The role of pore geometry in single nanoparticle detection. *ACS Nano* **6**, 8366–8380 (2012).

64. Ermann, N. *et al.* Promoting single-file DNA translocations through nanopores using electro-osmotic flow. *Journal of Chemical Physics* **149**, 163311 (2018).

65. Verschueren, D. v., Jonsson, M. P. & Dekker, C. Temperature dependence of DNA translocations through solid-state nanopores. *Nanotechnology* **26**, 234004 (2015).

66. Kowalczyk, S. W., Wells, D. B., Aksimentiev, A. & Dekker, C. Slowing down DNA translocation through a nanopore in lithium chloride. *Nano Letters* **12**, 1038–1044 (2012).

67. Rosenstein, J. K., Wanunu, M., Merchant, C. A., Drndic, M. & Shepard, K. L. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nature Methods* **9**, 487–492 (2012).

68. Firnkes, M., Pedone, D., Knezevic, J., Döblinger, M. & Rant, U. Electrically facilitated translocations of proteins through silicon nitride nanopores: Conjoint and competitive action of diffusion, electrophoresis, and electroosmosis. *Nano Letters* **10**, 2162–2167 (2010).

69. Plesa, C. *et al.* Fast translocation of proteins through solid state nanopores. *Nano Letters* **13**, 658–663 (2013).

70. Yusko, E. C. *et al.* Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nature Nanotechnology* **6**, 253–260 (2011).

71. Carlsen, A. & Tabard-Cossa, V. Mapping shifts in nanopore signal to changes in protein and protein-DNA conformation. *PROTEOMICS* 2100068 (2021) doi:10.1002/pmic.202100068.

72. Garibyan, L. & Avashia, N. Polymerase Chain Reaction. *Journal of Investigative Dermatology* **133**, 1–4 (2013).

73. Fologea, D., Ledden, B., McNabb, D. S. & Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Applied physics letters* **91**, 053901–1 (2007).

74. Li, W. *et al.* Single Protein Molecule Detection by Glass Nanopores. *ACS Nano* **7**, 4129–4134 (2013).

75. Bell, N. A. W. & Keyser, U. F. Specific Protein Detection Using Designed DNA Carriers and Nanopores. *Journal of the American Chemical Society* **137**, 2035–2041 (2015).

76. Gruszka, D., Schmid, S. & Dekker, C. Nanopores: a versatile tool to study protein dynamics. *Essays in Biochemistry* **65**, 93–107 (2021).

77. Madejski, G. R. *et al.* Monolithic Fabrication of NPN/SiNx Dual Membrane Cavity for Nanopore-Based DNA Sensing. *Advanced Materials Interfaces* **6**, 1900684 (2019).

78. Carlsen, A. T., Briggs, K., Hall, A. R. & Tabard-Cossa, V. Solid-state nanopore localization by controlled breakdown of selectively thinned membranes. *Nanotechnology* **28**, (2017).

79. Slinger, R. *et al.* Rapid PCR detection of group a streptococcus from flocked throat swabs: A retrospective clinical study. *Annals of Clinical Microbiology and Antimicrobials 2011 10:1* **10**, 1–5 (2011).

80. Qian, C., Wang, R., Wu, H., Ji, F. & Wu, J. Nicking enzyme-assisted amplification (NEAA) technology and its applications: A review. *Analytica Chimica Acta* **1050**, 1–15 (2019).

81. Piepenburg, O., Williams, C. H., Stemple, D. L. & Armes, N. A. DNA Detection Using Recombination Proteins. *PLOS Biology* **4**, e204 (2006).

82. Wong, Y.-P., Othman, S., Lau, Y.-L., Radu, S. & Chee, H.-Y. Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. *Journal of Applied Microbiology* **124**, 626–643 (2018).

83. Yang, S. & Rothman, R. E. PCR-based diagnostics for infectious diseases: Uses, limitations, and future applications in acute-care settings. *Lancet Infectious Diseases* vol. 4 337–348 (2004).

84. Abu Al-Soud, W. & Rådström, P. Capacity of nine thermostable DNA polymerases to mediate DNA amplification in the presence of PCR-inhibiting samples. *Applied and Environmental Microbiology* **64**, 3748–3753 (1998).

85. Wang, V., Ermann, N. & Keyser, U. F. Current Enhancement in Solid-State Nanopores Depends on Three-Dimensional DNA Structure. *Nano Letters* **19**, 5661–5666 (2019).