DNAZYME FOR NON-INVASIVE COLORECTAL CANCER DETECTION

EFFORTS TOWARDS FUNCTIONALIZING A DNAZYME FOR NON-INVASIVE COLORECTAL CANCER DETECTION

By DEVON MORRISON

B.Sc. University of British Columbia, Vancouver, BC, 2017

A Thesis Submitted to the Department of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University © Copyright by Devon Morrison, August 2020

MASTER OF SCIENCE (2020) McMaster University

(Biochemistry and Biomedical Science) Hamilton, Ontario

TITLE: Efforts Towards Functionalizing a DNAzyme for Non-Invasive Colorectal Cancer Detection

AUTHOR: Devon Morrison, B.Sc. (University of British Columbia)

SUPERVISOR: Dr. Yingfu Li

NUMBER OF PAGES: x, 81

ABSTRACT

The need for a non-invasive, accurate, easy-to-use, and cost-effective colorectal cancer (CRC) detection device is apparent in the low survival rates seen in late-stage diagnoses. Once CRC has progressed past stage I, the 5-year survival rate drops significantly, and treatment options become less favourable. The best way to treat CRC is to catch it early. The development of an RNA-cleaving fluorogenic DNAzyme (RFD) holds the potential to remediate this deficiency. A DNAzyme, called RFD-FN1, was identified from a synthetic random-sequence DNA library to selectively bind to an unknown target associated with *Fusobacterium nucleatum*, which has been found to be overabundant in pre- and cancerous colorectal tissue and stool. Target recognition by the DNAzyme induces the cleavage of a fluorogenic substrate and generates a fluorescent signal to indicate the presence of the bacterium. This thesis outlines the efforts made towards functionalizing the *F. nucleatum*-responsive probe in stool samples to create a non-invasive screening test.

RFD-FN1 is selective towards a heat-stable *F. nucleatum* protein, but its limit of detection is only 10^7 CFU/mL. Although able to detect spiked concentrations of *F. nucleatum* cells in processed stool samples, the use of heat, filtering, centrifugation, antibiotics, culturing or serial dilutions are not sufficient to detect the *F. nucleatum* that is naturally present in the diseased samples. Experiments designed to enrich the target concentration revealed that the target is not produced consistently in any growing condition tested.

Size exclusion chromatography and mass spectrometry analysis identified five potential targets that RFD-FN1 may be responding to. Three candidate targets were cloned and purified, but they failed to induce RFD-FN1's activity. Due to the COVID-19 pandemic, the purification of the

final two proteins was not completed. Purifying the two candidate targets and testing their ability to induce RFD-FN1 represents future research efforts. If the target for the DNAzyme is confirmed, a reselection for a more sensitive DNAzyme, that can function in human stool, can be attempted.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Yingfu Li, for his support and guidance throughout my post-graduate studies. Your passion for science is inspiring, and I have truly enjoyed being a part of your wonderful lab. You have given me the confidence to work independently, and you have helped me grow as a scientist and a person. I truly appreciate everything that you have done for me.

I would also like to thank my supervisory committee members Dr. Carlos Filipe, Dr. Yu Lu, and Dr. Bruno Salena for all of your support, encouragement, and insight during my studies. You all kept me motivated during the challenges that came with this project. Thank you, Dr. Lu, for your time, effort, and expertise for target identification; thank you Dr. Salena, for providing patient samples for this project. Thank you to the Surette lab for letting us use their anaerobic chamber at the start of the project; Dr. Lesley MacNeil for her Gram stain kit and microscope; and Susan McCusker for all of her help and patience with the SEC experiments.

I am so grateful for my mentors and friends in the Li Lab; this experience would not have been the same without all of you. I will miss our lunchroom talks, lab celebrations, and get-togethers. Thank you, Jim Gu, for all of your help; Jianrun Xia, for your help with collaborations and experiments; Erin McConnell, for your help and guidance with writing manuscripts. To Sandy Zakaria, my mentor, project partner, and friend throughout my graduate studies, thank you. You made every experiment, conference, presentation, and lab event fun. Thank you for making me excited to come into the lab every day (and night).

Finally, I would like to thank my mom and dad for their unconditional love and support. I would not be the person that I am today without the both of you. Thank you for always being my biggest fans and for pushing me to succeed. You are the world's best parents. *Love you lots.*

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

Δ	heat treatment
4-HB	4-hydroxybutyrate
ATP	adenosine 5’-triphosphate
BA	brucella agar
BMI	body mass index
CEM	crude extracellular mixture
CIM	crude intracellular mixture
CMB	cooked meat broth
CRC	colorectal cancer
CT colonography	computed tomography colonography
DCBE	double-contrast barium enema
ddH_2O	double-distilled water; nuclease free
DNA	deoxyribonucleic acid
DNAzyme	DNA enzyme
dNTP	2’-deoxynucleoside 5’-triphosphate
dPAGE	denaturing polyacrylamide gel electrophoresis
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
FIT	fecal immunochemical test
FNA	functional nucleic acid
FRET	fluorescence resonance energy transfer
FQ30	fluorogenic substrate
gFOBT	guaiac fecal occult blood test
IBD	irritable bowel disease
kD (kDa)	kilodaltons
LB	loading buffer
LC-MS/MS	liquid chromatography-mass spectrometry/mass spectrometry
LOD	limit of detection
nt (n)	nucleotide
OD	optical density

PCR	polymerase chain reaction
pI	isoelectric point
PI	protease inhibitor
RFD	RNA-cleaving fluorogenic DNAzyme
RNA	ribonucleic acid
SB	selection buffer
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
subsp.	subspecies
v/v	volume/volume

Chapter 1: INTRODUCTION

1.1 Colorectal cancer

Colorectal cancer (CRC) is the third most prevalent cancer in the world; it ranks second and third for cancer deaths in men and women, respectively.[1] The 5-year survival rate at stage I is 92%; that is when the cancer is most treatable because surgical excision is still effective.[2,3] However, once the cancer metastasizes to distant organs and tissues, the 5-year survival rate drops to 12%, and the treatment often becomes palliative.[2,3] Therefore, the best way to treat CRC is to detect it early, when the cancer is still localized and can be easily removed.[4]

1.1.1 Colorectal cancer screening

The most effective way to detect early stage CRC is to develop a screening method that encourages a high rate of compliance.[5] The current gold standard screening method is a colonoscopy, because of its accuracy in diagnosing CRC cases.[6] Through the insertion of a scope into the rectum and through the colon, doctors are able to get a visual perspective of the health of the tissue, and they can recognize any abnormalities.[7] During this exam, doctors have the ability to take biopsies of growths that are present and send them for pathology, and they can even completely remove polyps.[7] Despite the advantages of a colonoscopy, factors like risk, discomfort, and cost contribute to lower than recommended compliance rates.[5,8] Alternative screening methods include other direct visual examinations and stool-based tests. Like colonoscopy, direct visualization methods—flexible sigmoidoscopy, CT colonography, and double contrast barium enema—carry many of the same advantages as a colonoscopy, but reduce some of the invasiveness, bowel preparation and sedation used.[9] Unfortunately, these alternatives present their own disadvantages, including reduced scoping field and inability to remove polyps.[9] Stool-based tests—fecal occult blood test (gFOBT), fecal immunochemical test (FIT), and FIT

DNA—are more cost-effective, non-invasive, readily available and not limited by health resources, but usually come with higher false-positive rates, higher frequency of testing and poor detection of adenomas (**Figure 1**).[10]

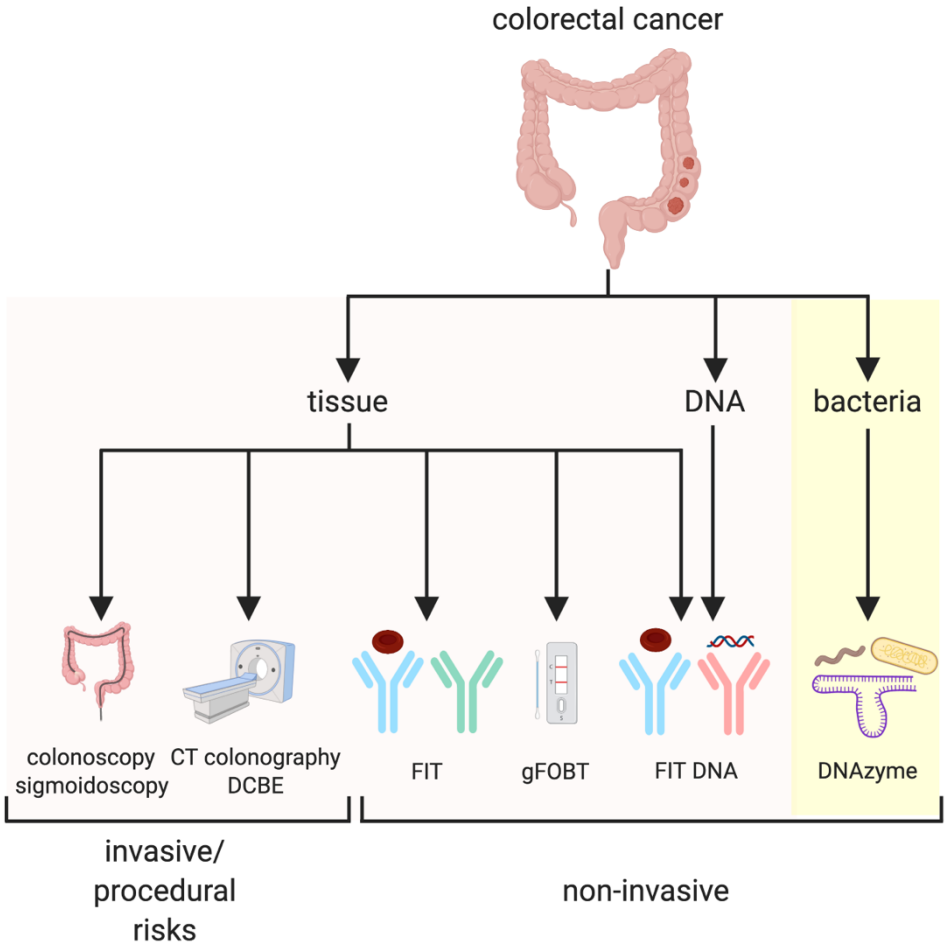


Figure 1. Schematic of CRC screening methods. Direct visual methods of CRC screening are associated with procedural risks including radiation and X-ray exposure, sedation, rigorous bowel cleansing, and intestinal perforations. Non-invasive methods detect biomarkers in patient stool samples but suffer from higher false-positive and decreased sensitivity. Current CRC biomarkers include DNA mutations and tissue distress markers like blood. The proposed DNAzyme screening test targets bacterial dysbiosis, which allows a sensitive, accurate, non-invasive alternative.

There is currently no screening method that can deliver reliable, non-invasive, cost-effective, at-home results. In an effort to increase compliance, and with it, the survival statistics of colorectal cancer, there is a need for the creation of such device.

1.1.2 Colorectal cancer risk factors

The risk factors associated with CRC have been rigorously studied in the last few decades, and it was found that irritable bowel disease (IBD) and CRC family history are two of the top factors associated with CRC development.[11] Factors that moderately increase a person's risk of developing CRC are "increased BMI [body mass index], red meat intake, cigarette smoking, low physical activity, low vegetable consumption, [and] low fruit consumption".[11] In addition to the hereditary, environmental and inflammatory risk factors listed above, there exists substantial evidence and research to support another factor—pathogens.[12]

In the last two decades, the human microbiota has received considerable attention from scientists, trying to discern its involvement in human diseases. The human microbiota is the collection of bacteria that live within the human body, which outnumbers our human cells by 10:1.[13,14] The shift from the bacteria that normally colonize a human body to more pathogenic species is referred to as bacterial dysbiosis, and these disruptions have been associated with many diseases and cancers. More specifically, bacterial dysbiosis that occurs in the gut has been linked to gastric cancers,[15] stomach ulcers,[16] IBD,[17] Crohn's disease,[18] and *Clostridium difficile* infection.[19] In the last decade, the hypothesis that CRC is also associated with disruptions in the gut microbiota has gained immense traction.[20] In addition to the over-colonization of pathogenic bacteria, the decreased abundance of commensal and mutualistic metabolic bacteria disrupts the host's cancer suppressing abilities.[20,21] Essentially, this dysbiosis interrupts a very complex interplay between bacteria and their host, and the effects of this can have serious consequences

on the health of the host.[22] Bacteria including *Streptococcus bovis*, *Enterococcus faecalis*, *C. septicum*, and *Fusobacterium nucleatum* are a few of the most common bacteria implicated in CRC development and progression.[23] However, whether these bacteria cause it, exacerbate it, or are a by-product of it, has not yet been proven.

1.2 Pathogenic *Fusobacterium nucleatum* in human disease

Fusobacterium nucleatum is a gram-negative anaerobe that was first identified as a ubiquitous member of the oral microbiome; it has been regularly detected in oral biofilms and saliva.[24,25] Notably, *F. nucleatum* has been identified as the most common oral bacterial species present in infections outside of the mouth.[26] The association of *F. nucleatum* with periodontitis, preterm birth, preeclampsia, IBD, appendicitis, cerebral aneurysm, Lemierre's Syndrome, Rheumatoid arthritis, and Alzheimer's disease, to name a few, have been investigated.[27] Although *F. nucleatum* is a commensal bacterium, certain changes in the host environment can promote an overabundance that results in pathogenicity.[27] The pathogenicity of *F. nucleatum* is largely caused by virulence factors, adhesins and invasins, that ultimately stimulate chronic inflammation. The most well-studied virulence factor, FadA, binds to E-cadherin on the intestinal epithelial cells to activate the ß-catenin signalling pathway and compromises the integrity of the epithelial barrier.[28] Thanks to these virulence factors, *F. nucleatum* is a proficient pioneering species during the formation of biofilms.[26] *F. nucleatum* is responsible for promoting associations between gram-positive and gram-negative species to expand the biofilm, and to also create an environment that is oxygen-deficient for other anerobic pathogens to proliferate.[26,29] The remodeling of the microbiome can drive pro-inflammatory conditions and might be a key player in a multitude of other human diseases as well.[30]

1.2.1 F. nucleatum in colorectal cancer

The association between the dysbiosis of *F. nucleatum* in the human gut and colorectal cancer has been the focus of research for many laboratories over the past 10 years. More recently, scientists are providing evidence to support that *F. nucleatum* contributes to the disease progression of CRC with the help of its virulence factors.[29] Rubinstein *et al*.[28] noted the concurrent increase of FadA, and subsequently *F. nucleatum*, across the progression of healthy colonic tissue, to adenomas, to carcinomas, and that FadA levels are around 10−250× higher in adenoma and carcinoma tissue than in healthy patient tissue.[31] Another well-studied correlation is the increase of *F. nucleatum* and the concurrent decrease of 4-hydroxybutyrate (4-HB) in colorectal cancer patients.[32] The metabolite, 4-HB, is involved in butyrate-production, which has been found to work in tumour-suppressing and anti-inflammatory pathways.[33] *F. nucleatum* has been predicted to interfere with this pathway, and possibly contribute to the progression of the tumour microenvironment. Other studies suggest that the enrichment of *F. nucleatum* may simply be that the opportunistic pathogen colonizes in the favourable conditions created by the colonic tumour.[34] Interestingly, the increased abundance of *F. nucleatum* in carcinomas compared to normal tissue is also observed in adenomas, indicating that the dysbiosis is present in the pre-cancerous stages as well.[29] Although the role of *F. nucleatum* in cancer cannot be confidently confirmed, the overabundance of this pathogen in diseased tissue, in many cases, is certain. Therefore, targeting this bacterium as a biomarker for CRC detection is promising, especially because it is also detectable in the early stages, when the prognosis is most favourable. Bacteria present their own set of advantages for use as biomarkers, compared to genomic or cellular targets, and with the right system, create a valuable detection assay. Therefore, the

detection of bacterial species using functional nucleic acids has become increasingly attractive in recent years.

1.3 Functional nucleic acid biosensors

In 1989, Thomas Cech and Sidney Altman were awarded the Nobel Prize in Chemistry for their work on 'the discovery and catalytic properties of RNA'.[35] This novel breakthrough spurred researchers to continue their efforts to select for nucleic acids that could perform functions, much like proteins. Functional nucleic acids (FNAs) are a specialized subset of molecules that exist within the nucleic acid world, but whose functions exist outside of the preeminent genetic role.[36] A defining property of FNAs is that they are single stranded and are therefore capable of folding into complex secondary structures that allow them to bind to targets and perform catalysis.[37] FNAs are divided into two main categories: molecules that selectively bind targets, aptamers and riboswitches; and molecules that perform a catalytic function upon cognate target recognition, ribozymes and deoxyribozymes (DNAzymes).[38] Aptamers and riboswitches are the nucleic acid version of antibodies, and their affinity for targets like other nucleic acids, proteins, protein fragments, toxins, metal ions and small molecules gives them a competitive edge.[39] DNAzymes and ribozymes are the nucleic acid version of a protein enzyme. They can be engineered to perform a number of catalytic activities upon target recognition, supported here by studies that detail porphyrin metallation,[40] RNA/DNA-cleavage,[41,42] RNA/DNA-ligation,[43,44] DNA phosphorylation,[45] DNA peroxidation,[46] DNA glycosylation,[47] and amide hydrolysis.[48] These activities have led to the use of DNAzymes in novel, complex systems like *in vivo* gene silencing,[49] live cell imaging,[50] tumour cell capture,[49] and neurodegenerative disease progression studies.[51] The combination of an aptamer and a nucleic

acid enzyme system creates an FNA subset called aptazymes[52]; by amalgamating two independent FNAs, researchers can fine-tune the most desirable properties, and tailor the output result to a very specific need.[53,54] The field of functional nucleic acids continues to evolve, with new systems constantly being pioneered for biosensing, drug discovery and therapeutic applications.[55-58]

1.3.1 RNA-cleaving fluorogenic DNAzyme

For the purpose of biosensing, DNAzymes are one of the more attractive molecular options. DNA is an extremely stable molecule compared to proteins and RNA; it can still function after being heated, frozen, put into solution, immobilized onto a surface, and chemically modified.[59] Additionally, DNA is simpler and considerably less expensive to prepare than protein enzymes.[60] The cost of synthesizing DNA is approximately one thousandth of the cost of antibodies and does not suffer from batch-to-batch variation.[59] However, much like proteins, precise target recognition is all that is required to trigger a specific catalytic activity. Thanks to their enzymatic abilities, DNAzymes can be constructed to possess an intrinsic signal transduction system. After target recognition, a cascade of reactions produces a signal which would indicate the presence of the cognate target to the user. One well-studied class of biosensing DNAzymes is called RNA-cleaving fluorogenic DNAzymes (RFD) (**Figure 2**).[61]

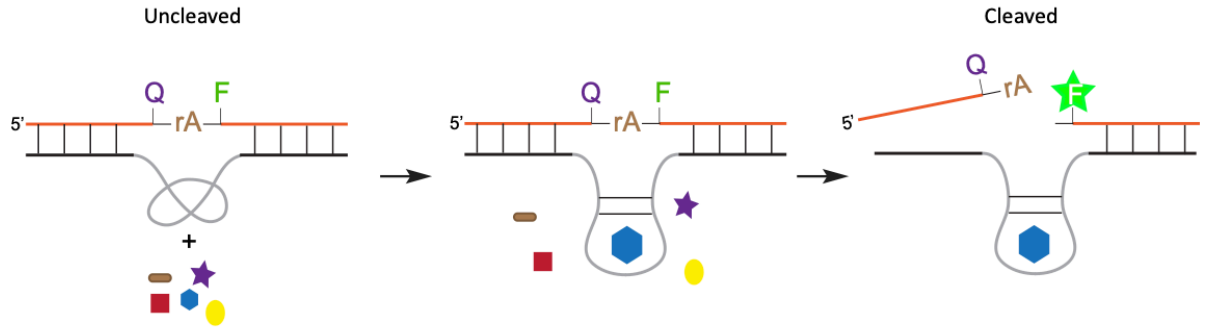


Figure 2. Schematic of *trans* RNA-cleaving fluorogenic DNAzyme (RFD). The *trans*-form of the DNAzyme is not ligated to the fluorescent substrate, it hybridizes in solution. The fluorescent substrate, in

orange, contains an adenosine ribonucleotide that separates a fluorescein (F)-modified deoxyribothymidine and a DABCYL (Q)-modified deoxyribothymidine. Upon target recognition, the DNAzyme undergoes a conformational change that cleaves the substrate at the ribonucleotide site, releasing the quencher fragment and generating an increase in fluorescence. The fluorescence intensity is measured at 521 nm.

In this system, a fluorogenic DNA substrate is used for signal generation. The substrate contains a single RNA base that separates a fluorophore molecule from a quencher. In close proximity, the quencher absorbs the majority of the fluorescence that the fluorophore gives off due to fluorescence resonance energy transfer (FRET).[62] The dissociation of the quencher-modified oligonucleotide fragment, after RNA cleavage, generates an increase in fluorescence that can be observed. An RNA linkage within the DNA substrate is used because it is ~100 000 times less stable than its deoxy counterpart.[63] The 2'-hydroxy group on the ribose is in an ideal position to perform a nucleophilic attack on the adjacent phosphodiester bond.[64] Most RNA-cleaving DNAzymes capitalize on this favourable mechanism to direct their DNAzyme's catalysis.[65] The donation of electrons, by the 2'-hydroxy to the phosphate, releases the 5'-hydroxy of the neighbouring deoxyribonucleotide. The result of this reaction is the generation of two separate DNA strands, with a 2'-3'-cyclic phosphate terminus on the ribonucleotide. Target recognition is what induces the conformational change needed to facilitate RNA-cleavage. As mentioned earlier, the single stranded state of FNAs is what allows them to adopt complex secondary structures and interact with an infinite number of molecules.

1.3.2 In vitro selection library design for functional nucleic acids

In vitro selection is the most common method used to select for DNAzymes. It is also often referred to as SELEX (Systematic Evolution of Ligands by EXponential Enrichment) because of its unique discovery. Discovered simultaneously by three independent labs—Tuerk & Gold,[66] Ellington & Szostak,[67] Robertson & Joyce[68]—*in vitro* selection proved to be a logical synthetic mimic for nature's evolutionary selection of enzymes. Although there have been a few

alternatives and advancements to the technique, it remains largely the same, reinforcing the power of this method to select for FNAs. The theory behind *in vitro* selection is that in the vast amount of possible DNA sequences, there should be, at the very least, one that is able to bind to a specific target and undergo the correct conformational change that enables it to perform a catalytic activity.[67,69] A selection scheme is optimized in order to pull out this perfect sequence. To select for an RFD, the first step is to start with a randomized library of DNA sequences that are all flanked by a specific primer pair, necessary for the amplification step later. In order to create the starting library, a balance between the sequence coverage and structural diversity must be achieved. The starting library is limited by the amount of DNA that can be synthesized; the practical limit for the number of DNA sequences in *in vitro* selection experiments is ~10^{16} sequences.[70] The sequence coverage is calculated by 4^n where n is the number of nucleotides within each DNA sequence; in order to create a library that covers every nucleotide possibility, the random sequence becomes limited to 28 nucleotides in length.[71] While small random region libraries still produce functional DNAzymes, having longer lengths does increase the probability of finding a successful, novel DNAzyme.[72,73] Degeneracy, caused by there being only four nucleosides, can contribute to redundancy in secondary structures—vastly different sequences can form the same structure—and, therefore, it is not completely necessary to achieve complete coverage of a random library.[72] Another instance where highly similar structures are created is within an area containing many G-nucleotides; these G-rich regions can form into highly ordered G-quadruplex structures which are even more stable than double-stranded, hybridized DNA structures.[74] There are numerous examples of these well-known structures occurring in selections.[75] For the best chance of finding an optimal DNAzyme, the library should represent the largest sample of all the possible sequence combinations that is within reason. Consensus

suggests that a library containing 10^{14}–10^{16} (10^{14} = 200 pmol) sequences, or 100−1000 pmol, with 30-120 nucleotides is a reasonable compromise between sampling sequence diversity and coverage of the sequence space.[71,72,76,77]

1.3.3 Selection of an F. nucleatum-responsive DNAzyme

For an RFD selection, the library is ligated to the fluorogenic substrate at the beginning of each round (**Figure 3**).

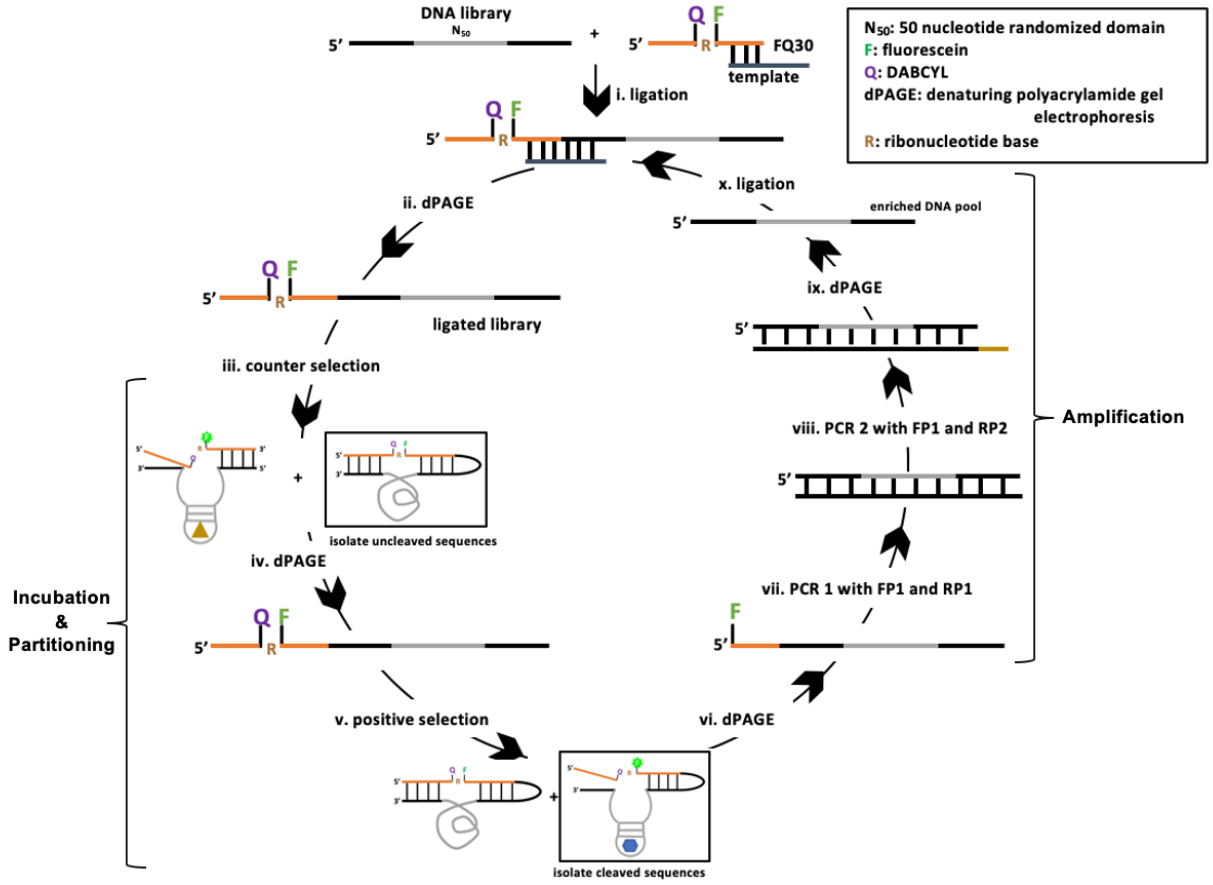


Figure 3. Schematic of RFD *in vitro* selection. The iterative process begins with ligating the DNA library to the fluorescent substrate. After incubating the library with the positive or negative (counter) selection mixtures, the desired DNAzymes are partitioned away from the unwanted sequences. The DNAzymes are enriched by PCR amplification, and re-ligated to fluorescent substrate to participate in another round.

The fluorogenic substrate is an important tool, both during and after the selection: it identifies DNAzymes from the pool of DNA sequences; separates positive from non-specific DNAzymes; contributes to the secondary structure of the DNAzyme for specificity; and signifies the presence and relative abundance of the specific target in a matrix. After ligation are the selection steps, which are used to identify and isolate sequences that are activated by the target of interest. In the selection designed to isolate an *F. nucleatum*-specific DNAzyme to identify CRC, two different selection steps were employed.[78] The first was counter-selection; the DNA library was incubated in a solution containing anything that the intended DNAzyme should not react to, including commensal bacteria, the selection buffer used, and healthy patient samples. Any molecules that were activated in the presence of these negative targets were discarded. The positive selection that followed was an incubation with the intended target, *F. nucleatum* subsp. *nucleatum*. Active sequences were isolated and enriched through multiple rounds until the remaining library was dominated by sequences with the ideal phenotype. Enrichment was done by PCR amplification facilitated by the primer binding sequences that flank the random nucleotide region. The enriched library was re-ligated to the fluorogenic substrate at the start of each round. In selection procedures, certain variables can be changed in each round in order to challenge the library; typically, the goal is to isolate the DNAzyme with the highest affinity to the specific target and the fastest cleavage rate. Challenging the library can be done by decreasing the incubation time and target concentration in order to enrich for the most competitive candidate—the one with the highest binding efficiency and subsequent cleavage.[79]

An interesting advantage to this selection procedure is that a specific biomarker does not have to be predetermined for the selection of the DNAzyme. By incubating the DNA library with bacterial culture components, the *in vitro* selection essentially choses the best biomarker: the

cellular component that shares the highest affinity with one of the oligonucleotide enzymes present. Counter-selection is used to ensure that the biomarker is specific to the bacterium of interest, and that the DNAzyme is not reacting to a molecule in other bacteria or solutions.

1.4 Detection in biological matrices

One of the more difficult challenges in this project is the detection of the bacterial target in the complex biological matrix, stool. The literature is considerably more sparse here than for DNAzymes that function in a simpler media. However, the last five years has seen a few labs focus their interests on using DNAzymes for sensing in complex matrices. Recent studies have applied RNA-cleaving DNAzymes to detect bacterial, metal ion, or nucleic acid targets in drinking/pond/lake water,[80-82] blood,[83] serum,[84] and stool.[85] DNA has shown promising stability in these environments despite the presence of nucleases, high and low pHs, metal ions, small molecule inhibitors, nucleic acid binding proteins, and chemicals.

1.4.1 Detecting F. nucleatum in stool to indicate colorectal cancer

F. nucleatum is a promising candidate for CRC screening in stool. The overabundance of the bacterial species that is observed in cancerous tissue persists in the stool samples of the patients as well. In patients with confirmed colonic adenomas or carcinomas, *F. nucleatum* levels are significantly higher in the stool samples, compared to healthy controls.[29,86] We hypothesize that we can select for a DNAzyme that can respond to a molecular signature—a protein, carbohydrate, small molecule or metabolite—that is associated with the *F. nucleatum* present in stool. We can later investigate the identity of the bacterial biomarker that induces cleavage, if necessary. Our lab has demonstrated the successful identification of the target that activates a *C. difficile*-responsive DNAzyme, RFD-CD1.[87]

The use of stool samples as a non-invasive detection matrix has proven to be successful in instances where the biomarker is separated from the stool material, like DNA extraction and PCR.[88] However, since the ultimate goal of this project is to develop a screening test that can be done non-invasively at home, we wanted to push the boundaries and develop a system that functions directly in the stool sample. A previous graduate student in the Li Lab investigated the feasibility of selecting for an RFD by using a CRC-positive stool sample as the positive selection matrix. To drive the specificity of the probe, healthy stool samples were used as the counter selection matrix. Two independent attempts failed to produce a DNAzyme that could differentiate between the cancerous and non-cancerous controls.[78] In the first selection, the candidate DNAzymes experienced non-specific degradation, likely due to the nucleases present in the stool. The second selection employed heat to inhibit the nucleases, but the candidate DNAzymes targeted a constituent of the selection buffer, rather than a molecular signature in the stool. A third selection, briefly described earlier, was successful in isolating a DNAzyme, DT4, that was responsive to *F. nucleatum*, although it was unable to indicate the presence of the bacterium in a CRC stool sample; the selection was performed with a crude extracellular mixture (CEM) of *F. nucleatum* culture instead of stool. In an effort to create an at-home, non-invasive screening test for CRC, we focused on functionalizing this DNAzyme in the stool samples of CRC-positive patients.

1.5 Thesis objectives

The goal of this thesis is to develop a DNAzyme that can screen for colorectal cancer in stool samples. The project focuses on using an *F. nucleatum*-responsive DNAzyme, that was selected in our lab, to detect an overabundance of *F. nucleatum* in CRC-patient stool samples.

Initial experiments showed that the DNAzyme was not inherently functional in stool, so we chose to investigate 1) optimizing the stool environment to promote DNAzyme:target interaction and 2) enriching the target in the stool to encourage detection. The second focus of the project is to identify the target that the DNAzyme is activated by. The target identity would allow for reselection of a more sensitive DNAzyme and provide us with an interesting *F. nucleatum* biomarker to explore for future experiments.

Chapter 2: MATERIALS AND METHODS

2.1 Bacterial strains, culture media, and growth conditions

2.1.1 Bacterial strain and culture media

Fusobacterium nucleatum subsp. *nucleatum* (ATCC 25586), purchased from American Type Culture Collection (ATCC, Manassas, USA), was cultured and maintained in our laboratory. The -80°C freezer stock was stored in a final concentration of 10% skim milk. Cooked Meat Broth (CMB) powder and Brucella Agar with hemin and vitamin K1 were purchased from Sigma-Aldrich (St. Louis, USA), and defibrinated sheep blood was purchased from Cedarlane (Burlington, CA). Water (ddH_2O) was purified with a Milli-Q Synthesis A10 purification system (EMD Millipore, Massachusetts, USA) and autoclaved for use in reactions and buffers.

2.1.2 Bacterial growth conditions

F. nucleatum colonies were streaked onto Brucella Agar plates supplemented with 5% defibrinated sheep blood (BA) and incubated inside a BactronEZ Anaerobic Chamber (Sheldon Manufacturing Inc., Cornelius, USA) at 37°C for 48 hours under anaerobic conditions (80% N_2, 10% CO_2, 10% H_2). All media was pre-reduced in the anaerobic chamber overnight, before culturing. Liquid cultures were inoculated from an isolated colony and grown to an optical density (OD_{600}) of 0.7; often, a subculture step was required to reach this mid-log phase. The subculture was incubated for 48 hours, unless otherwise stated.

2.1.3 16S rRNA sequencing for bacterial identification

F. nucleatum culture was grown to 0.7 OD_{600} for 16S rRNA sequencing to confirm its identity. From the cultures, 200 µL was centrifuged in a refrigerated compact micro centrifuge (VWR, Radnor, USA) at 9300 g, at 4°C for 10 min. The pellet was resuspended in 100 µL of 5% Chelex-100 (BioRad, Hercules, USA) and 5 µL of Proteinase K (800 U/mL; New England Biolabs, Ipswich, USA). The slurry was placed in a thermoshaker (VWR, Radnor, USA) at 1000 rpm, at 55°C for 30 min, and then at 95°C for 15 min. The suspension of crude DNA extraction was used directly in PCR amplification. The reaction components (Qiagen, Germantown, USA) were template DNA (1 µL), 10× PCR buffer (5 µL), 2 mM dNTP mix (5 µL), 10 µM 16S-8f (1 µL), 10 µM 16S-926r (1 µL), 0.25 U Taq DNA polymerase (0.25 µL), and ddH_2O (to 50 µL). Note that the sequences of all the oligonucleotides are provided in Table 1 below. The PCR was first run at 95°C for 5 min, then 30 cycles of 95°C for 30 s, 61°C for 30 s, 72°C for 45 s, and finally, at 72°C for 2 min. To verify proper amplification, 5 µL of PCR product was mixed with 1 µL of 6× loading dye (50% glycerol, 100 mM EDTA pH 8.0, 1% SDS (w/v), 0.01% bromophenol blue (w/v), 0.01% xylene cyanole FF (w/v)) and run on a 2% agarose gel. The fluorescence was measured on an Amersham Typhoon 9410 Imager (GE Healthcare Life Sciences, Marlborough, USA). The amplification product of *F. nucleatum* subsp. *nucleatum* is 893 base pairs. After confirmation of amplification, 5 µL of primer (1 µM) and amplification product (9 ng/µL) were sent to McMaster MOBIX lab (Hamilton, Canada) for Sanger sequencing.

2.1.4 Gram stain for bacterial identification

For a rapid, rudimentary method to confirm the identity of *F. nucleatum* colonies, a Gram stain was performed. *F. nucleatum* cells were heat fixed to a glass slide and stained with crystal violet stain for 1 min. The stain was rinsed with water, flooded with iodine mordant for 1 min, and again rinsed with water. Decolorization was conducted with 95% ethanol (v/v) for 10 s, and the cells were counterstained with Safranin and rinsed with water. The Gram stain was imaged on a Nikon Eclipse II microscope. *F. nucleatum* Gram-stains as pink, long, spindle-shaped rods.

2.2 Preparation of crude extracellular mixture (CEM) and crude intracellular mixture (CIM)

F. nucleatum culture was grown to an OD_{600} of 0.7 in the anaerobic chamber and brought out into aerobic conditions. Next, 1 mL volumes of culture were pelleted at 10 000 g, at 4°C for 10 min, and the supernatant was recovered and filtered through a luer lock 0.22 µm syringe filter (VWR, Radnor, USA). The filtered supernatant was henceforth referred to as the crude extracellular mixture (CEM). To prepare the crude intracellular mixture (CIM), the pelleted cells were resuspended in 150 µL 1× selection buffer (SB) (50 mM HEPES pH 7.5, 150 mM NaCl, 15 mM $MgCl_2$, 0.01% Tween 20). The resuspended pellet was heated at 90°C for 15 min and centrifuged at 11 000 g, at 4°C for 5 min. The recovered supernatant is the CIM. Both the CEM and CIM were flash frozen in liquid nitrogen and stored at -80°C, if not used immediately.

2.3 RFD constructs and reporter assays

2.3.1 Oligonucleotides

DNA oligonucleotides were purchased from IDT DNA Technologies (Coralville, USA), and the fluorescent reporter substrate FQ30 was purchased from Keck Oligo Synthesis Facility (Yale University, New Haven, USA). The oligonucleotides were synthesized with solid phase-phosphoramidite addition synthesis and purified by 10% denaturing (8M urea) polyacrylamide gel electrophoresis (dPAGE). DNA concentration was calculated based on OD_{260} and measured using a NanoVue Plus (GE Healthcare) UV-vis spectrophotometer. The oligonucleotide sequences are listed below.

Table 1: Sequences of oligonucleotides used in this study

Name	Sequence (5' – 3')
DT4 (39 nt)	GGT AGT GAT AAA TTT TAA TTT TTC ATA TAT TGA GTT CAT
tRFD-FN1 (44 nt)	CTT GGT AGT GAT AAA TTT TAA TTT TTC ATA TAT TGA GTT CAT AG
RFD-FN1 (73 nt)	CTA TGA ACT GAC QRF GAC CTC ACT ACC AAG CCG GTA GTG ATA AAT TTT AAT TTT TCA TAT ATT GAG TTC ATA G
cRFD-FN1 (43 nt)	CCG GTA GTG ATA AAT TTT AAT TTT TCA TAT ATT GAG TTC ATA G
FQ30 (30 nt)	CTA TGA ACT GAC QRF GAC CTC ACT ACC AAG
LT1 (30 nt)	ATT TAT CAC TAC CGG CTT GGT AGT GAG GTC
16S rRNA Forward Primer; 16S-8f (20 nt)	AGA GTT TGA TCC TGG CTC AG
16S rRNA Reverse Primer; 16S-926r (20 nt)	CCG TCA ATT CCT TTR AGT TT

Note: F: fluorescein-dT, Q: DABCYL-dT, R: adenine ribonucleotide

2.3.2 RFD-FN1 generation

The *cis*-acting RFD-FN1 was generated by ligating FQ30 and cRFD-FN1.

Phosphorylation of *c*RFD-FN1: The phosphorylation reaction contained 900 pmol of cRFD-FN1, 30 units of T4 polynucleotide kinase (PNK; ThermoFisher Scientific, Massachusetts, USA), and 1× PNK buffer A (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine) and 5 mM ATP. The reaction volume was 60 µL. The mixture was incubated at 37°C for 1 h and heated at 90°C for 5 min to deactivate PNK. An ethanol precipitation was performed to recover the DNA. The pellet was resuspended in 50 µL of ddH_2O.

Ligation of FQ30 to cRFD-FN1: The reaction contained 900 pmol of phosphorylated cRFD-FN1 prepared above, 900 pmol of FQ30, 900 pmol of LT1, 15 units of T4 DNA ligase and 1× ligation buffer (40 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP). The reaction volume was 400 µL. A pre-reaction mixture containing cRFD-FN1 and FQ30 was heated at 90°C for 2 min and cooled to room temperature before the addition of the ligation buffer and T4 DNA ligase. The entire reaction mixture was then incubated at room temperature for 4 h. The DNA was recovered by ethanol precipitation, and the ligated construct was purified on 10% dPAGE. The final concentration of RFD-FN1 was determined by measuring the OD_{260} on the Nanovue Plus. As-prepared RFD-FN1 probe was stored at -20°C.

2.3.3 DNAzyme cleavage assay

The DNAzyme cleavage assay was used to identify the presence of *F. nucleatum* in test samples. Reaction volumes were 30 µL, unless otherwise stated, and contained 15 µL 2× SB, 13.8 µL of sample/ddH_2O, and 1.2 µL of 1.02 µM DNAzyme probe. For the *trans*-acting DNAzyme cleavage probe, tRFD-FN1 and FQ30 were mixed in a 50:1 ratio, and in later experiments, at a 1:1 ratio. The change of ratios did not have any effect on cleavage activity. To

set up the cleavage assay, the DNAzyme was heated at 90°C for 1 min, cooled at room temperature for 2 min, and added to the other reaction components. The reaction mixture was incubated at room temperature for 4 h and quenched using a final concentration of 69 mM EDTA. The reactions were either stored at -20°C before continuing, or ethanol precipitated immediately. Ethanol precipitation was performed by adding 2.5 volumes of chilled 100% EtOH to the reaction mixture, and chilled at -20°C for 30 min. The mixture was centrifuged at 21 000 g at 4°C for 20 min. The supernatant was discarded, and the DNA pellet was resuspended in 2.5 volumes of chilled 70% EtOH and centrifuged at 21 000 g at 4°C for 15 min. The supernatant was discarded, and the pellet was dried in a speedvac. The DNA was resuspended in 13 µL ddH_2O and 13 µL 2× loading buffer (LB; 0.5 mM sucrose, 0.9 mM Tris base, 0.9 mM boric acid, 10 mM EDTA, (10% w/v) SDS, 0.4 µM Bromophenol Blue, 0.5 µM Xylene Cyanole FF, 18 mM urea), heated at 90°C for 2 min, and 11 µL of the resuspension was run on a 10% dPAGE gel. The fluorescent DNA bands in the gel were imaged on an Amersham Typhoon 9410 Imager and quantified using Image Quant 5.2. All DNAzyme cleavage assays were performed using tRFD-FN1 unless otherwise stated.

2.3.4 DNAzyme cleavage assay variations

The majority of the experiments performed were variations of the DNAzyme cleavage assay. In order to test the characteristics of the probe, and learn more about the unknown target, certain experimental conditions were modified in different assays. All assays follow the same experimental protocol and any modifications made to them are detailed. The experiments are listed as follows: (A) *Heat stability;* CEM and CIM were heated on a heat-block at 90°C for 5, 10, 20, 30 and 60 min before cooling at room temperature for 10 min and tested. (B) *Size*

filtration; to determine the approximate size of the *F. nucleatum* target that tRFD-FN1 recognizes, CIM was filtered through different sizes of NanoSep Omega Amicon Ultra centrifugal filters (Pall Corporation, Port Washington, USA). For each MW cut-off column—100, 50, 30, 10, 3 kD—200 µL of CIM was filtered using a centrifuge at 3500 g. The filtrates were collected for the DNAzyme cleavage assay. (C) *Sensitivity of tRFD-FN1;* to test the limit of detection of tRFD-FN1, *F. nucleatum* culture was grown to 0.7 OD_{600} and serially diluted 10-fold. For each dilution, 100 µL was plated in triplicate on a BA plate and incubated anaerobically at 37°C for 48 h. The DNAzyme cleavage assay was performed on each dilution, and the number of cells was calculated from the plate count. The bacterial concentration was determined using the dilution that contained between 30 and 300 colony forming units (CFUs). (D) *Proteinase K;* to test if the unknown target was a protein, Proteinase K was used to degrade any proteins present in the CIM. A working concentration of 870 µg/mL of Proteinase K was incubated with the CIM at room temperature for 15 min. After incubation, the CIM was tested with the DNAzyme cleavage assay. (E) *RNase inhibitor;* this assay was used to determine if the DNAzyme responds to an RNase present in the *F. nucleatum* culture. To start, 1 µL of Invitrogen SUPERase•In (20 U/µL; ThermoFisher Scientific, Massachusetts, USA) and NxGen RNase inhibitor (40 U/µL; Lucigen, Middleton, USA) were each incubated with three different CIM samples at room temperature for 15 min. SUPERase inhibits RNases A, B, C, I and T1, and NxGen inhibits RNases A, B and C. After incubation, the DNAzyme cleavage assay was performed. (F) *Culture shaking during growth;* four sets of *F. nucleatum* cultures were grown in the anaerobic chamber and were shaken vigorously by hand every 1, 2 or 4h. One set of cultures were grown without shaking. Cleavage activities and OD_{600} measurements were taken at 0, 10, 16, 20, 24, 28, 32, 40 and 48 h post-subculture. At each timepoint, 500 µL was removed to

measure the OD_{600} and 1000 µL was used to prepare CIM. The DNAzyme cleavage assay was performed on the CIM.

2.3.5 Target degradation analysis

To test the stability of tRFD-FN1's target at -20°C storage conditions, a 12-month timepoint analysis was performed. *F. nucleatum* was grown to 0.715 OD_{600} and prepared into 300 µL of CEM and CIM. The CEM and CIM samples were flash frozen with liquid nitrogen and immediately stored at -20°C. At 1-day, 1-week, 1-month, 3-months and 1-year post-harvest, the CEM and CIM tubes were removed from the freezer, defrosted, and an aliquot was removed for the assay. The tubes were returned to -20°C storage. The same stock of 1.02 µM tRFD-FN1 was used and returned to -20°C storage during each timepoint.

2.4 Stool samples

2.4.1 Stool sample biobank

Healthy and colorectal cancer (CRC) patient samples were collected by Dr. Bruno Salena, a Gastroenterologist at McMaster University Medical Centre (Hamilton, Canada). Patients underwent a colonoscopy to confirm pathology of the stool specimen. Stool samples were collected by patients and brought to the endoscopy clinic, which were then delivered to the laboratory and stored at -80°C, unless otherwise stated.

2.4.2 Processed stool samples

Healthy and CRC stool samples were processed to help with DNAzyme cleavage experiments. Eight CRC and eight healthy control stool samples were defrosted from -80°C

storage, and 5 g of each sample was aliquoted into separate falcon tubes. Next, 10 µL of 1× SB and 1× cOmplete EDTA-free protease inhibitor cocktail mixture (Roche Diagnostics, Indianapolis, USA) were added to each stool sample. The samples were rigorously vortexed to break apart the fecal matter and were placed on a shaker at room temperature for 2 h. Each sample was centrifuged at 5000 g at room temperature for 15 min to collect the aqueous fraction of the supernatant without large particulate. The centrifugation and supernatant collection steps were repeated three times to collect as much processed stool as possible. Stool aliquots of 1 mL were mixed with 500 µL of 40% glycerol to create 20% glycerol stocks. The samples were flash frozen with liquid nitrogen and stored at -80°C until use.

2.4.3 Solid stool samples

Stool samples that were not processed were also assayed to determine if tRFD-FN1 can recognize its target in the solid stool sample state. To prepare, 0.5g of stool sample was mixed in 4.5 mL of CMB in order to be able to pipette and work with the samples.

2.4.4 Stool sample assay conditions

Similar to the various conditions tested on *F. nucleatum* cultures and CIM, a variety of different assays were conducted on the stool samples to promote recognition of target by the DNAzyme. All reactions were performed using tRFD-FN1 and the standard DNAzyme cleavage assay protocol with stool as the sample. The differences in each assay are as follows: (A) *Heat-sterilized stool samples;* stool samples are weighed out and diluted 10% w/v into CMB. The stool samples are then sealed and heated in a 90°C water bath for 2 h. Sterilized stool samples were then transferred into the anaerobic chamber at 37°C for further manipulation, bacterial

spiking, dilutions or tests. (B) *CEM/CIM or culture-spiked stool;* in 30 µL reactions, the respective percentages of stool—0, 5, 10, 20, 40%; processed or unprocessed—were added to 13.8 µL of *F. nucleatum* CEM, CIM or culture in the DNAzyme cleavage assay. (C) *Stool culture;* stool samples were heat-sterilized for 2 h and brought into the anaerobic chamber. Five colonies of *F. nucleatum*, grown for 48 h on BA plates in the anaerobic chamber, were inoculated into the 10% w/v stool samples. After the inoculated stool samples were incubated for 48 h, 500 µL volumes were serially diluted into 4.5 mL to 10^{-8}. Prior to the DNAzyme cleavage assay, each dilution was heated at 90°C for 15 min. The cleavage reaction was allowed to proceed for 8 h. (D) *Fresh stool assay;* to determine if the target is better recognized by tRFD-FN1 in fresh stool samples, an experiment was performed on a freshly delivered specimen. Four tubes of 0.5 g of stool were mixed with 4.5 mL of CMB, with 2× protease inhibitor (PI) added to two of the mixtures. From one of each mixture (+/- PI), 0.5 g was serially diluted into 4.5 mL to 10^{-7} and incubated anaerobically for 48 h. The remaining two mixtures (+/- PI) were inoculated with five *F. nucleatum* colonies and incubated for 48 h anaerobically, before being serially diluted like above. All samples were heated at 90°C for 10 min before they were assayed with tRFD-FN1.

2.4.5 Lyophilization of stool samples

Healthy stool samples were resuspended 10% w/v in 1× SB without Tween 20 to a final volume of 10 mL. Three of the four stool resuspensions were heated at 90°C for 2 h; one of the two CIM samples was heated at 90°C for 15 min. Each stool mixture was poured into a BA 6040/STR strainer bag (Seward Limited, Worthing, UK) and homogenized in a Stomacher 80 Biomaster (Seward Limited, Worthing, UK) for 10 s. Stool filtrate was transferred to sterilized

glass vials, in 1 mL aliquots, for the stool and heated stool controls. For the spiked assays, 500 µL of stool filtrate and 500 µL of the heated or unheated CIM were mixed. All samples were completely frozen at -80°C and transported to the FreeZone Plus 4.5 L Cascade benchtop freeze dry system (Labconco, Kansas City, USA) on dry ice. All caps were removed, and the samples were placed into a glass jar and attached to the vacuum lid to pressurize. The samples were retrieved when completely dry after 24 h. The stool samples were resuspended in 100 µL of ddH_2O, and the CIM samples were resuspended in 20 µL. The DNAzyme cleavage assay was performed on the 10× concentrated stool and 5× concentrated stool.

2.4.6 Kanamycin, vancomycin assay

Two sets of 5 mL cultures were initiated with an isolated colony of *F. nucleatum.* Kanamycin, vancomycin, and a combination of both antibiotics were added to one set of the cultures at the 0 h timepoint. To the 5 mL cultures, 5 µL of 50 mg/mL stock concentrations were added to achieve a working concentration of 50 µg/mL. All cultures were incubated anaerobically at 37°C, with timepoints being taken at 0, 8, 24 and 48 h. The antibiotics were added to the second set of cultures at 24 h incubation. The DNAzyme cleavage assay was performed as per protocol. For the stool test, 1mL of *F. nucleatum* culture was added to 1 mL of 20% w/v stool-CMB, and 2 µL of 25 mg/mL kanamycin and vancomycin were added for a working concentration of 50 µg/mL. The stool cultures were then incubated for a total of 48 h, with timepoints taken at 0, 8, 24 and 48 h. The DNAzyme cleavage assay was performed as described above.

2.5 Saliva samples

2.5.1 Saliva-DNAzyme cleavage assay

To test tRFD-FN1 in saliva samples, three different saliva specimens were collected in separate, autoclaved 2 mL tubes. In each reaction, 11 µL of saliva sample was used, and 1.7× final concentration protease inhibitor (1 µL) was added to half of the reactions. To one set of the saliva samples from each specimen, 13.8 µL of *F. nucleatum* CIM was spiked in. The reactions were all 30 µL total volume, and 10× SB was used instead of 2× SB.

2.5.2 Heated saliva assay

The first specimen sample from the above saliva assay was used to test target recognition in the complex after heating. The reactions were 30 µL and used 10× SB, 11.1 µL of saliva and decreasing percentages of *F. nucleatum* CIM—46%, 35%, 25%, 15%, 5%. There were two reactions tested for each condition; heated and non-heated. In the heated reactions, the saliva was heated at 90°C for 10 min and cooled to room temperature before the addition of the rest of the reagents.

2.5.3 Saliva LOD

To determine the LOD of tRFD-FN1 in saliva samples, 500 µL of fresh 0.628 OD_{600} *F. nucleatum* culture was diluted serially into 4.5 mL of CMB, and 100 µL of each dilution was plated in triplicate on BA plates. The plates were incubated for 48 h, anaerobically. Each dilution was used to test the LOD of the probe. Reactions contained 10 µL of *F. nucleatum* culture, 13.8 µL of heated saliva (90°C, 10 min) and 10× SB.

2.6 Target identification

2.6.1 Size exclusion chromatography

F. nucleatum CIM was prepared according to the method above and two separate volumes of 600 µL were filtered through a 0.22 µm syringe filter (VWR) and a 30K Amicon Ultra centrifugal filter (Pall Corporation). One volume was prepared as 10× CIM; the pellet was resuspended in 15 µL of 1× SB. Tween 20 was omitted from all 1× SB used so that it wouldn't interfere with the fractionation column. The 1× SB without Tween 20 was also used as the mobile phase to elute the CIM in 1.5 column volumes (32 mL) across the 96-well plate, using an ÄKTA Explorer Protein Purification System (GE Healthcare Life Sciences, Marlborough, USA). Next, 500 µL of CIM was injected into the system to run through a Superdex 200 10/300 Size Exclusion Column, and 64 fractions of 500 µL volumes were collected. The 10× CIM was run immediately after the 1× CIM gel filtration. To determine where the target was present in the eluted fractions, groups of six fractions were pooled and the DNAzyme cleavage assay was applied to each pool. The pool that produced a cleavage signal was then subjected to the cleavage assay on each individual fraction. The amount of cleavage activity observed was interpreted to be directly correlated to the amount of target present in each fraction.

2.6.2 SDS-PAGE and silver stain for protein visualization

In order to visualize the SEC fraction proteins, the fractions were run on an SDS-PAGE gel and silver stained for sensitive detection. In a 30 µL volume, 15 µL of SEC fractions and 15 µL of 2× LB were heated at 90°C for 6 min and run on a 12% SDS-PAGE. The stacking layer was removed, and the gel was washed in ddH_2O for 5 min, twice. The gel was stained with the

Pierce Silver Stain Kit (ThermoFisher Scientific) reagents, according to their manual. The gel was imaged using a smartphone camera.

2.6.3 Qubit protein quantification

Protein quantification on the SEC fractions was done using Invitrogen Qubit Protein Assay Kit (ThermoFisher Scientific, Massachusetts, USA). The protein standards were set up using the assay's protein standard vials: 0 ng/µL, 50 ng/µL, 100 ng/µL, 200 ng/µL, 400 ng/µL. To each standard well, 190 µL of working solution and 10 µL of standard were added and vortexed on a MixMate plate mixer (Eppendorf, Hamburg, DEU) for 2-3 s. To each sample well, 198 µL of working solution and 2 µL of SEC sample were added and vortexed on the plate mixer for 2-3 s. The mixtures were incubated at room temperature for 15 min before being measured on Infinite M1000 (Tecan, Männedorf, CH) microplate reader. The excitation was 470 nm and the emission was 570 nm. The resulting Qubit standard curve was plotted, and the equation of the line was used to calculate the concentrations of the individual SEC fractions.

2.6.4 Mass spectrometry analysis

Mass spectrometry analysis was performed by Dr. Yu Lu, at McMaster University. Protein were digested into tryptic peptides using 200 µL volumes of each SEC fraction (C5-D2). The peptides were then labeled with TMT-10plex, to help identify and compare the abundance of each protein present. Liquid chromatography-mass spectrometry/mass spectrometry analysis was performed by injecting 1 µL from a 120 µL-pool of the peptide fractions. The results were matched to the sequenced database of *Fusobacterium nucleatum* subspecies *nucleatum* ATCC 25586 (Proteome ID UP000002521/ Organism ID 190304) on UnitProt.org. In order to

determine the mostly likely candidates to be the target protein of RFD-FN1, the abundances were compared to the cleavage assay of individual fractions.

2.7 Protein expression and purification

Using the information from the mass spectrometry data, the five potential targets were cloned into an expression vector in order to purify each target. Jianrun Xia, our lab technician, worked on cloning the proteins in pET28 with a TEV site, transforming them in BL21DE3, expressing and purifying the target. After expression, purification and removal of the 6×His-tag, the proteins will be assayed using the DNAzyme cleavage test to determine if any are the target of interest. *Due to the SARS-CoV-2 pandemic, the lab was shut down in the middle of this procedure.

Chapter 3: RESULTS

3.1 Functionalize tRFD-FN1 to detect *F. nucleatum* in CRC-confirmed, patient stool samples

3.1.1 Optimization of F. nucleatum-activated DNAzyme

The main challenge of the project is to functionalize the DNAzyme RFD-FN1 in stool samples to be able to diagnose CRC. DT4 is an RFD that was isolated from the *in vitro* selection performed by Qian Feng (**Figure 4A**). Preliminary tests revealed that DT4 is unable to detect *F. nucleatum* naturally present in CRC-positive stool samples. This thesis project began by modifying DT4 with extra bases to ensure proper hybridization in solution, since the *trans*-acting form is not ligated to the substrate sequence in the reaction mixture (Figure 4**B**). Two bases were added to the 3'-end and three bases were added to the 5'-end of the DNAzyme to form complete hybridization with the substrate strand. The extra bases did not affect the cleavage activity of the DNAzyme compared to the previously demonstrated results. The *cis*-acting sequence of the probe, RFD-FN1, is ligated to FQ30 and performs self-cleavage at the RNA site in the fluorescent substrate (Figure 4**C**). The modified version of DT4 will be henceforth referred to as RFD-FN1 (RNA-cleaving fluorogenic DNAzyme for *F. nucleatum*) and tRFD-FN1 (*trans*). The *trans* version of RFD-FN1 was used in the DNAzyme cleavage assays, unless otherwise stated.

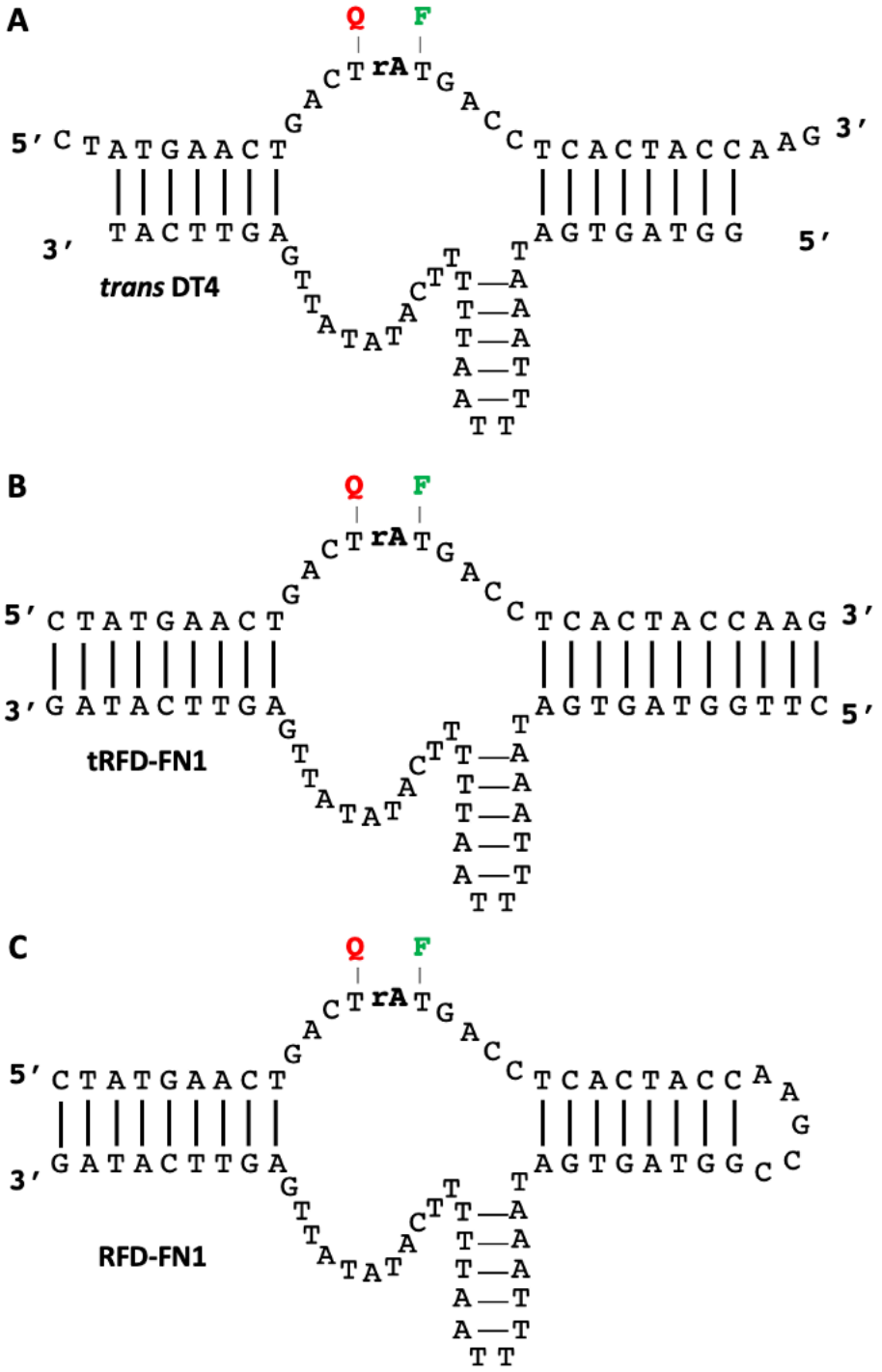


Figure 4. DNAzyme probe sequences. A. *Trans* DT4 sequence selected by Qian Feng. *Trans* DT4 hybridizes to the fluorescent substrate (FQ30) in solution. **B.** tRFD-FN1 is the modified version of *trans* DT4. Deoxyribonucleotides A and G were added to the 3'-end, and C and T were added to the 5'-end to form complete hybridization with FQ30. **C.** RFD-FN1 is *cis*-acting; the substrate is ligated to the 5'-end of the probe strand. FQ30; Q: quencher, DABCYL-modified thymine; F: fluorophore; fluorescein-modified thymine; rA: adenine ribonucleotide.

3.1.2 Validation of F. nucleatum-specific DNAzyme by polyacrylamide gel electrophoresis

The analysis of the dPAGE gels is done by calculating the RNA-cleavage (%), or cleavage activity, of the DNAzyme. The top (uncleaved substrate fraction) and bottom (cleaved substrate fraction) bands are analyzed based on their fluorescence intensities. The dissociation of the fluorophore from the quencher (cleaved substrate) generates a fluorescent signal that is six times stronger than the uncleaved substrate. The percentage of cleavage activity is calculated by dividing the fluorescence of the cleaved fraction, by the total fluorescence in that reaction. All gels were analyzed using this method.

The gels in **Figure 5** validate the results that were observed by Qian, which were initial characterization tests of the *F. nucleatum*-responsive DNAzyme that was selected. Briefly, a subculture step was introduced so that the *F. nucleatum* culture could reach 0.7 OD_{600} during incubation, as this was not being achieved per Qian's protocol.

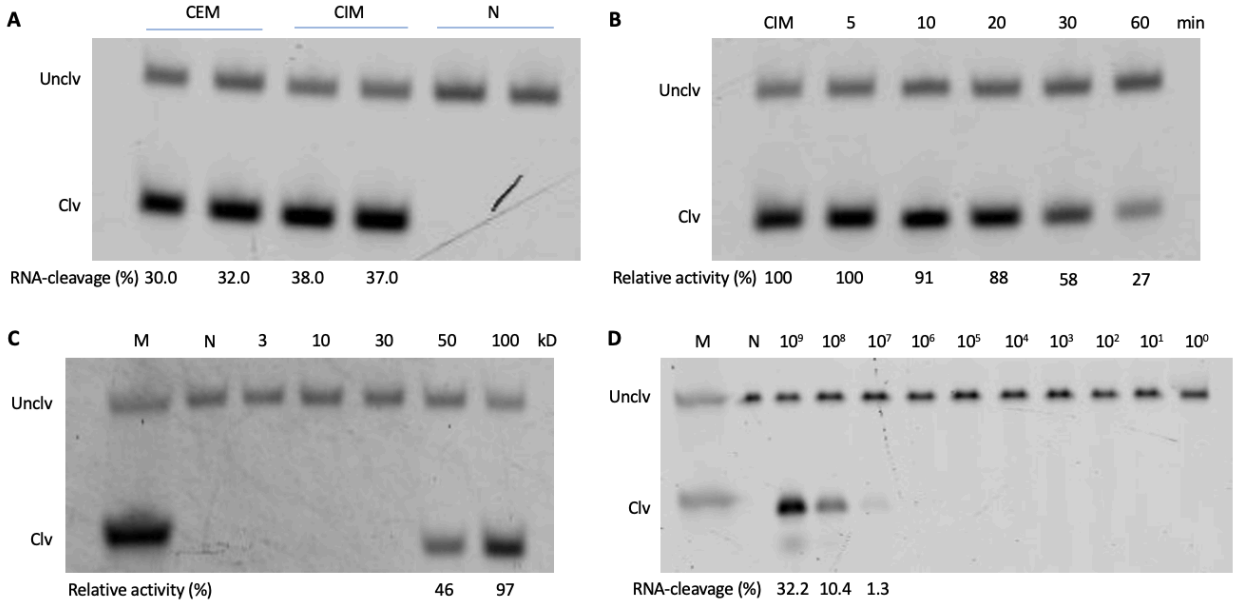


Figure 5. RFD-FN1 characterization tests validate preliminary findings. A. tRFD-FN1 activity in CEM and CIM. **B.** tRFD-FN1 activity of CIM heated at 90°C. **C.** Estimation of target size; filtrate of CIM passed through molecular size filters tested for tRFD-FN1 activity. **D.** Limit of detection; LOD estimated by testing

tRFD-FN1 in 10-fold diluted *F. nucleatum* cultures prepared into CIM. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; Unclv: uncleaved FQ30; Clv: cleaved FQ30. RNA-cleavage (%): ((clv/6)/ (unclv + (clv/6))) × 100%; Relative activity (%): RNA-cleavage divided by RNA-cleavage observed in unheated CIM (A).

The extracellular and intracellular mixtures of *F. nucleatum* induce 30-32% and 37-38% cleavage, respectively (Figure 5**A**). The top set of bands are the uncleaved fraction of the DNAzyme, and the bottom set of bands are the cleaved fraction; taken together, they represent the affinity of the DNAzyme to that amount of target. The negative lanes demonstrate that the DNAzyme is not responding to the metal ions present in the selection buffer (SB) and confirms that it is specifically activated by the bacterial species *F. nucleatum*.

A heating experiment was performed to analyze the stability of the target. After heating the CIM at 90°C for 30 min, the cognate target of tRFD-FN1 is still able to induce over 50% of the activity observed when the sample is not heated (Figure 5**B**). The relative activity calculation is based on the 38% cleavage activity of the CIM that was observed in Figure 5A. After heating the CIM, the DNAzyme is still able to recognize a viable target and perform its catalytic activity. As the heating time increases, the relative cleavage activity decreases, but a large percentage of the target is presumably still intact and recognized by the DNAzyme. Remarkably, the CIM that was heated for 60 min generated 27% of the relative DNAzyme cleavage activity.

To estimate the molecular size of the target, the CIM was passed through a series of molecular weight cut-off membrane filters (Figure 5**C**). The collected filtrate from each size filter suggests that the target is anywhere between 30-100 kD. The filtrate from the 100 kD filter demonstrated almost 100% of the activity that was observed in Figure 5A, indicating that all of the target can pass through that membrane size. Around 46% relative activity was observed in the filtrate of the 50 kD filter, which suggests that the target is smaller than 50 kD. The absence

of cleavage in the filtrates of the 3-30 kD filters means that no target could pass through a membrane that was 30 kD or less.

The limit of detection (LOD) of tRFD-FN1 was assessed to be 10^7 CFU/mL (Figure 5**D**). *F. nucleatum* culture was grown to 0.7 OD_{600}, serially diluted 1/10 to 10^{-9}, and plated in triplicate. The concentration of the original culture was calculated using the colony counts and was determined to contain ~1.07×10^9 CFU/mL. Each dilution was prepared into CIM and assayed. The gel shows that tRFD-FN1 can detect down to 10^7 CFU/mL, but the lanes that contain less than 10^7 CFU/mL do not have enough bacteria to generate cleavage. The DNAzyme cleavage assays validated the results that were initially demonstrated by Qian.

3.1.3 Characterization of DNAzyme in diluted, spiked and non-spiked stool samples

The biggest challenge of this project is to be able to use tRFD-FN1 to detect the overabundance of *F. nucleatum* in stool samples to indicate CRC—especially in the early stages. Preliminary data from Qian's thesis suggests that the probe is not currently able to indicate the presence of *F. nucleatum*, either naturally occurring or spiked-in. To validate this, CRC stool samples were homogenized in cooked meat broth media (CMB) and used in a DNAzyme cleavage assay. The reactions contained either 20% or 40% stool, and mixtures were spiked with *F. nucleatum* culture and CIM to determine if low *F. nucleatum* concentration in the stool was the reason for the lack of a cleavage band. The inability of tRFD-FN1 to detect any *F. nucleatum* that may be present in the diluted stool samples that had not been spiked, as well as in the samples that were spiked with the target, is confirmed (**Figure 6)**. The only lanes that showed DNAzyme cleavage activity were the pure *F. nucleatum* culture and CIM replicates.

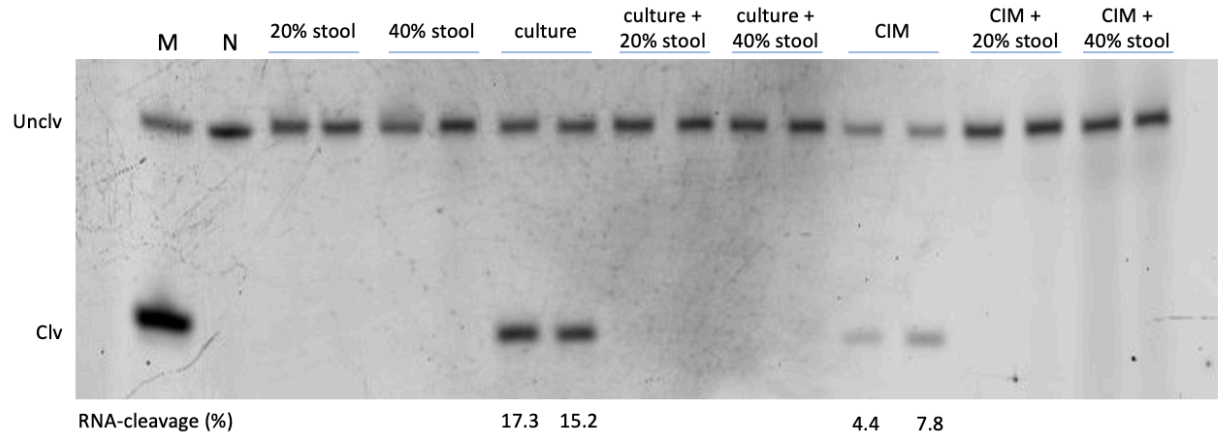


Figure 6. DNAzyme cannot detect *F. nucleatum* in non-spiked, culture-spiked and CIM-spiked stool. Activity of tRFD-FN1 in diluted stool, and diluted stool spiked with *F. nucleatum* culture and CIM. Stool samples were diluted 20% w/v and 40% w/v in CMB. Reactions were spiked with 46% v/v of culture or CIM. All mixtures were heated at 90°C for 15 min before the addition of DNAzyme. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

3.1.4 DNAzyme is able to detect F. nucleatum in saliva samples

Stool is an especially challenging matrix, and so we challenged the DNAzyme to determine if it could detect *F. nucleatum* in a relatively less complex biological matrix. Saliva was used as an alternative test sample because *F. nucleatum* is common in the oral microbiome. Like CRC, the bacterium is overabundant in cases of periodontal disease, but is still present in non-diseased patients in smaller numbers. An *F. nucleatum*-specific probe would be useful to detect an overabundance of the species in a patient's saliva, which could indicate oral disease and potentially be used to signify a higher risk for developing CRC. To test the functionality of tRFD-FN1 in saliva, three independent saliva samples were collected and aliquoted into six tubes each. Protease inhibitor was added to half of the tubes, *F. nucleatum* was added to two tubes (+/- PI), and two tubes were run without DNAzyme to ensure that no autofluorescence was present at the location of the cleavage band. tRFD-FN1 was able to detect *F. nucleatum* spiked into saliva samples (**Figure 7A**).

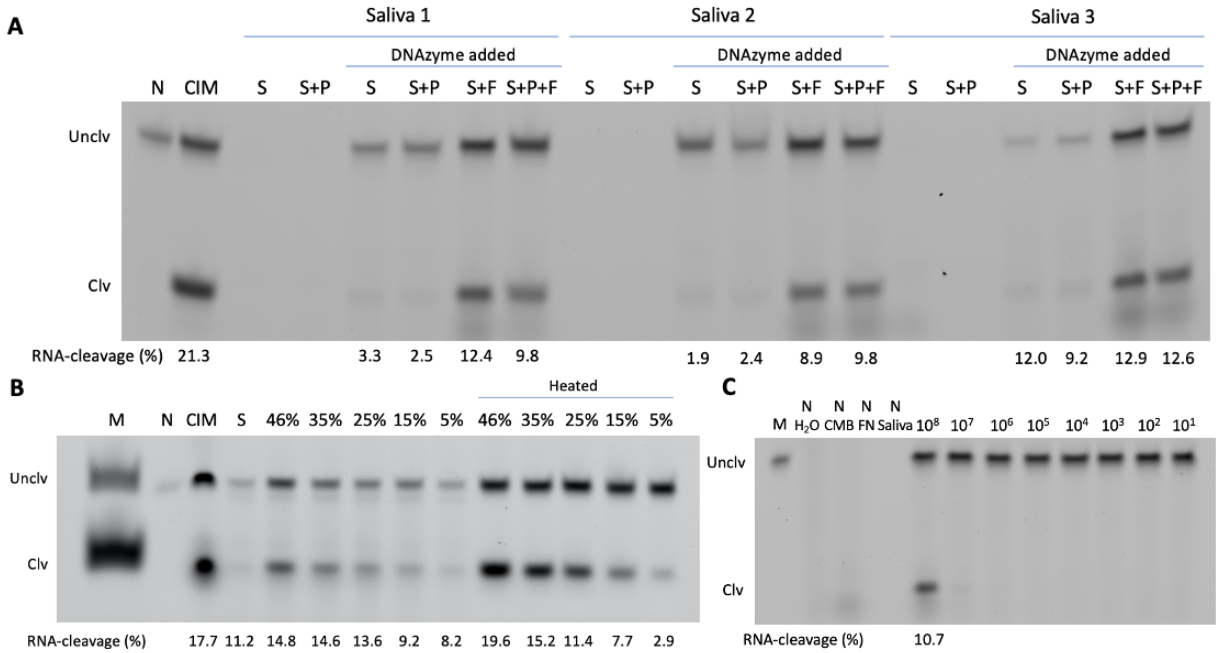


Figure 7. DNAzyme can detect *F. nucleatum* within spiked saliva samples. A. Activity of tRFD-FN1 in non-spiked and culture-spiked saliva samples. S: saliva; P: protease inhibitor; F: *F. nucleatum* CIM. **B.** Activity of tRFD-FN1 in decreasing amounts of heated and non-heated *F. nucleatum* CIM. Saliva sample 1 was heated at 90°C for 10 min, prior to the addition of CIM and DNAzyme. **C.** LOD of tRFD-FN1 in saliva sample. N H_2O, N CMB, N FN and N saliva lanes contain only the respective sample to indicate if autofluorescence is present. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control contains DNAzyme, SB and ddH_2O; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

The reactions that did not contain DNAzyme did not generate any autofluorescence that could interfere with the cleavage signal. The DNAzyme reactions that contained saliva samples 1 and 2, but no additional *F. nucleatum* cells, generated a weak cleavage band. The weak cleavage indicates that either a small amount of *F. nucleatum* are naturally present in the saliva, or that there is non-specific cleavage of the DNAzyme at the RNA site. The addition of *F. nucleatum* culture to the saliva induced a 7-10% increase in cleavage activity over the non-spiked samples. Conversely, sample 3 does not demonstrate a significant increase in the normal sample versus the spiked sample. The addition of protease inhibitor does not appear to have a significant difference on the cleavage activity, as observed by the similar levels of cleavage in the lanes with and without it.

Since preliminary evidence suggests that the target is heat stable, we heated the saliva to confirm if the DNAzyme cleavage observed in the non-spiked reactions was caused by nuclease degradation. Heating the saliva at 90°C for 10 min, before adding the reaction components, does not appear to affect the target, which was anticipated (Figure 7**B**). The cleavage activities decreased from 14.8-8.2% in the non-heated reactions containing decreasing percentages of *F. nucleatum* (46-5%), and the respective heated reactions decreased from 19.6-2.9%. The reactions that were not heated do confirm some non-specific cleavage was occurring, indicated by the faint banding pattern that is observed directly under the uncleaved band. However, the non-specific cleavage was abolished in the reactions that contained the heated saliva, and the observed cleavage activities were not significantly different than the unheated samples.

We determined the LOD for tRFD-FN1 in saliva using the same method as described in Figure 2D, except that *F. nucleatum* culture was used instead of CIM (Figure 7**C**). Here, each reaction contained 46% v/v of heated saliva and 33% v/v of the respective serial dilutions. tRFD-FN1 was not able to detect any dilution below the initial culture's concentration of ~7.6×10^7 CFU/mL. The LOD of tRFD-FN1 is the same in saliva as it was for pure *F. nucleatum* culture. The control lane containing CMB reveals autofluorescence in the liquid media that is present below the cleavage band. The saliva assay results confirm that tRFD-FN1 can function in a biological matrix and not just a pure culture, and it is still a good candidate to try to functionalize in stool. Upon these preliminary findings, we realize that DNAzyme:target recognition in stool is not without challenges; however, the inherent stability of DNA in the stool is a promising characteristic to work with.

3.1.5 F. nucleatum detection is possible in spiked, processed stool

The next challenge in the project was to establish a method that enabled tRFD-FN1 to perform its cleavage activity in response to *F. nucleatum* recognition in stool. The first attempt was to remove as much organic material as possible from the stool to eliminate possible non-specific competition or inhibition. From eight random samples from the CRC patient stool biobank, 5 g of stool was weighed out into separate falcon tubes and mixed with 10 mL of 1× cOmplete protease inhibitor cocktail + SB. The samples were vortexed to break apart the particulate matter and shaken on a desktop shaker set to high, for 2 h at room temperature. The samples were centrifuged to collect the aqueous fraction of the stool. For all tests in **Figure 8**, the eight processed stool samples were pooled together, heated at 90°C for 30 min, and diluted to the appropriate concentration to challenge the DNAzyme's detection ability.

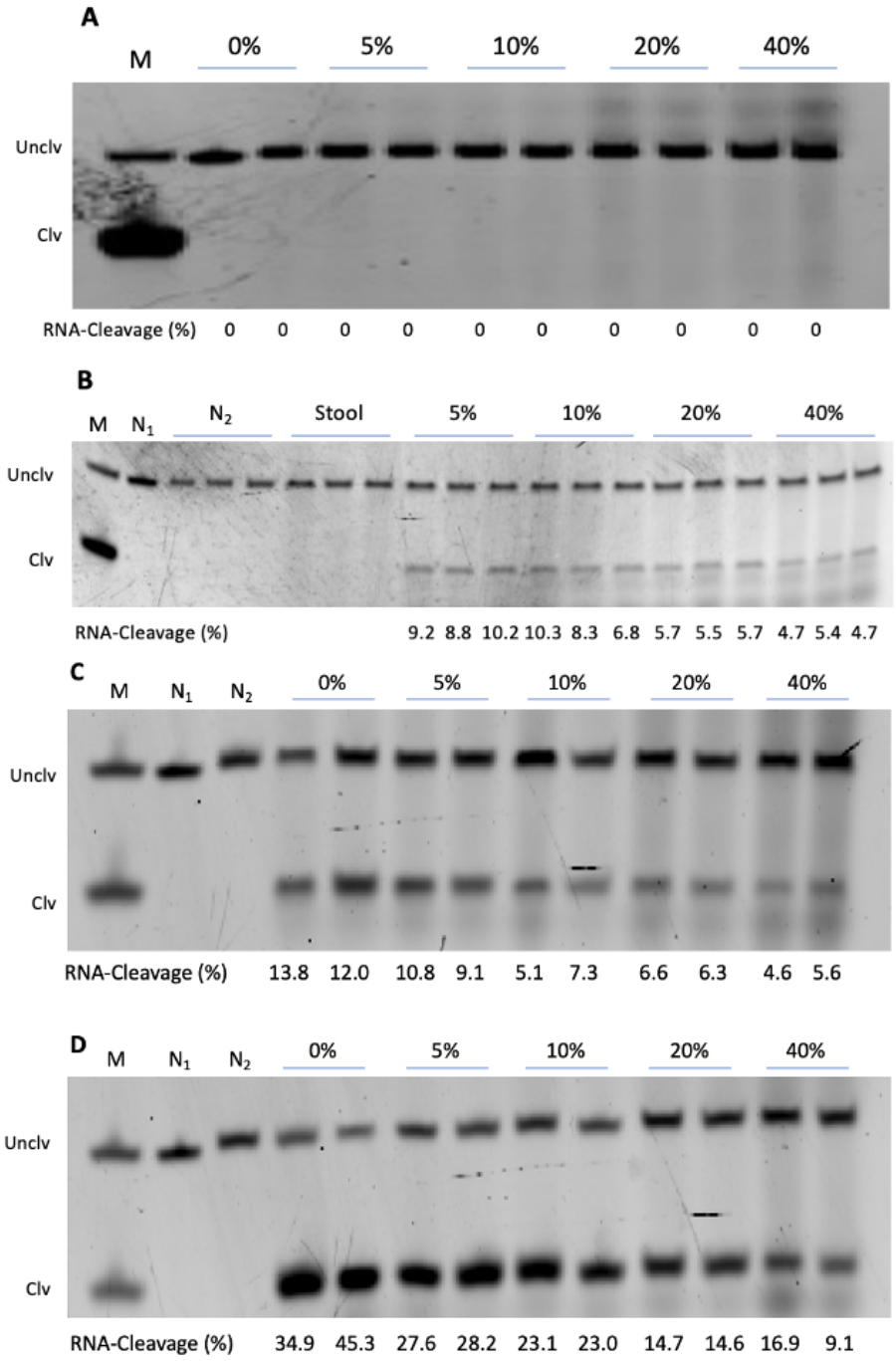


Figure 8. DNAzyme can detect spiked culture, CEM and CIM in aqueous fraction of stool. A. Activity of tRFD-FN1 in processed stool sample. Activity of tRFD-FN1 in culture-spiked (**B)**, CEM-spiked (**C**) and CIM-spiked (**D**) processed stool samples. Spiked samples all contained 46% v/v target (culture, CEM or CIM), and the percentage in each lane indicates amount of stool added to reaction. Stool was heated at 90°C for 30 min prior to the addition to the reaction. M: marker contains FQ30 treated with 0.2 M NaOH; N_1: FQ30 stock; N_2: negative control that contains DNAzyme, SB and ddH_2O; Stool: contains 46% v/v stool, SB and DNAzyme; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

tRFD-FN1 did not detect any *F. nucleatum* that may be naturally present in the processed stool of CRC patients, in stool concentrations that were of 0-40% of the reaction (Figure 8**A**). However, when the processed stool sample reactions were spiked with *F. nucleatum*, the cleavage results indicated that the probe could recognize its target in 40% v/v heated, processed stool (Figure 8**B-D**). Culture (Figure 8B) and CEM (Figure 8C) induced comparable amounts of cleavage activity in 5-40% v/v stool, whereas CIM induced 2-2.5× the amount of cleavage in the respective volumes (Figure 8D). In the gel with culture-spiked stool (Figure 8B), lanes 6-8 contained stool, SB and DNAzyme, which confirms that no naturally present *F. nucleatum* was being detected. In the following lanes, the stool volume was increased from 5-40%, and the gel demonstrates a proportional decrease in cleavage activity compared to the increase in processed stool. At the highest concentration of stool, the cleavage activities are still around 5%. The gels with CEM and CIM, with the reactions performed in duplicate, reveal a similar pattern to what was observed in the culture gel. The no-stool (0%) controls indicate the starting amount of target available, and a steady decrease in cleavage is seen as the stool volume increases from 5% to 40%. The darker smears present underneath the cleavage bands in the stool samples is caused by autofluorescence in the stool; it is not present in all patient samples to the same degree.

3.1.6 Lyophilization does not promote the detection of the target in stool

Clearly, tRFD-FN1 is able to function in the aqueous fraction of stool samples, however, this centrifugal processing might also be removing endogenous *F. nucleatum* cells. The same challenge stands: how can the stool samples be processed to be more suitable for the DNAzyme, but not disturb the presence of the *F. nucleatum* target? Lyophilization was implemented to preserve the integrity of the bacterial target and simultaneously concentrate it. The structural

integrity of the bacterial cells, and their proteins and nucleic acids, should remain viable, but hopefully whatever interferes with the assay is eliminated. The stool samples that were used were healthy controls and likely contained little to no endogenous *F. nucleatum*, so we spiked the samples with culture to ensure the presence of the bacterial species for target recognition. **Figure 9A** shows the stool and CIM samples after the lyophilization process, and after they were reconstituted with water.

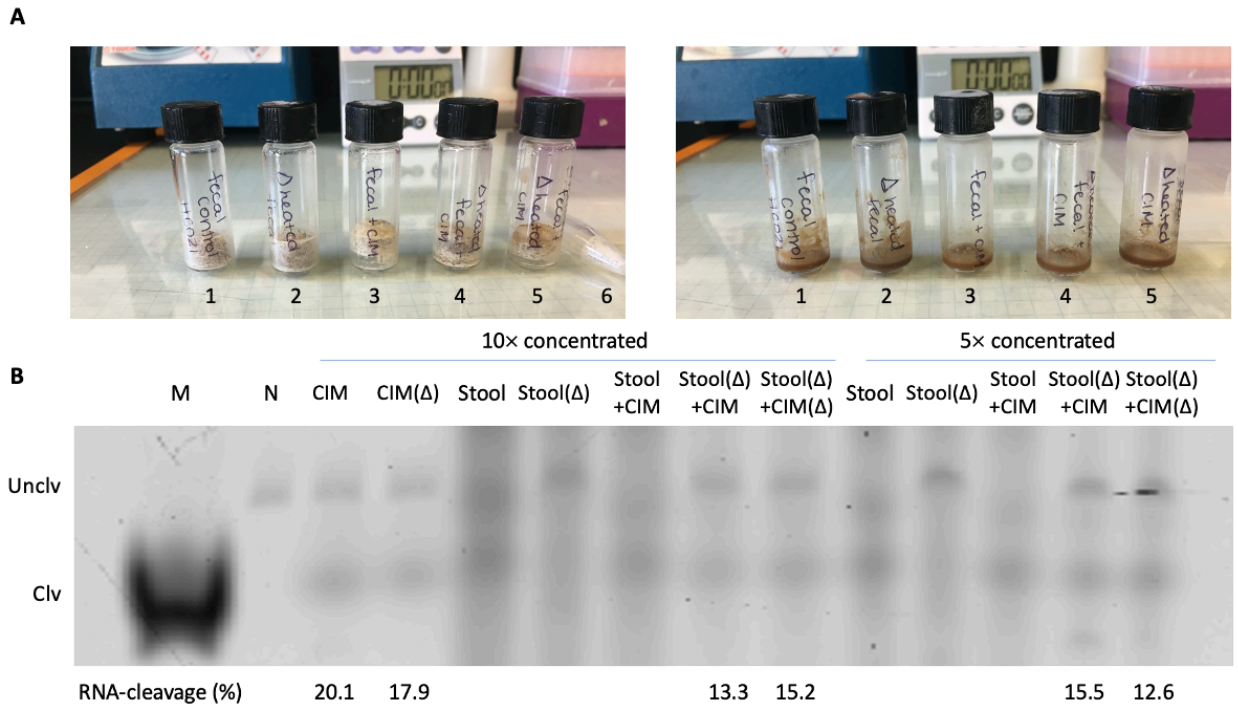


Figure 9. Lyophilization of stool does not improve DNAzyme cleavage. A. Stool and CIM samples after lyophilization and after being reconstituted in ddH_2O to achieve 10× concentrated samples. 1: Stool; 2: Stool (Δ); 3: Stool + CIM; 4: Stool (Δ) + CIM; 5: Stool (Δ) + CIM(Δ); 6: CIM. **B.** Activity of tRFD-FN1 in 10× and 5× concentrated lyophilized stool. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; Δ: stool was heated at 90°C for 2 h, CIM was heated at 90°C for 15 min; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

While there is potential cleavage indicated by a darker intensity at the cleavage band location, the concentrated stool induces too much non-specific degradation on the gel to confidently confirm the cleavage activity (Figure 9**B**). The 10× concentrated CIM and the

heated, spiked stool samples generated darker smears at the cleavage band location, but they were similar to the unheated samples, and they also lacked the distinct band pattern usually seen in the dPAGE gels. Heating the stool at 90°C for 2 h decreased the smearing slightly, but it also decreased the cleavage band intensity. Therefore, the cleavage cannot be confidently attributed to the recognition of *F. nucleatum* by tRFD-FN1 in the lyophilized stool.

3.1.7 Kanamycin and vancomycin do not promote the growth of F. nucleatum in stool for DNAzyme detection

Lastly, the addition of bactericidal, gram-positive antibiotics was used to reduce the bacterial load in cultured stool samples. By eliminating other competing bacterial species, the *F. nucleatum* that is present in the stool samples could be able to proliferate within the matrix. To investigate this, we first determined that kanamycin and vancomycin do not inhibit the growth of gram-negative *F. nucleatum* (**Figure 10A**).

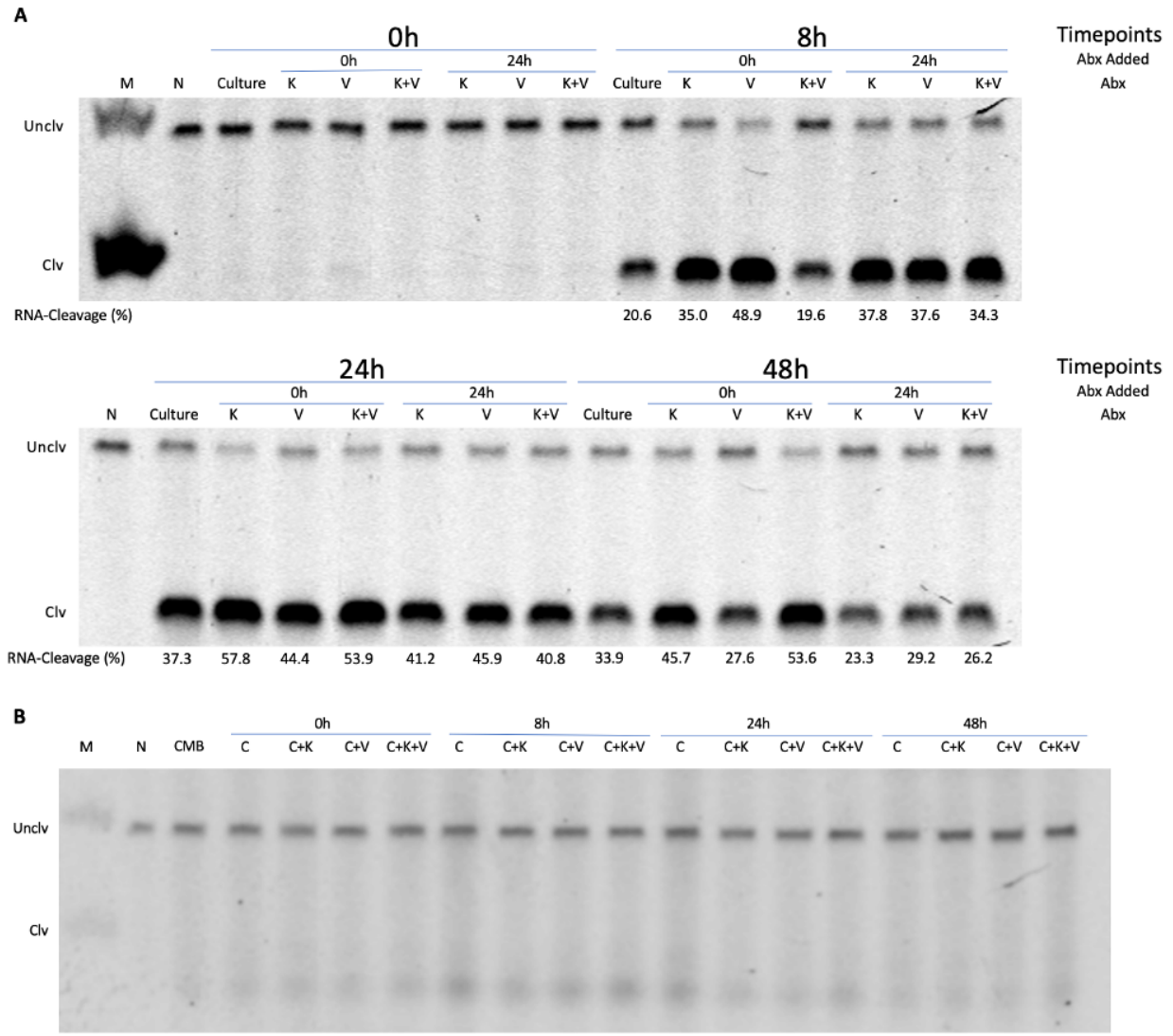


Figure 10. Kanamycin (K) and vancomycin (V) do not inhibit the growth of *F. nucleatum,* but do not improve DNAzyme cleavage in cultured, spiked stool. A. Activity of tRFD-FN1 in stool samples that have been spiked with K and V at 0 h and 24 h post-culture and cultured for 48 h. **B.** Activity of tRFD-FN1 in stool samples that have been spiked with K and V and cultured. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; K: kanamycin, 50 µg/mL working concentration; V: vancomycin, 50 µg/mL working concentration; C: *F. nucleatum* culture; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

The kanamycin and vancomycin were used at a working concentration of 50 µg/mL. Timepoints were measured at 0, 8, 24 and 48 h, for the antibiotics that were added at either 0 or 24 h of incubation. The 0 h timepoints record no cleavage activity, because that is when the culture was initiated. The target production is already quite prolific at 8 h, indicated by cleavage

activities between 20.6-48.9%, and after 24 and 48 h, reaches a maximum of 57.8%. The culture controls demonstrate the expected cleavage activities for each timepoint based on normal culturing conditions. The high cleavage activities at 8 h indicate that the cultures do not experience a significant lag phase with the addition of the antibiotics. In addition, in all timepoints, the individual or combined antibiotic reactions generated similar or higher cleavage activities than the culture controls, meaning that the antibiotics did not interfere with *F. nucleatum* growth or target production. The observed fluctuations in cleavage activity are a common product of sampling or experimental error. After determining that the antibiotics did not interfere with *F. nucleatum* target production, three CRC stool samples were resuspended 20% w/v in CMB, and 50 µg/mL working concentration of the respective antibiotics were added. The samples were brought into the anaerobic chamber and were incubated at 37°C—aliquots were removed at 0, 8, 24 and 48 h for the DNAzyme cleavage assay. No cleavage activity is present at any timepoint (Figure 10**B)**. The destruction of gram-positive bacterial species by the addition of the broad-spectrum antibiotics did not improve the ability of tRFD-FN1 to detect its cognate target.

The challenge posed by using a DNAzyme in the stool matrix was not overcome by centrifugal filtering, lyophilization or antibiotics, and the concentration of the endogenous *F. nucleatum* in the stool samples must be considered to be a limiting factor.

3.1.8 Stool culture-based techniques to amplify target require heating step to eliminate interfering molecule

The aim to create a rapid, non-invasive, at-home diagnostic test with tRFD-FN1 is not possible with the current DNAzyme probe. Instead, we decided to employ culture-based

techniques to enrich the number of endogenous *F. nucleatum* cells present. The previous results have already suggested that tRFD-FN1 is capable of detecting the bacterial target in stool when a competing factor is inhibited. Excess target available for the DNAzyme could be a successful method to overcome what could be an interfering molecule. **Table 2** summarizes the results from the stool culturing tests that were performed; for all tests, the incubations and dilutions were performed in the anaerobic chamber at 37°C, and the results are demonstrated by the cleavage activities from DNAzyme cleavage assays.

Table 2. Stool culturing experiments

Experiment	Result
Stool samples incubated for 48 h before 10-fold serial dilution	No cleavage activity; some non-specific degradation
Stool samples incubated for 48 h after 10-fold serial dilution	No cleavage activity; some non-specific degradation
Stool inoculated with *F. nucleatum* culture and incubated for 48 h before 10-fold serial dilution	No cleavage activity; some non-specific degradation
Stool inoculated with *F. nucleatum culture* and incubated for 48 h after 10-fold serial dilution	No cleavage activity; some non-specific degradation; one cleavage band observed in 10^{-6} dilution in one replicate
Fresh stool incubated for 48 h before 10-fold serial dilution; one set with protease inhibitor added. Samples were heated at 90°C for 10 min before DNAzyme cleavage test (**Figure 11A**)	Possible cleavage activity; non-specific degradation is most likely the reason for the same intensity bands across 10^{-2} to 10^{-7} dilutions; protease inhibitor does not improve cleavage activity

Fresh stool inoculated with five *F. nucleatum* colonies and incubated for 48 h before 10-fold serial dilution; one set with protease inhibitor added. Samples were heated at 90°C for 10 min before DNAzyme cleavage test (Figure 11**B**)	Possible cleavage activity but similar intensity to non-spiked fresh sample; non-specific degradation is most likely the reason for the same intensity bands across 10^{-2} to 10^{-7} dilutions; protease inhibitor does not improve cleavage activity
Stool heated at 90°C for 1, 2, 3, 4, 5 and 6 h and incubated for 48 h	No bacteria growth on BA plates after 2 h; no cleavage activity; non-specific degradation eliminated
Stool heated at 90°C for 2 h, inoculated with five *F. nucleatum* colonies and incubated for 48 h; DNAzyme cleavage reaction was 8 h (**Figure 12**)	Cleavage activity observed in 10^{0} (12.2%), 10^{-1} (9.8%) and 10^{-2} (2.5%) dilutions; non-specific degradation eliminated

Incubating the stool samples, before and after performing serial dilutions, did not enrich *F. nucleatum* enough for tRFD-FN1 detection. The level of non-specific degradation that was observed can also cause problems with interpreting the results of the assay. Spiking the stool samples with *F. nucleatum* culture, and then incubating them anaerobically, before and after serial dilutions, also did not produce conclusive cleavage results. The occasional cleavage band in a random dilution was observed, but there was no consistency to confidently indicate that the DNAzyme was accurately recognizing its target. Fresh stool samples were tested to determine if the target is degraded during storage conditions. The gels for 24 h and 48 h incubation of the serially diluted stool, that was not spiked with additional cells, reveal non-specific degradation in the form of multiple, weak bands appearing outside the expected tRFD-FN1 uncleaved and cleaved band locations **(Figure 11A)**.

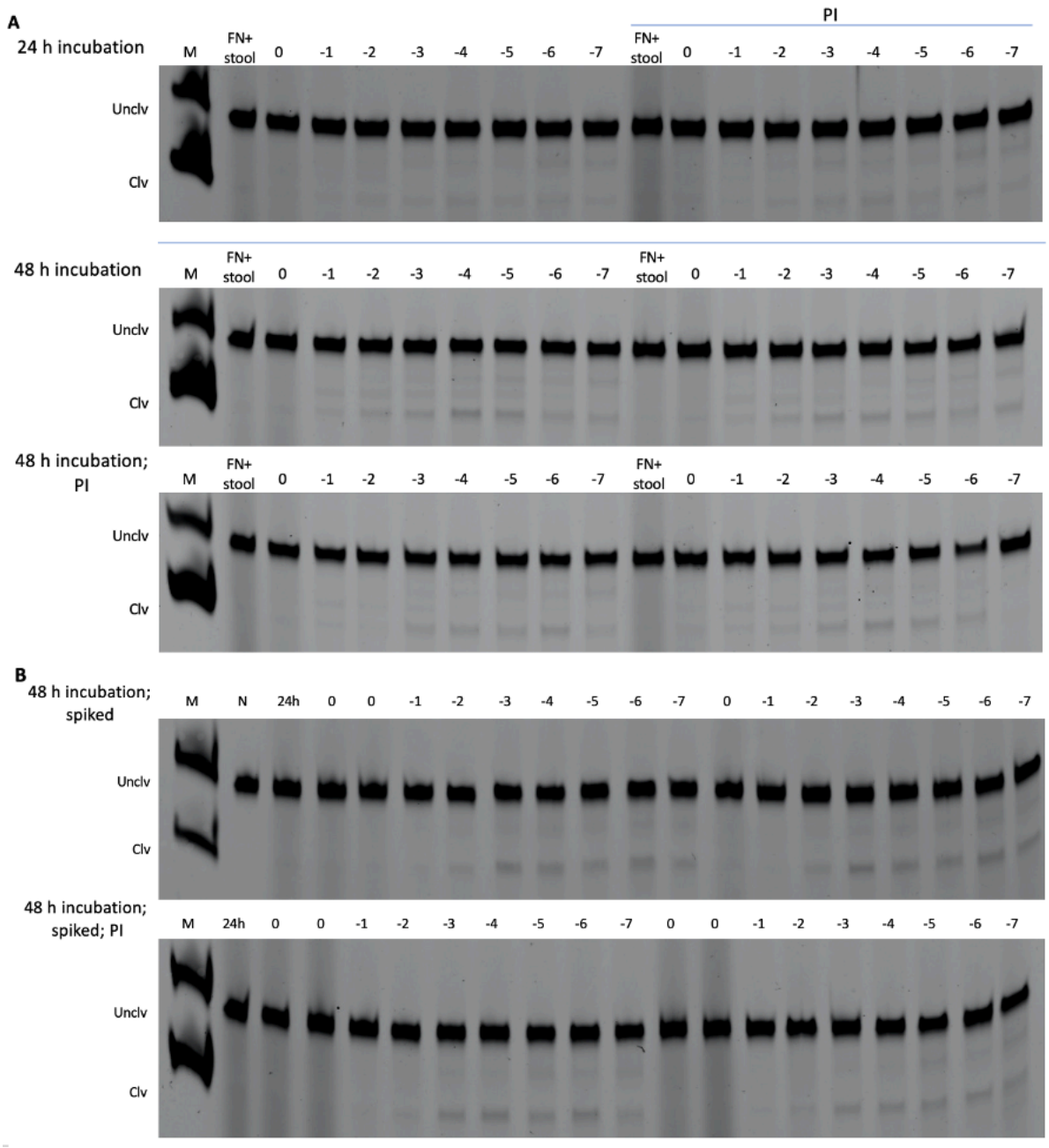


Figure 11. Fresh stool incubation does not improve DNAzyme cleavage. A. Activity of tRFD-FN1 in fresh stool samples that were serially diluted 10-fold and incubated for 24 and 48 h. **B.** Fresh stool samples were spiked with five *F. nucleatum* colonies, incubated for 48 h and serially diluted. Biological replicates were cultured for the 48 h reactions. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; 24h: timepoint of culture; PI: protease inhibitor; FN+ stool: stool sample spiked with *F. nucleatum* culture; lanes are labelled according to their dilution; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

The spiked stool lanes (FN+ stool) demonstrate nuclease degradation, indicated by the characteristic smearing pattern. However, the lanes containing the 10^{-2} to 10^{-7} dilutions have

reduced smearing, because the stool matter has been diluted out, and a faint but distinct cleavage band present. While the very faint bands could indicate the presence of endogenous *F. nucleatum* in the stool, we would expect a greater variation in band intensities across the dilutions to discredit non-specific degradation. The addition of protease inhibitor does not appear to increase the viability of the target. In addition to the above experiment, stool was also inoculated with *F. nucleatum* colonies, incubated, and then serially diluted (Figure 11**B**). The same banding pattern in Figure 11A is more pronounced in these gels—again, no difference observed with the addition of protease inhibitor. The banding pattern in the 10^{-3} to 10^{-7} dilutions is slightly more intense, but still does not exhibit a distinct increase or decrease across the dilutions. More investigation should be conducted whenever possible on fresh stool samples. Lastly, we investigated the use of heat on the stool samples to promote target recognition by the DNAzyme. Since the target is quite robust at 90°C, we wanted to see if it could remain viable after intense heating destroyed other stool components. After 2 h at 90°C, no bacteria were able to grow on the plates, and the DNAzyme cleavage assay on the sterilized stool sample shows no non-specific degradation of the probe. Unsurprisingly, the rigorous heating step did not promote *F. nucleatum* recognition by tRFD-FN1. However, the inoculation of five *F. nucleatum* colonies in heat-sterilized stool, similar to the fresh inoculated stool test described above, did produce cleavage bands in the 10^{0}, 10^{-1} and 10^{-2} dilutions (**Figure 12**).

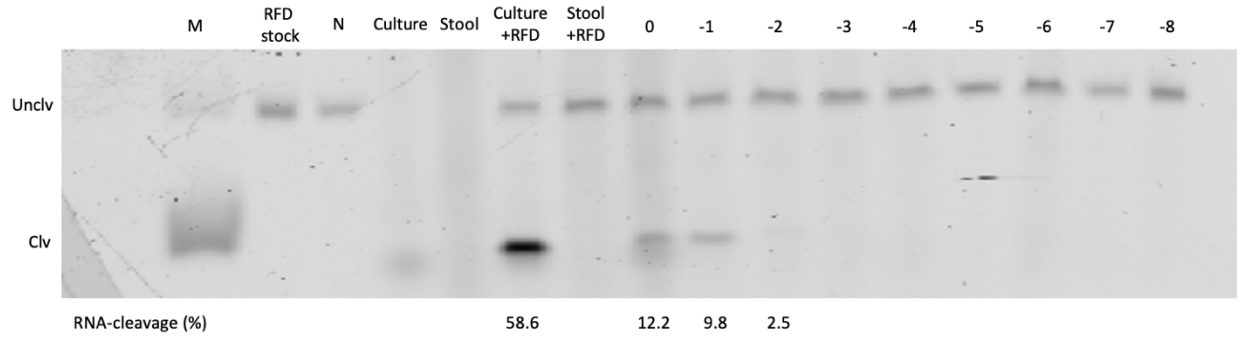


Figure 12. Heat-sterilized stool inoculated with *F. nucleatum* colonies did produce DNAzyme cleavage. Activity of tRFD-FN1 in stool heated at 90°C for 2 h, inoculated with five *F. nucleatum* colonies and cultured for 48 h. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; RFD stock: contains DNAzyme and SB; 24h: timepoint of culture; reactions labelled 0 through -8: respective 10-fold serial dilutions of stool culture; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

The culture+RFD lane is the control culture that was inoculated with five *F. nucleatum* colonies (into CMB), alongside the inoculated stool. The incubation in the 10^0-stool reported four-fold less cleavage than the pure culture. Cleavage is observed in the 10^{-1} and 10^{-2} dilutions, at 9.8% and 2.5% but not in the 10^{-3} or higher dilutions. No activity was present in the stool sample that was not inoculated with the colonies before incubation. While tRFD-FN1 is observed to be quite stable in unprocessed stool samples, the treatment of heat seems to be required for a clear detection signal on a dPAGE gel. Heat appears to be one method to promote the functionality of tRFD-FN1 in stool—possibly through the destruction of a heat-labile interfering molecule—however, sensitivity remains an issue. Next, we wanted to explore the possibility of selectively increasing target production by *F. nucleatum* to meet the sensitivity requirements of tRFD-FN1.

3.2 RFD-FN1 target identification

3.2.1 Target production was not consistent in cultures tested weeks apart

To approach the challenge from a different direction, we refocused our efforts on producing sufficient amounts of target to outcompete whatever is interrupting the DNAzyme-target interaction in stool. We explored some possible conditions to foster target upregulation in *F. nucleatum* cultures to outcompete the interference, which are summarized in **Table 3**.

Table 3. Conditions for target production

Variable	Experiment	Result
Media	Addition of Vitamin K1-Hemin to CMB	Extra supplements had no effect on the cleavage activity
Colony age	*F. nucleatum* colonies incubated 2-8 days; cultures inoculated with different aged colonies	Colony age does not affect cleavage activity
Growth phase	*F. nucleatum* cultures incubated anaerobically for 5 days; timepoints collected at 0, 10, 16, 20, 24, 28, 32, 40, 48, 72, 96, 120 h	Cleavage activities between 30-90%; early log phase demonstrated high initial cleavage activity, which decreased during log and reached maximum at late stationary and death phase
Growth condition—aeration	Shake cultures every 1, 2 or 4 h, or no shaking, during 48 h anaerobic incubation	Shaking conditions are similar; cleavage activity significantly higher in shaking conditions vs. no shaking
Growth condition—subculture	Subculture dilutions performed at 24 h and 48 h vs. no subculture	Contradicting cleavage activities between subcultured and non-subcultured conditions; no consensus

Handling and storage	Culture and CIM tested fresh, or stored at -20°C or -80°C before preparation and testing	Higher cleavage activity in cultures stored at -20°C before assay or being prepared into CIM
Reproducibility	Multiple experiments reproduced weeks to months apart	Experiments performed greater than three weeks apart demonstrated significant differences in cleavage activities. Activities were anywhere between 4-90% across the same experiments. Greatest differences observed in cultures performed Nov-Feb vs. May-Aug.

DNAzyme cleavage assays were performed on cultures that were incubated for 5 days, and timepoints were collected every 4-10 h for the first 2 days and every 24 h for the following 3 days. Cleavage activities were recorded to be anywhere between 30-90%, with no discernable pattern detected. Replicates also showed inconsistent cleavage activities. Incubation time of solid media, from 2-6 days; storage conditions of the culture before testing; shaking of liquid cultures during incubation; addition of Vitamin K1-Hemin supplement into liquid media; growth phase; subculture dilutions; no subculture; and culturing time were all investigated to determine if there was a reliable way to produce large amounts of the *F. nucleatum* target. Unfortunately, there was no condition that could provide consistent cleavage activities across replicates of the same experiment. Two identical timecourse assays that were performed 16 months apart demonstrate the inconsistencies that were observed (**Figure 13)**.

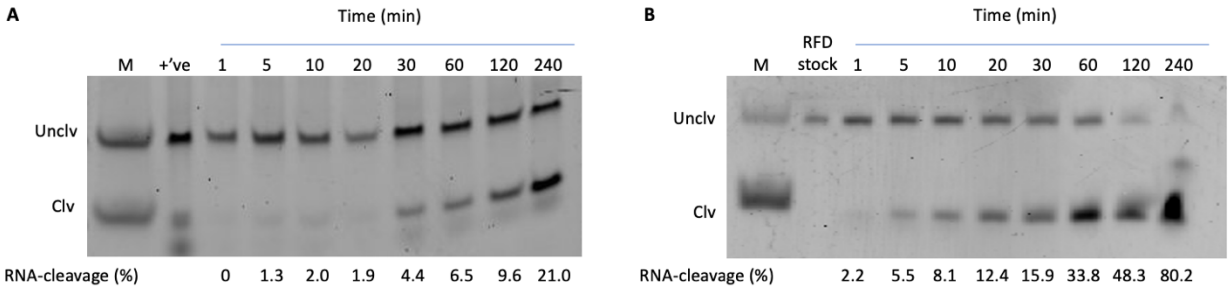


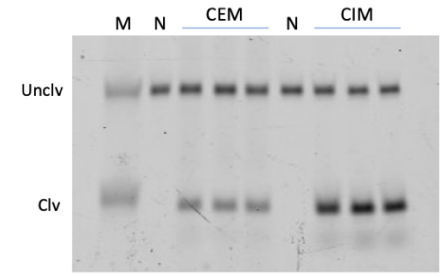
Figure 13. DNAzyme timecourse performed 16 months apart demonstrates inconsistent cleavage activities. Identical timecourse experiments using *F. nucleatum* CIM prepared Oct.18, 2017 (**A**) versus CIM prepared Feb. 27, 2019 (**B**). M: marker contains FQ30 treated with 0.2 M NaOH; RFD stock contains DNAzyme and SB; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

Despite the CIM being prepared only two weeks apart, the cleavage activities in the first timecourse (Figure 13**A**) were around 50% of the reported cleavages in the characterization assays performed in Figure 5. Surprisingly, the second timecourse (Figure 13**B**) produced some of the highest activities recorded by tRFD-FN1. To further confirm the inconsistent cleavage activities, CIM that was prepared approximately one month later reported CIM activity reduced by ~50%.

3.2.2 Target is stable during -20°C storage

Since the target displays remarkable stability in heat, we wondered about its stability in freezing storage conditions as well. To test this, both CEM and CIM were prepared, flash frozen with liquid nitrogen, and stored at -20°C. A master mix of DNAzyme probe was prepared and stored at -20°C as well. Interestingly, the ability of the target to activate tRFD-FN1 is not affected by freeze-thaw cycles, or storage at -20°C over the course of one year. The cleavage activities observed for the *F. nucleatum* CEM and CIM samples, tested at different timepoints during the experiment, did not significantly decrease over time or after re-freezing (**Figure 14)**.

A



B

% RNA-cleavage	CEM			CIM		
1 day	8.8	7.5	8.3	20.2	25.0	22.6
1 week	7.0	8.0	7.8	20.9	21.8	19.2
1 month	7.4	6.9	7.7	21.2	20.3	19.6
3 months	5.9	5.8	5.9	20.9	21.6	21.5
1 year	11.6	10.5	10.9	36.2	38.2	37.0

Figure 14. Target is still viable during -20°C storage after 1 year. A. Activity of tRFD-FN1 on CEM and CIM stored at -20°C for 1 day. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; Unclv: uncleaved FQ30; Clv: cleaved FQ30. **B.** Summary of RNA-cleavage (%) of CEM and CIM stored at -20°C for 1 day, 1 week, 1 month, 3 months and 1 year.

Over the first three months, the CIM cleavage remained the same, and the CEM cleavage decreased from around 8.2-5.9%. Surprisingly, the observed activity for CEM and CIM at one year is actually around 1.7-2.0× higher than it is at three months. The increase in cleavage activity could be the result of non-homogenized DNAzyme or CEM/CIM stocks, mechanical error in pipetting, or concentration of metal ions in 2× SB by evaporation.

3.2.3 Proteinase K test suggests that the target is a protein

The previous results demonstrated by the temperature stability assays suggest that the target might not be a protein, so *F. nucleatum* CIM was tested using Proteinase K to degrade any proteins that are present. Proteinase K was incubated with the CIM for 15 min before the reaction was performed to ensure that proteins in the CIM were destroyed. The positive CIM

control shows strong cleavage band intensities in the absence of Proteinase K; however, the cleavage activity of tRFD-FN1 is abolished in the reactions containing Proteinase K (**Figure 15**).

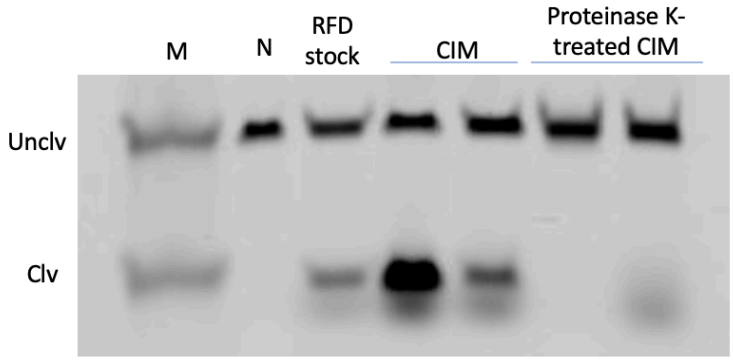


Figure 15. DNAzyme cleavage is abolish in the presence of Proteinase K-treated CIM. Activity of tRFD-FN1 in CIM treated with Proteinase K to destroy proteins. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; RFD stock contains DNAzyme and SB; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

The absence of the cleavage band indicates that the target was destroyed during Proteinase K treatment; because the target could no longer induce DNAzyme cleavage, it is likely a protein, or its ability to activate the DNAzyme is associated with a protein.

3.2.4 RNase inhibitor test indicates that the target is not RNase A, B, C, 1 or T1

Our lab has previously seen the selection of DNAzymes that were activated by or are responsive to bacterial RNases.[89] For that reason, we used two different RNase inhibitors to test tRFD-FN1's affinity for RNases; SUPERase inhibits RNases A, B, C, I and T1, and NxGen inhibits RNases A, B and C by noncovalently binding bacterial RNases and interrupting the hydrolysis of RNA. Similar to the Proteinase K reaction described above, the RNase inhibitor should inhibit the expected cleavage activity if RNases A, B, C, 1 and T1 are activating tRFD-FN1. In separate reactions, 20 U of SUPERase and 40 U of NxGen were incubated with three different CIM samples (prepared in different batches at different times), at room temperature for

15 min (**Figure 16**).

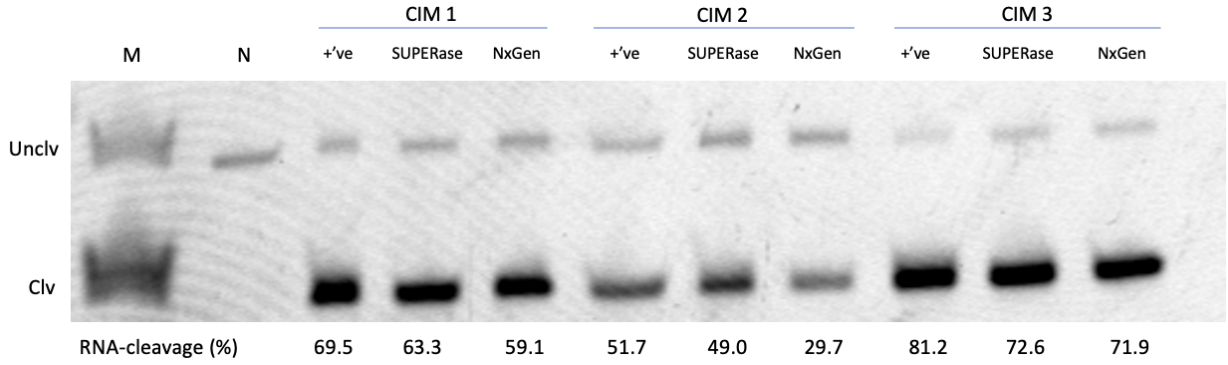


Figure 16. RNase inhibitor assay. Activity of tRFD-FN1 against CIM samples treated with SUPERase or NxGen RNase inhibitors. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; +'ve: positive control that contains CIM, DNAzyme and SB; Unclv: uncleaved FQ30; Clv: cleaved FQ30.

Treating the CIM with RNase inhibitor before testing with the DNAzyme did not abolish the probe's activity; in almost every reaction, the cleavage activity remained largely constant. In treating CIM 2 with NxGen, there was a larger reduction in activity that was likely a result of experimental error/inconsistency. Therefore, the identity of the protein is not RNase A, B, C, 1 or T1, and further experiments must be performed to elucidate the target.

3.2.5 Size exclusion chromatography to fractionate proteins

The *in vitro* selection experiment allowed us to create a DNAzyme probe that was highly selective to *F. nucleatum* without having to predetermine a suitable biomarker. However, because the sensitivity of the probe is not able to detect low concentrations of the target bacteria, we need to improve the affinity. One way to do this is to perform a reselection with the purified target; this is an attractive option for us because our target has robust properties that are ideal for enduring harsh environments such as stool. Size exclusion chromatography (SEC) (**Figure 17A**)

was used to isolate a smaller, more manageable fraction of proteins that contained the target, so that mass spectrometry could be used to ultimately help identify the target protein.

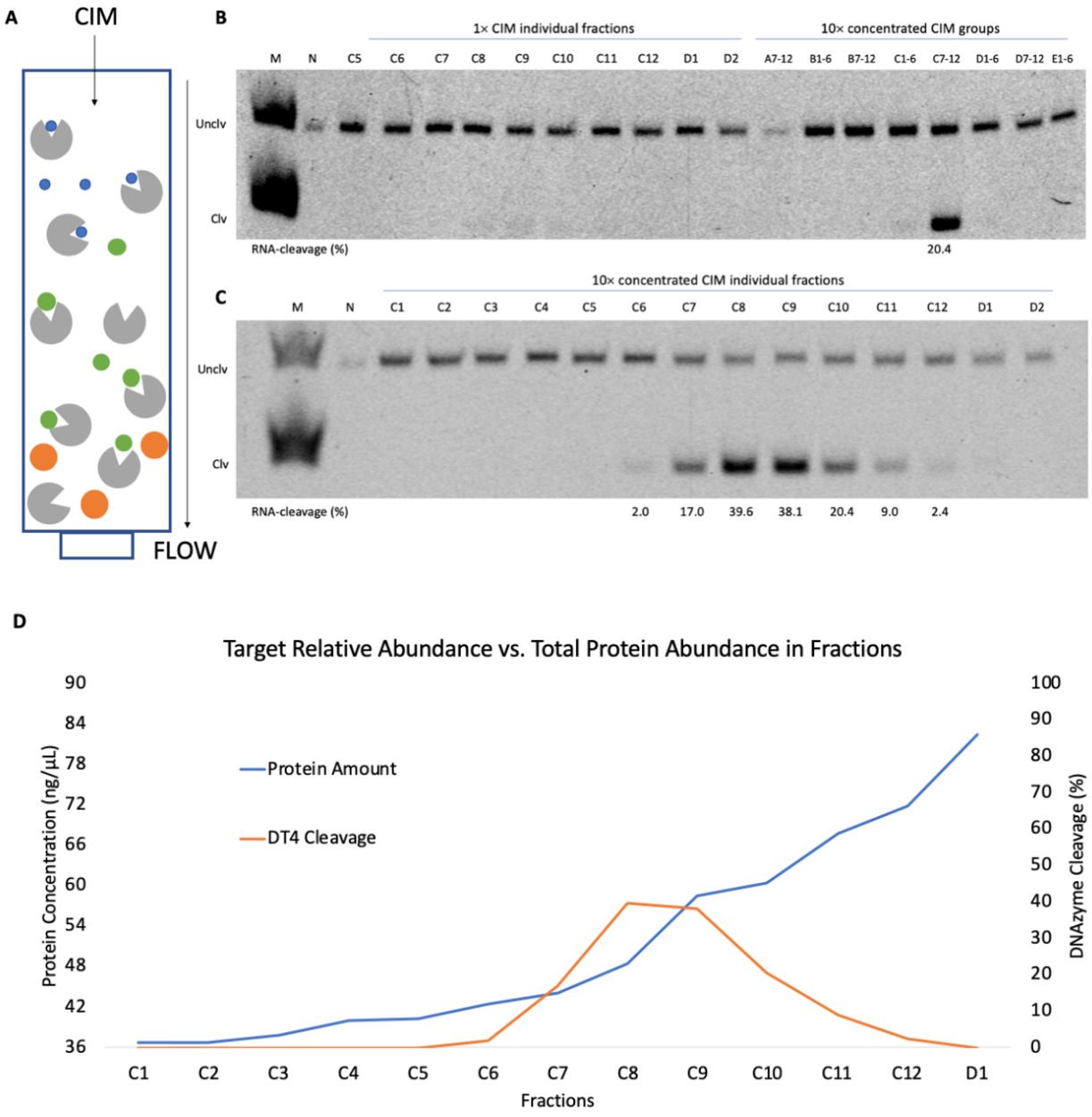


Figure 17. Size exclusion chromatography results. A. Schematic of CIM gel filtration through size exclusion column. **B.** Activity of tRFD-FN1 in C5-D2 1× SEC fractions and grouped 10× concentrated SEC fractions. **C.** Activity of tRFD-FN1 in C1-D2 10× concentrated SEC fractions. M: marker contains FQ30 treated with 0.2 M NaOH; N: negative control that contains DNAzyme, SB and ddH_2O; remaining lanes indicate isolated fractions of SEC filtrate; Unclv: uncleaved FQ30; Clv: cleaved FQ30. **D.** Graph of

the amount of protein present in each positive 10× concentrated SEC fraction versus the DNAzyme cleavage activity observed in the respective fractions.

To begin, 500 µL of each 1× and 10× CIM were prepared using 1× SB without Tween 20. The same selection buffer was used as the mobile phase of the gel filtration run. The CIM was passed through a Superdex 200 10/300 GL column and collected in 65 fractions of 500 µL volumes. To determine where the target was fractionated into, fractions were pooled into groups of six, and a DNAzyme cleavage assay was performed on each group. Upon a positive cleavage result, each fraction in that group, and a few of the neighbouring fractions, were tested for DNAzyme cleavage activity, individually. A strong presence of the target protein was observed in fractions C7-12, indicated by 20.4% DNAzyme cleavage in that lane; some of the target may also exist in fraction C6, indicated by a weak intensity band in lane C1-6 (Figure 17**B**). The 1× CIM individual fractions show extremely weak cleavage bands in the same fractions as the 10× CIM, but the 10× CIM individual fractions provided a more reliable signal; the 10× concentrated CIM SEC run was analyzed further. Based on the cleavage activities present in the individual fractions of the 10× concentrated CIM, the assumption is that the DNAzyme cleavage activity is directly proportional to the amount of target present (Slide 17**C**). The target abundance is highest in the middle two fractions (C8 and C9), with cleavage activities of 39.6% and 38.1%, respectively, and has the lowest observed abundance in C6 and C12 (2.0% and 2.4%, respectively). The protein abundance of the target, indicated by the DNAzyme cleavage activity, was compared to the total protein abundance in each fraction (Figure 17**D**). The total protein abundance, calculated by a Qubit Protein assay, shows an almost exponential increase in protein concentration; the comparison of this pattern to the DNAzyme cleavage pattern (target protein

abundance) creates a method to distinguish the cognate target from other non-target proteins in the positive fractions.

3.2.6 Mass spectrometry analysis of fractions containing target

Fractions C5-D2 were submitted to Dr. Yu Lu (Department of Biochemistry, McMaster University, Hamilton, CA) for liquid chromatography-mass spectrometry analysis (LC-MS/MS). Equal volumes of the fractions were digested into tryptic peptides, labelled, fractionated by size and loaded onto the mass spectrometry machine. The proteome database UniProt was used to analyze the proteins associated with *F. nucleatum* subsp. *nucleatum* strain ATCC 25586. Bioinformatic filters removed human contaminants like keratin. For the 82 proteins identified, the fractions in which each protein had a maximum abundance in was determined (**Figure 18A**); we focused on the proteins with the maximum abundance in fractions C8 and C9, as demonstrated in Figure 17D.

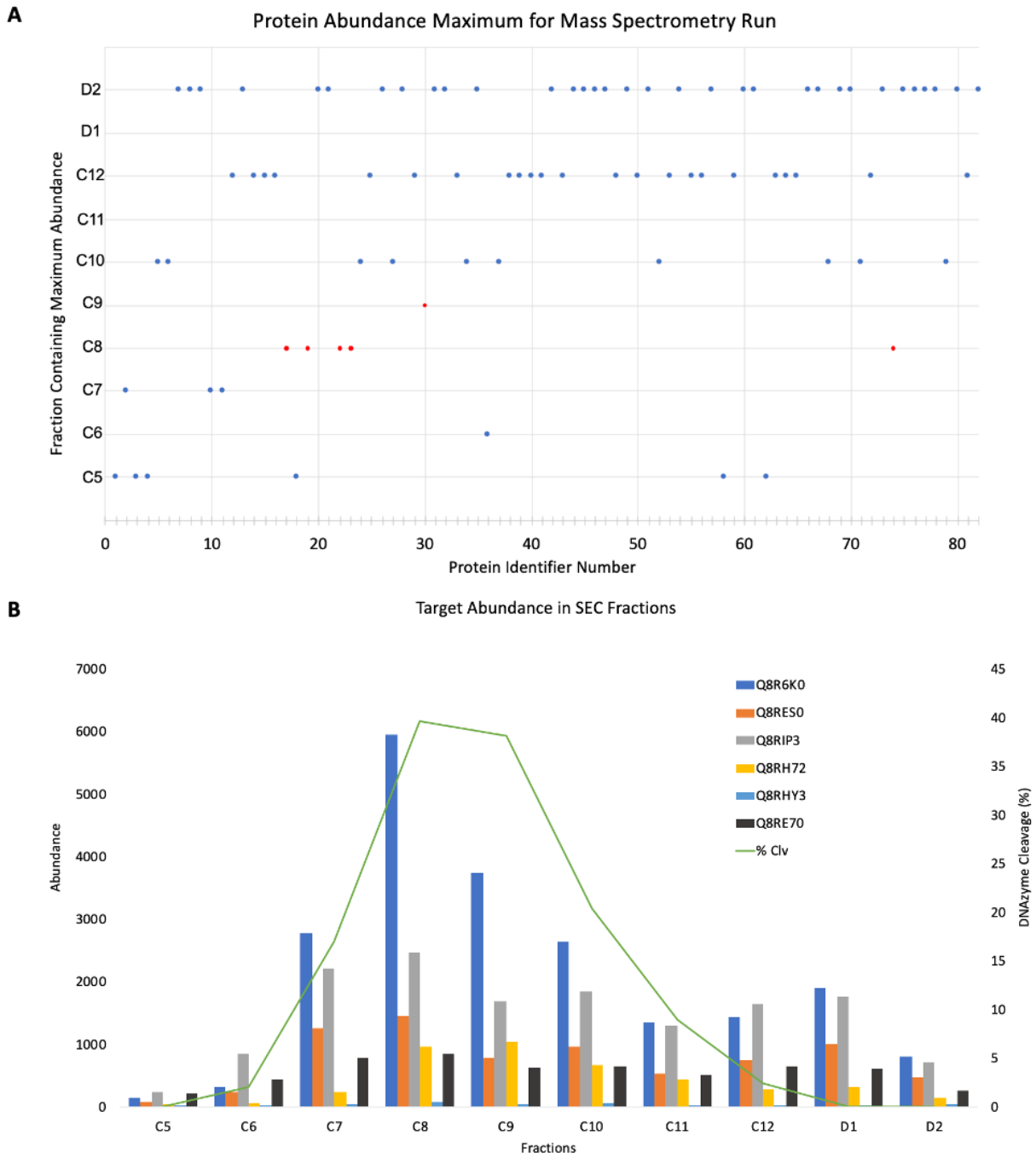


Figure 18. Potential protein targets from size exclusion chromatography. A. Points indicate the fraction that the 82 proteins had a maximum abundance in. The proteins that had the highest concentration in C8 or C9 are coloured red. **B.** Bar graph plotting the abundance of each protein in fractions C5-D2. Green line represents DNAzyme cleavage activity observed in the dPAGE of the 10× concentrated SEC fractions.

Six proteins met this criterion and each abundance pattern was compared to the distinct DNAzyme cleavage pattern (Figure 18**B**). All proteins, except Q8RE70, were determined to be a

candidate target based on how well their abundance pattern mimicked the DNAzyme cleavage pattern. Q8RH72 (T1), Q8R6K0 (T2), Q8RES0 (T3), Q8RIP3 (T4) and Q8RHY3 (T5) were all investigated further.

3.2.7 Cloning, expression and purification of potential targets

Table 4 is a summary of the five potential targets, identified by mass spectrometric analyses, that tRFD-FN1 could be activated by.

Table 4. Potential target summary

Target	UniProt Entry Name	Protein Name	Length (amino acids)	Molecular Weight (Da)	pI	% identity to proteins in FUSNP (E value); Query coverage	% identity to proteins in 9FUSO (E value); Query coverage
T1	Q8RH72 _FUSNN	Uncharacterized	100	12615.61	8.65	91.00% ($3e^{-60}$); 100%	39.33% ($1e^{-16}$); 88%
T2	Q8R6K0 _FUSNN	Uncharacterized—possible FadA relationship	119	15694.76	5.84	93.28% ($1e^{-74}$); 100%	53.92% ($4e^{-32}$); 85%
T3	Q8RES0 _FUSNN	DNA-binding protein HU	102	12603.78	10.00	99.02% ($3e^{-66}$); 100%	60.22% ($5e^{-35}$); 91%
T4	Q8RIP3_ FUSNN	Uncharacterized	118	14878.94	4.93	91.53% ($9e^{-58}$); 100%	33.60% ($5e^{-13}$); 88%
T5	Q8RHY3 _FUSNN	Acetoacetate: butyrate/ acetate coenzyme A transferase	217	24424.36	6.05	93.55% ($3e^{-151}$); 100%	83.26% ($7e^{-128}$); 99%

Note: pI= isoelectric point; FUSNN= *Fusobacterium nucleatum* subspecies *nucleatum;* FUSNP= *Fusobacterium nucleatum* subspecies *polymorphum;* 9FUSO= *Fusobacterium necrophorum* subspecies *necrophorum*

A bioinformatics search was performed to learn a little more about the potential targets. T1, T2 and T4 were all uncharacterized proteins—T2 had a strong familial relationship with FadA proteins—while T3 and T5 were DNA-binding protein HU and acetoacetate:butyrate/ acetate coA transferase, respectively. Blast results for the five potential targets against *F.*

nucleatum subsp. *polymorphum* (FNp) and *F. necrophorum* subsp. *necrophorum* (FNn) were compared to a specificity test performed by Qian (**Figure 19**).[78]

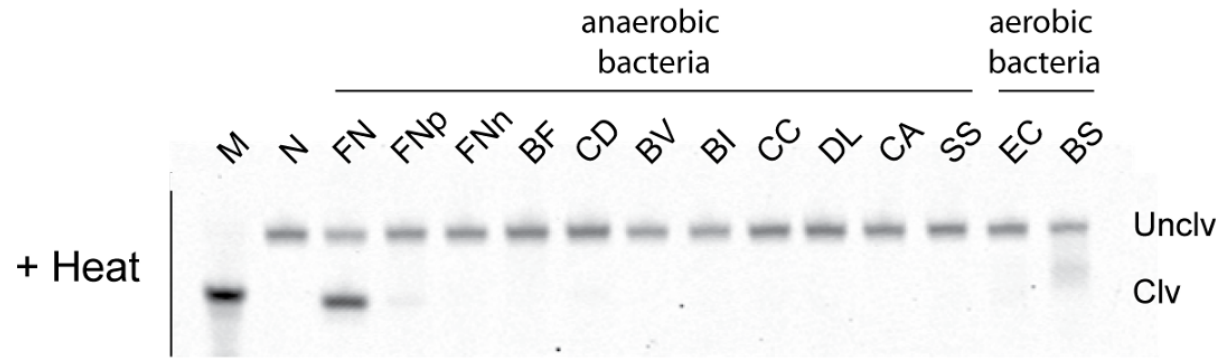


Figure 19. Specificity test indicates that the DNAzyme is specific to *F. nucleatum* subsp. *nucleatum*. "Reactivity of DT4 in the presence of CIM of some common anaerobic bacteria in human gut flora and in the presence of CIM of two common aerobic bacteria. Anaerobic bacteria: FN, *Fusobacterium nucleatum* subsp. *nucleatum*; FNp, *Fusobacterium nucleatum* subsp. *polymorphum*; FNn, *Fusobacterium necrophorum* subsp. *necrophorum*; BF, *Bacteriodes fragilis*; CD, *Clostridium difficile* NAP1; BV, *Bacteroides vulgatus*; BI, *Bacteroides intestinalis*; CC, *Coprococcus comes*; DL, *Dorea longicatena*; CA, *Collinsella aerofaciens*; SS, *Streptococcus salivarius*. Aerobic bacteria: EC, *Escherichia coli* K12; BS, *Bacillus subtilis*. Heat: CIM additionally heated 90ºC for 15 min." Figure from Qian Feng thesis.[78]

The specificity test shows that FNp incubation with the DNAzyme was able to generate a very slight cleavage band, which might indicate very low levels of target, or perhaps a mutant form of the *F. nucleatum* subsp. *nucleatum* protein. The closely related species, FNn, was not able to induce any cleavage, so it could either contain a mutated sequence that is not able to be recognized by the DNAzyme, or not have the same protein at all. All five potential targets were identified to have >90% identity to a protein in FNp for 100% query coverages. In comparison, the query coverages of the closest related FNn proteins were between 88-99%, and their identities were from 39-83%. These results seem consistent with the results observed in the specificity test.

The molecular weights of the potential targets are much smaller than the target size that was predicted with the Amicon filters; however, they agree with an SDS-PAGE performed on the fractions of an earlier SEC run, using 30× CIM (**Figure 20**).

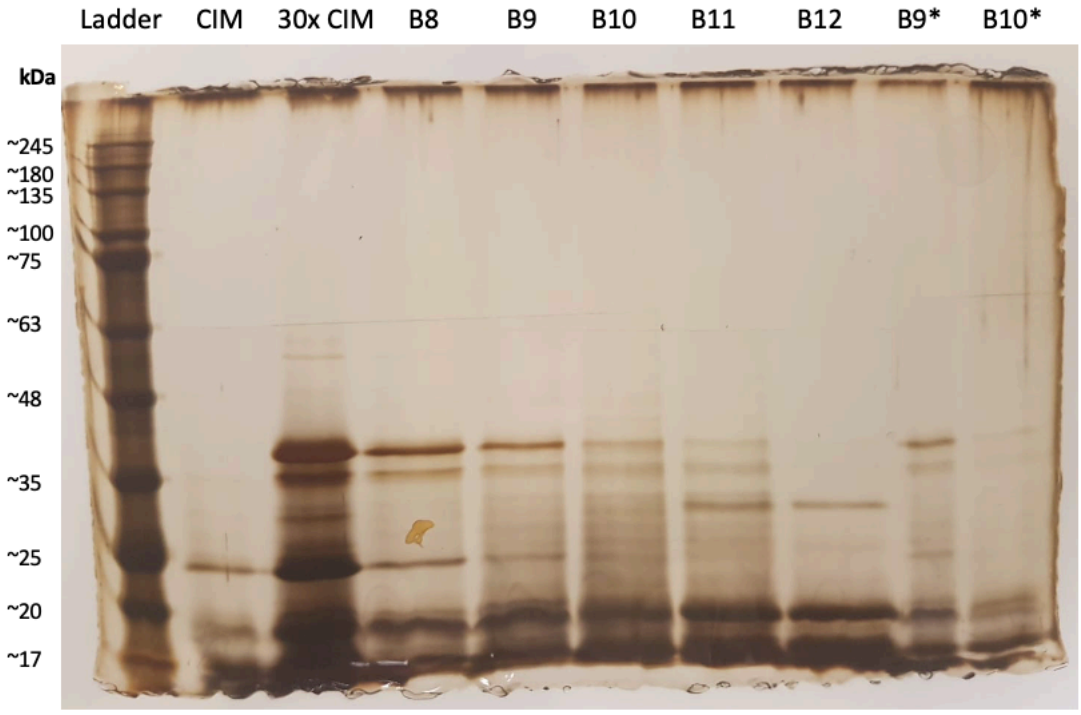


Figure 20. Bulk of proteins in silver stain of SEC fractions on SDS-PAGE occur within 17-25 kDa. A SEC run using 30× concentrated CIM was run on SDS-PAGE to separate and analyze potential target within positive fractions. Fractions B10 and B11 contain the largest amount of target, and fractions B8 and B12 contain the smallest amount of target, indicated by DNAzyme cleavage assay. * indicates fractions were 1/3 diluted.

The gel was silver stained to visualize the proteins that were present in the fractions that tested positive in a DNAzyme cleavage assay. The majority of the proteins present in these fractions appear to be between 17-25 kDa in size. The smallest size indicated by the ladder was 17 kDa, but a number of proteins appear to be smaller than that size as well. The discrepancy of this result to the Amicon filter size estimation could be because the protein might exist in a complex under non-denaturing conditions and could not pass through the molecular size cut-off of 30K. In addition, the isoelectric point (pI) of the target may also be a key reason why it was

not able to pass through the smaller membrane; at the pH of the selection buffer (pH 7.5), T1 and T3 carry strong positive charges, while T2, T4 and T5 carry negative charges. The stronger charges may cause the CIM to aggregate or bind to the membrane.

Each target was purified with the 6×His-tag and tested for DNAzyme cleavage activity. There was no activity observed from the culture or CIM of the induced, transformed bacteria. The samples were tested with and without RNase inhibitor and were treated with heat at 90°C for 10 min, or mechanical disruption to lyse the cells. Next, to test if a cofactor was needed for the target to be recognized by tRFD-FN1, *F. nucleatum* CIM was added to each sample. The addition of CIM did not induce an increased cleavage activity compared to the CIM control. The presence of the 6×His-tag may interfere with the DNAzyme:target interaction, so we decided to remove the His-tag and perform the cleavage assays again.

T1, T2 and T4 were successfully purified with the His-tag removed and tested using tRFD-FN1. T2 was tested at concentrations of 500, 50, 5, 0.5 and 0.05 ng/mL, and T1 and T4 were tested at 100, 50, 5, 0.5 and 0.05 ng/mL. We observed a faint band at the 500 ng/mL concentration of T2, but the band was not present when the target was heated at 90°C for 10 min. T1 and T4 did not exhibit any activity in the conditions tested. T3 and T5 still need to be purified without the His-tag and tested to see if they are the cognate target of tRFD-FN1.

Chapter 4: DISCUSSION

4.1 Functionalize tRFD-FN1 to detect *F. nucleatum* in CRC-confirmed, patient stool samples

The prognosis of CRC after a late-stage diagnosis has highlighted the need for a sensitive early-stage screening method with high compliance. To encourage compliance, the novel device would ideally be non-invasive, easy-to-use, accurate, rapid, and able to be performed at home by the patient. RNA-cleaving fluorogenic DNAzymes have the potential to meet all of these requirements. A number of RFDs, produced by our lab, are in various stages of development for use in on-site bacterial detection applications[61,85,87]; however, the search for an RFD that recognizes stool-borne bacterial biomarkers for cancer is among the first of its kind.

Although the selections done directly in stool before this project were unsuccessful, an *F. nucleatum* CEM-selection yielded a DNAzyme that could respond selectively to an unknown target associated with the CRC bacterium. Preliminary tests demonstrated that after incubation with a CRC-positive stool sample, the probe was unable to detect the presence of *F. nucleatum*. However, research supporting the marked overabundance of *F. nucleatum* in pre- and cancerous tissue and stool led us to believe that we could functionalize the existing DNAzyme to detect our unknown bacterial biomarker in stool samples.

The initial experiments were performed to characterize RFD-FN1's cleavage ability and its interaction with its target in a pure culture. The DNAzyme cleavage results for *F. nucleatum* CEM and CIM suggest that the mixtures could contain ~30-38% of their volume as target. Conversely, the cleavage activities could indicate a weaker affinity to the target by the DNAzyme. tRFD-FN1's LOD of 10^7 CFU/mL also supports a weak affinity—compared to other

studies that have recorded the selection of significantly more sensitive DNAzymes. tRFD-FN1 is four-fold less sensitive than similar RFDs developed in our lab.[87,90]

Our first approach to functionalize tRFD-FN1 in CRC samples was to manipulate the stool matrix to make it more conducive to DNAzyme detection. Working with stool is a relatively newer endeavour, and so we could not predict what was interrupting the DNAzyme:target interaction from proceeding as expected. The prevalence of non-specific substrate degradation by nucleases prompted us to heat the CEM and CIM prior to a cleavage assay. Remarkably, after 30 min at 90°C, tRFD-FN1 still demonstrated 58% cleavage relative to the non-heated assay. Since the unknown target displayed significant heat-stability, we hypothesize that its structure is impervious to high-heat denaturation during shorter periods of 90°C exposure—5-10 min. The advantage of having a heat-stable target here is two-fold: 1) we can apply heat before a DNAzyme cleavage assay to ensure that non-specific cleavage is not occurring, and 2) we can apply heat to the stool samples to try and eliminate any molecules that interfere with the DNAzyme's activity or the DNAzyme's target. Nucleases, which are prominent constituents in numerous testing environments, including stool, can cleave the DNAzyme substrate and generate a false-positive signal. The ability to simply use heat to inactivate them is convenient and provides a reliable result. In addition, although the positive- and counter-selection procedures deliver high specificity, heat treatment prior to a cleavage assay can further ensure that any potential cross-reactive molecules that may be present are inactivated. Finally, in trying to functionalize our DNAzyme in stool, we expected to encounter a number of unknown challenges due to the inestimable number of biologics that could be present. Having the ability to use heat-treatment as a tool to disable some of these biological molecules was valuable to have in our arsenal.

Although unable to detect any *F. nucleatum* cells that may be naturally present in the stool samples, we were able to detect bacteria that we spiked in after removing the organic component of the samples. The presence of *F. nucleatum* in the aqueous fraction of the stool allowed for DNAzyme detection to occur. Eliminating gram-positive bacteria had no effect on detection abilities, indicating that those organisms were not producing the interfering molecule, nor were they likely outcompeting the *F. nucleatum* in the stool cultures. Lyophilization of the samples did not produce a testing environment that was conducive to DNAzyme detection either. The freeze-drying process had no impact on the DNAzyme interference, and the subsequent concentration of stool made it hard to interpret any results from the dPAGE. Unfortunately, in using these processing methods, we risk compromising the endogenous *F. nucleatum* numbers that are present for each individual and lose the ability to diagnose potential disease.

The fluorescent signal readout method used for this probe also experienced some challenges working in the novel complex matrix. In gels containing stool samples, some lanes contained dark smears down the length of the lane and a concentrated signature underneath the cleavage band. The smears are caused by autofluorescence in the stool; it is not present in all patient samples to the same degree but varies from individual to individual. A similar autofluorescent signature, underneath the cleavage band, is also present in CMB media. The autofluorescence in respective samples is reduced after heating. Overall, the method is still effective for detection in stool, but signal interpretation was sometimes difficult in assays near the DNAzyme's LOD.

Despite the challenges encountered during testing in stool, tRFD-FN1 was able to function suitably in saliva samples. The LOD of tRFD-FN1 is the same for *F. nucleatum* in saliva and in pure cultures. As expected, heating the saliva samples did not affect the cleavage

activity, however, it did eliminate the banding pattern underneath the location of the cleaved band. The low cleavage activities observed in the saliva samples that were not spiked with *F. nucleatum* suggest that the samples contain around 10^7 CFU/mL, which is consistent with literature.[91,92] The DNAzyme test, done in saliva, may even be able to differentiate between CRC and healthy patients.[91]

Since we could not find a successful processing method for detecting *F. nucleatum* levels naturally present in stool, our second approach was to enrich the number of bacteria so that our DNAzyme would be able to detect the presence of its target. We chose to incubate the stool samples anaerobically, in the hopes that we could culture the *F. nucleatum* present, and subsequently produce target concentrations agreeable with tRFD-FN1. The only successful culture-based technique, that we observed, requires the stool to be heated at 90°C and inoculated with *F. nucleatum* colonies before a 48-h incubation.

Although difficult to coordinate, we were able to perform the stool-culturing tests on one fresh sample, to determine if the maximum amount of target existed in the sample before storage occurred. The sample was not treated with heat before the incubation took place. We observed possible DNAzyme activity in the 100-fold dilutions and higher, however, we could not rule out non-specific degradation as the reason for the cleavage bands, due to an abnormal band intensity pattern. While we should have seen increasing or decreasing cleavage intensities across the dilutions, all percentages were roughly the same. This suggests that an alternate factor in the stool could be initiating the cleavage signal that we observed in the gel. Whenever possible, more experiments should be performed on fresh stool samples; the reduced time from collection to testing increases the likelihood that the target is intact and has not been degraded by destructive molecules like nucleases or proteases, or even oxygen.

While our results have demonstrated that the DNAzyme is capable of detecting its target in the stool complex, the conditions that we must use to achieve that hinder our ability to detect any target that may be naturally present. The use of heat to promote tRFD-FN1's functionality in stool is a double-edged sword: the heat is necessary to make the stool environment conducive to target recognition and cleavage signal generation, but it is likely destroying the *F. nucleatum* population present in the diseased stool samples. We must find a way to functionalize the DNAzyme in stool without compromising the target and the respective diagnostic results.

4.2 RFD-FN1 target identification

While our DNAzyme is not powerful enough to detect its cognate target at the levels that it is naturally present in stool, further investigation regarding RFD-FN1's target could provide us with two advantages. The first advantage is the possibility to enrich for the target to levels detectable by RFD-FN1. The second advantage, if we can ascertain the identity of the target, would be to perform reselection of the DNAzyme with the purified target and increase the probe's sensitivity and affinity towards it.

A major challenge we encountered here was the consistency of producing the target in *F. nucleatum* cultures. Across six experiments, designed to test a number of culturing conditions to produce the maximum amount of target, we saw large discrepancies in cleavage activity in replicates performed at different times. Although we were able to achieve around 90% DNAzyme cleavage in a single test, the inconsistencies prevented us from determining the optimal conditions necessary to reproduce this result consistently.

The *F. nucleatum* target that RFD-FN1 is activated by is an interesting molecule. In addition to its stability at 90°C, it was also shown to be stable during multiple freeze-thaw cycles and storage at -20°C for one year. A proteinase assay revealed that the target is a protein, which

is surprising given its remarkable temperature stability. Adding to that idea, *F. nucleatum* are not necessarily considered extremophiles, nor are they spore-forming, so the function of this protein is intriguing. Most proteins are typically denatured at temperatures around 90°C, and those that are resistant to freeze-thaw cycles are usually more common in soil bacteria, which have learned to survive winter seasons.[93]

Of the 82 proteins that were identified in the positive gel filtration fractions, the top 5 potential candidates were chosen by their abundance relative to the DNAzyme cleavage percentages in each fraction. The potential targets were labelled T1 to T5 and were ordered according to how well they fit the cleavage pattern. While only T3 (Q8RES0) and T5 (Q8RHY3) were characterized proteins, T2 (Q8R6K0) was identified to have a strong familial relationship with FadA proteins. As previously mentioned, FadA is an *F. nucleatum* virulence factor that helps the bacterium with colonization, breaching epithelial barriers, and stimulating biofilm formation. Upon more investigation, T2 could be a promising biomarker to use in other *F. nucleatum*-detection systems, as well as in therapeutic situations. If *F. nucleatum* is identified to be a causative factor in certain diseases, aptamer sequences could be designed to bind to and block the function of this virulence protein. T3 is a DNA-binding protein, which is also a good candidate for the cognate target of the DNAzyme. Since RFD-FN1 is constructed primarily of deoxyribonucleic acids, it is reasonable that it would be able to recognize T3 with a higher affinity than other proteins.

The size of each protein is between 12.6–24.5 kDa, which agrees with the SDS-PAGE analysis performed on positive SEC fractions. The Amicon filter size estimation, performed on *F. nucleatum* CIM, indicated a much larger size, but these results could ultimately give us more clues towards the target's identity. Each protein has an isoelectric point that results in a strong

positive or negative charge in the pH 7.5 selection buffer, which means the target could have aggregated on the 30K Amicon filter membrane, rather than pass through. Additionally, the target could exist as part of a complex, or in dimer form, which would be one reason why the target remained on top of the filter.

To determine if any of the proteins are RFD-FN1's cognate target, a DNAzyme cleavage test was performed on the purified proteins. The proteins were initially purified with an N-terminal 6×His-tag still attached, and then tested with the DNAzyme. This method did not produce a positive result, likely because of the effect that the His-tag had on the proteins. Polyhistidine is a supercharged molecule and has the ability to influence the structure of other molecules, including a smaller protein it is attached to, or the DNAzyme trying to bind to its target. Therefore, we determined that the potential proteins, purified and with the His-tag removed, would give us a reliable idea of the identity of RFD-FN1's target. However, the method is not infallible, and the potential for a false-negative still exists: the purified protein might require a bacterial cofactor for recognition by the DNAzyme; the target might be a complex of proteins; the target may actually be a small molecule that binds to a protein *in vivo*.

Three of the five targets have been expressed, purified and tested with tRFD-FN1 to date. T1 and T4 did not demonstrate any cleavage activity when incubated with the DNAzyme. The gel for T2 revealed a weak band at a very high concentration of target, but this was eliminated in the reaction that contained heated target. No other concentration indicated any affinity towards tRFD-FN1. Targets T3 and T5 still need to be purified and tested. Unfortunately, T3, like T1, has a high pI value and is quite difficult to purify. The nature of T1 is quite interesting; T1 appeared to be toxic to the BL21 cells, as the overnight culture only reached 0.7 OD_{600}, compared to the uninduced OD_{600} of 1.7. The protein was also difficult to purify and concentrate using the

necessary columns. Similarly, the OD_{600} of the induced T3 cells is around 1.0 and is troublesome during purification. Efforts to obtain purified T3 to use in the DNAzyme test is ongoing. Due to the extraordinary circumstances that is the COVID-19 pandemic, we were not able to complete the purification and testing of the two remaining potential targets—T3 and T5.

4.3 Conclusion

The goal of this project was to functionalize the *F. nucleatum*-specific DNAzyme, RFD-FN1, in stool samples to diagnose CRC. While the current probe is able to detect *F. nucleatum* in heated, incubated stool, RFD-FN1 is not sensitive enough to detect the *F. nucleatum* cells at the concentration that is naturally present in the samples. While the application of RFD-FN1 in stool still requires more work, the DNAzyme can presently be used to detect the overabundance of the pathogen in saliva—a characteristic sign of periodontal disease. In order to be able to functionalize RFD-FN1 in stool samples, we believe that the sensitivity of the probe needs to be improved. We have already seen this sensitivity achieved by bacterial DNAzymes within our lab.

The most effective way to improve the DNAzyme's sensitivity is to perform reselection with the purified cognate target. The target that RFD-FN1 responds to is robust, heat-stable and specific to *F. nucleatum* subsp. *nucleatum*. For those reasons, this is an advantageous biomarker for pathogenic *F. nucleatum*, and its identity could be useful in future applications and for reselection to improve sensitivity. After *F. nucleatum* proteins were fractionated by size to provide a manageable pool of potential targets, we narrowed it down to five proteins that RFD-FN1 could be responding to. Three of the five proteins did not elicit a DNAzyme response and are likely not the cognate target, so the other two proteins must still be tested. Due to unusual circumstances, we were not able to purify all the proteins in time. The identification of RFD-

FN1's target is a priority for this project, and the results from this will hopefully be the catalyst that helps to functionalize the DNAzyme as a non-invasive CRC screening method.

REFERENCES

1. Canadian Cancer Society Statistics Advisory Committee. Canadian cancer statistics, 2019. cancer.ca/Canadian-Cancer-Statistics-2019-EN. Updated 2019. Accessed Mar 16, 2020.

2. National Cancer Institute. Colon cancer treatment (PDQ)-health professional version. https://www.cancer.gov/types/colorectal/hp/colon-treatment-pdq#_79. Accessed May 25, 2020.

3. Canadian Cancer Society Statistics Advisory Committee. Survival statistics for colorectal cancer. https://www.cancer.ca/en/cancer-information/cancer-type/colorectal/prognosis-and-survival/survival-statistics/?region=on. Accessed Mar 16, 2020.

4. Canadian Task Force on Preventive Health Care. Recommendations on screening for colorectal cancer in primary care. *Can Med Assoc J*. 2016;188(5):340-348.

5. Prince M, Lester L, Chiniwala R, Berger B. Multitarget stool DNA tests increases colorectal cancer screening among previously noncompliant medicare patients. *World J Gastroenterol*. 2017;23(3):464-471.

6. Tepus M, Yau TO. Non-invasive colorectal cancer screening: An overview. *Gastrointest Tumors*. 2020;7:62-73.

7. American College of Gastroenterology. Colonoscopy. https://gi.org/topics/colonoscopy/#tabs2. Accessed June 2, 2020.

8. Solbak NM, Xu J, Vena JE, et al. Patterns and predictors of adherence to colorectal cancer screening recommendations in Alberta's Tomorrow Project participants stratified by risk. *BMC Public Health*. 2018;18:177.

9. Riemann JF, Schröder C, Kallenbach M, Schmoll HJ. Benefits and risks of colorectal cancer screening. *Oncol Res Treat*. 2014;37(Suppl 3):11-20.

10. Niederreiter M, Niederreiter L, Schmiderer A, Tilg H, Djanani A. Colorectal cancer screening and prevention—pros and cons. *Memo*. 2019;12(3):239-243.

11. Johnson CM, Wei C, Ensor JE, et al. Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control*. 2013;24(6):1207-1222.

12. Antonic V, Stojadinovic A, Kester KE, et al. Significance of infectious agents in colorectal cancer development. *J Cancer*. 2013;4(3):227-240.

13. Sender R, Fuchs S, Milo R. Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell*. 2016;164(3):337-340.

14. Turnbaugh P, Ley R, Hamady M, Fraser-Liggett C, Knight R, Gordon JI. The human microbiome project: Exploring the microbial part of ourselves in a changing world. *Nature*. 2007;449(7164):804-810.

15. Ferreira RM, Pereira-Marques J, Pinto-Ribeiro I, et al. Gastric microbial community profiling reveals a dysbiotic cancer-associated microbiota. *Gut*. 2018;67(2):226-236.

16. Okamoto T, Hayashi Y, Mizuno H, et al. Colonization of an acid resistant *Kingella denitrificans* in the stomach may contribute to gastric dysbiosis by *Helicobacter pylori*. *J Infect Chemother*. 2014;20(3):169-174.

17. Strauss J, Kaplan GG, Beck PL, et al. Invasive potential of gut mucosa-derived *Fusobacterium nucleatum* positively correlates with IBD status of the host. *Inflamm Bowel Dis*. 2011;17(9):1971-1978.

18. Joossens M, Huys G, Cnockaert M, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut*. 2011;60(5):631-637.

19. Antharam VC, Li EC, Ishmael A, et al. Intestinal dysbiosis and depletion of butyrogenic bacteria in *Clostridium difficile* infection and nosocomial diarrhea. *J Clin Microbiol*. 2013;51(9):2884-2892.

20. Kostic AD, Gevers D, Pedamallu CS, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res*. 2012;22(2):292-298.

21. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol*. 2014;12(10):661-672.

22. Haque SZ, Haque M. The ecological community of commensal, symbiotic, and pathogenic gastrointestinal microorganisms – an appraisal. *Clin Exp Gastroenterol*. 2017;10:91-103.

23. Gagniere J, Raisch J, Veziant J, et al. Gut microbiota imbalance and colorectal cancer. *World J Gastroenterol*. 2016;22(2):501-518.

24. Nasidze I, Li J, Quinque D, Tang K, Stoneking M. Global diversity in the human salivary microbiome. *Genome Res*. 2009;19(4):636-643.

25. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol*. 2005;43(11):5721-5732.

26. Signat B, Roques C, Poulet P, Duffaut D. *Fusobacterium nucleatum* in periodontal health and disease. *Curr Issues Mol Biol*. 2011;13(2):25-36.

27. Han YW. *Fusobacterium nucleatum*: A commensal-turned pathogen. *Curr Opin Microbiol*. 2015;23:141-147.

28. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/ β-catenin signaling via its FadA adhesin. *Cell Host Microbe*. 2013;14(2):195-206.

29. Kostic AD, Chun E, Robertson L, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. 2013;14(2):207-215.

30. Sears CL, Pardoll DM. Perspective: alpha-bugs, their microbial partners, and the link to colon cancer. *J Infect Dis*. 2011;203(3):306-311.

31. Tahara T, Yamamoto E, Suzuki H, et al. *Fusobacterium* in colonic flora and molecular features of colorectal carcinoma. *Cancer Res*. 2014;74(5):1311-1318.

32. Wang Q, Ye J, Fang D, et al. Multi-omic profiling reveals associations between the gut mucosal microbiome, the metabolome, and host DNA methylation associated gene expression in patients with colorectal cancer. *BMC Microbiol*. 2020;20:1-13.

33. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol*. 2014;12(10):661-672.

34. Amitay EL, Werner S, Vital M, et al. *Fusobacterium* and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. *Carcinogenesis*. 2017;38(8):781-788.

35. Rothenbroker M, Zakaria S, Morrison D, Li Y. DNAzymes: Synthetic enzymes made of DNA. In: Husain Q, Ullah MF, eds. *Biocatalysis.* Springer, Cham; 2019:1-16.

36. Liu J, Cao Z, Lu Y. Functional nucleic acid sensors. *Chem Rev*. 2009;109(5):1948-1998.

37. Gold L, Polisky B, Uhlenbeck O, Yarus M. Diversity of oligonucleotide functions. *Annu Rev Biochem*. 1995;64(1):763-797.

38. Zhang J, Lan T, Lu Y. Molecular engineering of functional nucleic acid nanomaterials toward in vivo applications. *Adv Healthc Mater*. 2019;8(6):1801158.

39. Tombelli S, Minunni M, Mascini M. Analytical applications of aptamers. *Biosens Bioelectron*. 2005;20(12):2424-2434.

40. Li Y, Sen D. Toward an efficient DNAzyme. *Biochemistry*. 1997;36(18):5589-5599.

41. Breaker RR, Joyce GF. A DNA enzyme that cleaves RNA. *Chem Biol*. 1994;1(4):223-229.

42. Chandra M, Sachdeva A, Silverman SK. DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat Chem Biol*. 2009;5(10):718-720.

43. Flynn-Charlebois A, Wang Y, Prior TK, et al. Deoxyribozymes with 2'−5' RNA ligase activity. *J Am Chem Soc*. 2003;125(9):2444-2454.

44. Sreedhara A, Li Y, Breaker RR. Ligating DNA with DNA. *J Am Chem Soc*. 2004;126(11):3454-3460.

45. Li Y, Breaker RR. Phosphorylating DNA with DNA. *Proc Natl Acad Sci U S A*. 1999;96(6):2746-2751.

46. Cheglakov Z, Weizmann Y, Beissenhirtz MK, Willner I. Ultrasensitive detection of DNA by the PCR-induced generation of DNAzymes: The DNAzyme primer approach. *Chem Commun*. 2006(30):3205-3207.

47. Hesser AR, Brandsen BM, Walsh SM, Wang P, Silverman SK. DNA-catalyzed glycosylation using aryl glycoside donors. *Chem Commun*. 2016;52(59):9259-9262.

48. Zhou C, Avins JL, Klauser PC, Brandsen BM, Lee Y, Silverman SK. DNA-catalyzed amide hydrolysis. *J Am Chem Soc*. 2016;138(7):2106-2109.

49. Zhang K, Li Y, Liu J, et al. Y-shaped circular aptamer-DNAzyme conjugates for highly efficient in vivo gene silencing. *CCS Chemistry*. 2020:631-641.

50. Cui M, Li X, Xu J, Chen H. Acid-switchable DNAzyme nanodevice for imaging multiple metal ions in living cells. *ACS Appl Mater Interfaces*. 2020;12(11):13005-13012.

51. Kölhler T, Patsis PA, Hahn D, et al. DNAzymes as catalysts for l-tyrosine and amyloid β oxidation. *ACS Omega*. 2020;5(13):7059-7064.

52. Tang J, Breaker RR. Rational design of allosteric ribozymes. *Chem Biol*. 1997;4(6):453-459.

53. Hesselberth J, Robertson MP, Jhaveri S, Ellington AD. In vitro selection of nucleic acids for diagnostic applications. *J Biotechnol*. 2000;74(1):15-25.

54. Zhong G, Wang H, Bailey CC, Gao G, Farzan M. Rational design of aptazyme riboswitches for efficient control of gene expression in mammalian cells. *Elife*. 2016;5:e18858.

55. Silverman SK. Artificial functional nucleic acids: Aptamers, ribozymes, and deoxyribozymes identified by in vitro selection. In: Li Y, Yi L, eds. *Functional nucleic acids for analytical applications.* Springer; 2009:47-108.

56. Zhou W, Ding J, Liu J. Theranostic DNAzymes. *Theranostics*. 2017;7(4):1010-1025.

57. Ma H, Liu J, Ali MM, et al. Nucleic acid aptamers in cancer research, diagnosis and therapy. *Chem Soc Rev*. 2015;44(5):1240-1256.

58. Li Y, Lu Y. *Functional nucleic acids for analytical applications.* Springer; 2009.

59. Schlosser K, Li Y. Biologically inspired synthetic enzymes made from DNA. *Chem Biol*. 2009;16(3):311-322.

60. Peng H, Newbigging AM, Wang Z, et al. DNAzyme-mediated assays for amplified detection of nucleic acids and proteins. *Anal Chem*. 2018;90(1):190-207.

61. Ali MM, Aguirre SD, Lazim H, Li Y. Fluorogenic DNAzyme probes as bacterial indicators. *Angew Chem Int Ed Engl*. 2011;50(16):3751-3754.

62. Nazarenko IA, Bhatnagar S, Hohman R. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res*. 1997;25(12):2516-2521.

63. Li Y, Breaker RR. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group. *J Am Chem Soc*. 1999;121(23):5364-5372.

64. Fowler CC, Navani NK, Brown ED, Li Y. Aptamers and their potential as recognition elements for the detection of bacteria. In: Zourob M, Elwary S, Turner A, eds. *Principles of bacterial detection: Biosensors, recognition receptors and microsystems.* Springer; 2008:689-714.

65. Silverman SK. In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res*. 2005;33(19):6151-6163.

66. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. 1990;249(4968):505-510.

67. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature*. 1990;346(6287):818-822.

68. Robertson DL, Joyce GF. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature*. 1990;344(6265):467-468.

69. Carmi N, Shultz LA, Breaker RR. In vitro selection of self-cleaving DNAs. *Chem Biol*. 1996;3(12):1039-1046.

70. Peracchi A. DNA catalysis: Potential, limitations, open questions. *ChemBioChem*. 2005;6(8):1316-1322.

71. Vorobyeva MA, Davydova AS, Vorobjev PE, Pyshnyi DV, Venyaminova AG. Key aspects of nucleic acid library design for in vitro selection. *Int J Mol Sci*. 2018;19(2):470.

72. Pobanz K, Lupták A. Improving the odds: Influence of starting pools on in vitro selection outcomes. *Methods*. 2016;106:14-20.

73. Achenbach J, Chiuman W, Cruz R, Li Y. DNAzymes: From creation in vitro to application in vivo. *Curr Pharm Biotechnol*. 2004;5(4):321-336.

74. Lipps HJ, Rhodes D. G-quadruplex structures: In vivo evidence and function. *Trends Cell Biol*. 2009;19(8):414-422.

75. Li T, Shi L, Wang E, Dong S. Multifunctional G-quadruplex aptamers and their application to protein detection. *Chemistry*. 2009;15(4):1036-1042.

76. Nieuwlandt D. In vitro selection of functional nucleic acid sequences. *Curr Issues Mol Biol*. 2000;2:9-16.

77. Cowperthwaite MC, Ellington AD. Bioinformatic analysis of the contribution of primer sequences to aptamer structures. *J Mol Evol*. 2008;67(1):95-102.

78. Feng Q. *In vitro selection of DNAzymes for colorectal cancer detection.* [M.Sc. thesis]. Hamilton, ON: McMaster University; 2016.

79. Ihms HE, Lu Y. In vitro selection of metal-ion-selective DNAzymes. *Methods Mol Biol*. 2012;848:297-316.

80. Heaton I, Platt M. DNAzyme sensor for the detection of metal ions using resistive pulse sensing. *ChemRxiv*. 2020; Preprint. Accessed Jun 20, 2020

81. Gu L, Yan W, Wu H, et al. Selection of DNAzymes for sensing aquatic bacteria: *Vibrio anguillarum*. *Anal Chem*. 2019;91(12):7887-7893.

82. Saran R, Liu J. A silver DNAzyme. *Anal Chem*. 2016;88(7):4014-4020.

83. Shao G, Ji S, Wu A, et al. DNAzyme-based probe for circulating microRNA detection in peripheral blood. *Drug Des Devel Ther*. 2015;9:6109-6117.

84. Zhou W, Zhang Y, Ding J, Liu J. In vitro selection in serum: RNA-cleaving DNAzymes for measuring Ca^2 and Mg^2. *ACS Sensors*. 2016;1(5):600-606.

85. Ali MM, Wolfe M, Tram K, et al. A DNAzyme-based colorimetric paper sensor for *Helicobacter pylori*. *Angew Chem Int Ed Engl*. 2019;58(29):9907-9911.

86. Suehiro Y, Sakai K, Nishioka M, et al. Highly sensitive stool DNA testing of *Fusobacterium nucleatum* as a marker for detection of colorectal tumours in a Japanese population. *Ann Clin Biochem*. 2017;54(1):86-91.

87. Shen Z, Wu Z, Chang D, et al. A catalytic DNA activated by a specific strain of bacterial pathogen. *Angew Chem Int Ed Engl*. 2016;128(7):2477-2480.

88. Iijima Y, Asako NT, Aihara M, Hayashi K. Improvement in the detection rate of diarrhoeagenic bacteria in human stool specimens by a rapid real-time PCR assay. *J Med Microbiol*. 2004;53(7):617-622.

89. Chang D, Chang T, Salena B, Li Y. An unintentional discovery of a fluorogenic DNA probe for ribonuclease I. *ChemBioChem*. 2020;21(4):464-468.

90. Yousefi H, Ali MM, Su H, Filipe CD, Didar TF. Sentinel wraps: Real-time monitoring of food contamination by printing DNAzyme probes on food packaging. *ACS Nano*. 2018;12(4):3287-3294.

91. Guven DC, Dizdar O, Alp A, et al. Analysis of *Fusobacterium nucleatum*, *Streptococcus gallolyticus* and *Porphyromonas gingivalis* in saliva in colorectal cancer patients and healthy controls. *Biomark Med*. 2019;13(9):725-735.

92. Shet UK, Oh H, Kim H, et al. Quantitative analysis of periodontal pathogens present in the saliva of geriatric subjects. *J Periodontal Implant Sci*. 2013;43(4):183-190.

93. Walker VK, Palmer GR, Voordouw G. Freeze-thaw tolerance and clues to the winter survival of a soil community. *Appl Environ Microbiol*. 2006;72(3):1784-1792.