Summer 2020

# Characterization of Helicobacter pylori sRNAs HPnc2525, HPnc2600, and HPnc2645

Brandon M. Flatgard

# CHARACTERIZATION OF *HELICOBACTER PYLORI* SRNAS HPNC2525, HPNC2600, AND HPNC 2645

---

A Thesis

Presented To

Eastern Washington University

Cheney, Washington

---

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biology

---

By

Brandon M. Flatgard

Summer 2020

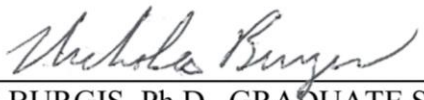
THESIS OF BRANDON M. FLATGARD APPROVED BY

DATE 8/6/2020

ANDREA CASTILLO, Ph.D., GRADUATE STUDY COMMITTEE

DATE 8-05-2020

JASON ASHLEY, Ph.D., GRADUATE STUDY COMMITTEE



DATE 8/5/2020

NICK BURGIS, Ph.D., GRADUATE STUDY COMMITTEE

# ABSTRACT

CHARACTERIZATION OF *HELICOBACTER PYLORI* SRNAS HPNC2525, HPNC2600, AND HPNC 2645

by

Brandon M. Flatgard

Summer 2020

*Helicobacter pylori* is a common microaerophilic gram-negative bacterium that infects approximately 50% of the human population. Although all *H. pylori* infections result in inflammation of the gastric epithelium, only 10-15% of infections are symptomatic and progress to severe gastric diseases such as gastric and duodenal ulcers, MALT lymphoma and gastric cancer. Different disease outcomes are due in part to genetic variations among *H. pylori* strains. *Helicobacter pylori* strains with a genomic region called the cytotoxin-associated pathogenicity island (*cag*PAI) are associated with an increased risk of severe disease. The *cag*PAI region encodes a type IV secretion system that transports the CagA effector into host gastric epithelial cells. Regulation of the *cag*PAI is a vital area of research. Previous studies on *H. pylori* and other bacteria have found that small RNAs (sRNAs) play a role in gene regulation. These transcripts are 50-300 nucleotides in length and act independently on expressed targets. sRNAs serve to regulate and fine-tune gene expression by interacting with target mRNA molecules to inhibit or accelerate gene translation or function. Understanding how *cag*PAI genes are

regulated is key to understanding how they promote disease. Herein, I investigate three *cag*PAI located sRNAs HPnc2525, HPnc2600, and HPnc2645 identified in previous studies. Using bioinformatics, conservation, and visual inspection, I predicted promoters and terminators for HPnc2525, HPnc2600, and HPnc2545. I devised an RT-PCR primer walking strategy to delimit the 5' end of HPnc2525 further. I analyzed the sRNA sequences defined by this and previous research using the TargetRNA2 program and predicted these sRNAs regulate virulence (e.g., *cagL*, *hopZ*, and *fliA*) and other genes. Due to CoVID-19, I was unable to verify sRNA interactions with their predicted targets using a plasmid-based GFP expression system.

## ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my mentor Dr. Andrea Castillo who has been tremendously helpful and inspiring through this research. I would like the thank Dr. Jason Ashley and Dr. Nick Burgis for help preparing this document. Finally, I would like to thank Veronica Albrecht for her collaboration during this study.

A Biology Graduate Mini-grant to support this project was awarded to Brandon Flatgard, a separate Biology Graduate Mini-grant Award to Veronica Albrecht, and a Start Something Big Foundation grant to Andrea Castillo.

## Table of Contents

# LIST OF TABLES

# LIST OF FIGURES

# PREFACE

Due to the global pandemic of COVID-19, time in the lab was cut short. Instead, the research took on more a bioinformatic approach to elucidate other aspects of my research. Much of the preliminary planning and preparation for the physical experiments were completed, and this will offer researchers taking on the next step a fantastic place to continue where I left off.

# CHAPTER I

## Introduction to *Helicobacter pylori*

*Helicobacter pylori* is a rod-shaped microaerophilic gram-negative Epsilonproteobacteria. It is approximately 2.4-4.0 μm by 0.5-1.0 μm and has two to six flagella that facilitate movement to penetrate gastric mucosa and counteract gastric efflux mechanisms[1]. *Helicobacter pylori* is a human-adapted gastric *Helicobacter* species located in the human gastric mucosa of approximately 50 percent of the human population, making it one of the most common infections in the world[1,2]. Other *Helicobacter* species are zoonotic and rarely can be found on Humans[3].

Coevolution between *H. pylori* and humans share a 58,000-year history with predominant strains located regionally along with human migration patterns[4]. The global distribution of *H. pylori* genetic diversity is a mirror of its human host, and simulations have also predicted the spread of *H. pylori* from east Africa to around the world over the same time scale of humans[4]. It has been shown that human populations located in areas with notably higher documented intestinal parasite load were found to have an immune response against *H. pylori* toward a T helper type 2 (Th2) (anti-inflammatory), thereby preventing T helper type 1 (Th1) responses associated with mediated microbicidal action, gastric inflammation, and corpus atrophy[5,6]. This interaction may, in turn, affect the selection of *H. pylori* strains in the host, favoring the loss of *cag*PAI, which is known to be associated with Th1 type responses[5].

Although the *H. pylori* infection rate is high worldwide, its prevalence varies dramatically in different regions; prevalence is higher in developing than developed countries, and in developed countries, infection is inversely correlated with access to

good sanitation[1,2]. Regions of Africa and South America were found to have the highest rates of infection at 70.1% and 69.4% respectively, and Nigeria was found to have the highest prevalence with 87.7% of the population infected[2]. In more developed regions such as North America, the population infection rate is 37.1%[2]. However, in these developed countries, it has been found that certain genetic groups have a disproportionate infection rate compared to the region's average infection rate[2]. In the United States, prevalence among non-Hispanic whites ranges from 18.4% to 26.2%. Non-whites conversely range from 34.5% to 61.6% (Fig. 1)[7].

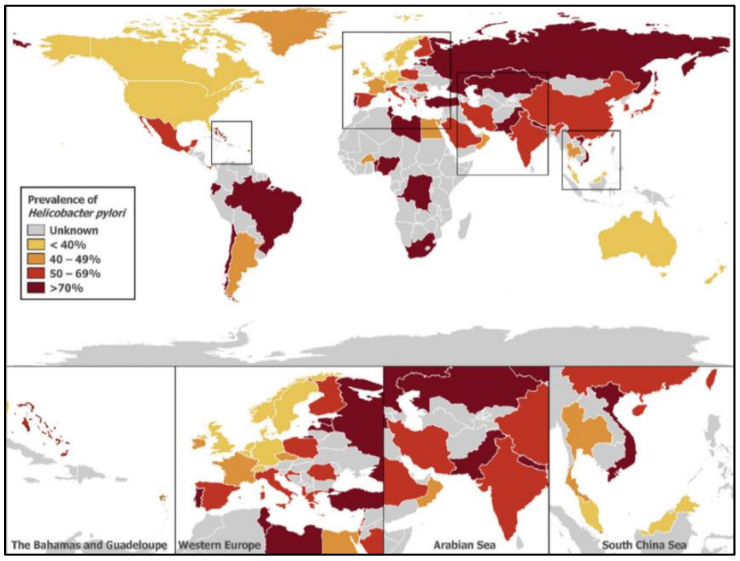



**Figure 1.** *Helicobacter pylori* Prevalence Worldwide. (Hooi, J. et al., 2017)

*Helicobacter pylori* is mainly spread within close family groups who have diminished access to clean water and are of lower socioeconomic status. However, infection rates are lower in children and adolescents, although infection tends to occur at a younger age and is lifelong without antibiotic treatment[2,8]. The main modes of transmission are through oral/oral, gastro/oral, and fecal/oral routes[1,9]. Oral-oral and gastro-oral transmission occur via direct contact with an infected individual, such as kissing or sharing unwashed drinking containers[10]. *Helicobacter pylori* infection

prevalence has also been found to be significantly higher among intimate partners via an oral/oral route[10]. One is four times as likely to be infected if their partner has *H. pylori*-related gastroesophageal reflux. However, the infection can still be spread from asymptomatic partners but to a lesser extent[10]. Fecal-oral transmission takes a more indirect route. Infected individuals can pass *H. pylori* through fecally contaminated water. In areas with lower water sanitation standards, *H. pylori* can be transmitted to a new host either by drinking contaminated water directly or eating food that has been contaminated with the unclean water[9,11].

## ***Helicobacter pylori*** **Caused Gastric Disease**

Although *H. pylori* infections are common, only 10-15% results in disease[12]. All *H. pylori* infections result in some level of chronic active gastritis, but most do not result in apparent symptoms or other complications[13]. Symptomatic infections include gastric cancer, acute and chronic gastritis, gastric and duodenal ulcers, functional dyspepsia, and mucosa-associated lymphoid tissue (MALT) lymphoma.

Successful colonization of the host gastric mucosa by *H. pylori* always results in acute and chronic gastritis, but these are not always symptomatic[13]. Acute gastritis is associated with nonspecific dyspeptic symptoms and also with hyperchlorhydria, the increased production of acid[14]. The main dyspeptic symptoms associated with *H. pylori* infection are reflux-like with heartburn and regurgitation. Chronic gastritis, because of permanent colonization, affects acid production and, when coupled with cellular disruption, can promote different diseases. A combination of *H. pylori* virulence factors, host factors, and host environmental factors plays a part in the acid production of the stomach and, in turn, where it colonizes[14]. An environment of low acidity leads to more

broad colonization of the stomach, in the antrum (lower), corpus (middle), and fundus (upper), which is associated with gastric carcinoma[15]. Conversely, an environment with higher acid production leads to more limited stomach colonization, mainly in the gastric antrum, which has a lower number of acid-secretory cells. This antral-predominant gastritis from *H. pylori* colonization is the leading cause of peptic ulcer disease (PUD) (Fig. 2)[14].

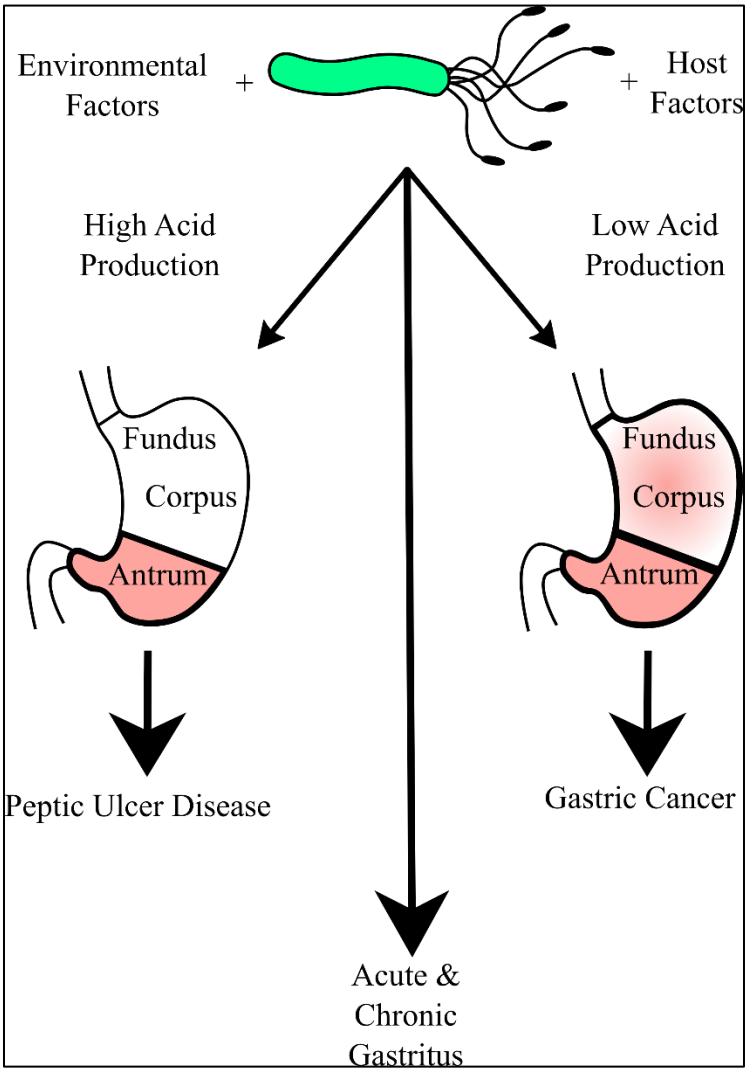



**Figure 2.** Locations of *Helicobacter pylori* Colonization
With a combination of host and environmental factors, *H. pylori* can colonize distinct locations in the stomach leading to different disease states. Pink areas represent increased acid production and colonization.

*Helicobacter pylori* was discovered in 1983 to be a cause of PUD and is still regarded as the primary pathogen associated with it[2,3,16]. PUD is caused by defects in the gastric mucosa that occur along with the transition from the corpus to antrum[17]. *Helicobacter pylori* presence has been found in about 82% of duodenal ulcers and 68% of gastric ulcers[18]. The risk of PUD is 3 to 10 times greater in people with *H. pylori* infection than without[19]. Eradication of the infection by antibiotic treatment or triple therapy (oral proton pump inhibitor, clarithromycin, and amoxicillin) has been found to reduce the risk of this disease[20].

Chronic infection with *H. pylori* is linked to gastric adenocarcinoma and MALT lymphoma, which accounts for 25% of cancers associated with infectious agents[21]. MALT lymphoma was first associated with *H. pylori* in 1991 and is now known to be responsible for 92-98% of all gastric MALT lymphoma[21]. This type of lymphoma is a non-Hodgkin B-cell neoplasm, which originates from the marginal zone of lymphoid follicles[22]. Although not contributing to increased risk, the presence of the *H. pylori* virulence protein cytotoxin-associated gene A (CagA) in tumor-associated B cells correlated with earlier pathogenic remission times, possibly due to the translocation of CagA into malignant B cells[23]. Eradication of *H. pylori* leads to remission in 80% of MALT lymphoma patients, but other cancers have been linked to *H. pylori* infection, such as Pseudomyxoma peritonei (PMP), laryngeal/pharyngeal cancer, and other lymphomas[12,24]. Pre-surgery *H. pylori* eradication with these other cancers has shown an increased length of patient survival over non-treated patients[12,24]. *Helicobacter pylori* is currently identified as a Group I carcinogen by the International Agency for Research on Cancer[25].

Functional Dyspepsia, or non-ulcer Dyspepsia, is one of the most common sensorimotor disorders involving the gastrointestinal tract[26]. It is defined as the presence of one or more of the following symptoms: bothersome postprandial fullness (full stomach feeling), bothersome early satiation (inability to eat), bothersome epigastric pain (upper abdomen), or bothersome epigastric burning (heartburn), and no evidence of structural disease (including upper endoscopic findings) to explain the symptoms[27]. The main symptoms of *H. pylori* are reflux-like with heartburn and regurgitation. It may be similar to ulcer-related dyspepsia, causing pain and vomiting. Of patients that exhibit the symptoms associated with Functional Dyspepsia, it is estimated that 30-60% carry *H. pylori*[28].

Our initial understanding of the role of *H. pylori*-induced carcinoma was believed to be the result of a long-term inflammation process. Whereby a Th1 type immune response led to the apoptosis of gastric epithelial cells and thus increased cell growth to compensate[29,30]. This damage led to errors during mitosis and accumulated mutations in the host cells, which is further compounded by a deficiency of the DNA mismatch repair mechanism in the host cells induced by *H. pylori* infection[31]. Recent studies have suggested that the CagA protein may also have a carcinogenic effect via interactions with proteins of the tight junctions and adherent junctions, leading to a breakdown of these junctions and uncontrolled activation of β-catenin, which will lead to tumor production[3,32,33]. Conclusive evidence for this interaction and carcinogen nature was shown in transgenic mouse experiments[34].

## *Helicobacter pylori* Caused Extragastric Disease

Extragastric diseases with which *H. pylori* infection may also be associated include iron deficiency anemia, idiopathic thrombocytopenic purpura (ITP), skin diseases, pregnancy diseases, liver and gallbladder, pulmonary, ocular, neurodegenerative, cardiovascular, autoimmune, neuromyelitis optica, multiple sclerosis, insulin resistance, diabetes, and ear nose and throat[12]. *Helicobacter pylori's* presence in these different areas of the body shows that it is not just isolated to the stomach, as was initially thought. Other than the stomach, *H. pylori* has been cultured and observed histologically in the nasal cavity, coronary arteries, liver, and peritoneum[12]. The relationships and their significance of *H. pylori* in these extragastric locations are debated and are still being investigated.

## Factors Affecting Disease Outcome

Human factors affecting disease outcomes include age, alcohol intake, nicotine intake, and salt intake[7]. These factors all affect *H. pylori* antibody serum levels but in separate ways. For example, age, salt, and nicotine intake are associated with a marked increase in serum antibody presence[35-37]. Nicotine causes vasoconstriction body-wide, thus reducing mucosal blood flow, epidermal growth factor secretion, and mucus secretion[36]. High salt concentration in the stomach was found to alter the *H. pylori* exoproteome, increasing the secretion of the *H. pylori* virulence factor VacA[35]. Alcohol, however, has an inverse effect on antibody serum levels, which is believed to be the result of lower pH in the stomach due to increased stimulation of gastric acid secretion[38].

Human genetic diversity can also lead to differing immune responses, which in turn can affect disease states. Increased gastric cancer risk can stem from functional

polymorphisms in the genes encoding interleukins (ILs). ILs are a group of cytokines, proteins, and other signaling molecules, which are an essential part of the immune system. In particular, functional polymorphism in IL-1β (responsible for antimicrobial resistance and inflammation) and IL-1 receptor antagonists (the natural inhibitor of the proinflammatory effects of IL-1β), as well as differing the amount of IL-1β secreted from the microbial response, have shown to increase gastric cancer risk[39]. An increase in colony density in patients lacking blood group antigen $Le^b$ has been noted as well[40]. These individuals are lacking $Le^b$ due to the fact they genetically have no Le gene or they have an Le gene but no Se gene[41]. Other polymorphisms exist in genes, increasing proinflammatory cytokines or reducing anti-inflammatory cytokines. *Helicobacter pylori* primarily mediates a Th1 response[42]. This response is associated with proinflammatory cytokine release, such as the cytokine important for innate and adaptive immunity against viral, some bacterial and protozoal infections, interferon-gamma (IFN-γ), as well as the activation of macrophages. This strong Th1 response leads to worse outcomes due to chronic inflammation and gastric atrophy. However, a Th2 response can mediate this effect[43]. There is a noted Th1 bias in patients from Japan and the United Kingdom who also present a higher level of disease[44].

Conversely, human populations in Africa present a lower instance of disease levels with a marked increase of Th2 response instead of Th1[44]. The mechanism for controlling Th response in humans is comprised of regulatory T-cells (Tregs), which are a specific population of T-cells that can maintain homeostasis and self-tolerance by suppression of the immune response[45]. The gastric mucosa of *H. pylori*-positive individuals contains high levels of immunosuppressive cytokines IL-10 (anti-

inflammatory) and transforming growth factor beta 1 (TGF- β1), which controls proliferation and differentiation of many cell types[46]. Tregs also suppress IL-8 (induces chemotaxis and stimulates phagocytosis) production from gastric epithelial and memory cell responses in the presence of *H. pylori*[46]. *Helicobacter pylori* can induce a Treg response, which may be a reason for its relatively benign coexistence with the human host. However, patients with PUD have decreased Tregs in their gastric mucosa, which highlights the importance of a balanced regulatory response[47]. Coevolution between *H. pylori* and humans spans 58,000-years, and while the human genetic diversity plays a role in disease, so does the genetic diversity of *H. pylori* itself[4].

## Bacteria Genetic Diversity Promoting Disease

There are four steps crucial for *H. pylori* colony establishment and pathogenicity, and the presence of genes required for these steps varies among *H. pylori* strains. When entering the host stomach, *H. pylori* must be able to survive in the acidic stomach. Survival is facilitated by the production of urease, which hydrolyzes host produced urea to ammonia and carbonic acid, serving as a buffer to the microenvironment enabling colonization (Fig. 3).

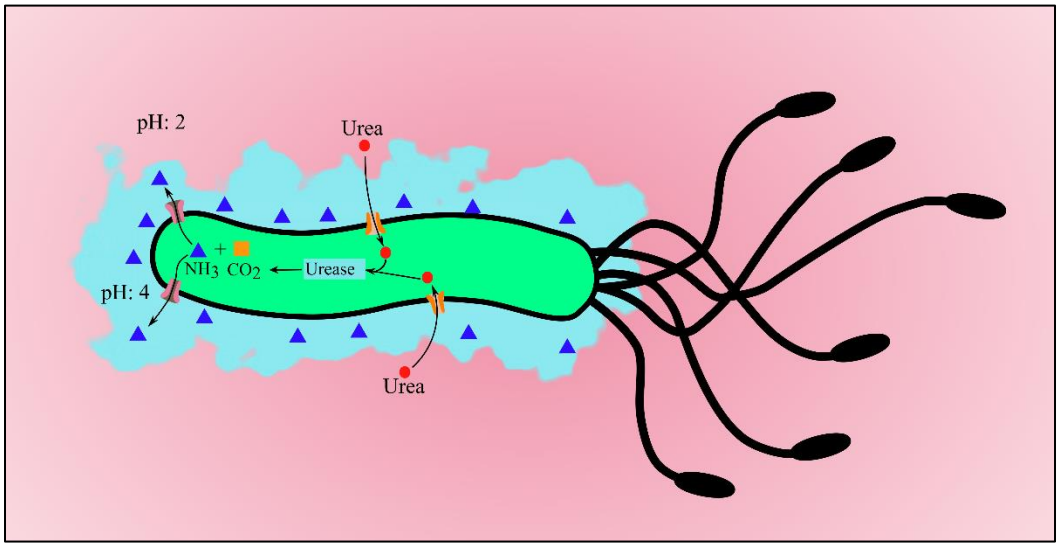



**Figure 3.** *Helicobacter pylori* can Neutralize Stomach pH
To successfully colonize the gastric mucosa, *H. pylori* neutralizes urea with urease. The pink area represents a high acid environment (≤ pH 2), and the blue area represents a lower acidity environment (≥pH4) due to the effects of urease. Urea from the environment enters through proton-gated channels (orange) on the surface of the bacterium. Urease breaks down the urea creating $CO_2$ (orange square) and $NH_3$ (blue triangles). The latter is excreted through proton-gate channels on the surface of the bacterium, raising the pH in the vicinity.

The urease gene cluster contains seven genes, including catalytic subunits (*ure*A/B), an acid-gated urea channel (*ure*I), and accessory assembly proteins (*ure*E-H)[48]. To prevent lethal alkalization, this activity is regulated by the proton-gated urea channel UreI. This regulation functions by only being active under acidic conditions, such that UreI channels in the inner membrane are closed at pH 7.0 and fully open at pH 5.0[49]. The surface of *H. pylori* also contains urease, which allows the bacteria to pass through the gastric environment unscathed[49]. Furthermore, urease has been found to regulate macrophage interactions by modulating phagosome pH and megasome formation and reduce the viscosity of gastric mucin[50,51].

While traversing the gastric mucosa, the next crucial step in colony establishment is the ability to move further toward epithelium cells. *Helicobacter pylori*, motility is

directly correlated with a rise in pH. *Helicobacter pylori* achieves motility by reducing the viscosity of its environment; it does not bore its way through the mucus gel, as previously thought[50]. It can then move through the gastric mucosa epithelium layer toward the basal layer when the pH is more neutral, and the bacteria have disrupted the tight cell junctions. Both its helical shape and flagella help it move into its microenvironment. *Helicobacter pylori* rotates as it swims, producing a helical trajectory, but the relatively slow body rotation rates make only a 15% contribution to propulsive thrust[52]. Thus, the flagella have more of an impact on motility than its helical shape. This flagella-mediated motility is one of the most complex systems in *H. pylori,* with more than 40 proteins involved in biosynthesis and operation[53]. More severe pathological outcomes have been observed in patients infected with *H. pylori* strains with higher motility[54]. There is still debate about whether flagella directly participate in cell adhesion. However, it was found that regulators controlling flagella genes may affect adhesin expression as well[55].

Neutrophil-activating protein A (*napA*), a gene encoding *H. pylori*-NAP (HP-NAP), is highly associated with chronic gastritis, as well as infiltration of neutrophils and mononuclear cells into the gastric mucosa[56]. HP-NAP is a member of the DNA-binding proteins from starved cells (Dps) family, which is highly similar to other bacterioferritin that store iron and bind and protect DNA[57]. However, HP-NAP differs as it induces large quantities of IL-12 and IL-8 via stimulation of both neutrophils and monocytes[58,59]. This stimulation, in turn, induces T-cells to secrete INF-γ and mediates the shift to a Th1 response[58]. This adhesion to neutrophils causes damage to local tissue by stimulating a

high production of oxygen radicals[60]. HP-NAP also facilitates another adhesin, SabA, by mediating the binding of sialylated antigens on the host cell surface[61].

Both variably expressed blood group antigen-binding adhesin (BabA) and sialic acid-binding adhesin (SabA) facilitate colonization of the epithelial layer of the antrum of the stomach[3,12]. BabA binds to fucosylated Lewis B blood-group antigen that is expressed on the gastric epithelium cells[62]. Interestingly, strains that are deficient in BabA resulting from mutation still express a chimeric BabB/A, which can bind to the $Le^b$ antigen by genetic recombination in certain conditions[63,64]. SabA, on the other hand, recognizes the sialyl-Lewis A antigens $sLe^x$ and $sLe^a$ [65,66]. Thus, the interaction of SabA and $sLe^x$ antigen enhances *H. pylori* colonization in patients with weak or no $Le^b$ expression.

After the pH has been addressed by urease and adhesion of *H. pylori* to epithelial cells occurs, the bacterium secretes or injects additional virulence factors into the cytoplasm to impact host cell physiology such as muramyl dipeptide (MDP) and Vacuolating toxin A (VacA) into their cytoplasm[67]. MDP, which is part of the peptidoglycan cell wall, is detected by the intracellular NOD receptors. These, in turn, activate the Nuclear factor-κB (NF-κB) pathway with the production of IL-8 and attract inflammatory cells[67]. Ultimately this immune response was found to be one-way epithelial cells that can differentiate pathogenic and non-pathogenic bacteria[67,68]. The presence ImaA, immunomodulatory autotransporter protein, can decrease this inflammation responses aiding in overall pathogenesis of *H. pylori*[69]. Also, VacA may induce many large vacuoles in host cells (Fig. 4).

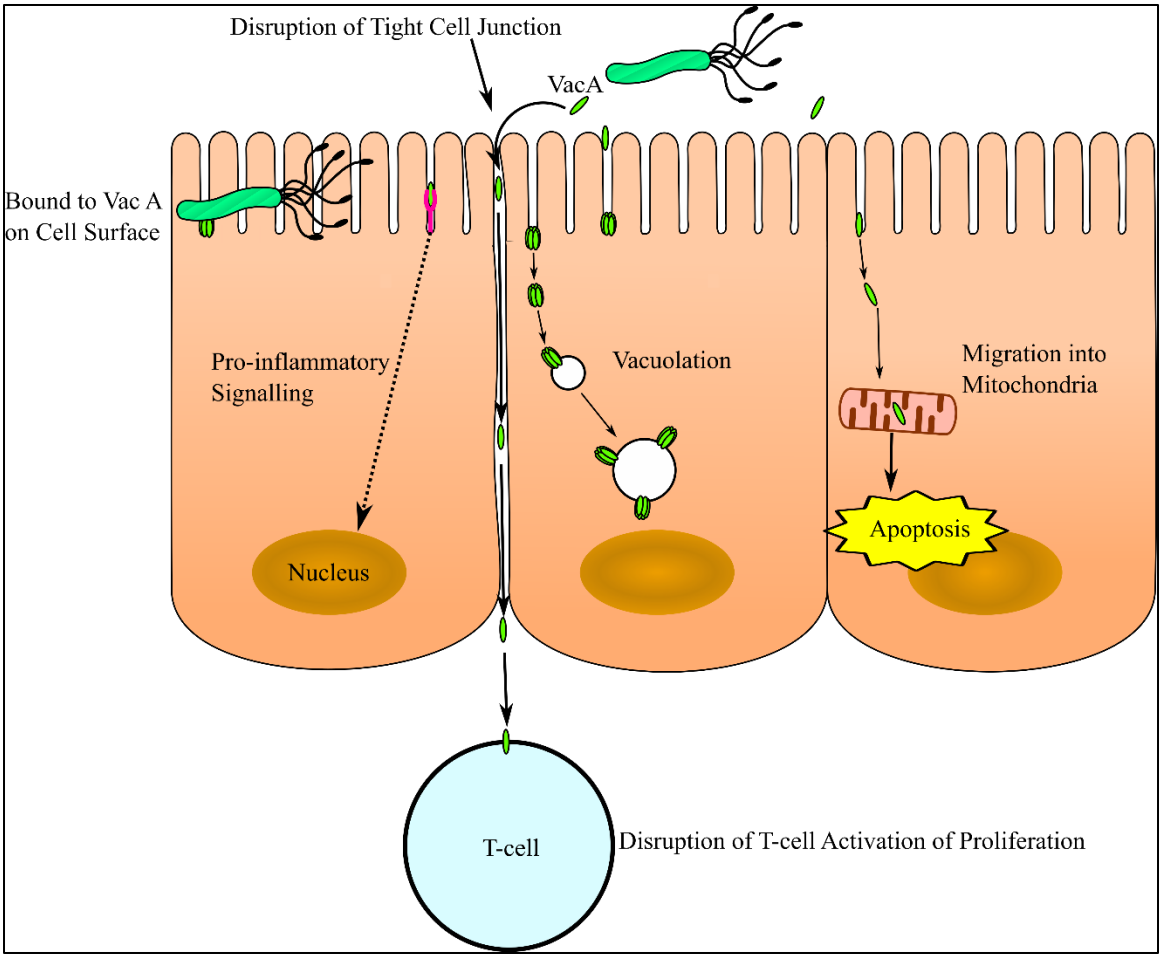



**Figure 4.** Virulence Factor Vacuolating Toxin A (VacA)
VacA enables *H. pylori* to remove nutrients from host cells when bound to the epithelium by an autotransporter system. VacA also promotes proinflammatory signaling, vacuolation (membranous spaces in the host cell cytoplasm), apoptosis, and after tight cell junction disruption, disruption of T-cell activation and proliferation.

There are multiple alleles of *vacA* with different activities and different levels of expression that impact disease outcome[70]. VacA forms an autotransporter structure to secrete itself from *H. pylori* without the need for host cell contact[12]. VacA has been shown to disrupt the barrier function of epithelial cells, allowing leakage of essential nutrients needed by *H. pylori* for growth[71]. VacA embeds into the host cell membrane and functions as an anion-selection channel releasing bicarbonate and organic anions in the host cytoplasm[72]. In addition to embedding, VacA can also be endocytosed and allow

anions to enter late endosomes, creating large vacuoles[71,73]. Acute inflammatory responses can also be achieved by inducing the host cell to release IL-8[74]. VacA also can cause mitochondrial dysfunction resulting in apoptotic death of the cell and T-cell blocking[71]. The presence of VacA can have a dramatic effect on epithelial cells; however, it is not clear whether VacA primarily causes a disease. Thus, it is believed that the ability of *H. pylori* to colonize a host successfully is merely enhanced via VacA[75].

The most significant and well researched *H. pylori* virulence factor is CagA. It will not function in isolation and works in concert with other virulence genes, those in the *cag*PAI genomic region[12]. This island is present in about half of the *H. pylori* strains, and its presence correlates with more severe disease states[3,67,76]. The *cag*PAI genes encode a type IV secretion system that enables the injection of CagA into host cells[71]. Once in the host cell cytoplasm, CagA is phosphorylated, which inhibits Src dephosphorylation, regulation of cellular matrix, and other cytoskeleton regulators in epithelial cells[68,77]. These will phenotypically change the cell and cause high markers of gastric cancer stem cells, the so-called hummingbird phenotype (Fig. 5)[3,71,78].

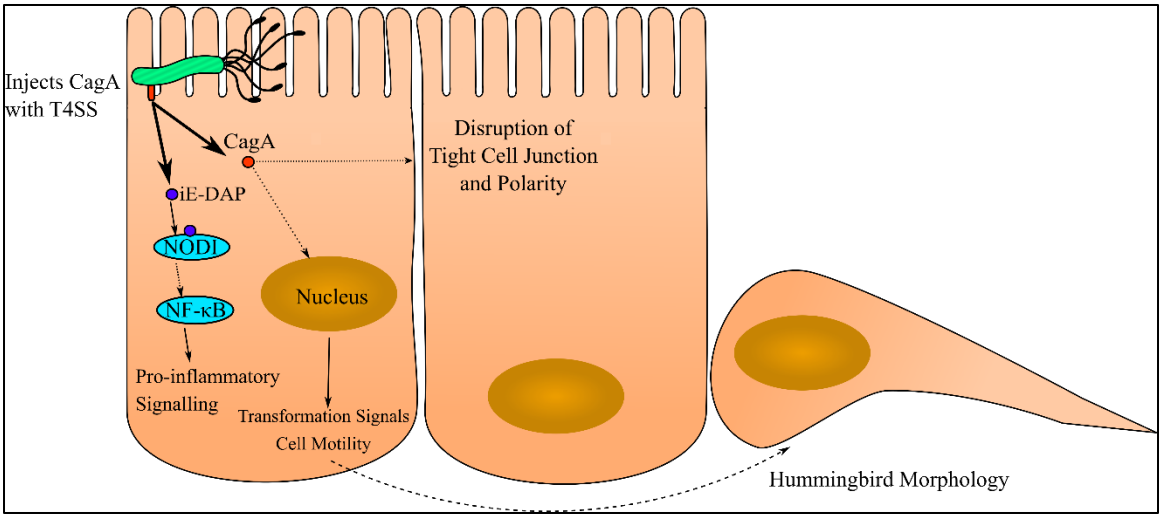



**Figure 5.** The Type IV Secretion System Injects CagA into the Epithelial Cell CagA disrupts the tight cell junction and disrupts the cellular matrix resulting in a hummingbird morphology. NOD1 also gets activated by iE-DAP, which in turn activates NF-κB, which subsequently produces proinflammatory signaling.

Africa has been shown to have the highest prevalence of infection, even though the reported cases of gastric cancer were lower than regions with lower infection prevalence[2,79]. The most prevalent strain of *H. pylori* in South Africa, shpAfrica2, was found to lack the *cagA* gene[2]. Further studies have shown that increased pathogenicity correlated with the more virulent strains to create morphological changes, vacuole formation, and progressive epithelial cell degeneration[13]. There is a noted positive correlation with the severity of disease, and *cag*A-positive *H. pylori* strains in both humans and animals[13,80].

Not all virulence factors are present in each strain of *H. pylori*. Although, all known *H. pylori* strains contain *vacA* genes, not all the strains produce a functional VacA protein due to genetic polymorphisms[81]. BabA is almost always present in CagA positive strains, and the presence of *cagA* usually correlates with the presence of *vacA*, and proinflammatory outer membrane protein (*oipA*)[12,56,82]. The pathogenicity of *H. pylori* is

a result of several genes that are responsible for the production of these virulence factors, each playing their part.

# CHAPTER II

## Small Regulatory RNAs

Regulatory RNAs (riboregulation) have been linked to bacteria pathogenicity[83]. *Helicobacter pylori* was previously believed to lack riboregulation, but recent studies revealing sRNA activity in *H. pylori* have challenged previous conclusions[84,85]. RNA regulators in bacteria are transcripts that use various mechanisms to change expression and promote a range of physiological responses[86]. These regulators come in different classes, such as riboswitches, small transcript binding proteins, CRISPR RNAs, and sRNAs[86]. The sRNA molecules are post-transcriptional regulators found in all kingdoms of life. They are short transcripts roughly 50-300 nucleotides in length and act on expressed mRNA targets through complementary base-pairing[87]. These interactions serve to regulate and fine-tune gene expression by inhibiting or accelerating mRNA translation through impacting access to the ribosome binding site and by impacting mRNA stability[86]. Of the various sets of sRNA, the most extensively studied are regulators that act through base pairing with mRNAs to affect their translation[86]. *Cis*-encoded antisense sRNAs are located in or near the same gene locus as their mRNA target and share complete base-pairing with their target, usually within the untranslated 5' region (Fig. 6a)[88]. Alternatively, *trans*-encoded sRNA, whose target is genetically distant loci, has only partial and interrupted base pairing with its target (Fig. 6b)[88]. The last known interaction is *Cis*-encoded antisense acting in *trans*, which involves two sRNA interactions, one in *cis* and one or more in *trans* (Fig. 6c)[88].

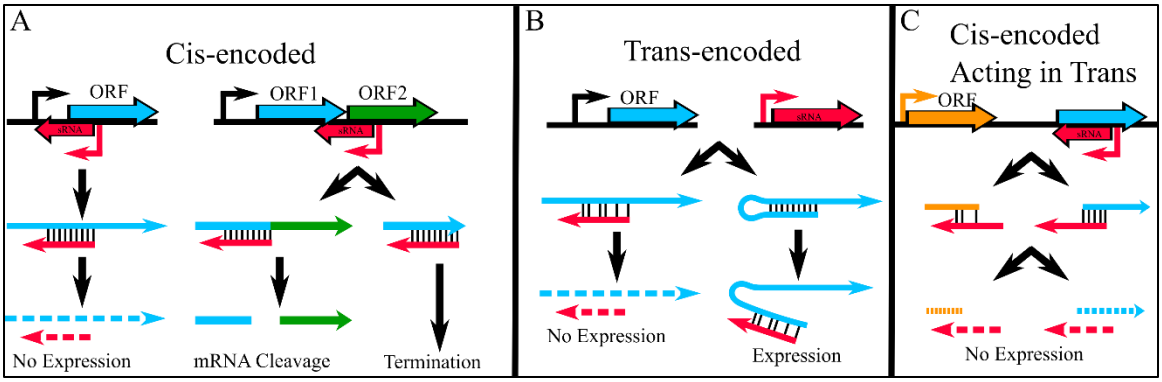



**Figure 6.** sRNA Can Post-transcriptionally Regulate Genes in Diverse Ways (A) *Cis*-encoded antisense sRNA with complete base pairing with mRNA can form a complex of the coding region that gets degraded. (B) Genetically distant *trans*-encoded sRNA has only partial and interrupted base pairing that can enable upregulation by binding with the mRNA and opens access to the ribosome binding site. (C) *Cis*-encoded antisense acting in *trans* sRNAs involves two interactions, one in *cis* and one in *trans*. The sRNA can downregulate expression by binding to the ribosome binding site, thus blocking translation

The majority of known sRNAs act by the *cis*-encoded antisense acting in trans mechanism whereby the *cis*-encoded sRNA acts both on the mRNA target it overlaps and mRNA targets at genetically distant loci[86,88]. The sRNAs can globally affect traits resulting from stress responses, adaptive metabolic changes, virulence, stalled ribosomes, uncharged tRNAs, elevated temperatures, or small molecule ligands at the post-transcriptional level[89]. Hence, sRNAs have evolved as an additional form of gene regulation[90].

In some instances, proteins aid the sRNA to facilitate regulation. Helper proteins like Hfq often play critical roles in sRNA function. However, Epsilonproteobacteria, including *H. pylori*, have no apparent Hfq homolog[91,92]. Hfq is an RNA-binding protein that forms homohexamers of ~12 kDa subunits, which protects the sRNA to prevent degradation and allow the sRNA to bind to its mRNA target[92]. Although all forms of life contain sRNA, not all of them are the same, and it is not currently known how sRNAs work in *H. pylori* given the apparent lack of helper proteins.

Different methods, such as computer-aided target prediction and functional assays, have been developed to find the sRNAs targets. A biocomputational method (TargetRNA) was developed to predict targets of sRNAs using a calculated hybridization score for a given sRNA and a candidate mRNA candidate[93]. This tool was later refined to update efficiency and accuracy and is freely accessible on the web as TargetRNA2[93,94]. Reporter assays have been developed to analyze sRNA-directed gene control that can be used for co-expression of an sRNA and a 5' sequence of an mRNA target fused to a green fluorescent protein (GFP) reporter[95]. Using this method, expression of the mRNA directed by the sRNA can be visualized by cloning GFP plasmids[96]. A super folder variant GFP (sfGFP) was later created, which results in increased fluorescence of target fusions, which is beneficial when the proteins do not fold flawlessly but still affect expression[97].

## Role of *Helicobacter pylori* sRNA

The *H. pylori* genome encodes relatively few transcriptional regulators, so how does it regulate all of its gene expression to respond to changing environmental conditions[98]. A potential answer came with the publication of four studies that together identified in excess of 60 *H. pylori* sRNAs. Livney et al. (2006) used sRNAPredict2 to identify sRNAs in intergenic regions with regards to the conservation of transcriptional regulatory elements, Rho-independent terminators, promoters, and with a size constraint of 60-550 nt in length[99]. They identified a total of 130 among six different strains of *H. pylori*; 50 were unique, 49 were conserved in at least two strains, and 29 were at least conserved in 3 strains[99]. Xiao et al. (2009) predicted six sRNAs based on gene location, promoter search, and terminator search[100,101]. Intergenic regions were screened in *H.*

*pylori* strain 26695 using pftools 2.2 to identify promoters, RNAmotif to identify hairpins followed by UU nucleotide sequences, and finally, if a promoter and terminator pair on one strand between 45 and 350nt between them[100,101]. These potential sRNAs were analyzed for conservation, which identified six sRNAs, and two were experimentally confirmed by Northern blot and RT-PCR[100,101]. Even though many sRNAs have been identified, only four have been well characterized. For example, a novel *cis*-encoded antisense small RNA was identified, which downregulated expression of the urease component, UreB, at neutral pH[85]. It was found that 5'ureB-sRNA, targets the 5' end of *ureB* and promotes premature termination of transcription of *ureAB* mRNA[85]. With this mechanism, a limited amount of 5'ureB-sRNA is sufficient to regulate the relatively high level of *ureAB* transcript[102]. This process effectively prevents urease from being produced when the environment does not demand it (Fig. 8a).

The second characterized sRNA, a *trans*-encoded sRNA *repG*, was shown to target a homopolymeric G-repeat in the mRNA leader of an acid-sensing chemotaxis receptor, TlpB[84]. This regulation of *tlpB*, which decreases expression could be necessary for *H. pylori* colonization, and research is still ongoing (Fig. 8b)[84]. Another *H. pylori* sRNA, CncR1, which is located in the *cag*PAI, has also been characterized[76]. CncR1 was found to target flagellar functions, specifically the flagellar checkpoint gene *fliK*. This downregulation, in turn, affects motility and adhesion (Fig. 8c)[76].

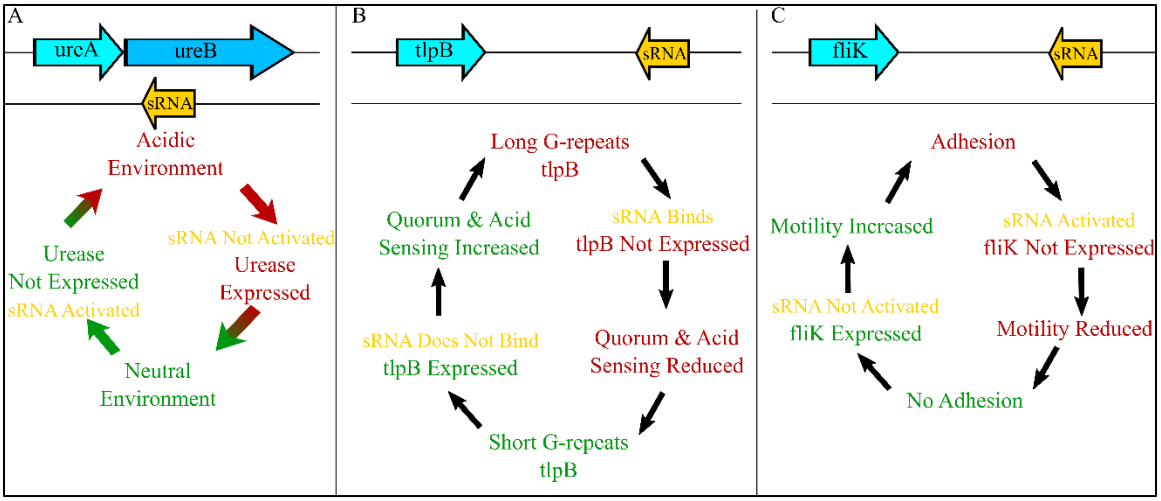



**Figure 7.** Different regulation of *Helicobacter pylori* sRNAs. (A) Urease expression changes based on the acidic environment. 5'ureB-sRNA, a *cis*-encoded antisense small RNA downregulates expression of the urease component, UreB, at neutral pH. (B) sRNA *repG* targets a homopolymeric G-repeat in the mRNA leader of an acid-sensing chemotaxis receptor, TlpB. This regulation of *tlpB* decreases the expression of quorum and acid-sensing. (C) sRNA CncR1 targets the flagellar checkpoint gene *fliK* when adhered to epithelial cells. This downregulation, in turn, negatively affects motility.

The most recent sRNA characterized is HPnc4160, discovered by Sharma et al. (2010), was recently found to regulate *cagA*, which was confirmed by comparing RNA-seq and iTRAQ analysis between wild-type and HPnc4160 deficient mutant strains[103]. HPnc4160 is expressed lower in patients that develop gastric cancer, showing that in the absence of HPnc4160, CagA can be expressed more readily, increasing virulence[103].

## Characterizing *H. pylori cag*PAI sRNAs

I am interested in sRNAs because of their intricate relationship with gene regulation. I want to learn about them in general and how they contribute to the regulation of gene expression and specifically virulence gene expression. Two previous studies (Sharma et al., 2010 and Ta et al., 2012) identified sRNAs in the clinically relevant *cag*PAI. I focused on analyzing two sRNAs, HPnc2620 (Sharma et al., 2010) and HPnc2525 (Ta et al., 2012), from this region in hopes of finding sRNAs with *cis*-

regulation of *cag*PAI genes. I chose to analyze one sRNA, HPnc2660, whose full length was identified by Sharma et al. (2010) and a second sRNA, HPnc2525, for which a partial transcript was identified by Ta et al. (2012) and Baber and Castillo unpublished). Focusing on the *cag*PAI region, one sRNA was selected from the database generated by Sharma et al. (2010) to be characterized, HPnc2600.

Additionally, three more *cag*PAI sRNAs that were identified by their promoter regions by Ta et al. (2012) were selected. These promoters were not located upstream of known *cag*PAI genes but rather were positioned to direct expression opposite and to overlap known *cag*PAI genes *cagE*, *cagY*, and *cagN*, suggesting that these regions may encode sRNAs[104]. Reverse transcription-PCR determined that these promoters did, in fact, direct expression of a transcript and these transcripts[104]. Preliminary work by C. Baber and A. Castillo cloned regions of HPnc2525, and HPnc2645 but did not identify the entire sRNAs (Table 1). My goal was to analyze and characterize sRNAs HPnc2525 and HPnc2645 from Ta et al. (2012) as well as HPnc2600 discovered by Sharma et al. (2010).

**Table 1.** sRNA Sequences (C. Baber and A. Castillo.)
Sequence discovered by TaKaRa SMARTer 3'/5'RACE highlighted in yellow. Extended sequence discovered by walking primers highlighted in green.

| **sRNA** | **Sequence (5' end)** | **mG27 Genomic Position** |
|---|---|---|
| HPnc2525 | TATACCTGTGCTACTTTAGAAGTCAGAGT GATTTCAATAGGGGTGTATTGCGCTAAAA CAAATGTGGGATCACTCTTGCCTATAAAG GCCTGTTTGACAGGGCTTGTTTCATCTTGT TTTTCTTCTTTCTTTTTATCGTCAATGGGA TTACCATTTTCATCTACAAAATCCCCACT CTTTTCTTTTTTGTTCTTAAGAGCGATTTC ACTGATATTCTTAGCAGTTTCTGCGACAT CTTTGT | 521426-521677 (antisense to *cagY*) |
| HPnc2645 | CTCTAGTTGTTTTTTGATCTAATAATTTTG ATCGCTATTTTTTCCATGAGCGATGCTTTG TTCTATTTGATTAATGACG | 539120-539199 (sense to *cagU*) |

# CHAPTER III

## Methods

Bacterial media and culture condition

*Helicobacter pylori* G27 was grown on Columbia blood agar plates (Appendix) using the CampyGen 2.5L Atmosphere Generation System and incubated at 37$^{o}$C. *Escherichia coli* media consisted of LB Broth and Agar grown incubated at 37$^{o}$C overnight. Plasmids were selected by addition of the appropriate antibiotics, ampicillin (100ug/ml), chloramphenicol (20ug/ml) to liquid and solid media.

RNA isolation

*Helicobacter pylori* strain G27 was cultured as above for three days, then passed for two days and then a third time for one day to generate a large number of cells in the log phase of growth. Bacterial growth from ½ plate was collected and resuspended in 300µl Trizol and mixed thoroughly. Samples were stored at -20$^{o}$C until used for extraction. For extraction, a Zymo RNA Mini-prep Plus kit was utilized per instructions (Zymo). Two *H. pylori* + Trizol tubes were used per extraction, each added to the column after being spun. DNaseI was added during extraction, then eluted with 50µl RNase/DNase free water. The RNA concentration and purity were measured using NanoDrop. Additional DNase treatment was required to remove all contaminating genomic DNA from the RNA samples following the ThermoFischer protocol for DNaseI. The DNase treated RNA was ran through the RNA Mini-prep plus procedure again to remove the enzymatic activity. RNA was stored at -80$^{o}$C.

Strains used for the study

*Helicobacter pylori* G27was used for this study for analysis and culture. *Escherichia coli* strains DH5-a and MC1061 were used for propagating plasmids and the plasmid assay (Future Directions).

Reverse transcription-PCR to Determining 5' and 3' ends of sRNAs

RT-PCR was completed in two steps using SuperScript IV One-Step RT-PCR with ezDNase in 10µl reactions with modifications as follows from the manufacturer's protocol (ThermoFisher). Two steps were used to prevent transcription of *cagY* due to the intergenic location of HPnc2525. The first step (reverse transcription) included all reagents except for the forward primer. The forward primer was added for the second step, starting DNA amplification.

Oligonucleotides

The sequences and names of all oligonucleotides used for reverse transcription, PCR, and cloning in this study are included in table 2.

**Table 2.** Oligonucleotides Used in this Study

| **Name** | **Sequence (5' end)** |
|---|---|
| PIX-F | CACTCTTGCCTATAAAGGCC |
| PIX-F2 | AATGTGGGATCACTCTTGCC |
| PIX-F4 | GGGTGTATTGCGCTAAAACA |
| PIX-F8 | TATACCTGTGCTACTTTAGA |
| PIX-R | GAAACTGCTAAGAATATCAGTG |
| PIX-Rnew | TATCAGTGAAATCGCTCTTAAGAACA |
| PIX-R2 | AGATGTCGCAGAAACTGCTA |
| PIX-R4 | GAAAGCAATCCGATGACAA |
| PIX-R9 | GTGGCAATAAAAAAGATGACGATAAAG |

Promoter Prediction

Transcriptional promoter and terminator sequences define the DNA region transcribed to RNA. Here I analyzed DNA regions around and within HPnc2525 and HPnc2645 for promoter and terminator consensus sequences to define better the entire transcripts (identified in this study and by Ta et al., 2012). I also analyzed HPnc2600, an sRNA identified by Sharma et al. (2010) for which the entire transcript is known, but for which the promoter and terminator sequences have not been identified, to test my identification methods.

To transcribe their genes, prokaryotic species only have a single RNA polymerase. However, they can have multiple sigma factors making up a part of the RNA polymerase complex that directs binding to the promoter[105,106]. *H. pylori* has three known sigma factors, fliA ($\sigma^{28}$), RpoN ($\sigma^{54}$), and RpoD ($\sigma^{70}$)[107]. These sigma factors have a preference to bind distinct sets of promoters containing two short sequence elements, usually around 10 (called -10) and 35 (called -35) nucleotides upstream from the transcription start site (TSS, also called +1)[105]. There is no single consensus for these sequence elements in *H. pylori*, and, to elucidate what these elements may be, an in-depth sequence analysis with many *H. pylori* strains needed to be employed. For HPnc2600, the transcription start site (TSS) was identified (Sharma et al., 2010) and so I analyzed the 50bps upstream of the TSS for sequences that matched or were similar to known *H. pylori* promoter sequences (Table3)[107]. In addition to this, I used another program that predicts binding of this same region with Virtual Footprint[108]. This program looks for promoter motifs for -10 and -35 sites of different bacterial species. My search only included promoters for fliA ($\sigma^{28}$), RpoN ($\sigma^{54}$), and RpoD ($\sigma^{70}$) in *Escherichia coli*.

**Table 3.** Previously Discovered Promoter -10 Sequences
There is no clear consensus for -35 in *H. pylori*. Studies, where a -35 found, are listed. Studies where no -35 or drastically differing -35 sites and noted with a dash.

| Reg | -35 | -10 | Source |
|---|---|---|---|
| σ70 | - | TAAAAG | Forsyth & Cover, 1999 [109] |
| | - | AATAAT | Odenbreit, Wieland, & Haas, 1996 [110] |
| | - | TAAAAT | Spohn & Scarlato, 1999 [111] |
| | - | TATTAT | Beier et al., 1998 [112] |
| | - | TACAAT | Shirai et al., 1999 [113] |
| | - | TATAAT | Spohn & Scarlato, 1999 [111] |
| | - | TATAGT | Suerbaum,1997 [114] |
| | - | TAACA | Jones et al., 1997 [115] |
| | - | AATAG | Spohn & Scarlato, 1999 [111] |
| | - | TTCGGT | Ge & Taylor, 1996 [116] |
| | - | CTTCTT | Heuermann & Haas, 1995 [117] |
| | - | TACAAT | Pesci & Pickett, 1994 [118] |
| | CGAGTG | TGATAA | Spohn et al., 1997 [119] |
| | TAGTTT | GCACTA | McGowan et al., 2003 [120] |
| | GCGCGA | GCTAGC | McGowan et al., 2003 [120] |
| | GTCAAA | TCTAAA | McGowan et al., 2003 [120] |
| | - | TAGAAT | McGowan et al., 2003 [120] |
| | TAATTT | TGGAAT | McGowan et al., 2003 [120] |
| | ATTAAG | TATAAA | McGowan et al., 2003 [120] |
| | TAAGGT | GAAAAT | McGowan et al., 2003 [120] |
| | TTCAGC | TTTAAT | McGowan et al., 2003 [120] |
| | ATGCAT | TAAATT | McGowan et al., 2003 [120] |
| | - | TTTGCTT | Niehus et al., 2004[121] |
| σ54 | TTGGTA | TGCAA | Josenhans et al., 2002 [122] |
| | TTAGAA | TCGCA | Josenhans et al., 2002 [122] |
| σ28 | TaAA | cCGAT | Spohn & Scarlato, 1999 [123] |
| | | | Josenhans et al., 2002 [122] |

Another method I used to predict the sRNA promoter sequence was overall conservation among currently completely sequenced *H. pylori* strains was BLAST[124]. I reasoned that nucleotides that are part of a functional sequence would exhibit more conservation that those that were not[111]. I conducted a microbe BLAST consisting of 50nt upstream of the sRNA gene that was used to determine the conservation of the predicted promoter sequence[124]. These results were analyzed further using the multiple alignment tool, MAFFT (multiple alignment using fast Fourier transform) v7.464, with auto-

algorithm[125]. For HPnc2525 and HPnc2645, I believed the TSS of HPnc2525 and HPnc2645 might extend beyond the experimentally determined 5' ends. I analyzed sequences from the experimentally determined 5' end to the 5' end of their functionally identified promoters.

Terminator Prediction

Two types of bacterial transcriptional terminators control gene termination, Rho-independent (RIT) and Rho-dependent (RDT), either of which is located downstream from the codon, typically within approximately 50 nt of the 3' end of the transcript[126-128]. RIT destabilize and dissociate the elongating transcription complex of the transcript relying solely on interactions of DNA and RNA with RNA polymerase by an intrinsic secondary hairpin structure with minimum free energy ($\Delta G$) <-3.0 kcal/mol, a loop size of 3-10 nt, a stem is between 4-15 nt and a thymine-rich region should be within 20 nt of the hairpin structure[126,129]. Rho-dependent terminators use an RNA helicase, Rho, to help dissociate at specific sites on the transcript[39]. I used RhoTermPredict to find sequences RDT; this program searchers for RDT consensus sequences (Rho utilization site, RUT), a motif comprising of a 78 nt with regularly spaced C residues, higher C than G content, as well as a putative pause site for RNA polymerase (hairpin structures)[130]. I searched both RDT and RIT sites since *H. pylori* can contain either[126].

This method failed to predict any Rho-dependent terminators for HPnc2600 and HPnc2645., which may indicate a lack of Rho-dependent termination; thus, both were analyzed for intrinsic terminators. The criteria for intrinsic terminators outlined by Castillo et al. (2008) include a stem of 4–18bp, a stem of 9–18bp with a loop comprising

<20% of the stem length, or stem of 7–18bp with a 1 to 5 nucleotide bulge in either the 5′ or 3′ side of the stem[126].

sRNA Target Prediction

The sRNA target prediction program TargetRNA2 was used to analyze HPnc2600, HPnc2525, and HPnc2645 to identify candidate genes that may be regulated. The sequences used for prediction only included what was experimentally identified, not the sequences predicted with the in-silico analysis for the promoter and terminator extensions. Predicted targets were obtained using TargetRNA2 with default p-values $\leq$ 0.010 [94]. This software developed the secondary structure of the sRNA, the secondary structure of each candidate mRNA target, and the hybridization energy between the sRNA and each candidate mRNA target to determine potential targets[94]. Accurately predicted targets had a match of ≤0.005. were considered putative targets, except *cag*PAI targets or other interesting targets in HPnc2600. Unlike HPnc2525 and HPnc2645, the 5' and 3' ends are known. Meaning the full transcript of HPnc2600 has been identified, as discovered by Sharma et al. (2010). However, like HPnc2525 and HPnc2645, HPnc2600 is located in the clinically relevant *cag*PAI region. This location indicated that Hpnc2600 might regulate the expression of *cag*PAI genes due to the conservation of both the sRNA and *cag*PAI genes in multiple strains.

## Results

Experiments conducted by Ta et al. (2012) functionally identified promoters that regulate HPnc2525 and HPnc2645 and also an expressed transcript representing HPnc2645 [104]. Subsequently, TaKaRa SMARTer 5'RACE was used to tentatively identify the 5' and 3' ends of sRNAs HPnc2525 (202bp, located at 521475-521677) and HPnc2645 (79bp, located at 539120-539199, Castillo and Baber, unpublished results, Table 2, Fig 8).

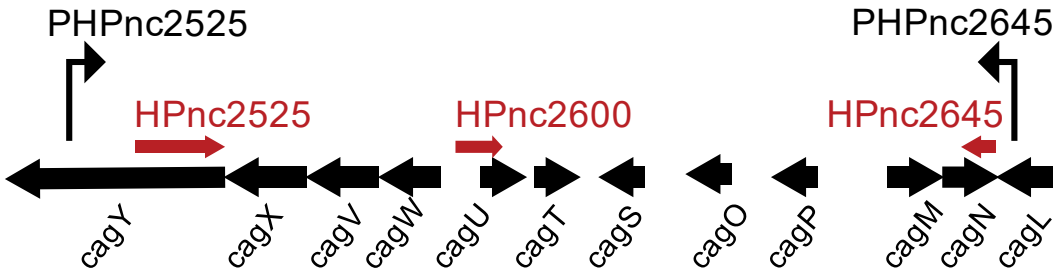



**Figure 8.** Subsection of *cag*PAI
Genes are represented by thick black arrows oriented in the direction of transcription. Promoters are represented by thin black flags pointing in the direction of transcription. Red arrows represent sRNA genes pointing in the direction of transcription. The lengths and spacing are approximately proportional.

Because the 5' end of HPnc2525 and HPnc2645 did not extend into the sequence that defines the functional promoters (PHPnc2525, 521375-521506; PHPnc2645, 539314-539259, Ta et al., 2012), we believe the TaKaRa SMARTer 5'RACE defined transcripts may be incomplete; this could result from the use of an RNA template that was degraded.

To determine if the 5' and 3' ends of HPnc2525 extends beyond what was identified by TaKaRa SMARTer 5' RACE, I isolated RNA from *H. pylori* strain G27 and used it as the template for an RT-PCR walking strategy. I used agarose gel electrophoresis to verify that the RNA was not degraded and would represent full-length transcripts free of impurities such as chromosomal DNA, proteins, and endotoxins. RNA

concentrations were between 98ng/µl and 133.6ng/µl, and A260/A280 readings were between 1.73 and 1.94. (Fig. 9).

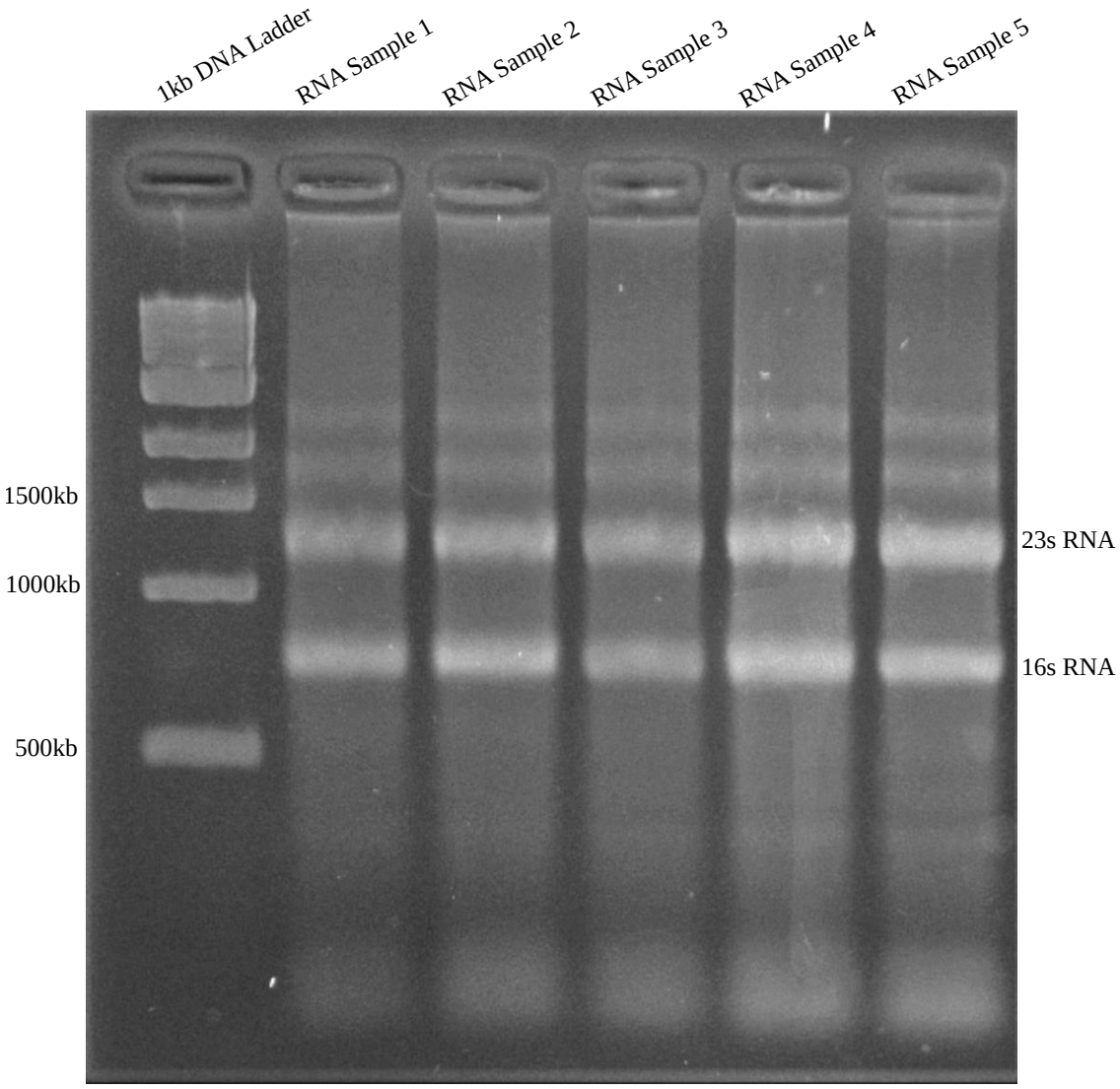



**Figure 9.** RNA Samples Used for RT-PCR Experiments
The RNA was not degraded. 5ug of the RNA samples were resolved on a 2% agarose gel. The 16S and 23S ribosomal RNA bands are visible in the intact RNA sample.

Oligonucleotides for the walking RT-PCR were designed with ten bp overlap upstream and downstream of those used to generate TaKaRa SMARTer 5′/3’ RACE products.

The upstream walk was conducted using PIX-Rnew for reverse transcription and then paired with PIX-F, PIX-F4, PIX-F6, or PIX-F8 in separate PCR reactions. A transcript was detected in each of these PCR reactions indicating that the sRNA was being transcribed. The 5' end of HPnc2525 extends at least to the 5' end of PIX-F8 (521426). This result puts the 5' end of HPnc2525 upstream of the 5' end of the cloned functional promoter. For the downstream walk, oligo PIX-Rnew, PIX-R, or PIX-R2 were used to conduct the reverse transcription step and then paired with PIX-F for PCR. This result indicates that the HPnc2525 3' end extends to at least the 3' end of oligo PIX-R2 (521643-521662) and the 5'end of PIX-F8 (521426-521445) (Fig. 10). Based on the RT-PCR experiments completed thus far, HPnc2525 is at least 241bp and is located at 521426-521677 (Table 1 and Fig. 10).

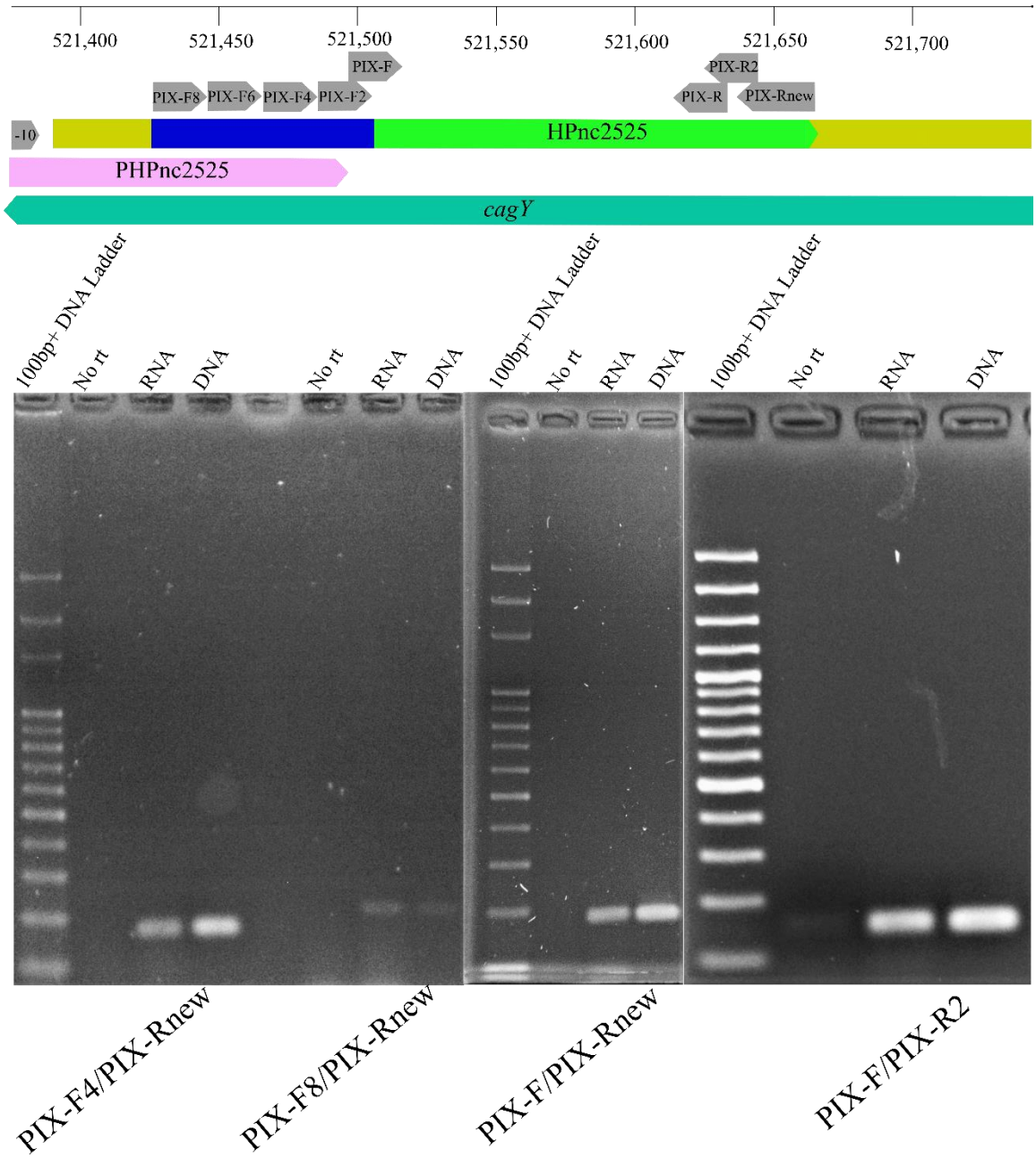



**Figure 10.** Walking Primer Strategy for HPnc2525
The green region corresponds to the putative sequence discovered previously by Ta et al. (2012). The blue line corresponds to the experimentally discovered 5' end. The yellow lines correspond to the sequence still yet to be experimentally verified. The gels shown have controls for no reverse transcription and positive DNA. All samples are 5µl loaded on 2% Agarose.

Predicted Transcriptional Promoters and Terminators

A sequence, 'TATAAT', that exactly matched a known promoter sequence for *cagA* was identified ~12nt from the TSS of HPnc2600[111]. There is no global consensus for the -35 site in *H. pylori*. Virtual Footprint also identified 'TATAAT' as a $\sigma^{70}$ promoter sequence, and no -35 sites were found. Based on this analysis, along with the BLAST and MAFFT results, I predict the promoter consensus sequence among *H. pylori* strains to be 'TATAA(T/C)' for the -10 site for HPnc2600 (Fig. 11a). Virtual Footprint predicted the promoter sequence 'CATAGT' for HPnc2525 located approximately 42bp upstream from the sequence confirmed by the walking RT-PCR experiment and within the PHpcnc2525 promoter region. This result was a close match to the known promoter 'TATAGT' (Fig. 11b)[114]. For HPnc2645, I predict that the promoter sequence is 'CAATAC', which is located approximately 86bp upstream from the 5'end sequence identified with the use SMARter 3'/5'RACE experiment and within the PHPnc2645 promoter region. This result was not close to a known promoter sequence, nor was this predicted by Virtual Footprint. However, by visual identification, a highly conserved 5nt sequence was found during the inspection of aligned sequences of PHpcnc2645 (Fig. 11c).

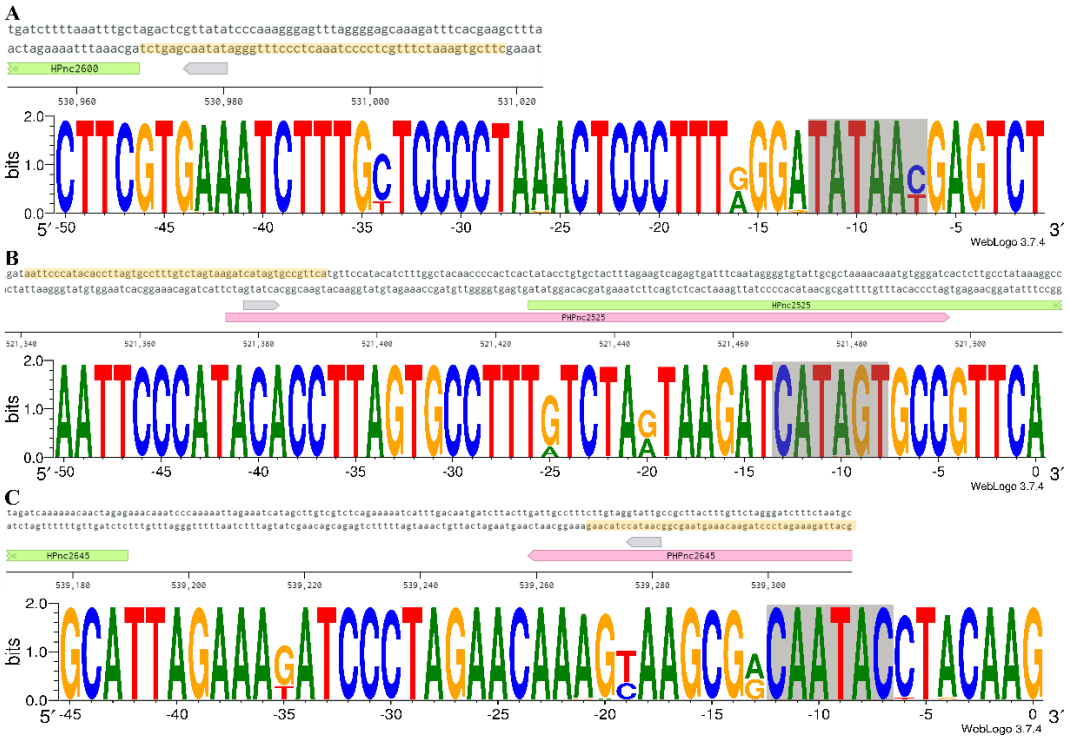



**Figure 11.** Promoters Found for HPnc2525, HPnc2600, and HPnc2645 (A) Conserved promoter sequence of HPnc2600. The -10 sequence TATAA(C/T). Prediction with Virtual Footprint – sigma70 – conservation shown with aligned microbe Blast+Mafft. The orange highlight is the -50 upstream from HPnc2600 identified by Sharma et al. (2010) with the -10 highlighted in grey. (B) Conserved promoter sequence of HPnc2525 -10 'CATAGT' prediction with Virtual Footprint – sigma70 – conservation shown with aligned microbe Blast+Mafft.HPnc2525. The orange highlight is upstream from the experimentally identified region of HPnc2525 with the -10 located in the functionally identified promoter region of PHPnc2525 in pink. (C) Conserved promoter sequence of HPnc2645 -10 'CAATAC' prediction with– sigma70 – conservation shown with aligned microbe Blast+Mafft. The orange highlight is upstream from the experimentally identified region of HPnc2645 with the -10 located in the functionally identified promoter region of PHPnc2645.

For HPnc2525, a putative RUT site was identified by RhoTermPredict approximately 38bp downstream from the experimentally identified sequence (Table 4 and Fig. 12a&d).

**Table 4.** Predicted Putative RUT Site as Predicted by RhoTermPredict for HPnc2525.

| **sRNA** | **Genomic sequence of putative RUT site** | **mG27 Genomic Position** |
|---|---|---|
| HPnc2525 | TTTTTACTTTTTTCTTTATCGTCATCTTTT TTATTGCCACCTAAAGCCTTAGCTAATTT AGCTTCTAAATTCTTATCC | 521705-521782 |

For both HPnc2600 and HPnc2645, a sequence 100nt upstream from sRNA 3'end was analyzed with RNAfold for structure and energy, which ΔG any of the stems had to have a ΔG < −3.0 kcal/mol (Fig. 12b,c,e,f)[126,131].

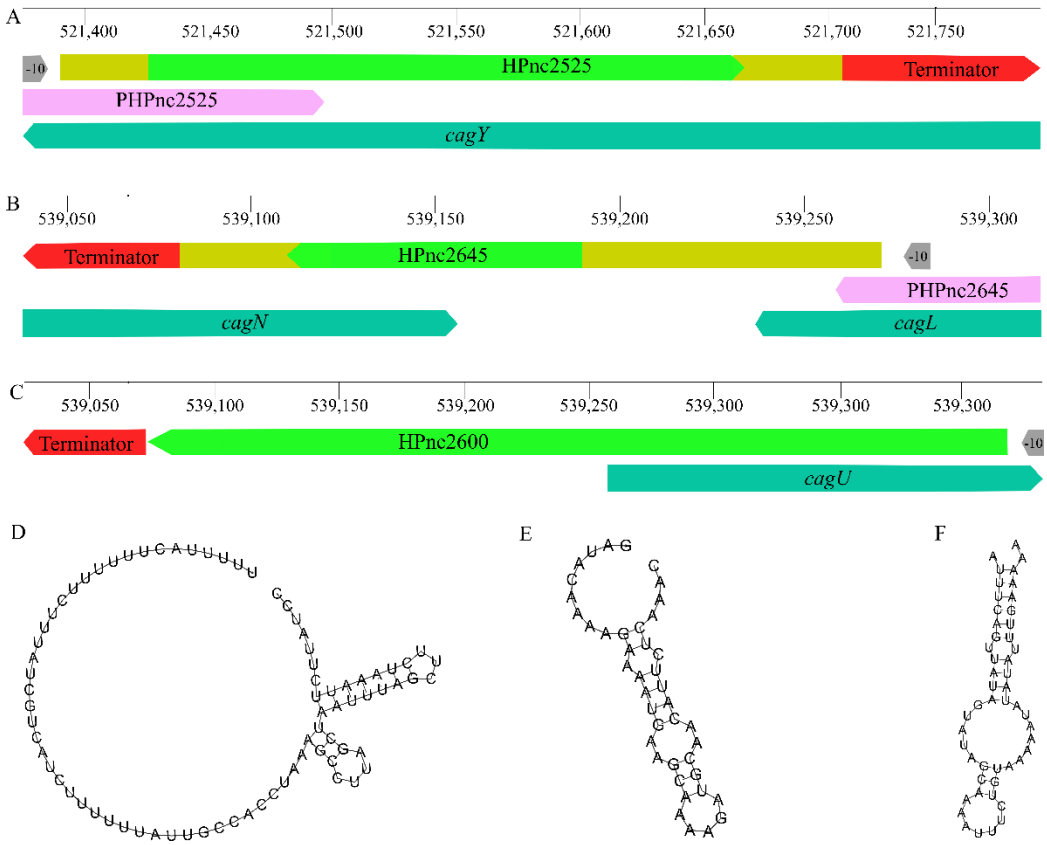



**Figure 12.** Terminators Found for HPnc2525, HPnc2600, and HPnc2645 (A)Putative RUT site of HPnc2525 predicted by RhoTermPredict. Predicted terminator (red) with respect to the experimentally identified sequence of Hpnc2525. (B) Predicted terminator (red) location for HPnc2645 with respect to HPnc2645 (partially cloned by Castillo & Baber, unpublished). (C) Predicted terminator (red) location with respect to HPnc2600, discovered by Sharma et al. (2010). (D) Stem-loop predicted by RhoTermPredict. (E) Stem-loop of HPnc2645 terminator, 14bp with a loop <20% of stem length, -11.20 kcal/mol and upstream from HPnc2645. (F) Stem-loop of HPnc2600 terminator, 19bp with a loop <20% of stem length, -5.71 kcal/mol and 2nt upstream from HPnc2600.

Identification of mRNA regulated by HPnc2525 and HPnc2600, and HPnc2645

The top 5 gene targets predicted by TargetRNA2 for HPnc2525, HPnc2600, and HPnc2645 are reported in Tables 5, 6, 7, and 8, except *hopZ* due to its interest in virulence being an outer membrane protein. Targets were predicted with TargetRNA2 in both *H. pylori* strain 26695 and G27. Sharma et al. (2010) discovered HPnc2600 in *H. pylori* strain 26695, and I am studying *H. pylori* strain G27. There was no identical conserved target, but they had similar targets being within *cag*PAI. In *H. pylori* strain 26695, cagU is a known translocator protein or pilus protein and part of the T4SS found in *cag*PAI[132,133]. In *H. pylori* G27, cagL, is also a translocator protein and which translocates cagA into host cells[134].

**Table 5.** Targets Predicted by Analysis of Experimentally Identified HPnc2525 Gene Sequence in TargetRNA2
I included targets that fell below the TargetRNA2 recommended p-value $\leq$ 0.010. Gene *fliA* was included due to the potential relation to virulence.

| **Rank** | **Gene** | **Synonym** | **Energy** | **P-value** | **Function** |
|---|---|---|---|---|---|
| 1 | *engA* | HPG27_793 | -16.01 | 0.000 | GTP-binding protein EngA |
| 2 | - | HPG27_647 | -15.85 | 0.000 | succinyl-CoA-transferase subunit A |
| 3 | - | HPG27_1140 | -14.97 | 0.001 | 30S ribosomal protein S7 |
| 4 | - | HPG27_825 | -14.77 | 0.001 | CDP-diacylglycerol pyrophosphatase |
| 14 | *fliA* | HPG27_396 | -12.76 | 0.004 | 8-amino-7-oxononanoate synthase |

**Table 6.** Targets Predicted by Analysis of sRNA HPnc2600 in *H. pylori* 26695 Gene Sequence in TargetRNA2
TargetRNA2 recommends using p-value ≤ 0.010.

| Rank | Gene | Synonym | Energy | P-value | Function |
|---|---|---|---|---|---|
| 1 | *hemC* | HP0237 | -19.17 | 0.000 | Porpgobilinogen Deaminase |
| 2 | *cagU* | HP0531 | -14.97 | 0.001 | *cag*PAI protein cagU |
| 3 | *Rnc* | HP0662 | -13.75 | 0.001 | Ribonuclease III |
| 4 | *ksgA* | HP1431 | -13.74 | 0.001 | Dimethyladenosine transferase |
| 5 | HP1375 | HP1375 | -13.69 | 0.001 | UDP-N-acetylglucosamine acyltransferase |

**Table 7.** Targets Predicted by Analysis of sRNA Hpc2600 in *H. pylori* G27 Gene Sequence in TargetRNA2
TargetRNA2 recommends using p-value ≤ 0.010. Gene *hopZ* was included due to the potential relation to virulence.

| Rank | Gene | Synonym | Energy | P-value | Function |
|---|---|---|---|---|---|
| 1 | *folE* | HPG27_877 | -16.44 | 0.000 | GTP cyclohydrolase I |
| 2 | HPG27_1320 | HPG27_1320 | -15.75 | 0.000 | UDP-N-acetylglucosamine acyltransferase |
| 3 | HPG27_571 | HPG27_571 | -15.26 | 0.000 | ABC transporter permease |
| 5 | *cagL* | HPG27_497 | -13.7 | 0.002 | *cag*PAI protein cagL |
| 22 | *hopZ* | HPG27_8 | -11.09 | 0.013 | Outer membrane protein |

**Table 8.** Targets Predicted by Analysis of Partial HPnc2645 Gene Sequence in TargetRNA2
TargetRNA2 recommends using p-value ≤ 0.010.

| Rank | Gene | Synonym | Energy | P-value | Function |
|---|---|---|---|---|---|
| 1 | *rpsM* | HPG27_1246 | -17.33 | 0.000 | 30S ribosomal protein S13 |
| 2 | *hydB* | HPG27_593 | -13.85 | 0.002 | quinone-reactive Ni/Fe hydrogenase |
| 3 | - | HPG27_139 | -13.23 | 0.003 | hypothetical protein |
| 4 | *rpsJ* | HPG27_1269 | -12.95 | 0.004 | 30S ribosomal protein S10 |
| 5 | - | HPG27_1198 | -12.68 | 0.005 | tryptophanyl-tRNA synthetase |

**Table 9.** HPnc2600 Sequence Found by Sharma et al. (2010) in *Helicobacter pylori* Strain 26695, and HPnc2600 in *H. pylori* Strain G27.
*H. pylori* G27 BLAST: 100% coverage, 93.3% identity, E-value=$4e^{-149}$ ($1e^{-50}$ is considered very high quality)

| **sRNA** | **Sequence (5' end)** | **Genomic Position** |
|---|---|---|
| HPnc2600 (26695) | AGCAAATTTAAAAGATCATTGCTCTGTG AGTTGCTAGGTGGTGGAGCGTTTAGCGG ATTAGGTCCATGATGCTCTGTTGTATCGT TCATGAAGCTCCTTTCAAGAATTAAATTG AGAAATTGTTTTGTTATTATACCATTCTC TCTCTGAGTTGTGATTGTCTTATCTCTTT AAATTAGGCGCTTCTAAAATTTCATTACT GGGTTACGACTGCTTACTCACTGCTCTTA CTTTTTGAGTTGCATCGTATTTCATCTTG CTTCTTGTTTGAAGCAATCCGCTACTTAC ATTTATTATAGAGAATCTTTGCTCAACAC CTTATCCAAAAAGATTCTTATTATAAGGT | 564469-564124 (antisense to *cag*V) |
| HPnc2600 (G27) | AGCAAATTTAAAAGATCATTGCTCTGTG AGTTGCTAGGTGGTGGGGCGTTTAGCGA ATTGGATCCATGATGCTCTGTTGTATCGT TCATGAAATTCCTTTCAAGAATTAAATTG AGAAATTGTTTTGATATTATACCATTCTC TCTCTGAGTTGTGATTGTCTTATCTCTTT GAATTAGGCGCTTCTAAAATTTCATTACT GATTACGACTGCTTACTTATTGCTCTTAC TTTTTGAGTTGCATCGTGTTTCATCTTGC TTCTTGTTTGAAGCAATCCGCTACCTTAC ATTTATTATAAGGAATCTTTGTTCAACGC CTTATCCAAAAAGGTTTTTATTAAAGGT | 530634-530978 (antisense to *cag*V) |

## Further Directions

Functional Assay to Verify Predicted mRNA Targets

To evaluate TargetRNA2 predicted targets for their regulation by HPnc2525 and HPnc2645, A Green-Fluorescent-Protein (GFP) plasmid-based gene expression reporter system would be constructed. This reporter requires co-expression of (HPnc2525, HPnc2600, or HPnc2645) and the predicted target, fused to GFP, from two different plasmids in *E. coli*. If the sRNA regulates expression of the predicted target, it is expected to see a change in fluorescence compared to control strains that do not express the sRNA (Fig. 13).

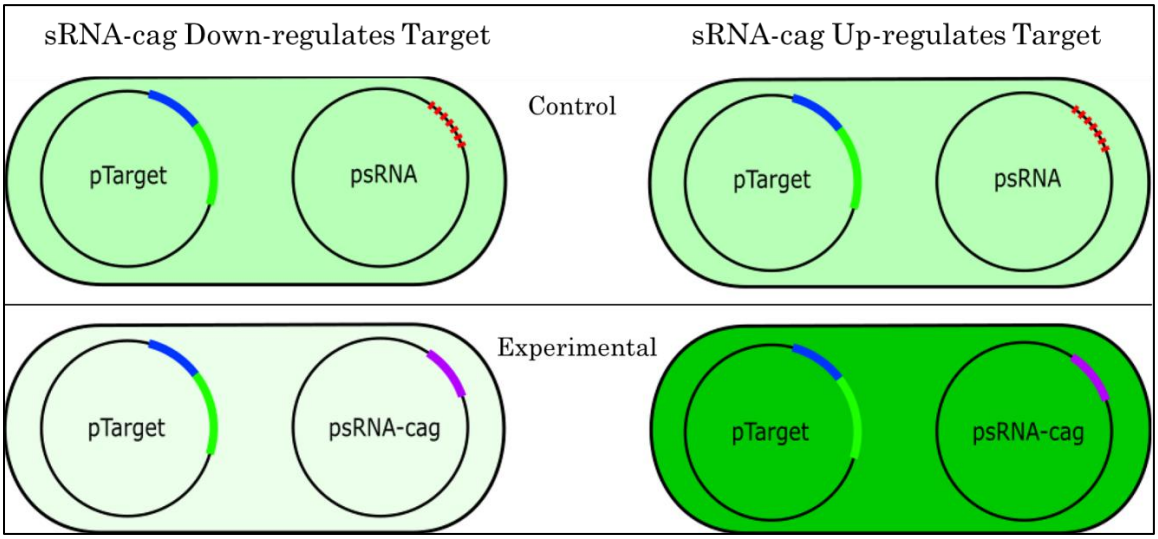



**Figure 13.** Plasmid Expression Reporter
Expected fluorescence intensity result of *E. coli* cells transformed with psRNA-random and pTARG compared to psRNA-cag and pTARG.

Plasmid psRNA would be constructed using an inverse PCR and Gibson assembly using pJV300 provided by Dr. Jörg Vogel from the University of Würzburg. The selectable marker for the constructed plasmids would be chloramphenicol.

Plasmid pTARG would be created for each predicted target that will include the target promoter and open reading frame cloned in-frame with GFP, which will be

accomplished by double digesting PxG10sf with NheiIHF and NsiIHF following the manufacturer's instructions to make the backbone vector.

Plasmid PxG10 was also obtained from Dr. Jörg Vogel from the University of Würzburg. The TargetRNA2 predicted sequences would be amplified via PCR from the *H. pylori* G27 genome from purified DNA. This amplification of the promoter regions and open reading frame will use oligonucleotides that will be ordered when the final TargetRNA2 targets are predicted from the completely experimentally identified sRNA sequences. These oligonucleotides will be designed to be ligated into the digested regions of PxG10sf. The target PCR product will then be ligated into the digest. The selectable marker for the constructed plasmids will be ampicillin.

The control plasmid used for this study is pJV300. Plasmid pJV300 is a built from pZE12luc and contains a random sRNA sequence (Fig. 14).

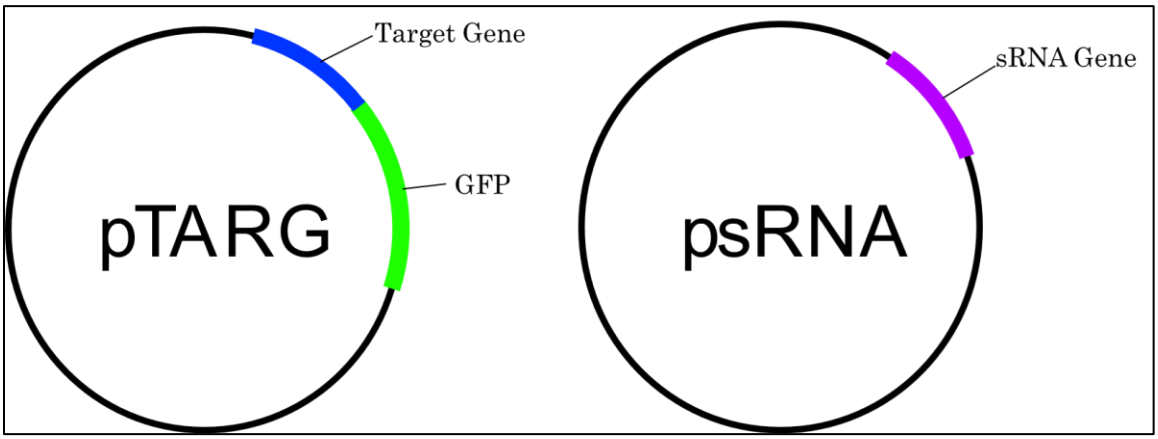



**Figure 14.** Plasmids pTARG and psRNA
Each sRNA would be cloned downstream of the constitutive promoter in plasmid psRNA to generate pHPnc2525, pHPnc2600, and pHPnc2645. The HPnc2525/2600/2645 genes must be moved from the cloning vector to a plasmid that is specific for this expression analysis. Each predicted sRNA target gene will be cloned downstream of a constitutive promoter and upstream and in-frame with GFP in pTARG.

Competent *E. coli* preparation and transformation

The cloned plasmids pHPnc2525, pHPnc2600, pHPnc2645, and pTARG would be sequentially transformed into *Escherichia coli* (strain DH5α) via electroporation. For the first transformation, competent *E. coli* will be prepared using a standard protocol available in the lab. In brief, cells are cultured in Luria-Bertani (LB, 500mL) media to an optical density (λ600) of approximately 0.5, then alternately centrifuged to pellet and washed, 2X with ice-cold water and 1X with 10% glycerol. After the final centrifugation, the cells are resuspended in 0.5mL ice-cold water, aliquoted to 50µl and stored at -80°C. For the second transformation, competent *E. coli* containing the first plasmid will be cultured in LB with an antibiotic (ampicillin 100µg/mL) to maintain that plasmid but otherwise prepared in the same way. For the electroporations, 50µl electrocompetent *E. coli* is mixed with 0.5 µl of the transforming DNA in a 2mm electroporation cuvette and electroporated using standard *E. coli* settings on a BioRad electroporator. Transformed *E. coli* would be cultured on LB media to select for transformation with the appropriate plasmid, ampicillin (100µg/mL) for, and for chloramphenicol (100µg/mL) (Fig. 15)[96].

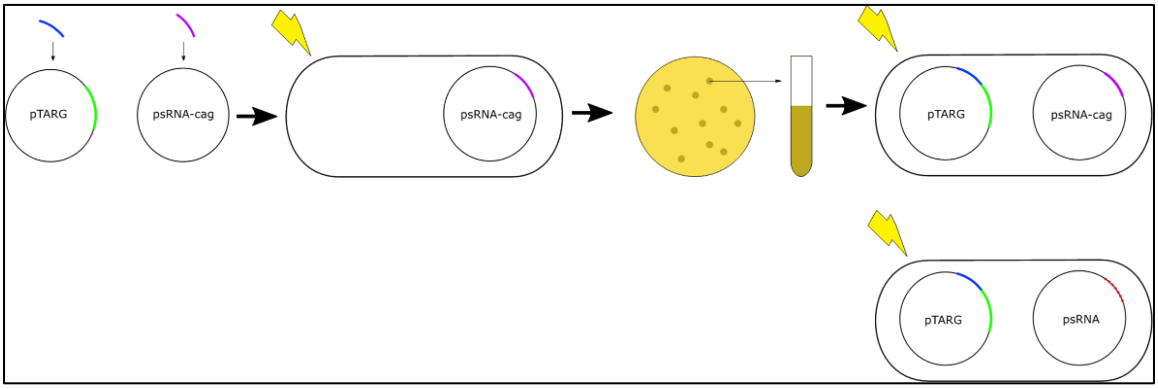



**Figure 15.** Sequential Transformation of *E. coli* with sRNA Containing Plasmid (psRNA-cag) and Target Gene Containing Plasmid (pTARG).

Quantitation of target-GFP expression

Fluorescence due to target-GFP expression will be quantitated using TECAN SpecTRFlourPlus Analysis. *E. coli* cultured in LB media plus selective antibiotics (ampicillin and chloramphenicol) for approximately 12 hours would be transferred to a 96 well plate and scanned by excitation at 488nm wavelength and emittance at 510nm wavelength. Fluorescence of *E. coli* would be compared with each psRNA and pTARG combination to a control, psRNA random, and pTARG. A minimum of three independent experiments with a minimum of 10 replicates each for each psRNA, and pTARG combination would be conducted. The mean fluorescence value of *E. coli* with each psRNA + pTARG would be compared to *E. coli* with psRNA-random+ pTARG (e.g., engA) using a two-tailed T-test.

## Current Preparation for Further Study

Plasmid preparation

Plasmid pXG10sf was generously provided J. Vogel at the University of Wurzburg, Germany. Due to the low copy number of pXG10sf, a more massive scale plasmid prep was conducted (Qiagen Midiprep, Qiagen). This prep yielded 997.6ng/µl. All other plasmids were prepared using small scale preparations, per the Qiagen miniprep protocol (Qiagen). These preps yielded approximately10 to 20ng/ul.

DNA representing the mRNA targets

*fli*A was amplified from 0.5 µl genomic DNA with oligonucleotides fliA-F and fliA-R, which add NheI-HF and NseI-HF restrictions sites and direct their in-frame cloning with the GFP in pXG10sf (Appendix). Digestion of mRNA target amplicons was

conducted using NheI-HF and NseI-HF (New England Biolabs) for 3 hours at 37$^{o}$C with no heat inactivation. The sample was cleanup via DNA precipitation (Appendix).

**Table 10.** Oligonucleotides used for further study.

| | |
|---|---|
| fliA-F | GTTTTTATGCATGTCTATTTTTATCGCTTGCACGGAT |
| fliA-R | GTTTTGCTAGCCTTGGCTTAAAATATCAGCCATGATC |

Assembly of the pXG10s vector and Target amplicons

Cloning of *fliA* was performed using OneTaq from NEB., however not yet assembled.

## Discussion

Due to the global pandemic of COVID-19, some benchwork was not able to be completed, and the experiments not able to be completed are mentioned in detail in the "Further Directions" section. Because of this, the 3' and 5' ends of HPnc2525 and HPnc2645 were not able to be finalized, thus preventing definitive structural analysis and target verification. Target verification for HPnc2600 was also incomplete.

The promoters and terminators of HPnc2525, HPnc2600, and HPnc2645 were searched for either computationally, visually, or consensus among multiple strains. These were not able to be definitively verified experimentally due to lab closures. Direct manipulation of HPnc2525, HPnc2600, and HPnc2645 will be needed to verify if these are the correct promoters and terminators in *H. pylori* G27.

There is no known -35 consensus in the promoter search, which is due to *H. pylori* not having a clear -35 consensus even with the same -10 sequence. The -35 sequence could also be experimentally discovered in future studies.

Although HPnc2600 is nearly identical between *H. pylori* strains 26695 and G27, their predicted targets are different. The interesting targets being *cagU* for strain 26695 and *cagL* in strain G27. A blast between strains for *cagU* shows an e-value of 0.00 and 97.67% identity and *cagL* shoes an e-value of 0.00 and 96.5% identity. It would be interesting to see if those targets hold true per strain. HPnc2600 may regulate *cagU* in 26695 but not *cagL* with the reverse in G27. It would be intriguing to see what is regulating the non-targeted *cag* gene in the respective strains. Another possibility is that the non-targeted *cag* gene may lack regulation in that strain. Further study would be needed.

It is also unknown whether *H. pylori* uses a protein chaperone like Hfq in order to protect the sRNA molecules. Searching for proteins with similar structural similarities to Hfq in *H. pylori* could be undertaken in order to find a homologue to Hfq. Another method could be to search evolutionarily. A BLAST search I completed yielded that many bacterial species have Hfq genes. Some classes included in Proteobacteria, excluding *Gammaproteobacteria*, have partial or similar genes to that of Hfq. Among the *Epsilonproteobacteria*, which *H. pylori* is a member of, *Campylobacter jejuni* and *Arcobacter sp.* both have partial Hfq sequences. However, from initial research the sRNAs in found *Epsilonproteobacteria* may not conform to the same rules as they do in *Gammaproteobacteria*[135]. It would be interesting to track the evolutionary changes to these genes and sRNAs to see what, if any, genes play a role in chaperoning sRNA molecules or how sRNA structural changes may negate the need for Hfq-like chaperones.

# REFERENCES


1 Brown, L. M. Helicobacter pylori: epidemiology and routes of transmission. *Epidemiol Rev* **22**, 283-297, doi:10.1093/oxfordjournals.epirev.a018040 (2000).

2 Hooi, J. K. Y. *et al.* Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology* **153**, 420-429, doi:10.1053/j.gastro.2017.04.022 (2017).

3 Megraud, F., Bessede, E. & Varon, C. Helicobacter pylori infection and gastric carcinoma. *Clin Microbiol Infect* **21**, 984-990, doi:10.1016/j.cmi.2015.06.004 (2015).

4 Linz, B. *et al.* An African origin for the intimate association between humans and Helicobacter pylori. *Nature* **445**, 915-918, doi:10.1038/nature05562 (2007).

5 Du, Y. *et al.* Helicobacter pylori and Schistosoma japonicum co-infection in a Chinese population: helminth infection alters humoral responses to H. pylori and serum pepsinogen I/II ratio. *Microbes Infect* **8**, 52-60, doi:10.1016/j.micinf.2005.05.017 (2006).

6 Fox, J. G. *et al.* Concurrent enteric helminth infection modulates inflammation and gastric immune responses and reduces helicobacter-induced gastric atrophy. *Nat Med* **6**, 536-542, doi:10.1038/75015 (2000).

7 Everhart, J. E., Kruszon-Moran, D., Perez-Perez, G. I., Tralka, T. S. & McQuillan, G. Seroprevalence and ethnic differences in Helicobacter pylori infection among adults in the United States. *J Infect Dis* **181**, 1359-1363, doi:10.1086/315384 (2000).

8 Nurgalieva, Z. Z. *et al.* Helicobacter pylori infection in Kazakhstan: effect of water source and household hygiene. *Am J Trop Med Hyg* **67**, 201-206, doi:10.4269/ajtmh.2002.67.201 (2002).

9 Aziz, R. K., Khalifa, M. M. & Sharaf, R. R. Contaminated water as a source of Helicobacter pylori infection: A review. *J Adv Res* **6**, 539-547, doi:10.1016/j.jare.2013.07.007 (2015).

10 Sgambato, D. *et al.* Prevalence of Helicobacter pylori infection in sexual partners of H. pylori-infected subjects: Role of gastroesophageal reflux. *United European Gastroenterol J* **6**, 1470-1476, doi:10.1177/2050640618800628 (2018).

11 Quaglia, N. C. & Dambrosio, A. Helicobacter pylori: A foodborne pathogen? *World J Gastroenterol* **24**, 3472-3487, doi:10.3748/wjg.v24.i31.3472 (2018).

12 Testerman, T. L. & Morris, J. Beyond the stomach: an updated view of Helicobacter pylori pathogenesis, diagnosis, and treatment. *World J Gastroenterol* **20**, 12781-12808, doi:10.3748/wjg.v20.i36.12781 (2014).

13 Kusters, J. G., van Vliet, A. H. & Kuipers, E. J. Pathogenesis of Helicobacter pylori infection. *Clin Microbiol Rev* **19**, 449-490, doi:10.1128/CMR.00054-05 (2006).

14 Smolka, A. J. & Schubert, M. L. Helicobacter pylori-Induced Changes in Gastric Acid Secretion and Upper Gastrointestinal Disease. *Curr Top Microbiol Immunol* **400**, 227-252, doi:10.1007/978-3-319-50520-6_10 (2017).

15 Uemura, N. *et al.* Changes in Helicobacter pylori-induced gastritis in the antrum and corpus during long-term acid-suppressive treatment in Japan. *Aliment Pharmacol Ther* **14**, 1345-1352, doi:10.1046/j.1365-2036.2000.00834.x (2000).

16 Marshall, B. & Warren, J. R. Unidentified Curved Bacilli in the Stomach of Patients with Gastritis and Peptic Ulceration. *The Lancet* **323**, 1311-1315, doi:10.1016/s0140-6736(84)91816-6 (1984).

17 Veldhuyzen van Zanten, S. J. O., Dixon, M. F. & Lee, A. The gastric transitional zones: Neglected links between gastroduodenal pathology and Helicobacter ecology. *Gastroenterology* **116**, 1217-1229, doi:10.1016/s0016-5085(99)70025-9 (1999).

18 Uppin, M. I. & Hannurkar, K. K. Prevalence of Helicobacter pylori in patients with peptic ulcer disease. *International Surgery Journal* **5**, doi:10.18203/2349-2902.isj20181011 (2018).

19 Nomura, S. *et al.* Effect of dietary anti-urease immunoglobulin Y on Helicobacter pylori infection in Mongolian gerbils. *Helicobacter* **10**, 43-52, doi:10.1111/j.1523-5378.2005.00290.x (2005).

20 Rauws, E. A. J. & Tytgat, G. N. J. Cure of duodenal ulcer associated with eradication of Helicobacter pylori. *The Lancet* **335**, 1233-1235, doi:10.1016/0140-6736(90)91301-p (1990).

21 Mbulaiteye, S. M., Hisada, M. & El-Omar, E. M. Helicobacter Pylori associated global gastric cancer burden. *Front Biosci (Landmark Ed)* **14**, 1490-1504, doi:10.2741/3320 (2009).

22 Du, M. Q. & Atherton, J. C. Molecular subtyping of gastric MALT lymphomas: implications for prognosis and management. *Gut* **55**, 886-893, doi:10.1136/gut.2004.061663 (2006).

23 Kuo, S. H. *et al.* Detection of the Helicobacter pylori CagA protein in gastric mucosa-associated lymphoid tissue lymphoma cells: clinical and biological significance. *Blood Cancer J* **3**, e125, doi:10.1038/bcj.2013.22 (2013).

24 Zullo, A. *et al.* Effects of Helicobacter pylori eradication on early stage gastric mucosa-associated lymphoid tissue lymphoma. *Clin Gastroenterol Hepatol* **8**, 105-110, doi:10.1016/j.cgh.2009.07.017 (2010).

25 Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. *IARC Monogr Eval Carcinog Risks Hum* **61**, 1-241 (1994).

26 Lacy, B. E. Functional dyspepsia and gastroparesis: one disease or two? *Am J Gastroenterol* **107**, 1615-1620, doi:10.1038/ajg.2012.104 (2012).

27 Stanghellini, V. *et al.* Gastroduodenal Disorders. *Gastroenterology* **150**, 1380-1392, doi:10.1053/j.gastro.2016.02.011 (2016).

28 Talley, N. J. & Hunt, R. H. What Role Does Helicobacter pylori Play in Dyspepsia and Nonulcer Dyspepsia? Arguments For and Against H. pylori Being Associated With Dyspeptic Symptoms. *Gastroenterology* **113**, S67-S77, doi:10.1016/s0016-5085(97)80016-9 (1997).

29 Moss, S. F., Calam, J., Agarwal, B., Wang, S. & Holt, P. R. Induction of gastric epithelial apoptosis by Helicobacter pylori. *Gut* **38**, 498-501, doi:10.1136/gut.38.4.498 (1996).

30 Yamaguchi, T. *et al.* Gastric epithelial cell proliferation and apoptosis in Helicobacter pylori-infected mice. *Aliment Pharmacol Ther* **14 Suppl 1**, 68-73, doi:10.1046/j.1365-2036.2000.014s1068.x (2000).

31 Kim, J. J. *et al.* Helicobacter pylori impairs DNA mismatch repair in gastric epithelial cells. *Gastroenterology* **123**, 542-553, doi:10.1053/gast.2002.34751 (2002).

32 Morin, P. J. β-catenin signaling and cancer. *BioEssays* **21**, 1021-1030, doi:10.1002/(sici)1521-1878(199912)22:1<1021::Aid-bies6>3.0.Co;2-p (1999).

33 Murata-Kamiya, N. *et al.* Helicobacter pylori CagA interacts with E-cadherin and deregulates the beta-catenin signal that promotes intestinal transdifferentiation in gastric epithelial cells. *Oncogene* **26**, 4617-4626, doi:10.1038/sj.onc.1210251 (2007).

34 Ohnishi, N. *et al.* Transgenic expression of Helicobacter pylori CagA induces gastrointestinal and hematopoietic neoplasms in mouse. *Proc Natl Acad Sci U S A* **105**, 1003-1008, doi:10.1073/pnas.0711183105 (2008).

35 Caston, R. R. *et al.* Effect of environmental salt concentration on the Helicobacter pylori exoproteome. *J Proteomics* **202**, 103374, doi:10.1016/j.jprot.2019.05.002 (2019).

36 Ogihara, A. *et al.* Relationship between Helicobacter pylori infection and smoking and drinking habits. *J Gastroenterol Hepatol* **15**, 271-276, doi:10.1046/j.1440-1746.2000.02077.x (2000).

37 Wawro, N. *et al.* Helicobacter pylori Seropositivity: Prevalence, Associations, and the Impact on Incident Metabolic Diseases/Risk Factors in the Population-Based KORA Study. *Front Public Health* **7**, 96, doi:10.3389/fpubh.2019.00096 (2019).

38 Brenner, H. *et al.* Alcohol consumption and Helicobacter pylori infection: results from the German National Health and Nutrition Survey. *Epidemiology* **10**, 214-218 (1999).

39 Hwang, I. R. *et al.* Effect of interleukin 1 polymorphisms on gastric mucosal interleukin 1beta production in Helicobacter pylori infection. *Gastroenterology* **123**, 1793-1803, doi:10.1053/gast.2002.37043 (2002).

40 Yamaoka, Y. Increasing evidence of the role of Helicobacter pylori SabA in the pathogenesis of gastroduodenal disease. *J Infect Dev Ctries* **2**, 174-181, doi:10.3855/jidc.259 (2008).

41 Mais, D. D. & American Society for Clinical Pathology. *Quick compendium of clinical pathology*. 3e. edn, (American Society for Clinical Pathology Press, 2014).

42 D'Elios, M. M. *et al.* T helper 1 effector cells specific for Helicobacter pylori in the gastric antrum of patients with peptic ulcer disease. *J Immunol* **158**, 962-967 (1997).

43 Bagheri, N., Salimzadeh, L. & Shirzad, H. The role of T helper 1-cell response in Helicobacter pylori-infection. *Microb Pathog* **123**, 1-8, doi:10.1016/j.micpath.2018.06.033 (2018).

44 Mitchell, H. M., Ally, R., Wadee, A., Wiseman, M. & Segal, I. Major differences in the IgG subclass response to Helicobacter pylori in the first and third worlds. *Scand J Gastroenterol* **37**, 517-522, doi:10.1080/00365520252903044 (2002).

45 Vignali, D. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nat Rev Immunol* **8**, 523-532, doi:10.1038/nri2343 (2008).

46 Lundgren, A. *et al.* Mucosal FOXP3-expressing CD4+ CD25high regulatory T cells in Helicobacter pylori-infected patients. *Infect Immun* **73**, 523-531, doi:10.1128/IAI.73.1.523-531.2005 (2005).

47 Robinson, K. *et al.* Helicobacter pylori-induced peptic ulcer disease is associated with inadequate regulatory T cell responses. *Gut* **57**, 1375-1385, doi:10.1136/gut.2007.137539 (2008).

48 Mobley, H. L., Island, M. D. & Hausinger, R. P. Molecular biology of microbial ureases. *Microbiol Rev* **59**, 451-480 (1995).

49 Weeks, D. L., Eskandari, S., Scott, D. R. & Sachs, G. A H+-gated urea channel: the link between Helicobacter pylori urease and gastric colonization. *Science* **287**, 482-485, doi:10.1126/science.287.5452.482 (2000).

50 Celli, J. P. *et al.* Helicobacter pylori moves through mucus by reducing mucin viscoelasticity. *Proc Natl Acad Sci U S A* **106**, 14321-14326, doi:10.1073/pnas.0903438106 (2009).

51 Schwartz, J. T. & Allen, L. A. Role of urease in megasome formation and Helicobacter pylori survival in macrophages. *J Leukoc Biol* **79**, 1214-1225, doi:10.1189/jlb.0106030 (2006).

52 Constantino, M. A., Jabbarzadeh, M., Fu, H. C. & Bansil, R. Helical and rod-shaped bacteria swim in helical trajectories with little additional propulsion from helical shape. *Sci Adv* **2**, e1601661, doi:10.1126/sciadv.1601661 (2016).

53 Lertsethtakarn, P., Ottemann, K. M. & Hendrixson, D. R. Motility and chemotaxis in Campylobacter and Helicobacter. *Annu Rev Microbiol* **65**, 389-410, doi:10.1146/annurev-micro-090110-102908 (2011).

54 Kao, C. Y. *et al.* Higher motility enhances bacterial density and inflammatory response in dyspeptic patients infected with Helicobacter pylori. *Helicobacter* **17**, 411-416, doi:10.1111/j.1523-5378.2012.00974.x (2012).

55 Kao, C. Y., Sheu, B. S. & Wu, J. J. CsrA regulates Helicobacter pylori J99 motility and adhesion by controlling flagella formation. *Helicobacter* **19**, 443-454, doi:10.1111/hel.12148 (2014).

56 Kao, C. Y., Sheu, B. S. & Wu, J. J. Helicobacter pylori infection: An overview of bacterial virulence factors and pathogenesis. *Biomed J* **39**, 14-23, doi:10.1016/j.bj.2015.06.002 (2016).

57 Smith, J. L. The physiological role of ferritin-like compounds in bacteria. *Crit Rev Microbiol* **30**, 173-185, doi:10.1080/10408410490435151 (2004).

58 Amedei, A. *et al.* The neutrophil-activating protein of Helicobacter pylori promotes Th1 immune responses. *J Clin Invest* **116**, 1092-1101, doi:10.1172/JCI27177 (2006).

59 Polenghi, A. *et al.* The neutrophil-activating protein of Helicobacter pylori crosses endothelia to promote neutrophil adhesion in vivo. *J Immunol* **178**, 1312-1320, doi:10.4049/jimmunol.178.3.1312 (2007).

60 Evans, D. J., Jr. *et al.* Characterization of a Helicobacter pylori neutrophil-activating protein. *Infect Immun* **63**, 2213-2220 (1995).

61 Petersson, C. *et al.* Helicobacter pylori SabA adhesin evokes a strong inflammatory response in human neutrophils which is down-regulated by the neutrophil-activating protein. *Med Microbiol Immunol* **195**, 195-206, doi:10.1007/s00430-006-0018-x (2006).

62 Ilver, D. *et al.* Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* **279**, 373-377, doi:10.1126/science.279.5349.373 (1998).

63 Backstrom, A. *et al.* Metastability of Helicobacter pylori bab adhesin genes and dynamics in Lewis b antigen binding. *Proc Natl Acad Sci U S A* **101**, 16923-16928, doi:10.1073/pnas.0404817101 (2004).

64 Sheu, S. M. *et al.* H. pylori clinical isolates have diverse babAB genotype distributions over different topographic sites of stomach with correlation to clinical disease outcomes. *BMC Microbiol* **12**, 89, doi:10.1186/1471-2180-12-89 (2012).

65 Mahdavi, J. *et al.* Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation. *Science* **297**, 573-578, doi:10.1126/science.1069076 (2002).

66 Boren, T., Falk, P., Roth, K. A., Larson, G. & Normark, S. Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens. *Science* **262**, 1892-1895, doi:10.1126/science.8018146 (1993).

67 Viala, J. *et al.* Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island. *Nat Immunol* **5**, 1166-1174, doi:10.1038/ni1131 (2004).

68 Backert, S., Neddermann, M., Maubach, G. & Naumann, M. Pathogenesis of Helicobacter pylori infection. *Helicobacter* **21 Suppl 1**, 19-25, doi:10.1111/hel.12335 (2016).

69 Sause, W. E., Castillo, A. R. & Ottemann, K. M. The Helicobacter pylori autotransporter ImaA (HP0289) modulates the immune response and contributes to host colonization. *Infect Immun* **80**, 2286-2296, doi:10.1128/IAI.00312-12 (2012).

70 Atherton, J. C. *et al.* Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. *J Biol Chem* **270**, 17771-17777, doi:10.1074/jbc.270.30.17771 (1995).

71 Palframan, S. L., Kwok, T. & Gabriel, K. Vacuolating cytotoxin A (VacA), a key toxin for Helicobacter pylori pathogenesis. *Front Cell Infect Microbiol* **2**, 92, doi:10.3389/fcimb.2012.00092 (2012).

72 Iwamoto, H., Czajkowsky, D. M., Cover, T. L., Szabo, G. & Shao, Z. VacA from Helicobacter pylori: a hexameric chloride channel. *FEBS Lett* **450**, 101-104, doi:10.1016/s0014-5793(99)00474-3 (1999).

73 Terebiznik, M. R. *et al.* Helicobacter pylori VacA toxin promotes bacterial intracellular survival in gastric epithelial cells. *Infect Immun* **74**, 6599-6614, doi:10.1128/IAI.01085-06 (2006).

74 Hisatsune, J. *et al.* Molecular characterization of Helicobacter pylori VacA induction of IL-8 in U937 cells reveals a prominent role for p38MAPK in activating transcription factor-2, cAMP response element binding protein, and NF-kappaB activation. *J Immunol* **180**, 5017-5027, doi:10.4049/jimmunol.180.7.5017 (2008).

75 Roesler, B. M., Rabelo-Goncalves, E. M. & Zeitune, J. M. Virulence Factors of Helicobacter pylori: A Review. *Clin Med Insights Gastroenterol* **7**, 9-17, doi:10.4137/CGast.S13760 (2014).

76 Vannini, A., Roncarati, D. & Danielli, A. The cag-pathogenicity island encoded CncR1 sRNA oppositely modulates Helicobacter pylori motility and adhesion to host cells. *Cell Mol Life Sci* **73**, 3151-3168, doi:10.1007/s00018-016-2151-z (2016).

77 Saadat, I. *et al.* Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. *Nature* **447**, 330-333, doi:10.1038/nature05765 (2007).

78 Bessede, E. *et al.* Helicobacter pylori generates cells with cancer stem cell properties via epithelial-mesenchymal transition-like changes. *Oncogene* **33**, 4123-4131, doi:10.1038/onc.2013.380 (2014).

79 Kodaman, N. *et al.* Human and Helicobacter pylori coevolution shapes the risk of gastric disease. *Proc Natl Acad Sci U S A* **111**, 1455-1460, doi:10.1073/pnas.1318093111 (2014).

80 Kim, S. S., Ruiz, V. E., Carroll, J. D. & Moss, S. F. Helicobacter pylori in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Lett* **305**, 228-238, doi:10.1016/j.canlet.2010.07.014 (2011).

81 Kim, I. J. & Blanke, S. R. Remodeling the host environment: modulation of the gastric epithelium by the Helicobacter pylori vacuolating toxin (VacA). *Front Cell Infect Microbiol* **2**, 37, doi:10.3389/fcimb.2012.00037 (2012).

82 Hennig, E. E., Mernaugh, R., Edl, J., Cao, P. & Cover, T. L. Heterogeneity among Helicobacter pylori Strains in Expression of the Outer Membrane Protein BabA. *Infection and Immunity* **72**, 3429-3435, doi:10.1128/iai.72.6.3429-3435.2004 (2004).

83 Papenfort, K. & Vogel, J. Regulatory RNA in bacterial pathogens. *Cell Host Microbe* **8**, 116-127, doi:10.1016/j.chom.2010.06.008 (2010).

84 Pernitzsch, S. R., Tirier, S. M., Beier, D. & Sharma, C. M. A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis receptor in Helicobacter pylori. *Proc Natl Acad Sci U S A* **111**, E501-510, doi:10.1073/pnas.1315152111 (2014).

85 Wen, Y., Feng, J., Scott, D. R., Marcus, E. A. & Sachs, G. A cis-encoded antisense small RNA regulated by the HP0165-HP0166 two-component system controls expression of ureB in Helicobacter pylori. *J Bacteriol* **193**, 40-51, doi:10.1128/JB.00800-10 (2011).

86 Waters, L. S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136**, 615-628, doi:10.1016/j.cell.2009.01.043 (2009).

87 Storz, G., Vogel, J. & Wassarman, K. M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43**, 880-891, doi:10.1016/j.molcel.2011.08.022 (2011).

88 Guillet, J., Hallier, M. & Felden, B. Emerging functions for the Staphylococcus aureus RNome. *PLoS Pathog* **9**, e1003767, doi:10.1371/journal.ppat.1003767 (2013).

89 Grundy, F. J. & Henkin, T. M. From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit Rev Biochem Mol Biol* **41**, 329-338, doi:10.1080/10409230600914294 (2006).

90 Wagner, E. G. H. & Romby, P. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv Genet* **90**, 133-208, doi:10.1016/bs.adgen.2015.05.001 (2015).

91 Rieder, R., Reinhardt, R., Sharma, C. & Vogel, J. Experimental tools to identify RNA-protein interactions in Helicobacter pylori. *RNA Biol* **9**, 520-531, doi:10.4161/rna.20331 (2012).

92 Sittka, A., Pfeiffer, V., Tedin, K. & Vogel, J. The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol Microbiol* **63**, 193-217, doi:10.1111/j.1365-2958.2006.05489.x (2007).

93 Tjaden, B. TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* **36**, W109-113, doi:10.1093/nar/gkn264 (2008).

94 Kery, M. B., Feldman, M., Livny, J. & Tjaden, B. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* **42**, W124-129, doi:10.1093/nar/gku317 (2014).

95 Ivain, L. *et al.* An in vivo reporter assay for sRNA-directed gene control in Gram-positive bacteria: identifying a novel sRNA target in Staphylococcus aureus. *Nucleic Acids Res* **45**, 4994-5007, doi:10.1093/nar/gkx190 (2017).

96 Urban, J. H. & Vogel, J. Translational control and target recognition by Escherichia coli small RNAs in vivo. *Nucleic Acids Res* **35**, 1018-1037, doi:10.1093/nar/gkl1040 (2007).

97 Corcoran, C. P. *et al.* Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* **84**, 428-445, doi:10.1111/j.1365-2958.2012.08031.x (2012).

98 Camilo, V., Sugiyama, T. & Touati, E. Pathogenesis of Helicobacter pylori infection. *Helicobacter* **22 Suppl 1**, doi:10.1111/hel.12405 (2017).

99 Livny, J., Brencic, A., Lory, S. & Waldor, M. K. Identification of 17 Pseudomonas aeruginosa sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Res* **34**, 3484-3493, doi:10.1093/nar/gkl453 (2006).

100 Xiao, B. *et al.* Identification of small noncoding RNAs in Helicobacter pylori by a bioinformatics-based approach. *Curr Microbiol* **58**, 258-263, doi:10.1007/s00284-008-9318-2 (2009).

101 Xiao, B. *et al.* Screening and identification of natural antisense transcripts in Helicobacter pylori by a novel approach based on RNase I protection assay. *Mol Biol Rep* **36**, 1853-1858, doi:10.1007/s11033-008-9390-5 (2009).

102 Wen, Y., Feng, J. & Sachs, G. Helicobacter pylori 5'ureB-sRNA, a cis-encoded antisense small RNA, negatively regulates ureAB expression by transcription termination. *J Bacteriol* **195**, 444-452, doi:10.1128/JB.01022-12 (2013).

103 Kinoshita-Daitoku, R. *et al.* Helicobacter small RNA regulates host adaptation and carcinogenesis. *bioRxIV*, doi:10.1101/2020.02.15.950279 (2020, PREPRINT).

104 Ta, L. H. *et al.* Conserved transcriptional unit organization of the cag pathogenicity island among Helicobacter pylori strains. *Front Cell Infect Microbiol* **2**, 46, doi:10.3389/fcimb.2012.00046 (2012).
105 Meysman, P. *et al.* Structural properties of prokaryotic promoter regions correlate with functional features. *PLoS One* **9**, e88717, doi:10.1371/journal.pone.0088717 (2014).
106 e Silva, S. d. A. & Echeverrigaray, S. in *Bioinformatics* Ch. Chapter 10, (2012).
107 Bhattacharyya, S., Go, M. F., Dunn, B. E. & Phadnis, S. H. in *Helicobacter pylori: Physiology and Genetics* (eds H. L. T. Mobley, G. L. Mendz, & S. L. Hazell) (2001).
108 Munch, R. *et al.* Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**, 4187-4189, doi:10.1093/bioinformatics/bti635 (2005).
109 Forsyth, M. H. & Cover, T. L. Mutational analysis of the vacA promoter provides insight into gene transcription in Helicobacter pylori. *J Bacteriol* **181**, 2261-2266 (1999).
110 Odenbreit, S., Wieland, B. & Haas, R. Cloning and genetic characterization of Helicobacter pylori catalase and construction of a catalase-deficient mutant strain. *J Bacteriol* **178**, 6960-6967, doi:10.1128/jb.178.23.6960-6967.1996 (1996).
111 Spohn, G. & Scarlato, V. Motility of Helicobacter pylori is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. *J Bacteriol* **181**, 593-599 (1999).
112 Beier, D., Spohn, G., Rappuoli, R. & Scarlato, V. Functional analysis of the Helicobacter pylori principal sigma subunit of RNA polymerase reveals that the spacer region is important for efficient transcription. *Mol Microbiol* **30**, 121-134, doi:10.1046/j.1365-2958.1998.01043.x (1998).
113 Shirai, M., Fujinaga, R., Akada, J. K. & Nakazawa, T. Activation of Helicobacter pylori ureA promoter by a hybrid Escherichia coli-H. pylori rpoD gene in E. coli. *Gene* **239**, 351-359, doi:10.1016/s0378-1119(99)00389-3 (1999).
114 Suerbaum, S., Josenhans, C. & Labigne, A. Cloning and genetic characterization of the Helicobacter pylori and Helicobacter mustelae flaB flagellin genes and construction of H. pylori flaA- and flaB-negative mutants by electroporation-mediated allelic exchange. *J Bacteriol* **175**, 3278-3288, doi:10.1128/jb.175.11.3278-3288.1993 (1993).
115 Jones, A. C. *et al.* A flagellar sheath protein of Helicobacter pylori is identical to HpaA, a putative N-acetylneuraminyllactose-binding hemagglutinin, but is not an adhesin for AGS cells. *J Bacteriol* **179**, 5643-5647, doi:10.1128/jb.179.17.5643-5647.1997 (1997).
116 Ge, Z. & Taylor, D. E. Helicobacter pylori genes hpcopA and hpcopP constitute a cop operon involved in copper export. *FEMS Microbiol Lett* **145**, 181-188, doi:10.1111/j.1574-6968.1996.tb08575.x (1996).
117 Heuermann, D. & Haas, R. Genetic organization of a small cryptic plasmid of Helicobacter pylori. *Gene* **165**, 17-24, doi:10.1016/0378-1119(95)00469-m (1995).
118 Pesci, E. C. & Pickett, C. L. Genetic organization and enzymatic activity of a superoxide dismutase from the microaerophilic human pathogen, Helicobacter pylori. *Gene* **143**, 111-116, doi:10.1016/0378-1119(94)90614-9 (1994).
119 Spohn, G., Beier, D., Rappuoli, R. & Scarlato, V. Transcriptional analysis of the divergent cagAB genes encoded by the pathogenicity island of Helicobacter pylori. *Mol Microbiol* **26**, 361-372, doi:10.1046/j.1365-2958.1997.5831949.x (1997).
120 McGowan, C. C., Necheva, A. S., Forsyth, M. H., Cover, T. L. & Blaser, M. J. Promoter analysis of Helicobacter pylori genes with enhanced expression at low pH. *Mol Microbiol* **48**, 1225-1239, doi:10.1046/j.1365-2958.2003.03500.x (2003).

121 Niehus, E. *et al.* Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of Helicobacter pylori. *Mol Microbiol* **52**, 947-961, doi:10.1111/j.1365-2958.2004.04006.x (2004).
122 Josenhans, C. *et al.* Functional characterization of the antagonistic flagellar late regulators FliA and FlgM of Helicobacter pylori and their effects on the H. pylori transcriptome. *Mol Microbiol* **43**, 307-322, doi:10.1046/j.1365-2958.2002.02765.x (2002).
123 Leying, H., Suerbaum, S., Geis, G. & Haas, R. Cloning and genetic characterization of a Helicobacter pylori flagellin gene. *Mol Microbiol* **6**, 2863-2874, doi:10.1111/j.1365-2958.1992.tb01466.x (1992).
124 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402, doi:10.1093/nar/25.17.3389 (1997).
125 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).
126 Castillo, A. R., Arevalo, S. S., Woodruff, A. J. & Ottemann, K. M. Experimental analysis of Helicobacter pylori transcriptional terminators suggests this microbe uses both intrinsic and factor-dependent termination. *Mol Microbiol* **67**, 155-170, doi:10.1111/j.1365-2958.2007.06033.x (2008).
127 Lesnik, E. A. *et al.* Prediction of rho-independent transcriptional terminators in Escherichia coli. *Nucleic Acids Res* **29**, 3583-3594, doi:10.1093/nar/29.17.3583 (2001).
128 d'Aubenton Carafa, Y., Brody, E. & Thermes, C. Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* **216**, 835-858, doi:10.1016/s0022-2836(99)80005-9 (1990).
129 Peters, J. M., Vangeloff, A. D. & Landick, R. Bacterial transcription terminators: the RNA 3'-end chronicles. *J Mol Biol* **412**, 793-813, doi:10.1016/j.jmb.2011.03.036 (2011).
130 Di Salvo, M., Puccio, S., Peano, C., Lacour, S. & Alifano, P. RhoTermPredict: an algorithm for predicting Rho-dependent transcription terminators based on Escherichia coli, Bacillus subtilis and Salmonella enterica databases. *BMC Bioinformatics* **20**, 117, doi:10.1186/s12859-019-2704-x (2019).
131 Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26, doi:10.1186/1748-7188-6-26 (2011).
132 Kumari, R., Shariq, M., Kumar, N. & Mukhopadhyay, G. Biochemical characterization of the Helicobacter pylori Cag-type IV secretion system unique component CagU. *FEBS Lett* **591**, 500-512, doi:10.1002/1873-3468.12564 (2017).
133 Johnson, E. M., Gaddy, J. A., Voss, B. J., Hennig, E. E. & Cover, T. L. Genes required for assembly of pili associated with the Helicobacter pylori cag type IV secretion system. *Infect Immun* **82**, 3457-3470, doi:10.1128/IAI.01640-14 (2014).
134 Cover, T. L. Role of Helicobacter pylori CagL in modulating gastrin expression. *Gut* **61**, 965-966, doi:10.1136/gutjnl-2012-302142 (2012).
135 Le, M. T. *et al.* Conservation of sigma28-Dependent Non-Coding RNA Paralogs and Predicted sigma54-Dependent Targets in Thermophilic Campylobacter Species. *PLoS One* **10**, e0141627, doi:10.1371/journal.pone.0141627 (2015).
136 Ewald, D. R. & Sumner, S. C. Blood type biochemistry and human disease. *Wiley Interdiscip Rev Syst Biol Med* **8**, 517-535, doi:10.1002/wsbm.1355 (2016).

# APPENDIX

**Table 11.** Relevant Interleukins and Their Functions in the Human Immune System Mentioned in this Paper.

| Name | Target cells | Function |
|---|---|---|
| IL-1 | T helper cells | Co-stimulation |
| | B cells | Maturation & Proliferation |
| | NK cells | Activation |
| | Macrophages, Endothelium, other | Inflammation<br>Tiny amounts induce acute phase reaction<br>Substantial amounts induce fever |
| IL-8 | Neutrophils, Basophils, Lymphocytes | Neutrophil chemotaxis |
| IL-10 | Macrophages | Cytokine production |
| | B cells | Activation |
| | Mast cells | |
| | Th1 cells | Inhibits Th1 cytokine production (IFN-γ, TNF-β, IL-2) |
| | Th2 cells | Stimulation |
| IL-12 | Activated T cells | Differentiation into Cytotoxic T cells with IL-2, ↑ IFN-γ, TNF-α, ↓ IL-10 |
| | NK cells | ↑ IFN-γ, TNF-α |

**Table 12.** Blood Groups and Disease Risks[136]

| Disease | Risk Factor | Blood Group/Antigens |
|---|---|---|
| Sickle cell anemia | Increased adhesion | Adhesion Molecules |
| Hemolytic disease of the newborn | Antibodies to RhD | RhD |
| Chronic and autoimmune hemolytic anemias | Rh null | Rh, RhAG |
| Vascular disorders, venous and arterial thromboembolism, coronary heart disease, ischemic stroke, myocardial infarction | Reduced clearance of von Willebrand factor and FVIII | Groups A > AB > B |
| Dementia, cognitive impairment | Coagulation factors | Groups AB > B > A |
| Plague, cholera, tuberculosis, mumps | Antigen profile | Group O |
| Smallpox, Pseudomonas aeruginosa | Antigen profile | Group A |
| Gonorrhea, tuberculosis, S. pneumoniae, E. coli, salmonella | Antigen profile | Group B |

| | | |
|---|---|---|
| Smallpox, E. coli, salmonella | Antigen profile | Group AB |
| N. meningitides, H. influenza, C. albicans, S. pneumoniae, E. coli urinary tract infections, S. pyogenes, V. cholera | Antigen profile | Non-secretors |
| H. pylori | Strain-dependent | Group A; 95% non-O |
| Peptic ulcers, gastroduodenal disease | Secretor status, H. pylori strain | All non-secretors; Group O |
| Norovirus | Strain-dependent | Secretors; groups O, A |
| P. falciparum malaria | Receptor/antigen profile | Knops antigens; groups A, B |
| P. vivax malaria | Antigen profile | Duffy FY antigens |
| Cholera | Severity differs by antigen profile | Lewis antigen; non-secretors; non-O groups |
| Bacterial Meningitis (N. meningitidis, H. influenza, S. pneumoniae) | Antigen profile | Non-secretors; A, AB, O blood groups |
| Cancer (tissue specific) | Increased tumor antigens and ligands | A, B, H antigens lost; "A- like" antigens gained |
| Leukemia and Lymphoma | RBC membrane changes | A, B, H antigens lost |
| Non-Hodgkin's central nervous system lymphoma (primary and secondary) | | Group O, B |
| Hodgkin's lymphoma | | Group B |
| Acute lymphoblastic leukemia | | Group O |
| Acute myeloid leukemia | | Group A |
| Stomach Cancer | H. pylori strain | Group A |
| Pancreatic Cancer | H. pylori strain | Group B > AB > A |
| Von Hippel-Lindau and Neuroendocrine | Multiple tumors | Group O |
| Multiple Endocrine Neoplasia Type 1 | Strongly associated | Group O |
| Colon/Rectum Cancer | Type 1 and 2 chains; Lewis antigens | Secretors; "A-like" antigens expressed |
| Hypertension | 3 phenotypes differ | Group B>A>AB |
| Hyperlipidemia | Low fat diet Ineffective; Intestinal ALP and apoB-48 vary by secretor status | LDL: Heterozygous MN; Group A, B; ALP/apoB-48: Group O and B secretors |
| Type 2 Diabetes | Rh group modifies | Group AB > B > A |
| Type 1 Diabetes | *FUT2* gene locus | Non-secretors |

**Table 13.** *cag*PAI Gene Names[104]

| **26695 ORF** | **Gene Number** | **Gene Name** |
|---|---|---|
| HP0520 | *cag1* | *cagC* |
| HP0521 | *cag2*/Hypothetical | Hypothetical |
| HP0522 | *cag3* | *cagA* |
| HP0523 | *cag4* | *cagγ* |
| HP0524 | *cag5* | *cagβ* |
| HP0525 | - | *cagα* |
| HP0526 | *cag6* | *cagZ* |
| HP0527 | *cag7* | *cagY* |
| HP0528 | *cag8* | *cagX* |
| HP0529 | *cag9* | *cagV* |
| HP0530 | *cag10* | *cagW* |
| HP0531 | *cag11* | *cagU* |
| HP0532 | *cag12* | *cagT* |
| HP0533 | Hypothetical | Hypothetical |
| HP0534 | *cag13* | *cagS* |
| HP0535 | *cag14* | *cagQ* |
| HP0536 | *cag15* | *cagP* |
| HP0537 | *cag16* | *cagM* |
| HP0538 | *cag17* | *cagN* |
| HP0539 | *cag18* | *cagL* |
| HP0540 | *cag19* | *cagI* |

| | | |
|---|---|---|
| HP0541 | *cag20* | *cagH* |
| HP0542 | *cag21* | *cagG* |
| HP0543 | *cag22* | *cagF* |
| HP0544 | *cag23* | *cagE* |
| HP0545 | *cag24* | *cagD* |
| HP0546 | *cag15* | *cagC* |
| HP0547 | *cag26* | *cagA* |

**Table 14.** SuperScript IV RT-PCR Temps

| ID | Sequence | Tm | Ta |
|---|---|---|---|
| PIX-F | CACTCTTGCCTATAAAGGCC | 57.4 | 61.2 |
| PIX-R new | TATCAGTGAAATCGCTCTTAAGAACA | 59.8 | |
| PIX-F | CACTCTTGCCTATAAAGGCC | 57.4 | 57.8 |
| PIX-R | GAAACTGCTAAGAATATCAGTG | 53.8 | |
| PIX-F | CACTCTTGCCTATAAAGGCC | 57.4 | 61.2 |
| PIX-R2 | AGATGTCGCAGAAACTGCTA | 58.5 | |
| PIX-F | CACTCTTGCCTATAAAGGCC | 57.4 | 54 |
| PIX-R4 | GAAAGCAATAAGATAGACAA | 49.7 | |
| PIX-F | CACTCTTGCCTATAAAGGCC | 57.4 | 61.2 |
| PIX-R9 | GTGGCAATAAAAAAGATGACGATAAG | 58.5 | |
| PIX-R new | TATCAGTGAAATCGCTCTTAAGAACA | 59.8 | 63.3 |
| PIX-F2 | AATGTGGGATCACTCTTGCC | 56.9 | |
| PIX-R new | TATCAGTGAAATCGCTCTTAAGAACA | 59.8 | 61.9 |

| ID | Sequence | Tm | Ta |
|---|---|---|---|
| PIX-F4 | GGGTGTATTGCGCTAAAACA | 58.1 | |
| PIX-R new | TATCAGTGAAATCGCTCTTAAGAACA | 59.8 | 55.8 |
| PIX-F8 | TATACCTGTGCTACTTTAGA | 51.6 | |

Blood Agar Plates for *Helicobacter pylori* growth *(from Antonello Covarcci, 11/95)*

1. Dissolve correct mass of Columbia Agar in $dH_20$ to 1 liter.
2. Autoclave on liquid cycle
3. WHILE AUTOCLAVING: Mix fresh ß Cyclodextrin.
4. Add ß-Cyclodextrin at 1 mL per 100 mL agar
5. *ß*-Cyclodextrin should be 200 mg/ml in DMSO
6. Cool media to about 55˚C
7. Mix blood and add 5 ml blood per 100 ml agar.
8. Add *ß-Cyclodextrin* 1 mL per 100 mL agar
9. Swirl gently and pour

LB Plates (Chloramphenicol and Ampicillin)

1. Dissolve correct mass of Luria Bertani broth (check container) in dH20 to 1 liter.
2. Add 25g LB powder to a flask
3. Mix to resuspend powder
4. Autoclave
5. Add either 20µg/ml Chloramphenicol or 100µg/ml Ampicillin
6. Gently pour plates

2% Agarose Gel

1. 50ml 1% TAE Buffer
2. 1g Agarose
3. Microwave for ~2min
4. Pour gel with a comb
5. Sit at room temperature for 15min
6. Cover gel with 1% TAE in the tray and run sample
7. When complete place gel in a tray covered with dH2O
8. Add 8µ EtBr to tray dH20
9. Place tray on a rocker for 15min
10. Image with UV light

# VITA

Author: Brandon M. Flatgard

Place of Birth: Windom, Minnesota

Undergraduate Schools Attended: Washington State University

Spokane Community College

Gonzaga University

Degrees Awarded: Bachelor of Arts, 2015, Gonzaga University

Associate of Arts, 2013, Spokane Community College

Honors and Awards: Graduate Assistantship, Ronald E. McNair Scholar Program, 2019-2020, Eastern Washington University

Travel Grant for presentation at AMS Microbe National Conference, Chicago, Illinois, 2020 (Canceled due to COVID-19)

EWU Research Grant, 2019, Eastern Washington University

Gonzaga Grant, 2013, Gonzaga University

Dussault Scholarship for Distinguished Academic Achievement, 2013, Gonzaga University

Phi Theta Kappa Scholarship, 2013, Gonzaga University

Professional Experience: Graduate Teaching Assistant, Eastern Washington University, Cheney, WA, 2019

Certified Clinical Hemodialysis Technician / Regulator and Compliance, Fresenius Medical Care, Spokane, Washington, 2010-2019

DVR/RMA Technician / Technical Support, PCO Inc., Liberty Lake, Washington, 2007-2010