Helicobacter pylori Population Dynamics *in vitro* and *in vivo*

By

Margaret Lorena Harvey

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbe-Host Interactions

January 31, 2022

Nashville, Tennessee

Approved:

Maria Hadjifrangiskou, Ph.D., Chair

James E. Cassat, M.D., Ph.D.

Keith T. Wilson, M.D.

Maureen A. Gannon, Ph.D.

Timothy L. Cover, M.D.

DEDICATION

For my late father, Bill Harvey – I love you mucho!

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PI, Tim Cover, for being such a fantastic mentor. Thank you for fostering such a supportive lab environment in which I could grow, not only as a scientist, but also as an individual. You've taught me to keep a positive outlook in any situation; arguably, one of the most valuable lessons in life. Thank you for everything! And as if I wasn't already lucky enough to have one amazing mentor, I was lucky enough to have three! Mark and John, thank you both for all your guidance and support. Sai, thank you for your friendship and for always being there for me! I'm so grateful we joined the lab together. To all my lab-mates, our daily banter, afternoon coffee breaks, and lunchroom talks have always been the highlight of my day. Thank you for always keeping a bubbly attitude even on rainy days. I will miss you all dearly!

Finally, I would like to thank my family. Dad, you were always my biggest supporter. Whether I wanted to become a ballerina, an artist, or a mad scientist, you were always finding new ways to help me pursue my passions. You taught me that to get ahead in life I must be strong and determined, an indispensable lesson in my pursuit of a Ph.D. Although you couldn't be here to see me earn my Ph.D., I know you and mom are doing your happy dance up in heaven. To my friends and family, thank you for believing in me! From sharing silly memes to beautiful memories, I want to thank you for supporting me throughout this Ph.D. journey. And finally, I'd like to thank my partner in crime, Cisco. You have helped me overcome so many obstacles and have been one of my biggest supporters. From late night tacos nights to sporadic waterfall hikes, you've always known just how to cheer me up. Thank you for everything!

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

OMPs	Outer Membrane Proteins
H. pylori	*Helicobacter pylori*
WHO	World Health Organization
CagA	Cytotoxin-associated gene A
VacA	vacuolating cytotoxin A
BabA	blood group antigen binding adhesin A
IL-1β	Interleukin-1 beta
TNF-α	Tumor necrosis factor alpha
IL-10	Interleukin-10
MATT	Microarray-Based Tracking of Transposon Mutants
STM	Signature Tagged Mutagenesis
Cy-3	Cyanine-3
Cy-5	Cyanine-5
NH_4^+	Ammonium
CO_2	Carbon dioxide
LPS	Lipopolysaccharide
BAM	Barrel-assembly machinery
Hop	Outer membrane porins
Hor	Hop-related
Hof	*H. pylori* OMP
Hom	*Helicobacter* outer membrane
DL1	Diversity loop 1
DL2	Diversity loop 2
CL2	Conserved loop 2
PUD	Peptic ulcer disease
GC	Gastric cancer
SabA	Sialic acid-binding adhesin
BabA	Blood group antigen-binding adhesin
OipA	Outer inflammatory protein A
SSM	Slipped strand mispairing
IL-8	Interleukin-8
AlpA	Adherence-associated lipoprotein A
CEACAM 1	Carcinoembryonic antigen-related cell adhesion molecule 1
CEACAM 3	Carcinoembryonic antigen-related cell adhesion molecule 3
CEACAM 5	Carcinoembryonic antigen-related cell adhesion molecule 5
CEACAM 6	Carcinoembryonic antigen-related cell adhesion molecule 6
NF-kB	Nuclear factor kappa B
MUC5AC	Mucin 5AC
LabA	LacdiNAc-binding adhesin
CFU	Colony Forming Unit
NGS	Next-generation sequencing

CHAPTER I

INTRODUCTION

Helicobacter pylori Prevalence and Disease

Helicobacter pylori is a Gram-negative bacterium that colonizes the gastric mucosa in about 50% of the world's population. The acquisition of *H. pylori* occurs most often during childhood (1). The mechanisms by which *H. pylori* is transmitted and acquired are incompletely understood. Humans are the only known reservoir for *H. pylori*, and it is proposed that transmission occurs through both oral-to-oral and fecal-to-oral routes. The prevalence of *H. pylori* infection is associated with factors such as age and socioeconomic status (2-4), and *H. pylori* infections are more prevalent in developing countries than in developed countries (5, 6).

H. pylori colonization of the stomach is a strong risk factor for the development of peptic ulcer disease and gastric cancer. As such, *H. pylori* has been classified as a Class I carcinogen by the World Health Organization (WHO) (7, 8). Gastric cancer is the fourth leading cause of cancer-related death worldwide, accounting for about 0.8 million annual fatalities (9). The incidence of *H. pylori*-related gastric cancer varies markedly throughout the world. In many regions of the world, rates of *H. pylori* infection and gastric cancer are concordant. Conversely, the prevalence of *H. pylori* infection in some African countries can be as high as 91% but these countries have a relatively low incidence of gastric cancer. This phenomenon is known as the "African Enigma" (10-12).

The factors that determine whether *H. pylori* infection remains asymptomatic or progresses into peptic ulcer disease or gastric cancer are not completely understood.

However, it is clear that disease progression is influenced by multiple factors, including characteristics of *H. pylori* strains, host genetic variation, and environmental exposures. For example, individuals infected with strains of *H. pylori* that produce CagA (a secreted effector protein encoded by *cagA*, cytotoxin-associated gene A), s1 forms of VacA (vacuolating cytotoxin A, a secreted toxin), or BabA (blood group antigen binding adhesin A, an outer membrane protein) have an increased risk of developing gastric disease compared to individuals infected with *H. pylori* strains that do not produce these proteins. Similarly, infected individuals who have variant forms of IL-1β, IL-10, and TNF-α genes have an increased risk of gastric cancer. Environmental exposures, such as smoking and the consumption of a high-salt diet, are also important risk factors for gastric cancer.

General information about the *H. pylori* outer membrane

A defining feature of Gram-negative bacteria is the presence of an outer membrane. The Gram-negative outer membrane is an asymmetrical bilayer with phospholipids in the inner leaflet and lipopolysaccharide (LPS) in the outer leaflet (13). LPS has three distinct domains: a lipid A anchor, a core oligosaccharide, and an O-antigen polysaccharide. Unlike most Gram-negative bacteria, which produce a hexa-acylated form of lipid A, *H. pylori* produces a modified tetra-acylated lipid A molecule. This alteration confers resistance to antimicrobial peptides and allows the bacteria to evade detection by TLR4 (14, 15).

The outer membrane is further distinguished by the presence of β-barrel outer membrane proteins (OMPs) (16-19). Generally, β-barrel OMPs are composed of eight or more anti-parallel β-strands that come together to form a β-barrel structure (20, 21). These β-barrels are folded and integrated into the outer membrane by the barrel-

assembly machinery (BAM) complex (22). β-barrel OMPs can have many functions, including roles as active/passive transporters, enzymes, receptors, or adhesins (20, 21, 23). Approximately 4% of genes in the *H. pylori* genome (64 genes in the prototype strain 26695) are predicted to encode outer membrane proteins (OMPs). The exact number of OMP-encoding genes varies among different *H. pylori* strains. These OMPs have been classified into several families, including the Hop (outer membrane porins), Hor (Hop-related), Hof (*H. pylori* OMP), and Hom (*Helicobacter* outer membrane) families, as well as iron-regulated OMPs and efflux pump OMPs. Phase variation is commonly used to regulate the production of *H. pylori* OMPs (21, 24, 25). Some OMPs are also regulated by two-component systems, the ferric uptake regulator (Fur), or other transcriptional regulatory systems. For the purposes of this study, we focused on the importance of members in the Hop family in *H. pylori* fitness *in vivo,* but not *in vitro.*

The Hop family of outer membrane proteins in *H. pylori*

In *H. pylori* strain 26695, 21 OMPs belong to the Hop family. All Hops have a highly conserved N-terminal motif (A↓EX[D,N]G, where the ↓ represents the signal sequence cleavage site) and are structurally characterized by a novel type 5-like interrupted β-barrel composed of 7 antiparallel β-strands at the C-terminus and one at the N-terminus (26). Conserved sequences at the 5’ and 3’ ends of the genes are sites of homologous recombination among different Hop-encoding genes (27, 28).

Hops are classified into two subgroups, F-Hops and Y-Hops, based on the identity of the C-terminal residue. Hop proteins that end with the characteristic phenylalanine residue are termed F-Hop proteins and those that end with tyrosine are termed Y-Hop proteins (29). Members of the Y-Hop subgroup share remarkable identity and are tightly

clustered on the phylogenetic tree, whereas members of the F-Hop subgroup display a higher level of diversity (29).

The extracellular variable domains located between the N-terminal and C-terminal β-strands display a relatively high level of sequence divergence compared to non-exposed regions and therefore are presumed to be determinants of each Hop protein's distinct functions. The extracellular Hop domains have a predominantly α-helical structure. One study identified three monophyletic groups with distinct architecture in the variable domains via structural and phylogenetic analysis (26). The first group, termed the Hop-I subfamily, is characterized by an extracellular region with both a head domain and coiled-coil stem domain. This group includes many Hops with predicted adhesive functions, such as SabA, BabA, and HopQ. Specifically, x-ray structures showed that BabA, SabA and HopQ ectodomains form a 3+4 helix bundle 'head' domain connected to a coiled-coil 'stem' domain (26, 30-33). The second (Hop-II) and third (Hop-III) monophyletic groups are highly divergent from the conserved topology seen in the Hop-I subfamily, lacking both the Hop-I head domain and coiled-coil stem domain.

Outer membrane protein-mediated adhesion

Adhesion is predicted to play a pivotal role in *H. pylori* colonization. *H. pylori* can potentially use a variety of adhesins to bind target receptors within the gastric mucosa. Therefore, multiple *H. pylori* Hops with adhesin activity may have redundant functions. Most *H. pylori* can be found free-swimming in the mucus, but some attach directly to gastric epithelial cells. This interaction is predicted to provide nutrient access to the bacteria, promote delivery of bacterial toxins to the host cells, and prevent mechanical clearance from the stomach.

Expression of several OMPs is regulated in response to changes in gastric pH and inflammation, thereby allowing *H. pylori* to adapt to environmental changes. A pH gradient is present between the gastric epithelial cell layer and the gastric lumen (~pH 6 to pH 2). One study reported that the *H. pylori* BabA adhesin is acid-responsive and BabA-mediated attachment is reversible, which enables *H. pylori* to escape from epithelial cells and mucus that are shed into the acidic lumen (34). *H. pylori* also uses adhesins to adapt to changes in the gastric microenvironment, such as inflammation-mediated changes in glycosylation patterns, to efficiently colonize the gastric mucosa. For instance, BabA binds to Lewis b (Le^b) blood group antigens expressed in glycoproteins from normal gastric tissue (35). *H. pylori* adhesion to gastric epithelial cells enables the translocation of bacterial virulence factors in host cells. This, in turn, stimulates a gastric mucosal inflammatory response. As a result, there is a shift in glycosylation patterns to express increased levels of sialylated glycans, which promotes SabA-mediated *H. pylori* attachment (35-37).

Role of OMPs in *H. pylori* colonization of the stomach.

Bacterial colonization of the gastric environment is a complex process. In order to establish a successful infection, *H. pylori* must be able to survive the acidic pH of the gastric lumen, swim towards the gastric mucosa and localize within the gastric mucus layer. OMPs have been shown to be important colonization factors in multiple bacteria, but little is known about the contribution of OMPs to *H. pylori* colonization of the stomach. As such, it is important to identify OMPs that contribute to *H. pylori* fitness in the gastric environment.

Previous studies have used Signature Tagged Mutagenesis (STM) and Microarray-Based Tracking of Transposon Mutants (MATT) to identify *H. pylori* genes essential for gastric colonization (38, 39). In a previous STM study, gerbils were infected with pools of 24 individually tagged *H. pylori* mutants for 21 days (38). Mutants present in the input pool but absent in the output pool were then identified using a tag-specific PCR method. Both *flaA* (major flagellin A; important for motility) and *ureA* (urease subunit a; important for acid resistance) were identified as essential for gastric colonization (38). Another study used MATT to identify genes that contribute to gastric colonization (39). In that study, mice were infected with pools of 48 *H. pylori* transposon mutants. For each experiment, the input pool was labeled with Cy-3 (green) and the output pool was labeled with Cy-5 (red). Microarray methodology was then used to quantify the red/green ratio for each gene. Numerous genes, including *ureA,* were essential for gastric colonization. The two studies described above, as well as others, suggest that FlaA and UreA are required for *H. pylori* colonization of the stomach (40-44).

Generation of comprehensive transposon mutant libraries is more challenging in *H. pylori* than in several other bacterial species, and the studies conducted thus far to identify *H. pylori* genes required for gastric colonization have had several limitations. The two previous studies described above relied on random transposon mutagenesis for generation of mutants to identify genes essential for gastric colonization. For example, the MATT study tracked 1,378 mutants corresponding to 758 different gene loci. Therefore, it can be presumed that numerous non-essential genes were not evaluated. Similarly, the mutant library analyzed in the previous STM study was probably not comprehensive (44). For example, the MATT study did not include analysis of *oipA, hopF,*

sabA, or *babA* mutants. As such, little progress was made to evaluate the importance of these OMPs, and possibly others, in *H. pylori* fitness *in vivo.* Furthermore, neither study used methods allowing a quantitative determination of relative levels of colonization among the pooled mutant strains. The genetic barcoding approach we used in the current study (discussed in Chapters II through IV) overcomes some of the limitations of previous transposon studies. This approach allowed us to evaluate mutant pools with known composition and to quantify the proportional abundance of individual mutants within a mixed population, thereby assessing whether mutagenesis of target genes leads to a fitness advantage, disadvantage, or neutral phenotype (relative to one or more barcoded control strains). An important goal of my project was to identify OMPs that contribute to *H. pylori* fitness *in vivo,* but not *in vitro.* The following sections provide a detailed description of many of the *H. pylori* OMPs analyzed in the current study, as well as control genes analyzed in the current study (*flaA* and *ureA*).

***H. pylori* Hops with Reported Adhesive Activity**

BabA/HopS

Blood group antigen-binding adhesin (BabA) is a Hop outer membrane protein with an approximate molecular weight of 75-80 kDa. Its structure includes two diversity loops (DL1 and DL2) and one conserved loop (CL2) (32). The BabA adhesin plays a pivotal role in *H. pylori* attachment to H-type 1 and ABO/Lewis b (Le^b) blood group antigens expressed by the human gastric epithelium (45). Previous studies have shown that most humans are blood group antigen [Le(a-b+)] secretors (46). Interestingly, Lewis antigens can also be components of *H. pylori* lipopolysaccharide. This is proposed to be a form of

molecular mimicry to evade host immune detection (47-49). Among *H. pylori* strains, there is substantial variation in *babA* alleles. For example, some *H. pylori* strains contain a *babA* gene and others do not. Some strains contain more than one copy of the *babA* gene. Some BabA proteins bind Lewis b and others do not (27, 50-52).

H. pylori attachment to gastric epithelial cells facilitates the delivery of toxins (such as CagA or VacA) to host cells and can stimulate inflammatory responses (52-56). Moreover, gastric colonization with *H. pylori* strains expressing *babA is* associated with increased disease severity in humans, including peptic ulcer disease (PUD) and gastric cancer (GC), compared to colonization with strains that do not express *babA* (53-60).

Interestingly, several studies have reported that BabA is selected against, either by phase variation or gene conversion, in various hosts, including rhesus macaques, Mongolian gerbils, and mice (51, 52, 61-63). Another study showed that the loss of BabA expression occurs in Le^b transgenic mice, similar to the loss of BabA in wild-type mice (52). To examine the effect of Le^b binding on stability of BabA expression in mice, the investigators generated an *H. pylori* mutant that produced a form of BabA that could not bind Le^b (BabACL2). Wild-type and transgenic mice that express Le^b were infected with wild-type or BabACL2 *H. pylori.* BabA expression was frequently lost in both *H. pylori* strains (wild-type and BabACL2) in both mice backgrounds (wild-type and transgenic mice that express Le^b) (52). Therefore, it was proposed that the loss of BabA expression is not dependent on presence or absence of Le^b nor the ability of BabA to bind Le^b.

SabA/HopP

Sialic acid-binding adhesin (SabA) binds primarily to sialyl-lewis X antigens found on the gastric epithelium. Like BabA, the expression of SabA is regulated in response to

changes in the gastric environment. *H. pylori*-mediated gastric inflammation has been shown to result in upregulated production of sialyl-Le^x antigens in the stomach, which promotes the SabA-mediated attachment to the gastric mucosa. SabA can also mimic selectin to activate host neutrophils, leading to the production of reactive oxygen species, which contributes to a persistent inflammatory response (64).

OipA/HopH

Outer inflammatory protein A (OipA), also known as HopH, has been reported to play a role in *H. pylori* adhesion to KATO-III cells (a human gastric carcinoma cell line)(65). The production of an intact OipA protein is regulated by slipped strand mispairing (SSM) (66). Murine infections with strains expressing *oipA* have also been associated with higher colonization density and severe neutrophil infiltration compared to strains not expressing *oipA* (67). The functional *oipA* "on" genotype has been associated with the capacity of *H. pylori* strains to stimulate increased IL-8 secretion by gastric epithelial cells *in vitro* and increased gastric inflammation *in vivo*. However, other studies reported different results (57, 68, 69). Nevertheless, many studies have shown that in-frame *oipA* is associated with the presence of *cagA* and with an increased risk for developing peptic ulcer disease and gastric cancer (70).

AlpA/B

Two OMPs, designated "adherence-associated lipoproteins" AlpA (HopC, or OMP20, ~56kDa) and AlpB (HopB, or OMP21, ~57 kDa), are encoded by highly homologous genes (*alpA* and *alpB*) within the same operon (71-73). These genes are expressed by most *H. pylori* isolates (57, 74). Both AlpA and AlpB have been shown to

contribute to *H. pylori* adhesion to the gastric mucosa by host laminin binding (75-78), and *Escherichia coli* expressing plasmid-born *alpA* or *alpB* exhibits a laminin-binding phenotype (77). AlpB has also been shown to play a role in biofilm formation (79). Strains lacking AlpA/AlpB have been reported to have defects in the ability to colonize the stomachs of guinea pigs and Mongolian gerbils and to poorly colonize the stomachs of mice (80, 81).

HopQ

HopQ contributes to *H. pylori* adhesion by binding to members of the carcinoembryonic antigen-related cell adhesion molecule family (CEACAM1, CEACAM5, and CEACAM6) expressed on the surface of epithelial cells and granulocytes (33, 82). Two families of *hopQ* alleles have been described (*hopQ-I* and *hopQ-II*) (83). Generally, HopQ-I is more prevalent in CagA- and s1-VacA-positive strains, whereas HopQ-II is more prevalent in CagA-negative s2-VacA-positive strains. *H. pylori* strains harboring CagA and s1-VacA have been associated with worse disease outcome. As such, HopQ-I is associated with an increased risk of peptic ulcer and gastric cancer development, compared to HopQ-II (83-86). HopQ-mediated binding enables CagA translocation and modulates inflammatory signaling (NF-kB activation and IL-8 induction) in host cells (87, 88). HopQ-CEACAM interactions also contribute to gastric colonization in rats that express CEACAM1 (87). Furthermore, a recent study showed that the interaction of HopQ with CEACAM1 expressed on immune cells may inhibit immune cell activities, thereby protecting developing tumors from immune cell attack (89).

HopD/LabA

One study reported that the outer membrane protein HopD (~55 kDa) acts as an *H. pylori* adhesin that binds specifically to the lacdiNAc (GalNAcbeta1-4GlcNAc) motif on the surface of MUC5AC mucins (90). As such, it was renamed LabA (lacdiNAc-binding adhesin). Structurally, LabA shares an L-shaped fold with SabA and BabA (91). The head region contains a 4 + 3 α-helix bundle, with a small insertion domain consisting of a short antiparallel beta sheet and unstructured region (91). LabA has been shown to form pores within a planar lipid bilayer model membrane system (92). One study showed that *H. pylori* co-localizes to lacdiNAc-MUC5AC mucins in gastric tissue samples, and that mutagenesis of *labA* reduced *H. pylori* binding to the surface of gastric mucosa (90). To further evaluate the specific interaction between LabA and lacdiNAc, an *H. pylori* wild-type strain and a *labA* mutant were preincubated with free lacdiNAc, and this led to reduced binding to tissues by a wildtype strain, but not the *labA* mutant (90). Another study did not detect lacdiNAc-mediated adhesion of *H. pylori* strains (93). Therefore, additional studies are needed to elucidate the role of LabA in *H. pylori* adhesion.

HopZ

Another putative *H. pylori* adhesin is HopZ, a 74 kDa protein. One study demonstrated that mutagenesis of *hopZ* leads to reduced binding to gastric epithelial cells, compared to binding of a wild-type strain (67). A host receptor for HopZ has not been identified. Expression of *hopZ* is regulated by slipped strand mispairing, which leads to phenotypic variation between strains (94). The switch status of HopZ has been observed to switch from “off” to “on” during human infection (95). HopZ switch status has also been shown to influence *H. pylori* colonization density in mice (67). However, the role of HopZ in adhesion and colonization *in vivo* is not yet clear. While *hopZ* mutants

were able to colonize guinea pig stomachs (80), they were less fit to colonize the stomachs of gnotobiotic transgenic mice (96).

Urease and Gastric Colonization

Urease, a multi-subunit metalloenzyme, is one of the most abundantly expressed proteins in *H. pylori.* Seven genes (*ureABIEFGH*) are necessary for the synthesis of an active urease enzyme. The urease enzyme is composed of two proteins, UreA (26.5 kDa) and UreB (60.3 kDa). Various accessory proteins interact with the apoenzyme and deliver nickel ions to the active site in an energy-dependent process. Functionally, urease hydrolyzes urea to ammonium (NH_4^+) and carbon dioxide (CO_2).

The ability of *H. pylori* to survive the low pH of the stomach might suggest that it is an extreme acidophile. However, that is not the case. In fact, *H. pylori* is a neutrophile that utilizes its urease activity to tolerate the acidity of the stomach environment when exposed to low pH in the gastric lumen (97). The importance of urease as a colonization factor in *H. pylori* is supported by numerous studies demonstrating that urease-negative mutants fail to colonize the gastric mucosa as efficiently as urease-positive *H. pylori* strains. This was first shown in studies of gnotobiotic piglets infected with urease-negative or urease-positive strains (40, 42-44). The urease-negative strain did not colonize any of the piglets, whereas the urease-positive strain colonized all the piglets and caused gastritis. Since urease has been shown to act as an important colonization factor, we included a *ureA* mutant as a control to assess whether our genetic barcoding approach could be used to identify mutants with known fitness disadvantages *in vivo.*

Flagella and Gastric Colonization

H. pylori flagella are required for *H. pylori* colonization of the gastric mucosa (98). Each bacterium possesses 2-8 unipolar flagella that aid in *H. pylori* motility. This rotatory nanomachine consists of three major components: the basal body, hook, and filament (99). The basal body is composed of several proteins, and it plays a role in providing the energy source for motility. The hook is composed mostly of FlgE, which links the basal body to the flagellar filament (100, 101). The flagellar filament is a copolymer of the major flagellin (FlaA) and the minor flagellin (FlaB), which are important for motility (102).

The importance of motility as a colonization factor in *H. pylori* was first shown in gnotobiotic piglets infected with motile or non-motile *H. pylori* variants (98). Motile *H. pylori* were recovered from piglets at both 6- and 21- days post-infection, whereas non-motile *H. pylori* were only recovered at 6 days post-inoculation. Other animal studies, using motility-deficient mutants, such as *motB* (103), *fliD* (104), *putA* (105), *flaA* and *flaB* (106, 107) mutants, provided further evidence indicating that motility is important for *H. pylori* gastric colonization. In our studies, we included a *flaA* mutant as an additional control to determine if our genetic barcoding approach could be used to identify mutants with known fitness disadvantages *in vivo.*

Animal models

Multiple animal models, including rodents, gnotobiotic piglets, dogs and primates have been used to study *H. pylori* infection. Rhesus macaques are most similar to humans, but large-scale studies with this animal model are impractical because of high cost and limited availability. As such, most researchers use Mongolian gerbils or mice to study *H. pylori* infection.

Similar to humans, Mongolian gerbils can develop ulcers and gastric cancer in response to *H. pylori* infection. Therefore, the gerbil model is especially useful when studying *H. pylori*-induced gastric diseases. However, Mongolian gerbils are an outbred species, which makes it difficult to study host factors that contribute to disease in this model, and relatively few immunologic tools are commercially available for analysis of gerbils.

Unlike gerbils, mice rarely develop ulcers and gastric cancer in response to *H. pylori* infection. On the other hand, *H. pylori* can colonize the mouse stomach and mice develop a gastric mucosal inflammatory response (gastritis). Numerous immunologic tools are available for studies of mice, and mice can be genetically manipulated. As such, mice are useful for studying bacterial genes or proteins that contribute to *H. pylori* colonization of the stomach and host factors that contribute to the gastric mucosal inflammatory response.

Gaps in Knowledge

Previous studies have analyzed adhesive properties of specific Hop proteins, but there has been relatively little effort to systematically assess the contributions of Hop proteins to *H. pylori* fitness. To investigate this topic, we developed methodology that allowed us to track changes in a population of *H. pylori* mutant strains over time. Chapter II discusses the development of this technique for *H. pylori* research. Chapters III and IV discuss the use of this method to evaluate the potential contributions of *hop* genes to *H. pylori* fitness during growth *in vitro* and in colonization of the mouse stomach and to test the hypothesis that specific members of this gene family contribute to *H. pylori* fitness *in vivo* but not *in vitro.*

CHAPTER II

ADAPTATION OF GENETIC BARCODING TO MONITOR CHANGES IN ABUNDANCE OF POOLED *HELICOBACTER PYLORI* MUTANT STRAINS

A modified version was previously published as:

Harvey ML, Lin AS, Sun L, Koyama T, Shuman JHB, Loh JT, Scott Algood HM, Scholz MB, McClain MS, Cover TL. Enhanced fitness of a *Helicobacter pylori babA* mutant in a murine model. Infect Immun. 2021 Jul 26:IAI0072520. doi: 10.1128/IAI.00725-20. Epub ahead of print. PMID: 34310886.

Introduction

Genetic barcoding has commonly been used to study and compare the fitness advantages and disadvantages of mutant strains, but there has been relatively little use of this approach for studies of *H. pylori* (44). Our goal is to adapt genetic barcoding paired with high throughput sequencing to be used in *H. pylori* research. We believe that this will provide a powerful tool to the *H. pylori* community that could be used to study many aspects of *H. pylori* biology, including *H. pylori* population dynamics, analysis of selective pressures that shape the evolution of *H. pylori* populations, and for studying the roles of *H. pylori* genes in various selective environments.

Materials and Methods

Bacterial strains and culture conditions

H. pylori strain J166 [GenBank accession CP007603] was originally isolated from a human with a duodenal ulcer (108, 109). This strain is motile, contains the *cag* pathogenicity island and exhibits Cag T4SS activity. We chose strain J166 for the experiments in this study because it has previously been shown to colonize multiple

animal species (mice, macaques, Mongolian gerbils, and guinea pigs) and achieves a lower colonization density in the mouse stomach compared to other mouse-colonizing strains (e.g., SS1 and PMSS1), which might facilitate the detection of subtle differences in colonization among multiple mutants. To generate a mouse-adapted wild-type J166 strain, we infected a mouse with strain J166 for 3 months. A single colony isolate that retained Cag T4SS activity, VacA activity, and motility was selected for use in this study. Mouse-adapted wild-type J166 and barcoded mutant strains (described below) were grown on trypticase soy agar plates supplemented with 5% sheep blood (TSA with 5% sheep blood) in room air supplemented with 5% CO_2 at 37^{o}C. All mutant strains were generated from mouse-adapted wild-type strain J166 (a single colony output from a 3-month mouse infection) by selective growth on medium containing chloramphenicol (5 μg/ml). *E. coli* strain DH5α, used for plasmid propagation, was grown on Luria-Bertani (LB) agar plates or in LB liquid medium supplemented with ampicillin (100 μg/ml) at 37^{o}C.

Barcode generation

The KAnalyze program was used to compile lists of overlapping 7-nucleotide sequences (7-mers), offset by a single nucleotide (110). A list of all possible 7-mers was compared to a set of 7-mers not present in the genome sequence of *H. pylori* strain J166 (nullomers). Three randomly selected nullomers then were concatenated to generate 21-mer barcodes that would differ from any 21-mer found in *H. pylori* J166 by at least 3 nucleotide positions.

Generation of barcoded *H. pylori* strains

The *mdaB-hydA* intergenic region (corresponding to HP0630-HP0631 in *H. pylori* 26695) was identified in *H. pylori* strain J166 (GenBank accession CP007603) and was used as a neutral site for introducing nucleotide barcodes. A control library was composed of *H. pylori* mutants (LH-1, LH-2, etc.) containing barcodes in the *mdaB-hydA* intergenic region. Similarly, we searched the J166 genome with known *hop* gene sequences from *H. pylori* strain 26695 to identify J166 orthologs with greater than 90% sequence identity. We identified 20 *hop* genes in J166, and successfully mutagenized 16 (Table 2.1). Two pairs of *hop* genes are duplicate copies (*hopJ/K* and *hopM/N*); one copy of each duplicate was mutagenized for this study. If no phenotype (such as a fitness advantage or disadvantage) was detected for these mutants, it could be attributed to the presence of a duplicate gene. Additionally, we did not mutate *alpB* and we did not successfully isolate a *hopZ* mutant. Altogether, the OMP mutant library is composed of 16 *hop* mutants, 1 neutral intergenic mutant (*mdaB-hydA:* a neutral locus), and 2 additional control mutants (*ureA* and *flaA,* previously reported to be essential for gastric colonization) (38-40, 106, 111) (38-40, 106, 111), totaling 19 uniquely barcoded *H. pylori* mutants. Independently generated mutants were obtained by once again transforming the mouse-adapted wild-type strain J166 with the appropriate plasmids. Hop-encoding genes and control genes selected for mutagenesis are shown in Table 2.1.

Barcoded *mdaB-hydA* control strains and *hop* mutant strains were generated by transforming mouse-adapted wild-type *H. pylori* strain J166 with a cassette comprised of a chloramphenicol acetyltransferase gene and unique 21-nucleotide barcode tags specific for each mutant, flanked by 500 base pairs homologous to the targeted insertion

site to facilitate the insertion of this cassette via homologous recombination into the targeted site (Figure 2.1). *H. pylori* Hop proteins have conserved sequences at the N- and C-terminus, and a region of diversity that is predicted to contribute to each Hop's distinct functions (72). Simple insertions of the barcoded cassette were targeted to this region of diversity, approximately 500 base pairs downstream of the transcriptional start site. Following transformation, *H. pylori* transformants were plated on sodium bisulphite-free Brucella (BSFB) agar plates supplemented with cholesterol and containing chloramphenicol (5 µg/ml) (BB5ChlC) (112, 113). Transformants were validated via sequencing of PCR amplicons generated with primers flanking the predicted barcoded cassette insertion site (about 600 base pairs upstream and 600 base pairs downstream of the insertion site).

Growth Curve Analysis

H. pylori strains J166 wild-type and barcoded mutant strains (*alpA, hopE, oipA, babA,* and *flaA* mutants) were grown overnight in sodium bisulphite-free Brucella broth supplemented with 10% FBS in room air supplemented with 5% CO_2 at 37^{o}C. These were sub-cultured at a starting OD_{600} of 0.1 and then measured every 6 hours for 30 hours, with a final measurement taken at 48 hours.

Results

Nucleotide barcoding of H. pylori strains

To evaluate the fitness of *H. pylori* strains under different conditions, we used a genetic barcoding approach (Fig. 2.1). We first generated a "control library" composed of 6 *H. pylori* strains (LH-1 – LH-6), each containing a unique nucleotide barcode inserted

in the *mdaB-hydA* intergenic region, predicted to be a neutral locus (unrelated to bacterial fitness) (114, 115) (Table 2.1). In the initial development of methodology for tracking changes in a population of barcoded strains, we conducted experiments to assess the validity and accuracy of the approach. We pooled the six strains in equal proportions, and we also created pools in which one control strain (LH-1) was reduced in proportional abundance by 100-fold or omitted from the pool. As expected, the experimental detection of barcodes matched the known composition of the pools (Fig. 2.2). Specifically, in an experiment in which the pool was composed of equal proportions of strains, barcodes corresponding to LH-1 accounted for 16.7% of total counts (Fig. 2.2A). When the initial proportion of LH-1 was reduced by 100-fold, barcodes corresponding to LH-1 accounted for 0.22% of the total counts, with barcodes for other strains increasing in proportional abundance (Fig. 2.2B). Finally, when LH-1 was excluded from the pool, we no longer detected barcodes from this mutant (Fig. 2.2C). These data indicate that genetic barcoding combined with next-generation sequencing (NGS) is a sensitive and quantitative approach that can be used to track changes in the proportional abundance of individual mutants within a pool.

Overall, Chapter II describes the development and optimization of genetic barcoding paired with next generation sequencing. This approach will be used in the following chapters to investigate *H. pylori* population dynamics and to investigate the role of Hops in *H. pylori* fitness *in vitro* and *in vivo.*

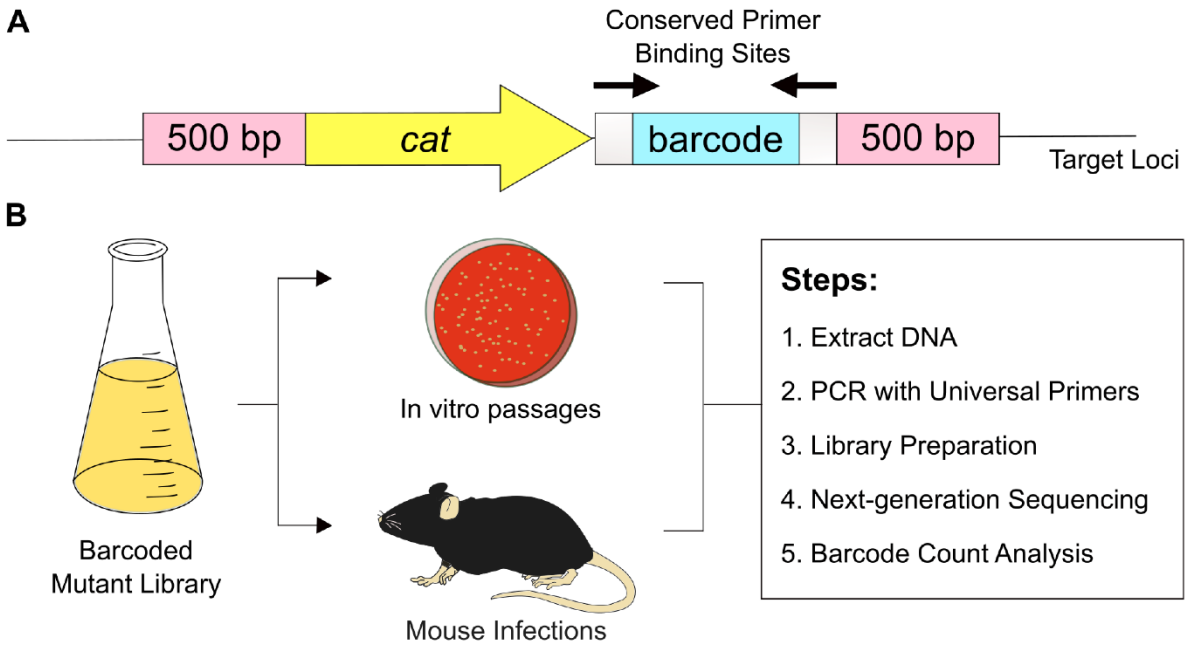

Figure 2.1: *H. pylori* barcoded mutant library design and workflow.
(A) Target genes were subjected to insertional mutagenesis with a cassette comprised of a chloramphenicol acetyltransferase (*cat*) gene and unique 21-nucleotide barcodes. Flanking segments (~500 bp) on either side of the *cat* cassette are derived from regions upstream and downstream of the desired site of insertion in the *H. pylori* chromosome. Conserved regions flanking each barcode allow unbiased parallel PCR amplification of barcodes from a pool of mutants. (B) The cassettes illustrated in panel A were introduced into selected sites in the *H. pylori* chromosome. Barcoded *H. pylori* mutants were pooled, and the resulting libraries were either passaged *in vitro* or administered to mice by oral gavage. At the end of the experiments, the relative abundance of barcoded strains in the populations was analyzed and compared to the abundance of barcoded strains in the input libraries. Input and output pools were analyzed by extracting *H. pylori* genomic DNA, amplifying barcodes via conventional PCR, prepping barcoded amplicons with NGS adapters, and performing Next-Generation Sequencing.

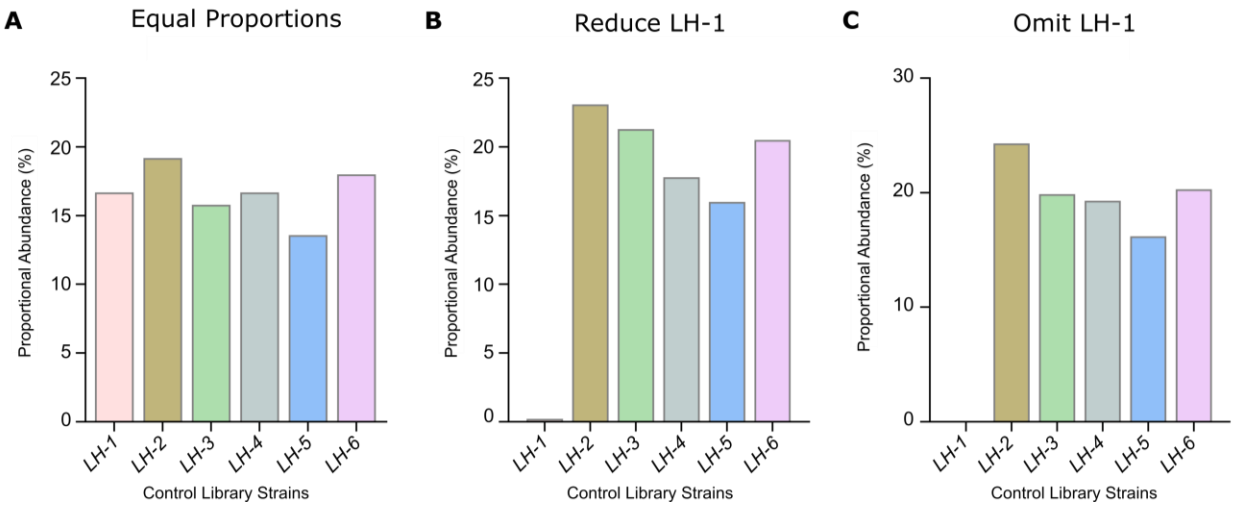

Figure 2.2: Alterations in barcode abundance are detectable via next-generation sequencing. We generated a control library composed of 6 isogenic mutants, barcoded at the *mdaB-hydA* intergenic locus (LH-1 through LH-6). These six barcoded control strains were pooled together at (A) equal proportions, (B) 100-fold reduced abundance of one strain (LH-1) compared to other strains, or (C) in the absence of one strain (LH-1). The detected proportion of barcode counts linked to LH-1 shifted from 16.7% to 0.22% to 0.00%, respectively.

Table 2.1: *H. pylori* strains used in this study

Strain name	**Description or genotype[a,b]**	**Annotation of mutated locus in strain J166**	**Alternate designation(s) of mutated locus[c]**	**Barcode Sequence (5' to 3')**
LH-1	J166 *mdaB::bc1::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	TGTACGGTCGAGTAACGTGAC
LH-2	J166 *mdaB::bc2::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	ACTGTACGTACAGTCCTCGGG
LH-3	J166 *mdaB::bc3::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	CGTACGGGTACAGTCGACGTA
LH-4	J166 *mdaB::bc4::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	TGTACAGCCCGAGGCCGTACG
LH-5	J166 *mdaB::bc5::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	TGTACGGCCTCGGGAACGTAG
LH-6	J166 *mdaB::bc6::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*	GTACAGTCGTCGAGTCGAGTA
LH-7	J166 *mdaB::bc7::hydA*	EG65_03855-EG65_03860	HP0630-HP0631, *mdaB-hydA*,	GTCACGTTCGAGTAGTCACGT
LH-8	J166 *ureA::bc8*	EG65_00400	HP0073, *ureA*	ACGTCGAGTCGGAAGTACAGT
LH-9$_a$	J166 *flaA::bc9*	EG65_04015	HP0601, *flaA*	CTCGACGGTCGACGGTCGGAA
LH-9$_b$	J166 *flaA::bc9,* independent	EG65_04015	HP0601, *flaA*	CTCGACGGTCGACGGTCGGAA
LH-10	J166 *hopA::bc10*	EG65_01210	HP0229, *hopA, omp6*	GTCACGTGGGACGTTCGACAA
LH-11$_a$	J166 *alpA::bc11*	EG65_04820	HP0912, *alpA, hopC, omp20*	AACGTAGCCACGTGGTCACGT
LH-11$_b$	J166 *alpA::bc11,* independent	EG65_04820	HP0912, *alpA, hopC, omp20*	AACGTAGCCACGTGGTCACGT
LH-12	J166 *hopD::bc12*	EG65_00150	HP0025, *hopD, omp2*	CGTACGGCGTACGGTCGAGTA
LH-13$_a$	J166 *hopE::bc13*	EG65_03490	HP0706, *hopE, omp15*	GTCGACGGTCACGTCCGTACA
LH-13$_b$	J166 *hopE::bc13,* independent	EG65_03490	HP0706, *hopE, omp15*	GTCGACGGTCACGTCCGTACA
LH-14$_a$	J166 *hopF::bc14*	EG65_04015	HP0252, *hopF, hopX, omp7*	CCTCGGGACTGTACCCGTACA
LH-14$_b$	J166 *hopF::bc14,* independent	EG65_04015	HP0252, *hopF, hopX, omp7*	CCTCGGGACTGTACCCGTACA
LH-15	J166 *hopG::bc15*	EG65_01335	HP0253, *hopG, hopY*	GTCGAGGGTCGGAAGGGACGT
LH-16$_a$	J166 *oipA::bc16*	EG65_03820	HP0638, *oipA, hopH, omp13*	AACGTAGGTCACGTCGTACGG
LH-16$_b$	J166 *oipA::bc16,* independent	EG65_03820	HP0638, *oipA, hopH, omp13*	AACGTAGGTCACGTCGTACGG
LH-17	J166 *hopI::bc17*	EG65_06180	HP1156, *hopI, omp25*	TCGACGTTCGACGTGTACAGT
LH-18	J166 *hopK::bc18*	EG65_04865	HP0923, *hopK, omp22*	CCGACGTCGTACGGCCGTACG
LH-19	J166 *hopL::bc19*	EG65_06185	HP1157, *hopL, omp26*	CGTACGGACGTCCCGTACAGT
LH-20	J166 *hopN::bc20*	EG65_01195	HP1342, *hopN, omp29*	CCGTACGTTCCGACCCCGAGG
LH-21	J166 *hopO::bc21*	EG65_03405	HP0722, *hopO, sabB, omp16*	GTACAGTACGTGACCCGACGT
LH-22	J166 *sabA::bc22*	EG65_03405	HP0725, *sabA, hopP, omp17*	ACGTGACTTCCGACCCCGAGG
LH-23	J166 *hopQ::bc23*	EG65_03420	HP1177, *hopQ, omp27*	CGACGTACCGTACGCGTCGAC
LH-24$_a$	J166 *babA::bc24*	EG65_04730	HP1243, *babA, hopS, omp28*	CCGACGTCGTCGACGTCGGAA
LH-24$_b$	J166 *babA::bc24,* independent	EG65_04730	HP1243, *babA, hopS, omp28*	CCGACGTCGTCGACGTCGGAA
LH-25	J166 *babB::bc25*	EG65_06705	HP0896, *babB, hopT, omp19*	GTCGGAACTCGACGCCTCGGG
LH-26	J166 *alpA::bc11 rdxA::alpA-bc26*	EG65_05040	HP0954, *rdxA*	TTGTCGACTGTACACCACGTG
LH-27	J166 *hopE::bc13 rdxA::hopE-bc27*	EG65_05040	HP0954, *rdxA*	TCGAGTATTGTCGACTGTACA
LH-28	J166 *flaA::bc2 rdxA::flaA-bc28*	EG65_05040	HP0954, *rdxA*	GTCGACGCTGTACATTGTCGA

[a] LH 1-28, Target loci in mouse-adapted strain J166 wild-type were mutated by inserting cassettes containing a chloramphenicol acetyltransferase gene and unique 21 nucleotide barcodes (bc), flanked by 500 base pair sequences adjacent to the targeted insertion site.
[b] LH 26-28, These strains are complemented mutants in which the *rdxA* gene was replaced with the intact gene of interest (*alpA, hopE,* or *flaA*).
[c] These correspond to gene numbers in strain 26695 or gene names.

CHAPTER III

HOP OUTER MEMBRANE PROTEINS AS DETERMINANTS OF *HELICOBACTER PYLORI* FITNESS *IN VITRO*

A modified version was previously published as:

Harvey ML, Lin AS, Sun L, Koyama T, Shuman JHB, Loh JT, Scott Algood HM, Scholz MB, McClain MS, Cover TL. Enhanced fitness of a *Helicobacter pylori babA* mutant in a murine model. Infect Immun. 2021 Jul 26:IAI0072520. doi: 10.1128/IAI.00725-20. Epub ahead of print. PMID: 34310886.

Introduction

Data presented in Chapter II showed that genetic barcoding was successfully adapted for *H. pylori* research. It allows for both the simultaneous disruption of target genes while introducing a genetic barcode. This methodology enables the quantification of changes in the proportional abundance of a mixed pool of *H. pylori* mutant strains. Our goal for the work presented in Chapter III was to determine if barcoded control strains (LH-1 through LH-7) have equal fitness to one another, which would indicate that a non-selective bottleneck is absent *in vitro.* Another goal was to identify Hops that play a role in *H. pylori* fitness *in vitro.* We are not currently aware of any studies investigating the role of Hops in *H. pylori* fitness *in vitro.* We consider this to be essential step prior to investigating the roles of Hops in *H. pylori* fitness *in vivo*.

Materials and Methods

Serial passaging of barcoded libraries *in vitro*

H. pylori input pools were generated by combining barcoded mutant strains at equal proportions in bisulfite-free Brucella broth, resulting in a final concentration of the

pooled mutants of 1.0 x 10^9 cells/ml. To circumvent a potential loss of bacterial viability that is sometimes encountered with growth of *H. pylori* in liquid cultures, we passaged the pooled bacteria on solid media. The control library input pool was plated on blood agar plates and passaged every two days for 21 days, for a total of 10 passages. The OMP mutant library input pool was plated on blood agar plates supplemented with 10% FBS or BSFB agar plates (0.5% or 0.9% sodium chloride) supplemented with cholesterol and passaged every two days for 21 days, for a total of 10 passages. When passaging bacteria from one plate to another, a zig-zagging streak was taken across the bacterial lawn to allow a comprehensive sampling.

Next-generation sequencing and sequence data analysis

Genomic DNA was extracted from *H. pylori* input and output pools (bacterial lawns from *in vitro* passages or pooled single colonies grown from mouse stomach lysates, >100 colonies/mouse). Subsequently, barcodes were amplified via conventional PCR with primers that bind to conserved regions flanking the barcodes. Specifically, we used 5' ATCTACTGCCGATATTTACG 3' as a forward primer and 5' TAAATCCACTGTGATATCCAT 3' as a reverse primer. Following a PCR clean-up step (Qiagen), amplicon libraries were generated using NEBNext Ultra II DNA Library Prep Kit for Illumina. Sequencing was performed using Paired-End (PE) 150 bp on the Illumina NovaSeq 6000 at approximately 10 million PE Reads/Sample. Flanking sequences were cleaved from barcodes *in silico* with the "Seqtk_Trimfq" Galaxy software tool (https://github.com/lh3/seqtk/). Barcodes were then sorted by barcode identity and quantified using "Barcode Splitter" Galaxy software tool. (http://hannonlab.cshl.edu/fastx_toolkit/). Individual barcode counts for each mutant

within a single sample were normalized to the same total standard count across all experiments [gene of interest (goi) normalized count]. For instance, for a given gene of interest:

$$goi_{normalized\ count} = \left(\frac{goi_{raw\ count}}{total_{raw\ count}}\right) * total_{standard\ count}$$

Additionally, the $\log_{10}$ ratio was calculated by taking the $\log_{10}$ of the output count divided by its respective input count for each strain. Due to the limitations imposed by $\log_{10}$ fold calculations, all barcode counts that fell to 0 were replaced by 0.01 for quantification purposes. As such, the limit of detection corresponded approximately to a $\log_{10}$ ratio of -8.0.

$$Log_{10}\ ratio = Log_{10}\left(\frac{goi_{output}}{goi_{input}}\right)$$

Small-scale Competition Assays

We generated two libraries containing 12 barcoded mutant strains at equal proportions, at a final concentration of 1.0 x 10^9 cells/ml in BSFB. The first library was composed of the original control strains (LH-1 through LH-7) mixed with the original barcoded *hop* mutants (*alpA, hopE, hopF, oipA,* and *babA* mutants) (LH-11_a, LH-13_a, LH-14_a, LH16_a, LH-24_a). The second library was composed of the original control strains (LH-1 through LH-7) mixed with the new independent barcoded *hop* mutants (*alpA, hopE, hopF, oipA,* and *babA* mutants) (LH-11_b, LH-13_b, LH-14_b, LH-16_b, LH24_b). Each library was plated and serially passaged on blood agar plates every two days for 21 days.

Genomic DNA was isolated from input and output pools and processed for next generation sequencing as described above.

Pairwise Competition Assays

The LH-7 control strain was combined 1:1 with original *flaA, alpA, hopE, hopF, oipA,* and *babA* mutants (LH-9_a, LH-11_a, LH-13_a, LH-14_a, LH-16_a, LH-24_a) or subsequently generated mutants (LH-9_b, LH-11_b, LH-13_b, LH-14_b, LH-16_b, LH-24_b) and passaged on blood agar plates every two days for 21 days. Subsequently, we extracted genomic DNA from the input and outputs using phenol-chloroform-based methods (116). Real-time qPCR analysis of DNA samples was performed with SYBR green fluorophore (iQ SYBR Green Supermix; Bio-Rad) on an ABI StepOnePlus machine. Primer sequences can be found in Table 3.5. A standard curve of each DNA target was generated using 10-fold dilutions starting at 50 ng/well. The abundance of individual strains was calculated using the appropriate standard curve for each DNA target.

Generation of Complemented Mutant Strains

Plasmids for complementation contained the relevant gene with its endogenous promoter along with a unique 21-nucleotide barcode, flanked by 500 base pairs homologous to sequence flanking the *rdxA* gene. These plasmids were designed to allow insertion of the relevant gene into the *rdxA* locus, rendering these strains metronidazole resistant (117).Complemented mutant strains (LH-26, LH-27, LH-28) were generated by transforming J166 *alpA* (LH-11_a)*, hopE* (LH-13_a)*,* and *flaA* (LH-9_a) mutants with the appropriate plasmids. Following transformation, *H. pylori* transformants were plated on BSFB agar plates supplemented with cholesterol and containing metronidazole (7.5

μg/ml). Transformants were validated via sequencing of PCR amplicons generated with primers flanking the *rdxA* gene (about 600 base pairs upstream and 600 base pairs downstream of the insertion site).

Competition Assays with Complemented Mutants

The LH-7 control strain was mixed 1:1:1 with a *hop* mutant strain [*alpA* (LH-11$_a$), *hopE* (LH13$_a$), or *flaA* (LH-9$_a$) mutants] and the corresponding complemented mutant (LH-26, LH-27, or LH-28) and passaged on blood agar plates every two days for 21 days, for a total of 10 passages. Genomic DNA was extracted from the input and output populations using phenol-chloroform-based methods (116). We used real-time qPCR methodology described above to assess the abundance of each strain in the output and input populations. Primer sequences can be found in Table 3.5.

Statistical analysis

For experiments analyzing OMP mutants *in vitro*, we used the following approach to identify mutants in the OMP mutant library that were significantly different from the control strain (LH-7) while accounting for the observed variability in the control library. All pairwise mean differences within the control library were estimated with 99% confidence intervals, and the furthest boundary (in either positive or negative direction) from 0 was selected as the boundary of no difference. Two-sample t-test without assuming equal variance was used to compare a mutant in the OMP mutant library with the control strain (LH-7). A 99% confidence interval (corresponding to type I error rate of 1%) was computed for each comparison (Figures 3.1, 3.2, 3.4E, and 3.4F). Strains with p value less than 0.01 (when comparing to LH-7) and also a 99% CI that does not overlap the

boundary of no difference (based on the control library) were considered significant. All confidence intervals and p-values are reported in supplemental tables (3.1, 3.2, 3.3, 3.4, 3.6).

Results

Tracking changes in composition of the control library *in vitro*

We first evaluated the stability of the control library during growth *in vitro.* We prepared three independent preparations of the control mutant library on three separate days, each containing equal proportions of the component strains, and we passaged the libraries on blood agar plates (TSA-5% sheep blood). Following 10 serial passages of the libraries over 21 days, we compared the composition of the passaged libraries (outputs) with the composition of the starting libraries (inputs), using PCR and NGS as described in Methods. As expected, at the start of the experiment each of the input libraries contained similar proportions of the six component strains (Fig. 3.1A). At the end of the experiment, all six strains remained present in the output libraries. When comparing the relative abundance of various strains in the output libraries, there were statistically significant differences in the numbers of barcode counts detected (Table 3.1), but the magnitude of changes was relatively small (less than 10-fold differences in the numbers of barcode counts) (Fig. 3.1B). These data provide a frame of reference for interpreting subsequent experiments investigating the contribution of specific *hop* genes to *H. pylori* fitness.

Tracking changes in composition of the OMP mutant library in vitro

To evaluate possible contributions of *hop* genes to *H. pylori* fitness, we generated an “OMP mutant library” containing a panel of barcoded strains with mutated *hop* genes (Table 2.1). This library also contained several barcoded control strains for comparison, including a strain containing the barcode in the *mdaB-hydA* intergenic region (LH-7*)* and two strains containing mutations in non-*hop* genes predicted to be required for *H. pylori* fitness *in vivo* but not required for fitness *in vitro* (*ureA* and *flaA,* encoding a subunit of urease and a flagellar component, respectively*)* (Table 2.1) (38-40, 106, 111) . We prepared three independent preparations of the OMP mutant library on three separate days, each containing equal proportions of the component strains. Following 21 days of serial passaging *in vitro*, we compared the composition of the passaged libraries with the composition of the input libraries. At the start of the experiment each of the input libraries contained similar proportions of the component strains (Fig. 3.2A). There were marked changes in the composition of the libraries over the course of the experiment, and the observed reduction in abundance of several mutants was much greater than what was observed in the previous experiment with control strains (compare Fig. 3.2B with Fig. 3.1B). We used multiple criteria to identify mutants that exhibited significant differences in fitness compared to control strains, as described in Methods. Neither the control strain nor the *ureA* mutant exhibited a fitness defect *in vitro*, but unexpectedly, the *flaA* mutant exhibited a fitness defect. Based on the combined results of two statistical analyses (t-tests and analysis of confidence intervals), *hopE and oipA* mutants (designated LH-13_a and LH-16_a) exhibited significant fitness defects and a *hopF* mutant (LH-14_a) exhibited a fitness advantage (Fig. 3.2C, Table 3.1). Growth curve analyses of the *hopE, oipA,* and

flaA mutant strains indicated that these mutant strains grew at rates similar to that of the J166 wild-type strain and several other mutants (Fig. 3.8).

We then conducted further experiments to evaluate if similar changes in the library composition would be observed after passage on TSA-5% sheep blood compared to two alternate media [BSFB-cholesterol agar plates containing either routine levels of supplemental sodium chloride (0.5%) or increased levels of sodium chloride (0.9%)]. Barcode counts from multiple *hop* mutant strains (including *hopE, oipA,* and *hopF* mutants) were significantly different compared to barcode counts from the control strain under all three conditions (Fig. 3.3 and Table 3.2).

Further evaluation of mutants *in vitro*

To address the possibility of unrecognized secondary mutations in mutant strains, we repeated the mutagenesis of mouse-adapted wild-type strain J166, thereby generating independently constructed replicates of the original mutants. Specifically, we generated newly constructed *hopE, oipA, and hopF* mutants (designated LH-13_b, LH-16_b, and LH-14_b) since the original *hopE, oipA, and hopF* mutants exhibited fitness defects in the initial experiments (Fig. 3.2 and Fig. 3.3). We also generated a second *flaA* mutant strain (LH-9_b) to determine if the unexpected fitness disadvantage of the *flaA* mutant observed in Figure 3.2 and Figure 3.3 would be recapitulated. For comparison, we also generated a newly constructed *alpA* mutant (LH-11_b) and *babA* mutant (LH-24_b). In the initial experiments (Fig. 3.2), the *alpA* mutant exhibited a non-significant trend toward a fitness defect and the *babA* mutant did not exhibit any detectable fitness defect.

We evaluated the fitness of the new *hop* mutants, along with the new *flaA* mutant, using small-scale competitions and pairwise competitions assays. Small-scale

competitions were conducted using two libraries composed of 12 barcoded mutant strains [7 control strains (LH-1 to LH-7) and the new *alpA, hopE, oipA, hopF,* and *babA* mutants] and the populations were analyzed via NGS, whereas pairwise assays were carried out with 1:1 mixtures of mutants competed with LH-7 and analyzed via qPCR. Consistent with results of the initial experiment (Fig. 3.2), we observed fitness defects of *hopE* and *oipA* mutants (LH-$13_{a,b}$ and LH-$16_{a,b}$) when testing independently constructed mutant strains (Fig. 3.3-3.5, Tables 3.2-3.4). The significance of the fitness defect observed for the *alpA* mutant compared to LH-7 was variable depending on the type of statistical analysis performed (Fig. 3.3-3.5, Tables 3.2-3.4). The properties of the original *hopF* mutant (LH-14_a) were not recapitulated by the newly constructed *hopF* mutant strain (LH-14_b). As expected, there were no significant differences when comparing the *babA* mutants (LH-24_a and LH-24_b) with the LH-7 mutant, confirming that *babA* does not contribute to *H. pylori* fitness *in vitro* (Fig. 3.3-3.5, Tables 3.2-3.4).

Among the 5 genes analyzed in the experiments described above, several are monocistronic (for example, *oipA* and *babA*), whereas *alpA* and *hopE* are each predicted to be transcribed within operons. This raises the possibility that mutagenesis of the latter genes might alter transcription of downstream genes (73). Therefore, we assessed if the fitness of *alpA* and *hopE* mutants could be restored via complementation. In parallel, we undertook complementation of the *flaA* mutant, which had an unexpected fitness defect. Complemented mutants were generated as described in Methods. qRT-PCR experiments showed that *alpA* and *hopE* were transcribed as expected in the complemented mutants (data not shown) and motility of the *flaA* mutant was successfully restored by the complementation (Fig. 3.6). To evaluate the fitness of the complemented strains

compared to the mutants from which they were derived, we did competition experiments in which three strains (LH-7 control strain, a mutant strain and the corresponding complemented mutant strain) were combined at equal proportions and passaged on blood agar plates for 21 days. Subsequently, the relative abundance of each strain compared to the control strain was analyzed by qPCR as described in Methods. As expected, the relative abundance of the *alpA, hopE,* and *flaA* mutants was reduced compared to the relative abundance of the control strain (Fig. 3.7). Unexpectedly, complementation did not rescue the fitness of these three mutant strains (Fig. 3.7). Potential explanations for why the fitness of these mutants was not restored by complementation are considered in the Discussion.

Discussion

As an initial step toward development of this methodology, we generated a barcoded control library in which all strains were genetically manipulated at a neutral locus, unrelated to bacterial fitness. Following 21 days of growth *in vitro,* we observed relatively low variability in composition of the control library, indicating that these strains were able to form a stable community and did not encounter any strong non-selective pressures *in vitro.* Next, we assessed a library of *hop* mutant strains and analyzed changes in its composition during 21 days of passage *in vitro.* The *hopE* and *oipA* mutants exhibited a fitness defect that was detected using three types of solid culture media and the results were corroborated when analyzing independently generated *hopE* and *oipA* mutants. Likewise, the *alpA* mutant exhibited a trend consistent with a fitness defect *in vitro.* Unexpectedly, we also observed that *flaA* mutants exhibited a fitness defect *in vitro.*

To further evaluate the fitness defects of these mutants, we generated complemented mutant strains. The *rdxA* locus, which was used as the site for introduction of these complemented genes, was previously reported to have no effect on *H. pylori* fitness. Complementation restored expression of *alpA* and *hopE,* and restored motility in the *flaA* mutant, but unexpectedly, complementation failed to reverse the fitness defects of these mutants. There are multiple possible explanations for the failure of complementation to restore fitness of the mutants. One possibility (relevant for *alpA* and *hopE*) is that the insertion of barcoded antibiotic resistance cassettes in these genes might have altered the transcription of other genes in the operon. Another possibility is that co-transcription of these operon-encoded genes is necessary for proper function, and in this scenario, complementation into a heterologous locus would not rescue fitness. As another possibility, insertional mutagenesis of OMP-encoding genes might have resulted in the production of truncated proteins, causing a detrimental effect on *H. pylori* fitness. Finally, it is possible that the mutagenesis may have altered production of sRNAs. *H. pylori* is known to produce hundreds of different sRNAs, and we speculate that there may be strain-specific differences in sRNA production in the strain used for the current study compared to a prototype strain used for previous studies of *H. pylori* sRNA production. Since the complementation experiments did not provide evidence indicating a contribution of *alpA* or *hopE* to *H. pylori* fitness, we did not undertake further studies to investigate the mechanistic basis for the fitness defects of these mutants *in vitro.* Overall, we demonstrated the utility of our barcoding approach to quantify changes in the composition of a barcoded library *in vivo.* In the next chapter, we will investigate *in vivo* pressures and the role of Hops in *H. pylori* fitness using this same genetic barcoding approach.

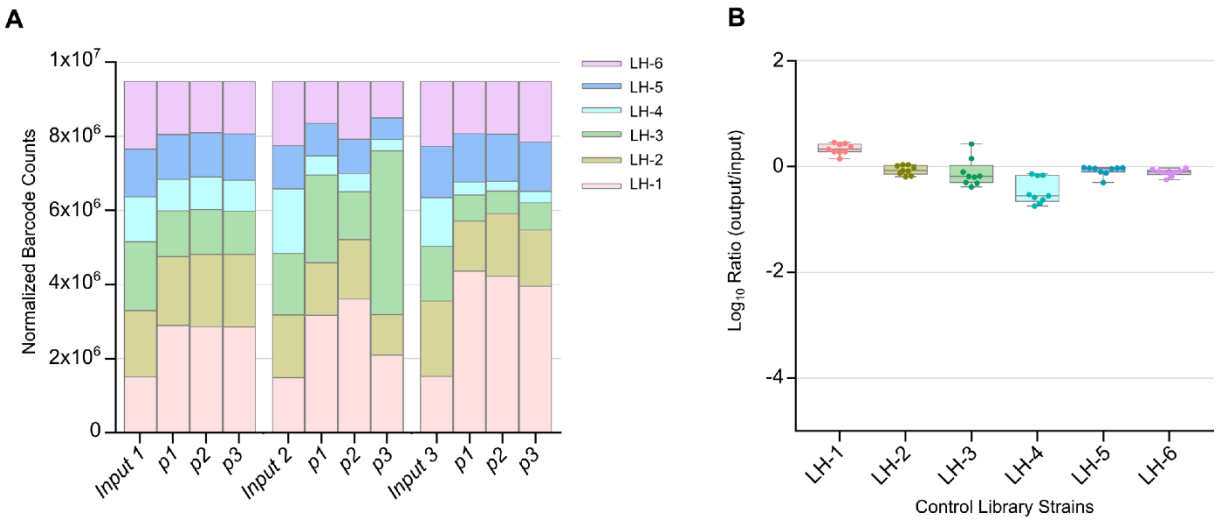

Figure 3.1: The control library proliferated as a stable community *in vitro*.
We generated control libraries composed of 6 strains, each containing a unique nucleotide barcode at the mdaB-hydA intergenic locus. In three independent experiments conducted on separate days, a single input pool was inoculated to 3 separate plates (p1-p3) which were passaged *in vitro* for 21 days. The composition of input (input 1-3) and output pools (p1-p3) was analyzed as described in Methods. The barcode counts of input and output pools were normalized to standardized total counts. Panel A shows relative proportions of barcode counts and panel B shows log10 ratios comparing output counts to the input counts (see details in Methods). Statistical analysis is shown in Table 3.1. The control library proliferated as a relatively stable community, without substantial changes in the proportional abundance of individual strains compared to the input pool.

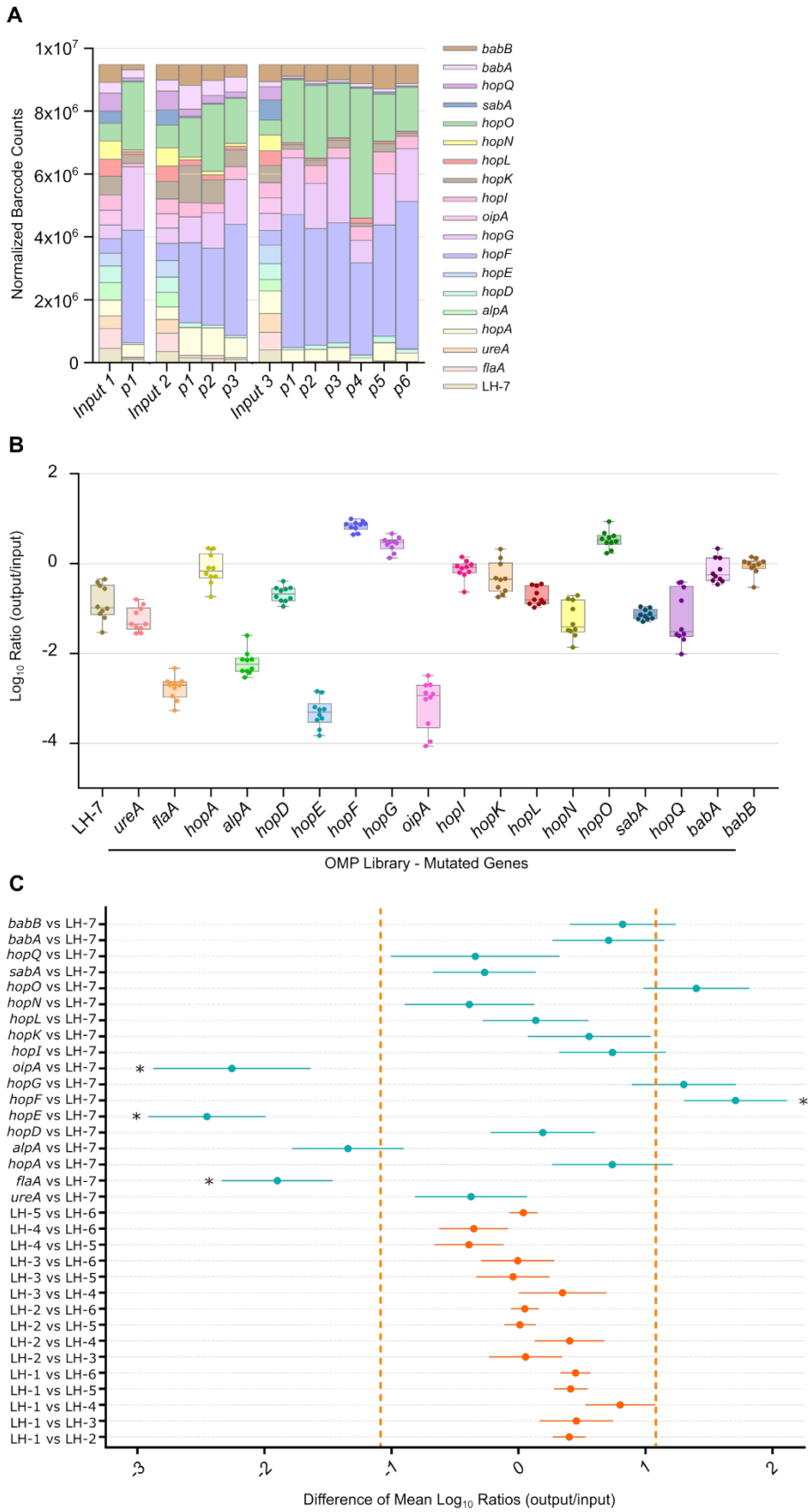
A
Normalized Barcode Counts
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
Input 1
p1
Input 2
p1
p2
p3
Input 3
p1
p2
p3
p4
p5
p6
babB
babA
hopQ
sabA
hopO
hopN
hopL
hopK
hopI
oipA
hopG
hopF
hopE
hopD
alpA
hopA
ureA
flaA
LH-7
B
Log10 Ratio (output/input)
2
0
-2
-4
LH-7
ureA
flaA
hopA
alpA
hopD
hopE
hopF
hopG
oipA
hopI
hopK
hopL
hopN
hopO
sabA
hopQ
babA
babB
OMP Library - Mutated Genes
C
babB vs LH-7
babA vs LH-7
hopQ vs LH-7
sabA vs LH-7
hopO vs LH-7
hopN vs LH-7
hopL vs LH-7
hopK vs LH-7
hopI vs LH-7
oipA vs LH-7
hopG vs LH-7
hopF vs LH-7
hopE vs LH-7
hopD vs LH-7
alpA vs LH-7
hopA vs LH-7
flaA vs LH-7
ureA vs LH-7
LH-5 vs LH-6
LH-4 vs LH-6
LH-4 vs LH-5
LH-3 vs LH-6
LH-3 vs LH-5
LH-3 vs LH-4
LH-2 vs LH-6
LH-2 vs LH-5
LH-2 vs LH-4
LH-2 vs LH-3
LH-1 vs LH-6
LH-1 vs LH-5
LH-1 vs LH-4
LH-1 vs LH-3
LH-1 vs LH-2
*
*
*
*
-3
-2
-1
0
1
2
Difference of Mean Log10 Ratios (output/input)

Figure 3.2: Fitness of a library of hop mutant strains during passage *in vitro*.
We generated “OMP mutant libraries” composed of 16 hop mutants and 3 non-hop control mutants. In three independent experiments conducted on separate days, a single input pool was inoculated to one or more separate plates (p1-p6) which were passaged *in vitro* for 21 days. The composition of input (input 1-3) and output pools (p1-p6) was analyzed as described in Methods. (A, B) Panel A shows relative proportions of barcode counts and panel B shows log10 ratios comparing the output pools with the input. (C) We estimated the variability observed in the control library (Fig. 3.1) by analyzing all pairwise differences of mean log10 ratios within the control library (see Table 3.1). For each of the 15 pairwise differences, 99% confidence intervals were computed. Among these, the furthest boundary from 0 is 1.075 (LH-1 vs LH-4). Thus, we used the -1.075 to 1.075 range as the boundary for “no difference”. In other words, each mutant in the OMP mutant library needs to be further from the control strain (LH-7) than this distance to be considered a significant change from the input. We computed 99% confidence intervals of the difference between each mutant to the control (LH-7). Additionally, all 16 pairwise comparisons within the control library are depicted (orange). If a comparison to the control strain (LH-7) within the OMP mutant library (blue) clears this threshold (99% confidence is wholly above or below the threshold) (*), we concluded that the mutant is significantly different from the control. Statistical analysis is shown in Table 3.1. The hopE (LH-13a), oipA (LH16a), and flaA (LH-9a) mutants had significantly reduced barcode counts based on both these criteria, whereas hopF (LH-14a) mutant had significantly increased barcode counts.

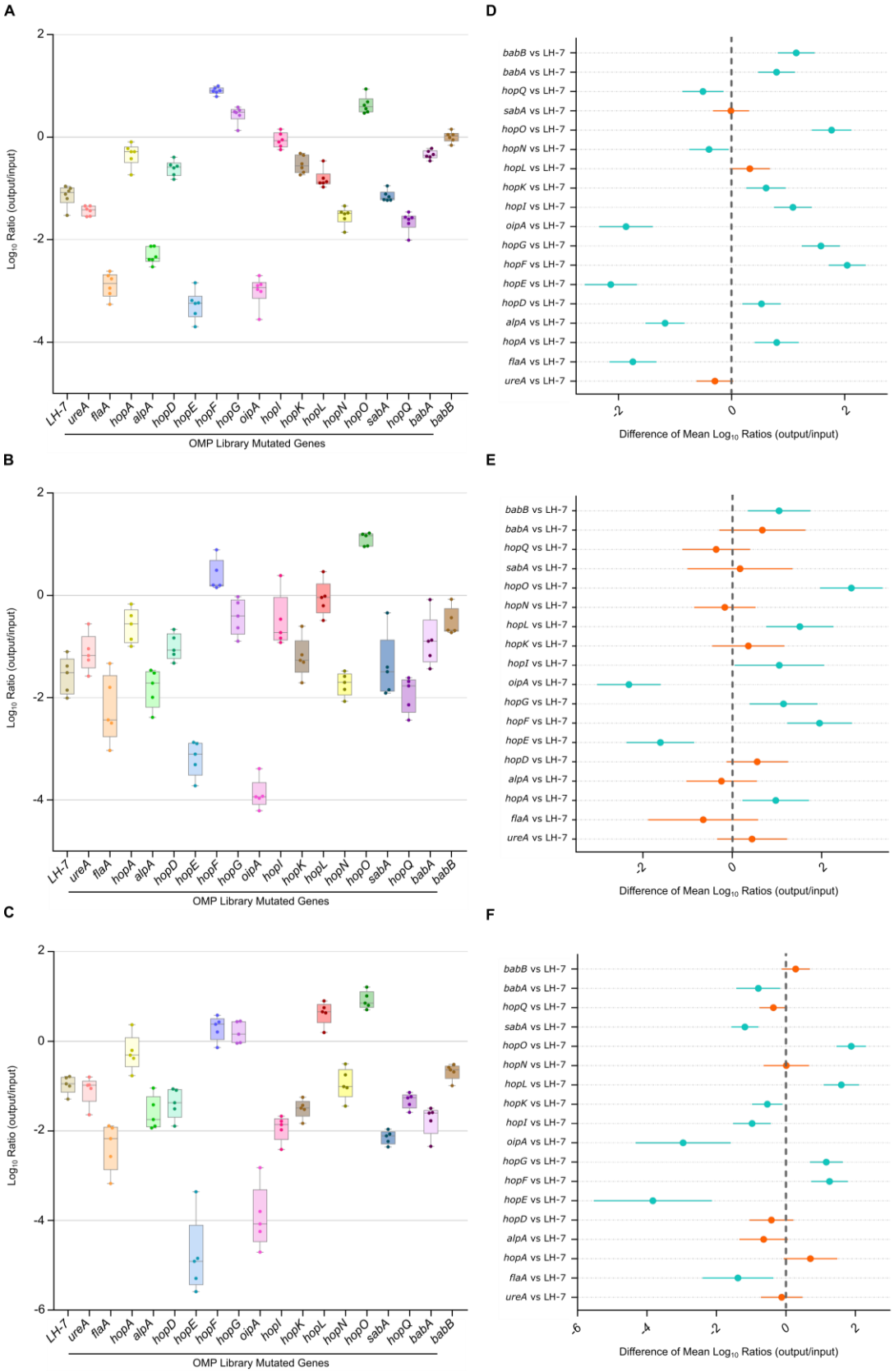
A
B
C
D
E
F
Log10 Ratio (output/input)
OMP Library Mutated Genes
LH-7
ureA
flaA
hopA
alpA
hopD
hopE
hopF
hopG
oipA
hopI
hopK
hopL
hopN
hopO
sabA
hopQ
babA
babB
babB vs LH-7
babA vs LH-7
hopQ vs LH-7
sabA vs LH-7
hopO vs LH-7
hopN vs LH-7
hopL vs LH-7
hopK vs LH-7
hopI vs LH-7
oipA vs LH-7
hopG vs LH-7
hopF vs LH-7
hopE vs LH-7
hopD vs LH-7
alpA vs LH-7
hopA vs LH-7
flaA vs LH-7
ureA vs LH-7
Difference of Mean Log10 Ratios (output/input)

Fig. 3.3. Mutagenesis of specific hop genes alters H. pylori fitness independent of media composition.
The OMP mutant library was passaged on (A,D) blood agar plates, (B, E) brucella agar plates containing 0.5% sodium chloride, or (C, F) brucella agar plates containing 0.9% sodium chloride for 21 days. The composition of input and output pools was analyzed as described in Methods. (A,B,C) These panels show log10 ratios (output/input) for each mutant. (D,E,F) Differences in mean log10 ratios (output/input) were estimated with a 99% confidence interval, and the mutants significantly different from the control were identified. Significant changes are in blue and non-significant changes are in orange. Statistical analysis is shown in Table 3.2. Strains with mutations in hopE and oipA exhibited significant fitness defects in all three conditions, whereas hopF, hopG, and hopO mutants had significant fitness advantages in all three conditions ($p < 0.01$).

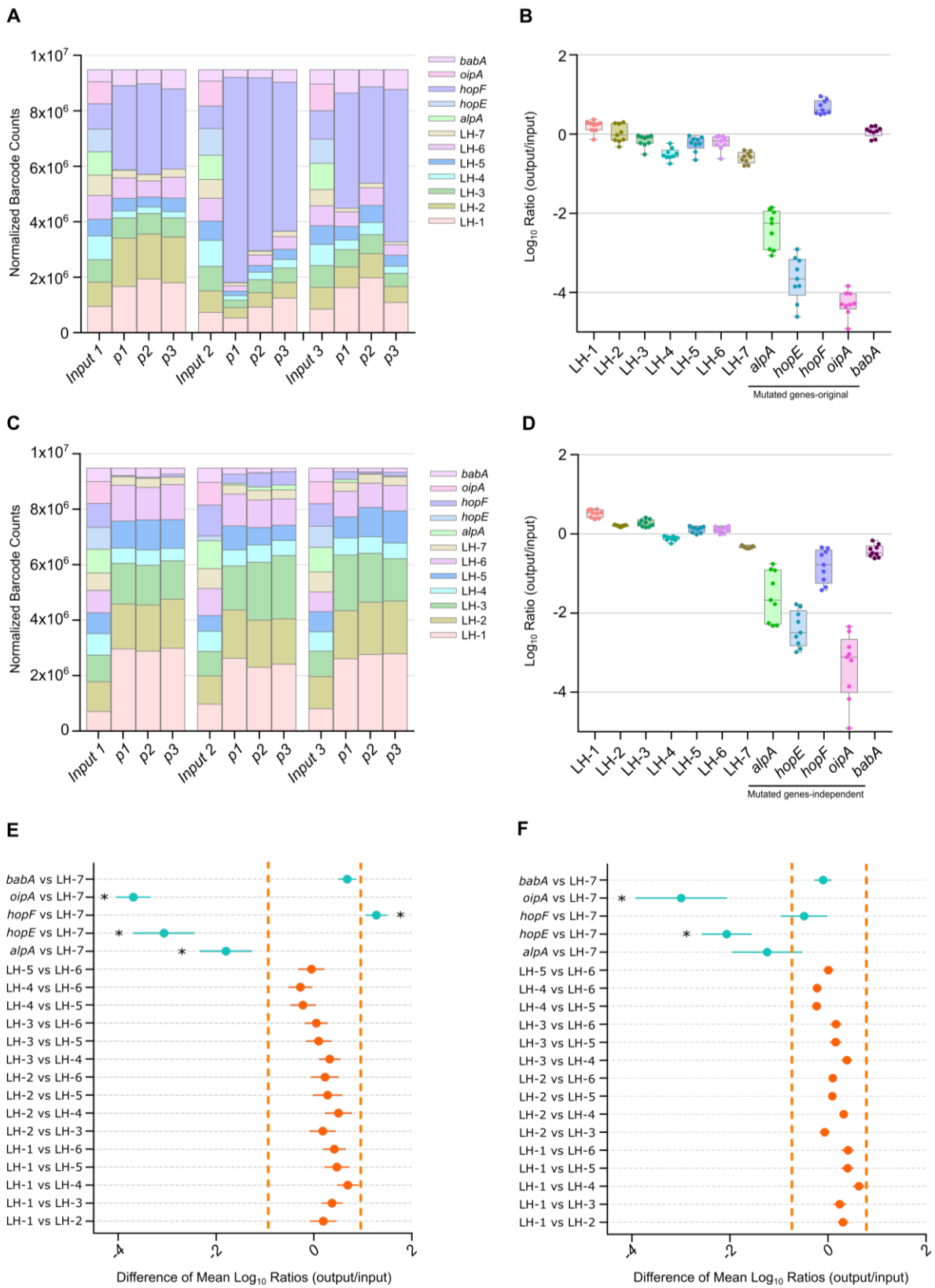
A
B
C
D
E
F
Normalized Barcode Counts
Log10 Ratio (output/input)
babA
oipA
hopF
hopE
alpA
LH-7
LH-6
LH-5
LH-4
LH-3
LH-2
LH-1
Input 1
Input 2
Input 3
p1
p2
p3
Mutated genes-original
Mutated genes-independent
babA vs LH-7
oipA vs LH-7
hopF vs LH-7
hopE vs LH-7
alpA vs LH-7
LH-5 vs LH-6
LH-4 vs LH-6
LH-4 vs LH-5
LH-3 vs LH-6
LH-3 vs LH-5
LH-3 vs LH-4
LH-2 vs LH-6
LH-2 vs LH-5
LH-2 vs LH-4
LH-2 vs LH-3
LH-1 vs LH-6
LH-1 vs LH-5
LH-1 vs LH-4
LH-1 vs LH-3
LH-1 vs LH-2
Difference of Mean Log10 Ratios (output/input)

Figure 3.4: Evaluation of mutant pools containing independently generated *hop* mutants. We generated two libraries composed of 12 barcoded mutant strains. Each library included the original control strains (LH-1 through LH-7) and either the (A, B) original *flaA, alpA, hopE, hopF, oipA,* and *babA* mutants (LH-9_a, LH-11_a, LH-13_a, LH-14_a, LH-16_a, LH-24_a) or (C, D) subsequently generated *flaA, alpA, hopE, hopF, oipA,* and *babA* mutants (LH-9_b, LH-11_b, LH-13_b, LH-14_b, LH-16_b, LH-24_b). The libraries were passaged *in vitro* for 21 days in three independent experiments. The composition of input (input 1-3) and output pools (p1-p3) was analyzed as described in Methods. Differences in mean $\log_{10}$ ratios (output/input) were estimated with a 99% confidence interval for the barcoded library composed of (E) original mutants and (F) subsequently generated mutants (blue). Additionally, all pairwise comparisons within the control library are depicted (orange). Statistical analysis is shown in Table 3.3. Based on the criteria described in Figure 3.2 and in Methods, *alpA, hopE,* and *oipA* mutants decreased in abundance relative to the control strains within the barcoded library composed of the original mutants, whereas the *hopF* mutant increased in abundance relative to the control strains. The relative abundance of the *babA* mutant remained similar to the control strains. The relative abundance of *hopE* and *oipA* mutants decreased relative to the control strains within the barcoded library composed of subsequently generated mutants. The relative abundance of the subsequently generated *hopF* and *babA* mutants remained similar to control strains. *comparison is wholly above or below boundary of no difference.

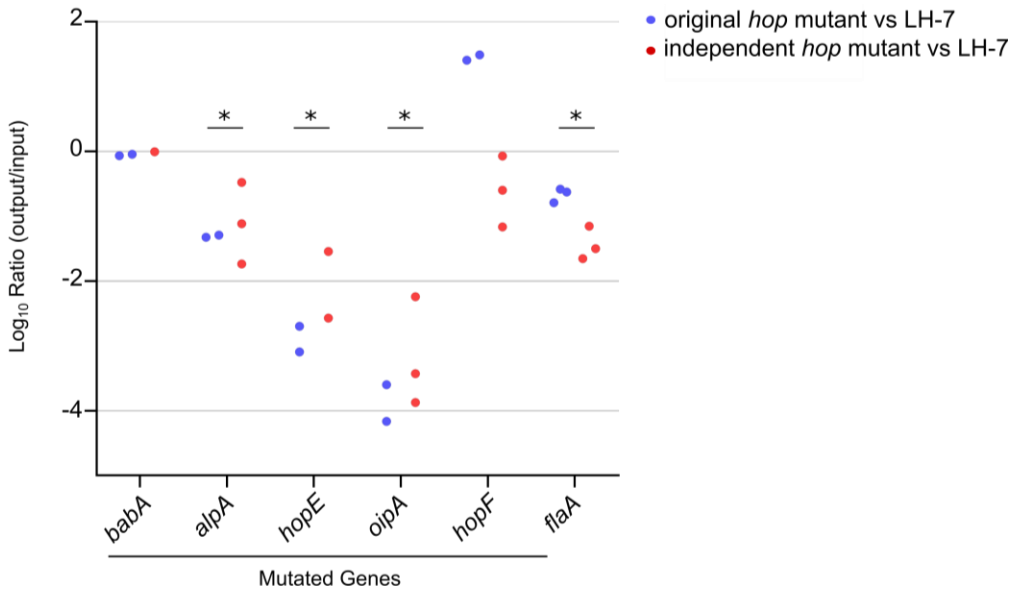

Figure 3.5: Fitness of newly generated *hop* mutants evaluated by qPCR.
The original *flaA, alpA, hopE, hopF, oipA,* and *babA* mutants (LH-9$_a$, LH-11$_a$, LH-13$_a$, LH-14$_a$, LH-16$_a$, LH-24$_a$) or subsequently generated mutants (LH-9$_b$, LH-11$_b$, LH-13$_b$, LH-14$_b$, LH-16$_b$, LH-24$_b$) were mixed with the control strain (LH-7) in a 1:1 ratio and passaged on blood agar plates every 2 days for 21 days. Original mutants are shown in blue and subsequently generated mutants are shown in red. We used qPCR to determine the abundance of each mutant in the input and output populations, based on comparison to a standard curve generated for each target. The $\log_{10}$ ratio (output/input) was calculated based on the relative abundance of mutant/LH-7 in the output and input populations. At the end of the experiment, the relative abundance of the original and second *alpA, hopE*, *oipA,* and *flaA* mutants was decreased relative to that of LH-7. The relative abundance of the original *hopF* mutant was increased compared to that of LH-7, whereas the second *hopF* mutant remained similar to LH-7. The relative abundance of the original and second *babA* mutants remained similar to that of LH-7. The changes in abundance of both *alpA, hopE*, *oipA,* and *flaA* mutants were statistically significant. Statistical analysis is shown in Table 3.4. *p-value <0.01.

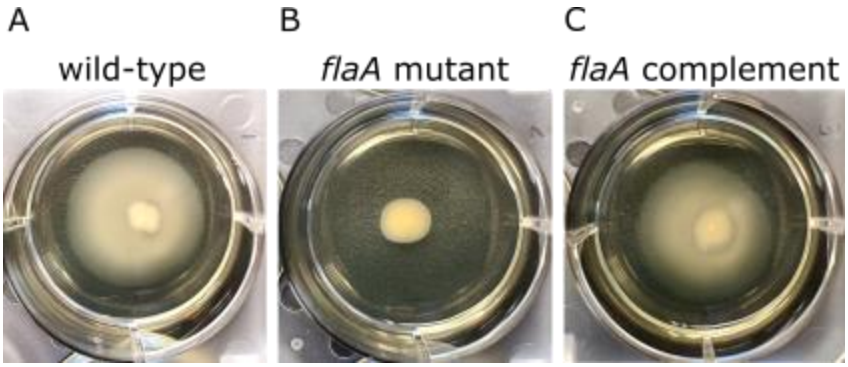

Figure 3.6: Restored motility in complemented *flaA* mutant.
Bacterial strains of A) J166 wild-type, B) J166 *flaA* mutant, and C) J166 complemented *flaA* mutant were inoculated into the center of soft agar gel. After 5 days, the diameter of motility was measured. Strains J166 wild-type and the complemented *flaA* mutant were motile, whereas the *flaA* mutant was non-motile.

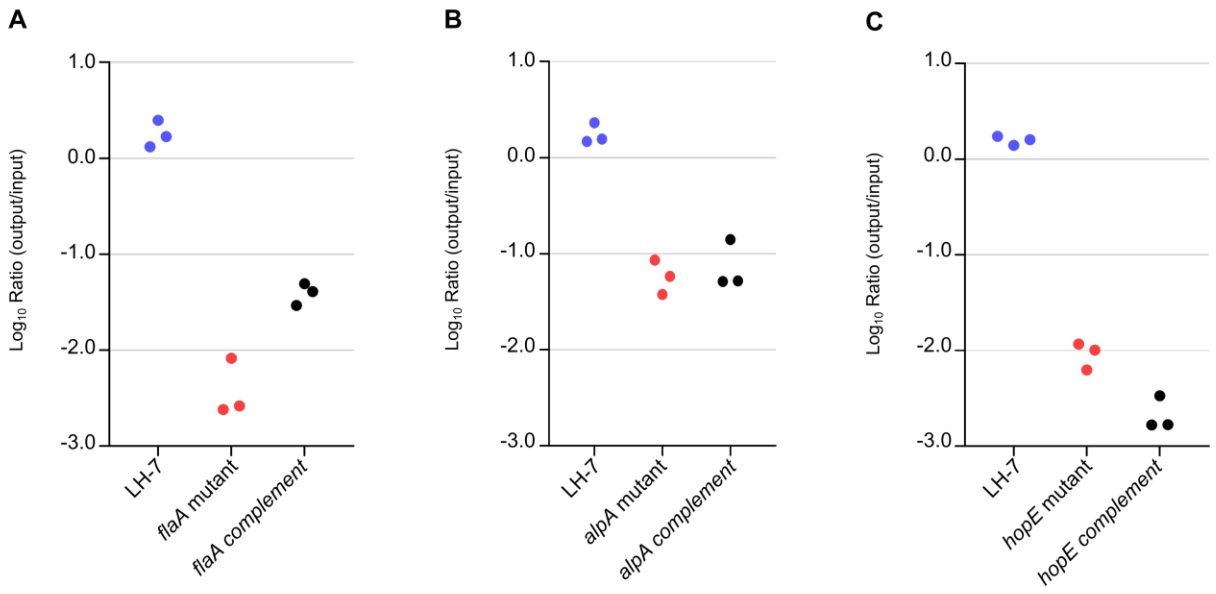

Figure 3.7: Complementation does not restore *in vitro* fitness of *alpA* or *hopE* mutants. Bacterial strains were mixed in a 1:1:1 ratio and passaged on blood agar plates for 21 days. These mixtures were composed of A) the control strain (LH-7), a *flaA* mutant (LH-9$_a$), and the complemented *flaA* mutant (LH-28), B) the control strain (LH-7), the *alpA* mutant (LH-11$_a$), and the complemented *alpA* mutant (LH-26), and C) the control strain (LH-7), the *hopE* mutant (LH-13$_a$), and the complemented *hopE* mutant (LH-27). We used qPCR to determine the abundance of each mutant in the input and output populations, based on comparison to a standard curve generated for each target. The $\log_{10}$ ratio (output/input) represents the abundance ratio of the mutant/LH-7 in the output and input populations. The abundance of *flaA*, *alpA*, and *hopE* mutants was reduced compared to LH-7. Likewise, the abundance of the complemented mutants was also reduced compared to LH-7.

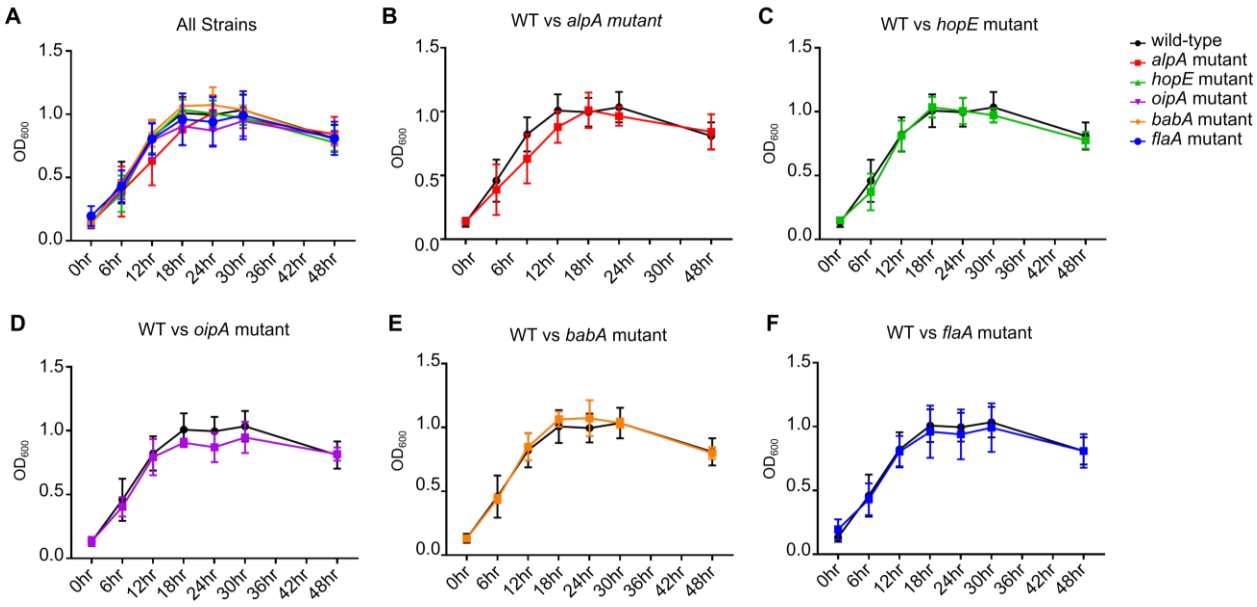

Fig. 3.8: Growth curve analysis of *omp* mutant strains.
The J166 wild-type strain and multiple mutant strains [*alpA* (LH-11$_a$), *hopE* (LH-13$_a$), *oipA* (LH-16$_a$), *babA* (LH-24$_a$), and *flaA* (LH-9$_a$) mutants] were cultured in liquid media and samples were collected every 6 hours for 48 hours. We used OD_{600} measurements to evaluate bacterial growth. Panel A shows the growth curves of all strains. Panels B-F show comparisons of wild-type strain J166 (black) with individual mutants for ease of interpretation. No significant differences were detected.

Table 3.1: Statistical analysis of results from experiments in which the control and OMP mutant libraries were passaged on blood agar plates for 21 days

comparison	[1]**mean diff**	[2]**99% CI**	[3]**p value**
LH-1 vs LH-2	0.399	[0.27, 0.530]	<0.001
LH-1 vs LH-3	0.455	[0.17, 0.74]	0.001
LH-1 vs LH-4	0.801	[0.53, 1.08]	<0.001
LH-1 vs LH-5	0.410	[0.28, 0.54]	<0.001
LH-1 vs LH-6	0.448	[0.33, 0.57]	<0.001
LH-2 vs LH-3	0.055	[-0.23, 0.34]	0.554
LH-2 vs LH-4	0.401	[0.13, 0.68]	0.001
LH-2 vs LH-5	0.011	[-0.11, 0.14]	0.797
LH-2 vs LH-6	0.049	[-0.06,0.16]	0.205
LH-3 vs LH-4	0.346	[0.00, 0.69]	0.010
LH-3 vs LH-5	-0.044	[-0.33, 0.24]	0.636
LH-3 vs LH-6	-0.007	[-0.29, 0.28]	0.942
LH-4 vs LH-5	-0.390	[-0.66, -0.12]	0.001
LH-4 vs LH-6	-0.352	[-0.63, -0.08]	0.002
LH-5 vs LH-6	0.038	[-0.07, 0.15]	0.334
ureA vs LH-7	-0.375	[-0.82, 0.07]	0.025
flaA vs LH-7	-1.899	[-2.34, -1.46]	<0.001
hopA vs LH-7	0.738	[0.26, 1.21]	<0.001
alpA vs LH-7	-1.343	[-1.78, -0.90]	<0.001
hopD vs LH-7	0.190	[-0.22, 0.60]	0.185
hopE vs LH-7	-2.453	[-2.92, -1.99]	<0.001
hopF vs LH-7	1.708	[1.30, 2.11]	<0.001
hopG vs LH-7	1.301	[0.89, 1.71]	<0.001
oipA vs LH-7	-2.257	[-2.88, -1.64]	<0.001
hopI vs LH-7	0.739	[0.32, 1.16]	<0.001
hopK vs LH-7	0.556	[0.07, 1.04]	0.004
hopL vs LH-7	0.135	[-0.28, 0.55]	0.350
hopN vs LH-7	-0.387	[-0.90, 0.12]	0.042
hopO vs LH-7	1.399	[0.98, 1.82]	<0.001
sabA vs LH-7	-0.268	[-0.67, 0.14]	0.063
hopQ vs LH-7	-0.341	[-1.01, 0.32]	0.152
babA vs LH-7	0.709	[0.27, 1.15]	<0.001
babB vs LH-7	0.819	[0.40, 1.24]	<0.001

Table 3.1 corresponds to the data presented in Figure 3.2C, which examines changes in composition of the control library and OMP library following 21 days of *in vitro* passaging on blood agar plates.
[1] Difference of mean $\log_{10}$ ratios of compared samples (output/input). Log_{10} ratios were calculated as described in Methods. For each comparison in the first column, the mean $\log_{10}$ ratio of the first mutant listed was subtracted from the mean $\log_{10}$ ratio of the second mutant. For instance, LH-1 vs LH-2 equals LH-1 minus LH-*2.*
[2] 99% confidence interval of the mean difference of $\log_{10}$ ratios (output/input).
[3] p-value from two sample t-test of pairwise comparisons.

Table 3.2: Statistical analysis of results from experiments in which OMP mutant libraries were passaged on multiple types of media *in vitro*

	Blood Agar			BSFB with Regular Salt			BSFB with High Salt		
Comparison	**[1]mean diff**	**[2]99% CI**	**[3]p value**	**[1]mean diff**	**[2]99% CI**	**[3]p value**	**[1]mean diff**	**[2]99% CI**	**[3]p value**
ureA vs LH-7	-0.29	[-0.62, 0.03]	0.015	0.45	[-0.34, 1.23]	0.092	-0.12	[-0.72, 0.48]	0.501
flaA vs LH-7	-1.75	[-2.16, -1.33]	<0.001	-0.65	[-1.88, 0.58]	0.101	-1.38	[-2.40, -0.36]	0.003
hopA vs LH-7	0.8	[0.41, 1.19]	<0.001	0.98	[0.23, 1.72]	0.002	0.71	[-0.06, 1.48]	0.014
alpA vs LH-7	-1.18	[-1.52, -0.83]	<0.001	-0.24	[-1.03, 0.55]	0.329	-0.64	[-1.34, 0.06]	0.015
hopD vs LH-7	0.53	[0.19, 0.87]	0.001	0.56	[-0.13, 1.25]	0.025	-0.42	[-1.05, 0.22]	0.052
hopE vs LH-7	-2.13	[-2.60, -1.67]	<0.001	-1.61	[-2.37, -0.85]	<0.001	-3.83	[-5.53, -2.13]	<0.001
hopF vs LH-7	2.04	[1.72, 2.37]	<0.001	1.96	[1.23, 2.68]	<0.001	1.26	[0.73, 1.79]	<0.001
hopG vs LH-7	1.58	[1.24, 1.92]	<0.001	1.15	[0.39, 1.91]	0.001	1.16	[0.69, 1.64]	<0.001
oipA vs LH-7	-1.86	[-2.34, -1.39]	<0.001	-2.32	[-3.03, -1.6]	<0.001	-2.96	[-4.33, -1.59]	<0.001
hopI vs LH-7	1.08	[0.75, 1.42]	<0.001	1.06	[0.05, 2.06]	0.008	-0.97	[-1.52, -0.43]	<0.001
hopK vs LH-7	0.61	[0.26, 0.96]	<0.001	0.36	[-0.45, 1.17]	0.174	-0.54	[-0.97, -0.10]	0.003
hopL vs LH-7	0.32	[-0.03, 0.68]	0.016	1.51	[0.76, 2.27]	<0.001	1.6	[1.09, 2.11]	<0.001
hopN vs LH-7	-0.4	[-0.75, -0.05]	0.005	-0.16	[-0.85, 0.52]	0.425	0.02	[-0.64, 0.67]	0.922
hopO vs LH-7	1.77	[1.42, 2.12]	<0.001	2.67	[1.97, 3.37]	<0.001	1.88	[1.46, 2.31]	<0.001
sabA vs LH-7	-0.01	[-0.33, 0.31]	0.917	0.17	[-1.00, 1.35]	0.614	-1.18	[-1.57, -0.79]	<0.001
hopQ vs LH-7	-0.51	[-0.87, -0.14]	0.001	-0.36	[-1.12, 0.4]	0.151	-0.36	[-0.76, 0.04]	0.017
babA vs LH-7	0.8	[0.47, 1.12]	<0.001	0.68	[-0.29, 1.64]	0.045	-0.79	[-1.43, -0.16]	0.003
babB vs LH-7	1.15	[0.82, 1.47]	<0.001	1.05	[0.35, 1.75]	0.001	0.29	[-0.12, 0.69]	0.046

Table 3.2 corresponds to the data presented in Figure 3.3, which examines the changes in composition of the OMP library following 21 days of *in vitro* passaging on various types of solid culture media (blood agar, BSFB with regular salt, or BSFB with high salt).
[1] Difference of mean $\log_{10}$ ratios of compared samples (output/input). For each comparison in the first column, the mean of $\log_{10}$ ratio of the first mutant listed was subtracted from the mean $\log_{10}$ ratio of the second mutant. For instance, *ureA* vs LH-7 equals *ureA* minus LH-7.
[2] 99% confidence interval of the mean difference of $\log_{10}$ ratios (output/input).
[3] p-value from two sample t-test of pairwise comparisons.

Table 3.3 Statistical analysis of results from experiments in which barcoded libraries composed of original or replicate *hop* mutants were passaged on blood agar plates for 21 days

	Barcoded Library with Original mutants			*Barcoded Library with Replicate Mutants*		
comparison	**[1]mean diff**	**[2]99% CI**	**[3]p value**	**[1]mean diff**	**[2]99% CI**	**[3]p value**
LH-1 vs LH-2	0.189	[-0.08, 0.46]	0.057	0.308	[0.20, 0.42]	<0.001
LH-1 vs LH-3	0.369	[0.16, 0.58]	<0.001	0.244	[0.12, 0.37]	<0.001
LH-1 vs LH-4	0.692	[0.47, 0.91]	<0.001	0.632	[0.52, 0.75]	<0.001
LH-1 vs LH-5	0.468	[0.21, 0.72]	<0.001	0.402	[0.28, 0.52]	<0.001
LH-1 vs LH-6	0.418	[0.18, 0.65]	<0.001	0.411	[0.29, 0.53]	<0.001
LH-2 vs LH-3	0.180	[-0.09, 0.45]	0.070	-0.064	[-0.17, 0.04]	0.069
LH-2 vs LH-4	0.502	[0.23, 0.78]	<0.001	0.324	[0.26, 0.39]	<0.001
LH-2 vs LH-5	0.279	[-0.02, 0.58]	0.015	0.093	[0.01, 0.18]	0.007
LH-2 vs LH-6	0.228	[-0.06, 0.51]	0.033	0.103	[0.02, 0.19]	0.003
LH-3 vs LH-4	0.323	[0.10, 0.55]	0.001	0.388	[0.28, 0.50]	<0.001
LH-3 vs LH-5	0.099	[-0.16, 0.35]	0.265	0.157	[0.04, 0.27]	0.001
LH-3 vs LH-6	0.049	[-0.19, 0.28]	0.553	0.167	[0.05, 0.28]	0.001
LH-4 vs LH-5	-0.224	[-0.49, 0.04]	0.023	-0.231	[-0.33, -0.13]	<0.001
LH-4 vs LH-6	-0.274	[-0.52, -0.03]	0.005	-0.221	[-0.32, -0.13]	<0.001
LH-5 vs LH-6	-0.050	[-0.32, 0.22]	0.589	0.010	[-0.10, 0.12]	0.792
alpA vs LH-7	-1.796	[-2.33, -1.26]	<0.001	-1.238	[-1.96, -0.52]	<0.001
hopE vs LH-7	-3.059	[-3.69, -2.43]	<0.001	-2.066	[-2.57, -1.56]	<0.001
hopF vs LH-7	1.276	[1.05, 1.50]	<0.001	-0.482	[-0.95, -0.01]	0.009
oipA vs LH-7	-3.685	[-4.04, -3.33]	<0.001	-2.992	[-3.93, -2.06]	<0.001
babA vs LH-7	0.684	[0.50, 0.87]	<0.001	-0.098	[-0.27, 0.08]	0.096

Table 3.3 corresponds to the data presented in Figure 4.1E and F, which examine the changes in composition of the barcoded libraries composed of original or replicate *hop* mutant strains following 21 days of *in vitro* passaging on blood agar plates.
[1] Difference of mean $\log_{10}$ ratios of compared samples (output/input). For each comparison in the first column, the mean of $\log_{10}$ ratio of the first mutant listed was subtracted from the mean $\log_{10}$ ratio of the second mutant. For instance, LH-1 vs LH-2 equals LH-1 minus LH-2.
[2] 99% confidence interval of the mean difference of $\log_{10}$ ratios (output/input).
[3] p-value from two sample t-test of pairwise comparisons.

Table 3.4: Statistical analysis of pairwise competitions of a control mutant with original or replicate *hop* mutants *in vitro*

comparison	**Mean**	**99% CI**	**p value**
alpA vs LH-7	-1.190	[-2.13, -0.25]	0.004
babA vs LH-7	-0.037	[-0.20, 0.13]	0.154
hopE vs LH-7	-2.477	[-4.41, -0.55]	0.005
hopF vs LH-7	0.212	[-2.25, 2.67]	0.711
oipA vs LH-7	-3.462	[-4.98, -1.94]	<0.001
flaA vs LH-7	-1.050	[-1.80, -0.30]	0.002

Table 3.4 corresponds to the data presented in Figure 3.5, which examines the changes in the relative abundance of the control mutant (LH-7) versus an original or replicate *hop* mutant following 21 days of pairwise competitions *in vitro*. These changes were examined using single-plex qPCR with mutant-specific primers.
[1] Mean of the relative abundance ratio of output to input of mutant/LH-7. A negative number indicates that the relative abundance of the *hop* mutant is less than the control (LH-7), whereas a positive number indicates that the relative abundance of the *hop* mutant is greater than the control (LH-7).
[2] 99% confidence interval of the mean.
[3] p-value from one sample t-test.

Table 3.5: Primer sequences used for qPCR analysis

Primer names	***Sequence (5' to 3')***
LH-7 *forward*	ACGTGACTACTCGAACGTGAC
LH-7 *reverse*	CAGAGCTCCCGAACGCTTTG
alpA forward	GCTCCCTAGCTTGACTGACAGG
Cassette reverse	GTGAATTGCATCATTTATCCTCCG
hopE forward	TGTACGGACGTGACCGTCGAC
hopE reverse	CATCAGGACCTTTAGCTTCAATG
hopF forward	GGTGGTGGATGTAGGGAAAGAAC
Cassette reverse	GTGAATTGCATCATTTATCCTCCG
oipA forward	CCGTACGACGTGACCTACGTT
oipA reverse	GTATTAGCGTCTAGCGTTCTGCC
babA forward	GACTTGCTGAATAAATAATTCCGACG
babA reverse	CTTTGAGCGCGGGTAAGCC

Table 3.6: Statistical analysis of results from experiments in which mice were infected with the control or OMP mutant libraries for 21 days

comparison	[1]**mean diff**	[2]**99% CI**	[3]**p value**
LH-1 vs LH-2	-2.129	[-5.19, 0.93]	0.061
LH-1 vs LH-3	-1.464	[-4.60, 1.68]	0.196
LH-1 vs LH-4	0.223	[-2.75, 3.20]	0.831
LH-1 vs LH-5	-1.822	[-4.97, 1.33]	0.113
LH-1 vs LH-6	-1.570	[-4.63, 1.49]	0.156
LH-2 vs LH-3	0.665	[-2.39, 3.72]	0.538
LH-2 vs LH-4	2.352	[-0.53, 5.23]	0.030
LH-2 vs LH-5	0.307	[-2.75, 3.37]	0.776
LH-2 vs LH-6	0.559	[-2.41, 3.52]	0.594
LH-3 vs LH-4	1.687	[-1.28, 4.65]	0.118
LH-3 vs LH-5	-0.358	[-3.50, 2.78]	0.746
LH-3 vs LH-6	-0.106	[-3.15, 2.94]	0.921
LH-4 vs LH-5	-2.045	[-5.02, 0.93]	0.063
LH-4 vs LH-6	-1.793	[-4.66, 1.08]	0.089
LH-5 vs LH-6	0.252	[-2.80, 3.31]	0.815
ureA vs LH-7	-3.281	[-4.31, -2.25]	<0.001
flaA vs LH-7	-3.366	[-4.38, -2.36]	<0.001
hopA vs LH-7	-0.048	[-1.46, 1.37]	0.928
alpA vs LH-7	-3.326	[-4.34, -2.31]	<0.001
hopD vs LH-7	-1.740	[-2.93, -0.55]	<0.001
hopE vs LH-7	-3.352	[-4.36, -2.34]	<0.001
hopF vs LH-7	-1.736	[-2.86, -0.61]	<0.001
hopG vs LH-7	-2.418	[-3.49, -1.34]	<0.001
oipA vs LH-7	-2.001	[-3.24, -0.76]	<0.001
hopI vs LH-7	-2.975	[-4.04, -1.91]	<0.001
hopK vs LH-7	-0.099	[-1.42, 1.23]	0.844
hopL vs LH-7	-1.006	[-2.22, 0.21]	0.031
hopN vs LH-7	-0.391	[-1.56, 0.78]	0.380
hopO vs LH-7	-2.676	[-3.76, -1.59]	<0.001
sabA vs LH-7	-2.798	[-3.88, -1.72]	<0.001
hopQ vs LH-7	0.151	[-1.12, 1.42]	0.755
babA vs LH-7	4.550	[3.35, 5.75]	<0.001
babB vs LH-7	-0.590	[-1.95, 0.77]	0.255

Table 3.6 corresponds to the data presented in Figure 4.2C, which examines the changes in composition of the control library and OMP mutant library after 21 days in mice.

[1] Difference of mean $\log_{10}$ ratios of compared samples (output/input). For each comparison in the first column, the mean of $\log_{10}$ ratio of the first mutant listed was subtracted from the mean $\log_{10}$ ratio of the second mutant. For instance, LH-1 vs *clt-2* equals LH-1 minus LH-2.

[2] 99% confidence interval of the mean difference of $\log_{10}$ ratios (output/input).

[3] p-value from two sample t-test of pairwise comparisons.

CHAPTER IV

HOP OUTER MEMBRANE PROTEINS AS DETERMINANTS OF *HELICOBACTER PYLORI* FITNESS *IN VIVO*

A modified version was previously published as:

Harvey ML, Lin AS, Sun L, Koyama T, Shuman JHB, Loh JT, Scott Algood HM, Scholz MB, McClain MS, Cover TL. Enhanced fitness of a *Helicobacter pylori babA* mutant in a murine model. Infect Immun. 2021 Jul 26:IAI0072520. doi: 10.1128/IAI.00725-20. Epub ahead of print. PMID: 34310886.

Introduction

Data presented in Chapter III demonstrated that the barcoded control strains have equal fitness in a competitive *in vitro* environment. One goal of the work presented in Chapter IV is to determine if *H. pylori* experiences a non-selective bottleneck *in vivo.* Changes in *H. pylori* colonization density in parallel with various stages of immune system activation suggests the presence of at least one bottleneck mediated by host-immune responses (118). Another goal is to identify Hops that play a role in *H. pylori* fitness *in vivo,* but not *in vitro.* Previous studies have shown that several *H. pylori* OMPs play a role in gastric colonization. Among these, AlpA/B have been shown to be ubiquitously expressed in isolates from both human and mouse models (57, 74). Mutagenesis of *alpA/alpB* has been shown to lead to a colonization deficiency in the stomachs of guinea pigs and Mongolian gerbils, and to poorly colonize the stomachs of mice (80, 81). Similarly, mutagenesis of *oipA* has been shown to lead to decreased colonization density and lessened neutrophil infiltration compared to strains expressing *oipA* in the murine model (67). The roles of other *H. pylori* Hops in gastric colonization are understudied.

Materials and Methods

Orogastric infection of mice with barcoded libraries

Male conventional C57BL/6 mice, 6-8 weeks old, were used in all studies (The Jackson Laboratory). *H. pylori* input pools were generated by resuspending individual barcoded mutants in pre-warmed BSFB broth. Each barcoded mutant was then combined at equal proportions; the final concentration of the combined mutants was 1.0×10^9 cells/ml in BSFB broth. Mice were anesthetized and infected with 0.5 ml of *H. pylori* input pools via oral gavage. Mice were sacrificed at 21- or 90-days post-infection. After euthanasia, the stomachs were excised, the forestomach removed, and the remaining stomach opened along the minor curvature. Stomach contents were washed away with sterile PBS and stomach tissue was homogenized in 500 µl of brucella broth for 10 minutes using a Bullet Blender (Next Advance). 100 µl of homogenized stomach tissue was cultured on TSA plates supplemented with 5% sheep blood, nalidixic acid [10 µg/ml], vancomycin [50 µg/ml], amphotericin [20 µg/ml], and bacitracin [100 µg/ml]. Samples were processed as described below.

Pairwise Competition Assays

The LH-7 control strain was combined 1:1 with original *babA* mutants (LH-24_a) or subsequently generated mutant (LH-24_b). Mice were infected with 1:1 mixtures of LH-7 competed with either the original or independent *babA* mutant (LH-24_a or LH24_b). Subsequently, we extracted genomic DNA from the input and outputs using phenol-chloroform-based methods (116). Real-time qPCR analysis of DNA samples was performed with SYBR green fluorophore (iQ SYBR Green Supermix; Bio-Rad) on an ABI

StepOnePlus machine. Primer sequences can be found in Table 3.5. A standard curve of each DNA target was generated using 10-fold dilutions starting at 50 ng/well. The abundance of individual strains was calculated using the appropriate standard curve for each DNA target.

Statistical analysis

For analysis of *in vivo* experiments, thresholds of -2 and 2 were used as a threshold for defining marked differences of mean $\log_{10}$ ratios, corresponding >100-fold differences (Fig. 4.2C). Strains with a p value less than 0.01 (when comparing to LH-7) and also a 99% CI that does not overlap the $\pm$ 2 threshold were considered significant. In experiments without a corresponding control library for comparison, we required a 99% confidence interval to be wholly above or below 0, which is equivalent to a multiplicity-controlled type I error rate of 1% (Fig. 3.3). Finally, for analysis of qPCR data, a one-sample t-test was used to estimate log-transformed OMP/LH-7 ratios with a 99% confidence interval (Fig. 4.3, Table 4.1). All confidence intervals and p-values are reported in supplemental table 4.1.

Results

Tracking changes in composition of the control library *in vivo*

We next undertook experiments to analyze the fitness of *H. pylori* strains *in vivo*. To assess changes in a population of control strains *in vivo*, we infected mice with the control library for 21 days (corresponding to the same time period analyzed for the *in vitro* experiments). Animals were euthanized, *H. pylori* strains were cultured from the stomachs, and the composition of the resulting populations was compared to the

composition of the input libraries (Fig 2.1B). As seen in Figure 4.1A, we detected barcodes corresponding to each of the strains in the population. These data confirm that multiple *H. pylori* strains can co-colonize the same stomach. Notably, the distribution of barcode counts in individual mice was highly variable. In each case, the composition of the *H. pylori* population recovered from individual mice was markedly different from the input population, and populations from individual mice were markedly different from populations of other mice (Fig. 4.1A). This variation occurred in a random manner, with no statistically significant differences detected when comparing individual control strains with other control strains (Fig. 4.1 and Table 3.6). Therefore, in contrast to the relatively stable composition of the control library observed during passage *in vitro*, we consistently observed marked changes in the composition of the control library when it was introduced into murine stomachs (compare Fig. 4.1B with Fig. 3.1B). These findings are consistent with the existence of a non-selective bottleneck *in vivo*.

Testing the fitness of hop mutant strains *in vivo*

We next infected mice with the OMP mutant library and analyzed output pools cultured from the murine stomachs at 21 days post-infection. Similar to the results of *in vivo* experiments with the control library, we observed a high variability in the distribution of barcode counts in individual mice, and in each case, the composition of the population recovered from individual mice was markedly different from the composition of the input population (Fig. 4.2). Consistent with previous reports that *ureA* and *flaA* are required for gastric colonization (38-40, 106, 111), we detected very low numbers of barcode counts for the *ureA* and *flaA* mutants. Interestingly, we did detect barcode counts specific for *ureA* or *flaA* mutants at levels substantially above the limit of detection (defined in

Methods) in several of the animals (Fig. 4.2B, Table 3.6). The *babA* mutant was the dominant strain (accounting for >50% of total barcode counts) in output pools from the majority of the mice (Fig. 4.2, Table 3.6). The same result was observed in mice infected with independent input pools. A similar dominance of the *babA* mutant was observed in mice infected for 90 days (Fig. 4.4). At this time point (90 days post-infection), barcode counts corresponding to the *babA* mutant accounted for >95% of counts from 4 out of 5 mice (Fig. 4.4). Statistical analysis was complicated by a very high level of variability in results for individual animals, which is presumably attributable to the bottleneck phenomenon observed in Fig. 4.1. In a statistical analysis using the approach employed for Fig. 3.2, none of the OMP mutants met the criteria for statistically significant changes. Nevertheless, we were able to identify *hop* mutants that exhibited marked changes in numbers of barcode counts compared to the control strain (LH-7). Specifically, there was a >100-fold reduction in the ratios of output/input barcode counts for *alpA* and *hopE* mutants compared to the control strain (Fig 4.2, Table 3.6). The numbers of barcode counts corresponding to *alpA* and *hopE* mutants were similar to the numbers of barcode counts corresponding to *ureA* and *flaA* mutants. The *babA* mutant exhibited an increased abundance of similar magnitude (>100-fold increase).

When comparing the results of experiments conducted *in vivo* with results of experiments conducted *in vitro* (compare Figure 3.2 with Figure 4.2), the *ureA* mutant exhibited a fitness defect *in vivo* but not *in vitro*, as expected. We did not identify any *hop* mutants that exhibited unequivocal fitness defects exclusively *in vivo* without corresponding fitness defects *in vitro*. Interestingly, the *babA* mutant exhibited a strong fitness advantage *in vivo* compared to control strains and other mutants, but the *babA*

mutant did not exhibit a corresponding fitness advantage *in vitro*. The *babA* mutant accounted for greater than 50% of nucleotide bar counts in bacterial populations recovered from 26/40 (65%) mice, but did not exhibit a similar fitness advantage in 10 experiments conducted *in vitro* ($p = 0.0002$). Therefore, the *babA* mutant exhibited a fitness advantage *in vivo* but not *in vitro.*

To further evaluate the *in vivo* fitness advantage observed with the *babA* mutant, we mixed the LH-7 control strain with either the original *babA* mutant (LH-24_a) or the subsequently constructed *babA* mutant (LH-24_b) in equal proportions and passaged these mixtures on blood agar plates for 21 days. In parallel, the mixtures were introduced into mouse stomachs via oral gavage, and the gastric *H. pylori* populations were assessed at 21 days post-infection. We then analyzed the relative abundance of the *babA* mutants compared to the control strain by qPCR as described in Methods. As expected, there were no significant differences when comparing the relative abundance of the *babA* mutants with the control strain *in vitro* (Fig. 4.3). In mice, however, we observed that both *babA* mutants outcompeted the control strain (Fig. 4.3), confirming that the *babA* mutant exhibited enhanced fitness *in vivo* but not *in vitro*.

Discussion

As an initial step in undertaking experiments *in vivo,* we orogastrically infected mice with the control library. Following 21 days of infection, we observed that all of the control library strains were detected in each stomach, indicating that multiple strains of *H. pylori* can colonize that same stomach. Notably, the proportional abundance of each mutant varied from mouse to mouse, suggesting that strong non-selective pressures influenced the population colonizing each mouse. This is consistent with the results of

previous studies, which suggested that *H. pylori* experiences a bottleneck early in infection. Thus, the animal experiments with the control library provided insight into the non-selective environmental pressures encountered *in vivo* and established a foundation for understanding the range of experimental variability that can be expected in the absence of functional strain-specific differences.

Next, we compared the fitness of *hop* mutants *in vivo* compared with the fitness of the corresponding mutants *in vitro.* We anticipated that we would identify *hop* mutants required for colonization of the stomach but not required for *H. pylori* fitness *in vitro.* A *ureA* mutant exhibited the expected behavior (essential for fitness *in vivo* but not *in vitro*), which indicated that the methodology could identify such mutants. Unexpectedly, we did not identify any *hop* mutants that exhibited fitness defects *in vivo* but not *in vitro.* Functional redundancy among Hop proteins might account for the failure to identify any mutants that exhibited these characteristics. Another possibility is that the colonization advantage of the *babA* mutant might have hindered efforts to detect mutant that have reduced fitness *in vivo.* In future studies, it might be informative to conduct experiments in which *babA* mutants are excluded. This is discussed as a future direction in Chapter 6. There were several other challenges associated with analysis of the barcoded libraries *in vivo.* PCR amplification of the barcoded regions from the mouse stomachs was not consistently successful, so we cultured *H. pylori* from the mouse stomachs and analyzed the cultured bacteria. Consequently, there could be bias related to detection of strains that had properties favoring efficient culture from the mouse stomachs. Conversely, a strength of this approach was that it assured that the barcodes detected originated from viable organisms. The second limitation was that the numbers of individual *H. pylori*

colonies cultured from stomachs varied considerably among mice. To reduce bias, we restricted our analysis to mice from which at least 100 *H. pylori* colonies were cultured.

Although we were not able to identify any mutants that exhibited a fitness defect *in vivo* but not *in vitro,* we found that the *babA* mutant exhibited a fitness advantage *in vivo* but not *in vitro.* Several previous studies observed that the BabA adhesin is selected against (via the loss of *babA* expression or loss of Lewis^b binding) during experimental *H. pylori* infection of mice, Mongolian gerbils, or rhesus macaques (51, 52, 61-63). The current results are consistent with such observations.

In *H. pylori* strains isolated from humans, *babA* may be either present or absent, and in some instances when the gene is present, the BabA protein is not produced or when produced, it does not bind to Lewis^b (27, 50-52). We presume that there are selective pressures in humans that favor retention of *babA,* at least under some circumstances. Previous studies have shown that BabA is important for initial *H. pylori* attachment to the human gastric mucosa and allows for reversible attachment as gastric cells are shed off into the lumen. Therefore, BabA might confer a benefit to *H. pylori* at specific stages of infection in humans but might be deleterious at other stages of infection. Alternatively, BabA may confer a benefit in individuals who produce Le^b antigen (the BabA receptor) but not in individuals who do not produce Le^b. The results of the current study, along with previous studies, suggest that the presence of *babA* is deleterious in the mouse model. The mechanisms leading to selection against *babA*-positive strains in mice (in contrast to retention of *babA*-positive strains in humans, at least under some circumstances) are not yet known. Previous studies have shown that BabA is selected against in stomachs of rhesus macaques, Mongolian gerbils, and mice, suggesting the

presence of a selective pressure against retention of BabA in the stomach (51, 61). One of the relevant differences between the human and mouse gastric environments is that mice lack Le^b (the BabA receptor). A previous study showed that the loss of BabA is not dependent on the expression of Le^b in the murine model, nor on BabA's ability to bind to Le^b (52). We presume that the presence of *babA* confers both benefits and costs to *H. pylori,* and retention or loss of the gene is dependent on the balance of benefits and costs in varying gastric environments.

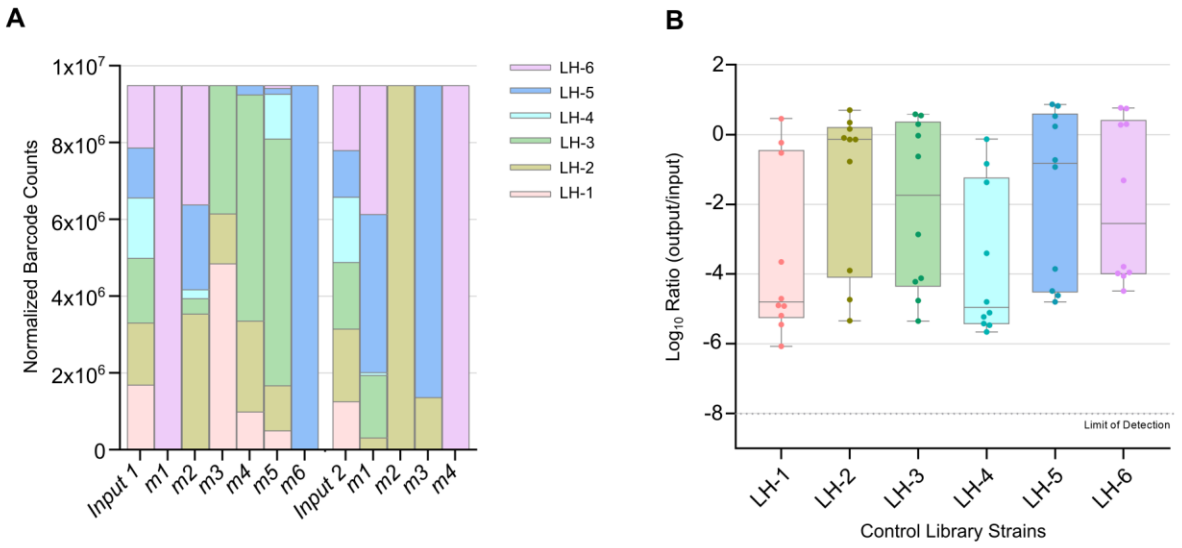

Figure 4.1: Multiple strains of *H. pylori* can colonize the same stomach.
C57Bl/6J mice were infected with control libraries for 21 days in two independent experiments. *H. pylori* was cultured from the stomachs and the composition of input and output pools was analyzed as described in Methods. The labels m1-m6 indicate results for individual mice in separate experiments. (A, B) Panel A shows relative proportions of barcode counts and panel B shows $\log_{10}$ ratios comparing the output pools with the input. Statistical analysis is shown in Table 3.6. All control strains (LH-1 through LH-6) within these libraries could be detected after 21 days, but the numbers of barcode counts corresponding to each mutant were highly variable from mouse to mouse.

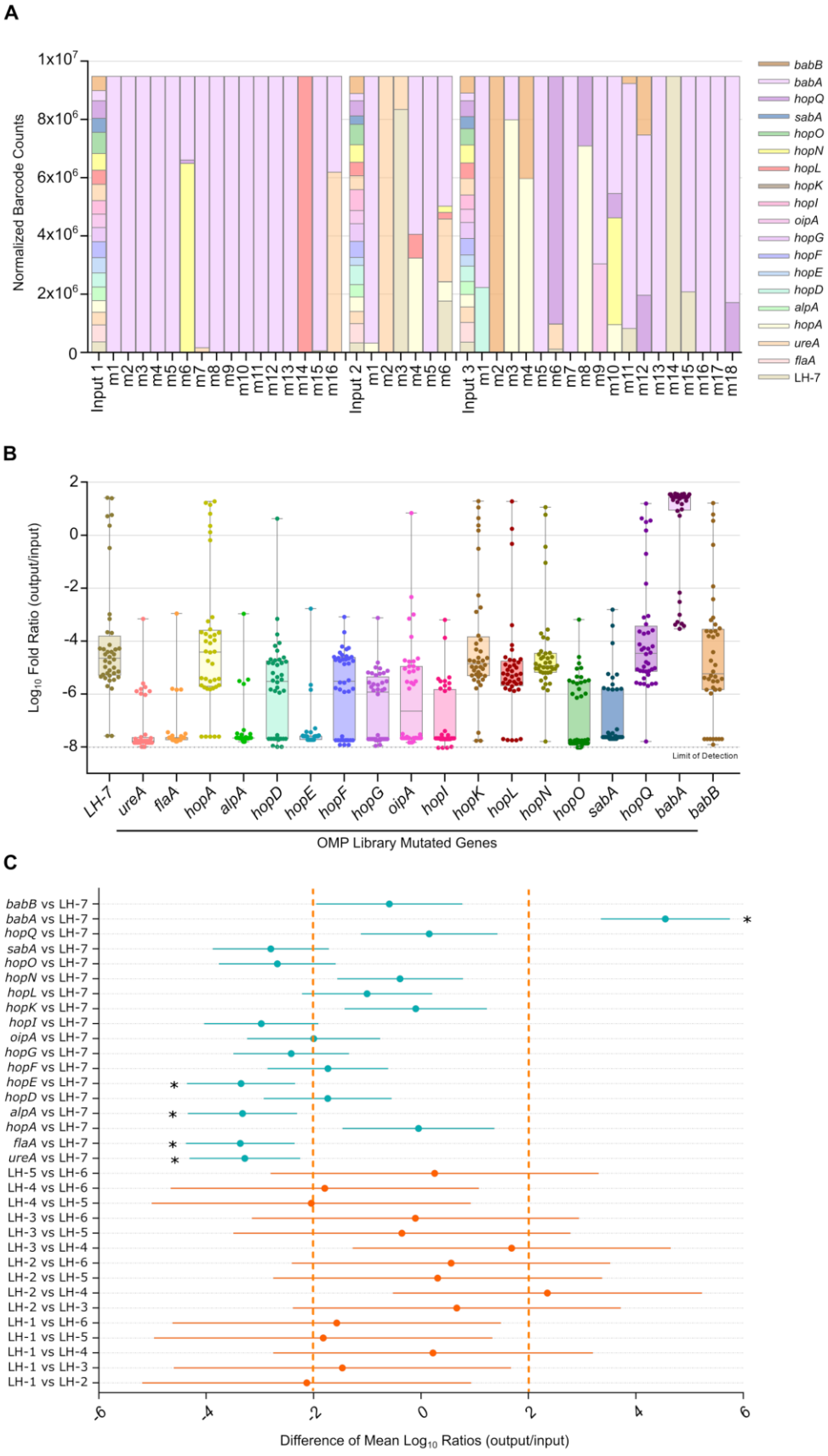
A
Normalized Barcode Counts
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
Input 1
Input 2
Input 3
babB
babA
hopQ
sabA
hopO
hopN
hopL
hopK
hopI
oipA
hopG
hopF
hopE
hopD
alpA
hopA
ureA
flaA
LH-7
B
Log10 Fold Ratio (output/input)
Limit of Detection
LH-7
ureA
flaA
hopA
alpA
hopD
hopE
hopF
hopG
oipA
hopI
hopK
hopL
hopN
hopO
sabA
hopQ
babA
babB
OMP Library Mutated Genes
C
babB vs LH-7
babA vs LH-7
hopQ vs LH-7
sabA vs LH-7
hopO vs LH-7
hopN vs LH-7
hopL vs LH-7
hopK vs LH-7
hopI vs LH-7
oipA vs LH-7
hopG vs LH-7
hopF vs LH-7
hopE vs LH-7
hopD vs LH-7
alpA vs LH-7
hopA vs LH-7
flaA vs LH-7
ureA vs LH-7
LH-5 vs LH-6
LH-4 vs LH-6
LH-4 vs LH-5
LH-3 vs LH-6
LH-3 vs LH-5
LH-3 vs LH-4
LH-2 vs LH-6
LH-2 vs LH-5
LH-2 vs LH-4
LH-2 vs LH-3
LH-1 vs LH-6
LH-1 vs LH-5
LH-1 vs LH-4
LH-1 vs LH-3
LH-1 vs LH-2
Difference of Mean Log10 Ratios (output/input)

Figure 4.2: Fitness advantage of a *babA* mutant *in vivo*.
C57Bl/6J male mice were infected with OMP mutant libraries for 21 days in three independent experiments. *H. pylori* was cultured from the stomachs and the composition of input and output pools was analyzed as described in Methods. The labels m1-m18 indicate results for individual mice in separate experiments. (A,B) Panel A shows relative proportions of barcode counts and panel B shows $\log_{10}$ ratios relative to the input. At the end of the experiment, half of the infected mice had >50% of total counts originating from the *babA* (LH-24_a) mutant barcode. (C) If the 99% confidence interval was wholly above 2 or below -2 in the $\log_{10}$ scale, then the OMP mutant strain was considered significantly different from the control strain (LH-7). Statistical analysis is shown in Table 3.6. Using this approach, we determined that the relative abundance of *alpA* (LH-11_a), *hopE* (LH-13_a), *ureA* (LH-8), and *flaA* (LH-9_a) mutants was decreased, whereas the relative abundance of *babA* mutant was increased relative to LH-7. * indicates >100-fold difference in ratio of output compared to input.

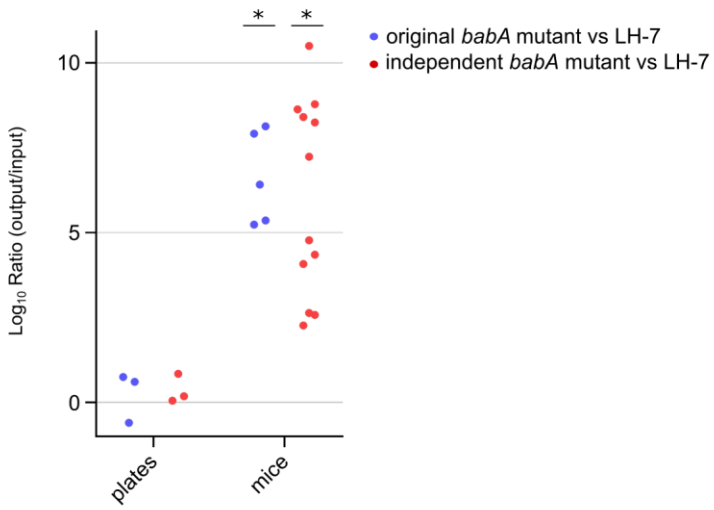

Figure 4.3: Fitness of *babA* mutants *in vitro* and *in vivo* evaluated by qPCR.
Each *babA* mutant [original (LH-24$_a$) and replicate (LH-24$_b$) mutants) was mixed with the control strain (LH-7) in a 1:1 ratio. The mixtures were passaged on blood agar plates for 21 days and also used for infection of mice, which were euthanized at 21 days post-infection. We used qPCR to determine the relative abundance of each strain in the input and output populations, based on comparison to a standard curve generated for each target. The $\log_{10}$ ratio (output/input) was calculated based on the relative abundance of each mutant compared to the control strain in the output and input populations. At the end of the experiment, the relative abundance of the original (LH-24$_a$) and independent (LH-24$_b$) *babA* mutants remained similar to that of LH-7 *in vitro* but was increased in mice. The changes in abundance of the *babA* mutants in mice were statistically significant (asterisks indicate p value <0.001). Statistical analysis is shown in Table 4.1.

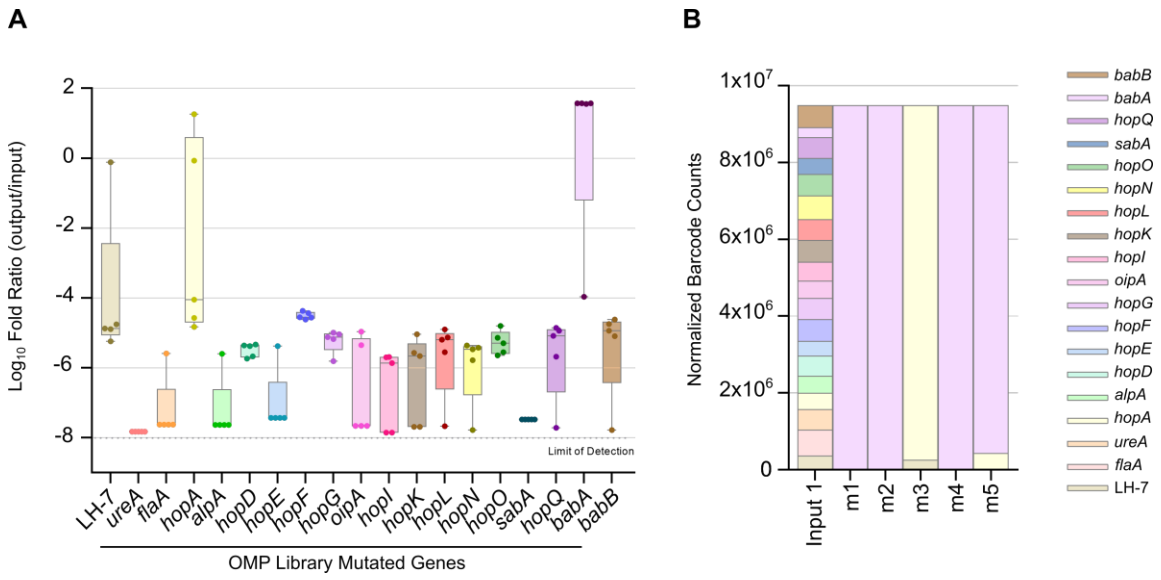

Figure. 4.4: Fitness advantage of a *babA* mutant *in vivo* after 90-day infection.
In parallel with studies done in Figure 5, five C57Bl/6J mice were infected with the first OMP mutant input library (Input 1) and euthanized at 90 days post-infection. *H. pylori* was cultured from the stomachs and the composition of input and output pools was analyzed as described in Methods. (A,B) and Panel A shows $\log_{10}$ ratios (output/input) and panel B shows the relative proportions of barcode counts individual animals. At the end of the experiment, 4 out of 5 mice had >95% of total barcode counts originating from the *babA* mutant (LH-24_a).

Table 4.1: Statistical analysis of pairwise competitions of a control mutant with original or replicate *babA* mutants *in vitro* and *in vitro.*

Comparison in vitro	**[1]mean diff**	**[2]99% CI**	**[3]p value**
babA vs LH-7	0.304	[-0.59, 1.20]	0.229
***Comparison* in mice**	**[1]mean diff**	**[2]99% CI**	**[3]p value**
babA vs LH-7	6.206	[4.43, 7.98]	<0.001

Table 4.1 corresponds to the data presented in Figure 6, which examines the changes in the relative abundance of the control mutant (LH-7*)* versus an original or replicate *babA* mutant following 21 days of pairwise competitions *in vitro* and in mice. These changes were examined using single-plex qPCR with mutant-specific primers.
[1] Mean of the relative abundance ratio of output to input of mutant/LH-7. A negative number indicates that the relative abundance of the *babA* mutant is less than the control (LH-7), whereas a positive number indicates that the relative abundance of the *babA* mutant is greater than the control (LH-7).
[2] 99% confidence interval of the mean.
[3] p-value from one sample t-test.

CHAPTER V

PROJECTS IN PROGRESS

There are multiple projects that are extensions of work described in the previous chapters.

Fitness of *babA* mutants in multiple *H. pylori* strains

Previous studies have shown that the BabA adhesin is selected against during experimental *H. pylori* infection of animal models, including mice, Mongolian gerbils, and rhesus macaques (62). Data presented in Chapters III and IV build upon these findings by demonstrating that mutagenesis of *babA* leads to a fitness advantage *in vivo,* but not *in vitro.*

To further evaluate the fitness advantage of a *babA* mutant, we infected C57Bl/6J mice with monocultures of a J166 control mutant (LH-7) or a J166 *babA* mutant for 21 days. After 21 days, mice were sacrificed, and stomachs were plated for CFU counts. We found that mice infected with either J166 LH-7 or the *babA* mutant had similar colonization densities (Figure 5.1). Therefore, the *babA* mutant outcompeted other strains in mixed infections, but it did not achieve a higher colonization density than the control strain when the strains were tested individually.

Next, we investigated whether the fitness advantage rendered by mutagenesis of *babA* in *H. pylori* strain J166 would be recapitulated in alternate *H. pylori* strain backgrounds. To test this, we generated a control mutant (LH-7) and a *babA* mutant in *H. pylori* strain G27. C57Bl/6J mice were infected with monocultures of G27 wild-type, G27 LH-7 and the G27 *babA* mutant for 16 days. We found that both G27 wild-type and G27 LH-7 were unable to colonize mice. Interestingly, the G27 *babA* mutant strain was able

to colonize mice to levels similar to J166 (Figure 5.2). Follow-up experiments to investigate multiple aspects of the role of BabA in *H. pylori* fitness are described in Future Directions in Chapter VI.

Biofilm studies

Biofilms are surface-attached microbial communities. Biofilm formation is initiated by attachment of single cells to a surface, and bacterial adhesins have been shown to play an important role in biofilm formation. Previous literature has reported that *alpB* and *homB* contribute to *H. pylori* biofilm formation *in vitro* (on glass cover slips and in tissue culture treated plates) (79, 119). Flagella have also been reported to be important for *H. pylori* biofilm formation (120-122). Our goal was to identify specific Hops that contribute to biofilm formation.

To test this, we generated a barcoded control library and OMP mutant library in the parental wild-type strain G27 using the barcoded plasmids described in previous chapters. *H. pylori* strain G27 was chosen because it is capable of biofilm formation *in vitro,* unlike strain J166*.* Next, we assessed each mutant's ability to form biofilm *in vitro* by growing monocultures of each mutant strain in brucella broth supplemented with 10% fetal bovine serum in tissue culture-treated plates. Cultures were grown for 4 days in room air supplemented with 5% CO_2 at 37ºC. Subsequently, planktonic bacteria were washed away, and the surface-attached biofilm was stained with crystal violet. OD_{600} measurements were taken to quantify the total biofilm formed by each mutant compared to a G27 wild-type control. As expected, the G27 control mutants were able to form biofilm at a level similar to that of G27 WT (Figure 5.3). Additionally, we found that a *flaA* mutant was no longer able to form biofilm, which concurs with the literature showing that flagellar

mutants have decreased biofilm-forming ability (Figure 5.3). An *alpA* mutant also showed decreased biofilm formation (Figure 5.3). This could be a consequence of polar effects on *alpB* (since *alpA* is in an operon with *alpB*), which is consistent with literature showing that *alpB* is important for biofilm formation (79). We also identified several additional OMPs that may play a role in biofilm formation. Mutagenesis of *oipA, babA, sabA, sabB, hopD,* and *hopN* led to significant decreases in biofilm formation (Figure 5.3). Follow-up experiments are further described in Future Directions in Chapter VI.

Optimization of Genetic Barcoding Approaches

Viability PCR (v-PCR)

In the research presented in Chapter IV, the composition of the output libraries was assessed based on analysis of bacteria cultured from stomachs. This could potentially impose a bias, since culturing enriches for bacteria that are able to efficiently transition from a gastric environment to a culture plate environment. Direct PCR from the stomach would overcome this limitation. However, one potential pitfall would be that bacterial DNA amplified directly from stomach contents can originate from both viable and non-viable bacteria. Viability PCR, developed by Biotium, uses a PMAxx dye, which is a cell membrane-impermeable and photoreactive dye that has a high affinity for double stranded DNA (dsDNA). This dye intercalates into dsDNA and forms a covalent linkage upon exposure to intense visible light. When a sample containing both live and dead bacteria is treated with the PMAxx dye, only dead bacteria with compromised cell membranes are susceptible to DNA modification by the PMAxx dye. In a qPCR reaction, we expect that dead cell DNA will show delayed amplification and higher Ct values than

live cell DNA. To test the efficiency of this methodology, purified *H. pylori* DNA was treated with or without PMAxx, exposed to intense light, the highly conserved *ureA* gene was amplified in a qPCR reaction, and Ct values were determined via real-time qPCR. We found that PMAxx inhibits amplification of purified DNA by approximately 20 cycles (Figure 5.4A). We also tested live *H. pylori* or heat-treated *H. pylori* treated with PMAxx and found that PMAxx inhibits amplification from heat-treated bacteria (approximately 6 cycles difference in cycle threshold compared to amplification from live bacteria) (Figure 5.4B). Additional studies will be needed to determine the optimal concentration of PMAxx dye for *in vitro* experiments and for use with mouse stomachs.

Optimization of methodology for analyzing *H. pylori*-AGS cell co-cultures

H. pylori express multiple adhesins that bind various receptors on host cells. The Hop family encodes at least 8 reported adhesins (AlpA/B, BabA, SabA, HopQ, OipA, LabA, and HopZ), but the function of the remaining Hop family members remains elusive. Since adhesion involves direct contact of bacteria with the host, adhesins are typically exposed on the bacterial surface. At least six additional Hops meet multiple criteria for surface exposure (HopF, HopJ/K, BabB, HopI, and HopL) (123). We hypothesize that surface-exposed Hops play a role in *H. pylori* adhesion. To elucidate which Hops play a dominant role in adhesion to gastric epithelial cells (AGS), we must first optimize an adhesion assay to study AGS-*H. pylori* co-cultures.

One commonly used assay for evaluating *H. pylori* adhesion to cultured cells is performed by incubating monocultures of *H. pylori* with AGS cells, then washing away the non-adherent *H. pylori,* and finally separating the *H. pylori* from the AGS cells for CFU count analysis by selectively lysing the AGS cells with a detergent, such as saponin.

Several steps within this methodology have the potential to be optimized further. As one example, the detergent conditions might not be optimal for high-throughput studies of adherent *H. pylori* within a co-culture. To test this, we carried out the same protocol described above, but lysed the AGS cells with either saponin or Triton X-100 or detached the cells from the plate by scraping. We found that cell-scraping and saponin treatment of co-cultures enriched for intact bacteria better than treatment with Triton X-100, resulting in a higher abundance of *H. pylori* DNA from AGS co-cultures quantified by qPCR (Figure 5.4C). Another step that can be optimized is determining the optimal concentration of saponin added to co-cultures. Saponin forms pores in membrane bilayers, leading to cell lysis. Some studies suggest that saponin preferentially disrupts eukaryotic cell membranes. To determine the optimal concentration of saponin for use with AGS-*H. pylori* co-cultures, we incubated either bacteria only or an AGS co-culture with various concentrations of saponin (0%, 0.01%, 0.1%, or 1%). We found that saponin concentrations of 0.01% were well tolerated by *H. pylori* in the presence or absence of AGS cells (Figure 5.4D), whereas concentrations of 0.1% and above were cytotoxic to *H. pylori* (Figure 5.4D). I also observed that AGS co-cultures were easier to resuspend for CFU counts when treated with 0.01% saponin than untreated samples that were scraped off the plate. Moving forward, adhesion assays used to quantify *H. pylori* CFUs will be performed with 0.01% saponin.

Another application of the adhesion assay methodology would incorporate a high-throughput approach to test the role of Hops in adhesion. In this case, AGS cells would be incubated with a barcoded *H. pylori* library and the composition of the adherent population would be compared to that of the input population. As part of this approach,

non-adherent *H. pylori* would be washed away, followed by lysis of AGS cells with saponin, and centrifugation to separate the adherent population of *H. pylori* from lysed AGS cells. Genomic DNA would then be extracted from this pellet and used for high-throughput sequencing. One limitation of this approach is that there could be DNA carryover from lysed AGS cells that could, in turn, decrease the relative abundance of *H. pylori* DNA in any given sample following DNA isolation. Another limitation could be that a portion of the *H. pylori* DNA sequenced could originate from non-viable bacteria. To address these limitations, we tested whether PMAxx would increase the abundance of *H. pylori* DNA from viable bacteria by inhibiting the amplification of DNA from non-viable cells (AGS and *H. pylori*). To test this, we compared the efficiency of amplification from the input pool compared to amplification from preparations of adherent *H. pylori* treated with both saponin and PMAxx or with PMAxx only. As expected, we found that the combined treatment of AGS co-cultures with saponin and PMAxx increases the yield of bacterial DNA recovered from viable *H. pylori* compared to recovery from samples treatment with PMAxx only(Figure 5.4E).

We also sought to determine the limit of detection of barcoded *H. pylori* DNA within an AGS co-culture. To test the sensitivity of this assay, we infected AGS cells with various multiplicities of *H. pylori* infection (MOIs) and analyzed the abundance of DNA recovered from each co-culture. We found that we could detect *H. pylori* DNA under conditions with an MOI of 1 *H. pylori* cell per 10,000 AGS cells *in vitro* (Figure 5.4F).

As an alternative approach to identifying Hops that are important for *H. pylori*-adhesion to gastric epithelial cells, we could use immunofluorescence methodology. To determine the feasibility of this approach, AGS cells were incubated with either a wild-

type strain or a Hop mutant strain (*alpA, babA,* or *sabA* mutants). AGS-adherent *H. pylori* were imaged, and bacterial DNA intensity was quantified by outlining each AGS cell, and subtracting DNA signal from the AGS nucleus (Figure 5.5). This preliminary study was done using DAPI to quantify bacterial DNA intensity. In the future, we would perform immunofluorescence with an *H. pylori*-specific antibody to facilitate analysis procedures. Overall, we did not identify any differences in the attachment of the mutant strains to AGS cells compared to attachment of the wild-type strain. One possibility is that *H. pylori's* adhesins have redundant functions. Another possibility is that *H. pylori* adhesins have cell type-specificity, and that the adhesins tested thus far are not important for adhesion to AGS cells. Additional studies will be needed to characterize the role of Hops in adhesion.

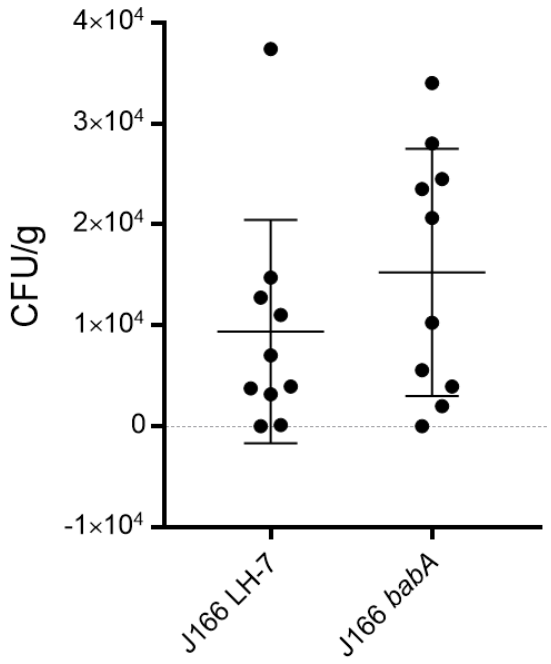

Figure 5.1 *H. pylori* colonization densities of mice infected with wild-type strain J166 or a *babA* mutant.
C57Bl/6J male mice were infected with a monoculture of J166 mutant strains LH-7 or a monoculture of the J166 *babA* mutant (LH-24$_a$) for 21 days. *H. pylori* was cultured from the stomachs and the numbers of colony forming units (CFU) were determined. Statistical analysis was performed using a t-test. There were no significant differences in CFU counts.

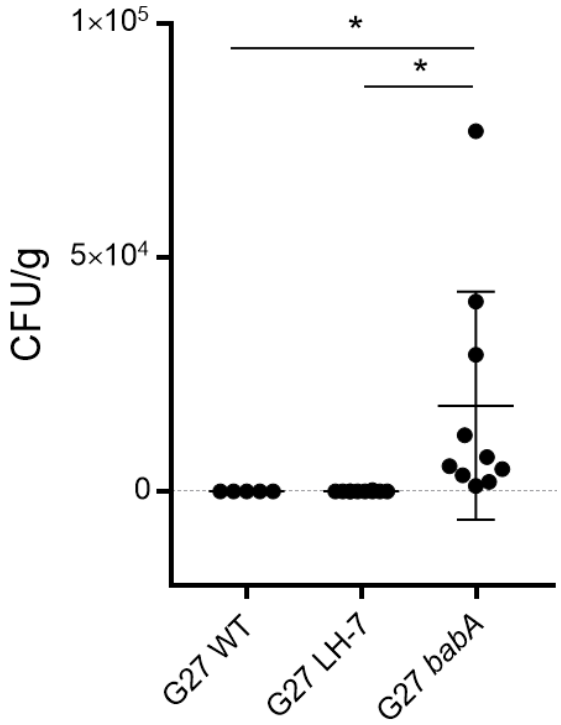

Figure 5.2 *H. pylori* colonization densities in mice infected with wild-type strain G27 or a *babA* mutant.

C57Bl/6J male mice were infected with monocultures of G27 wild-type, G27 control mutant strain LH-7, or a G27 *babA* mutant for 16 days. *H. pylori* was cultured from the stomachs and colony forming units (CFU) counts were determined. Strains G27 wild-type and G27 LH-7 were not able to colonize mice, whereas the G27 *babA* mutant was able to colonize mice. Statistical analysis was performed using a one-way ANOVA with Dunnett's multiple comparisons test. *p-value <0.0001 relative to G27 WT (100%).

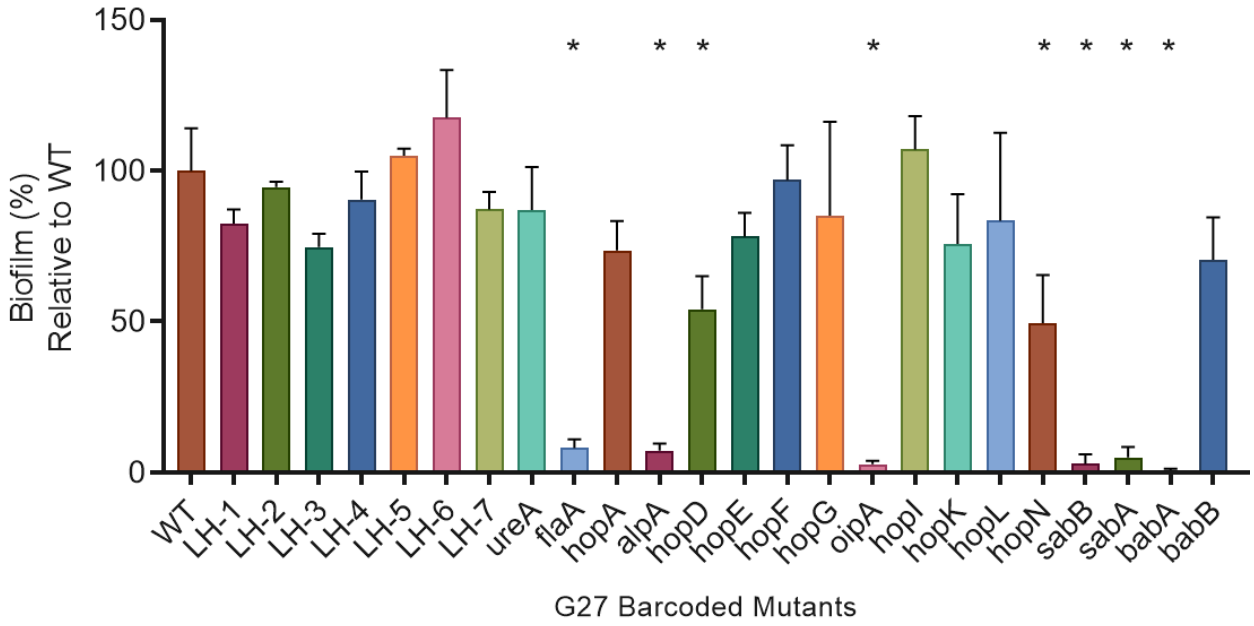

Figure 5.3 Biofilm quantification of G27 barcoded mutants.
H. pylori strain G27 was used to generate a panel of barcoded control mutants (LH-1 through LH-7) and a panel of OMP mutants (*ureA* through *babB*). Monocultures of G27 wild-type (WT) and mutant strains were grown for 4 days. Biofilm was quantified using crystal violet staining. Relative levels of biofilm formation (%) by each strain relative to the G27 wild-type strain are shown. Control strains LH-1 through LH-7 produced biofilm at similar levels as WT. Mutants of *alpA, hopD, oipA, hopN, sabA, sabB,* and *babA* genes showed significantly decreased biofilm formation compared to wild-type. Statistical analysis was performed using a one-way ANOVA with Dunnett's multiple comparisons test. *p-value <0.0001 relative to G27 WT (100%).

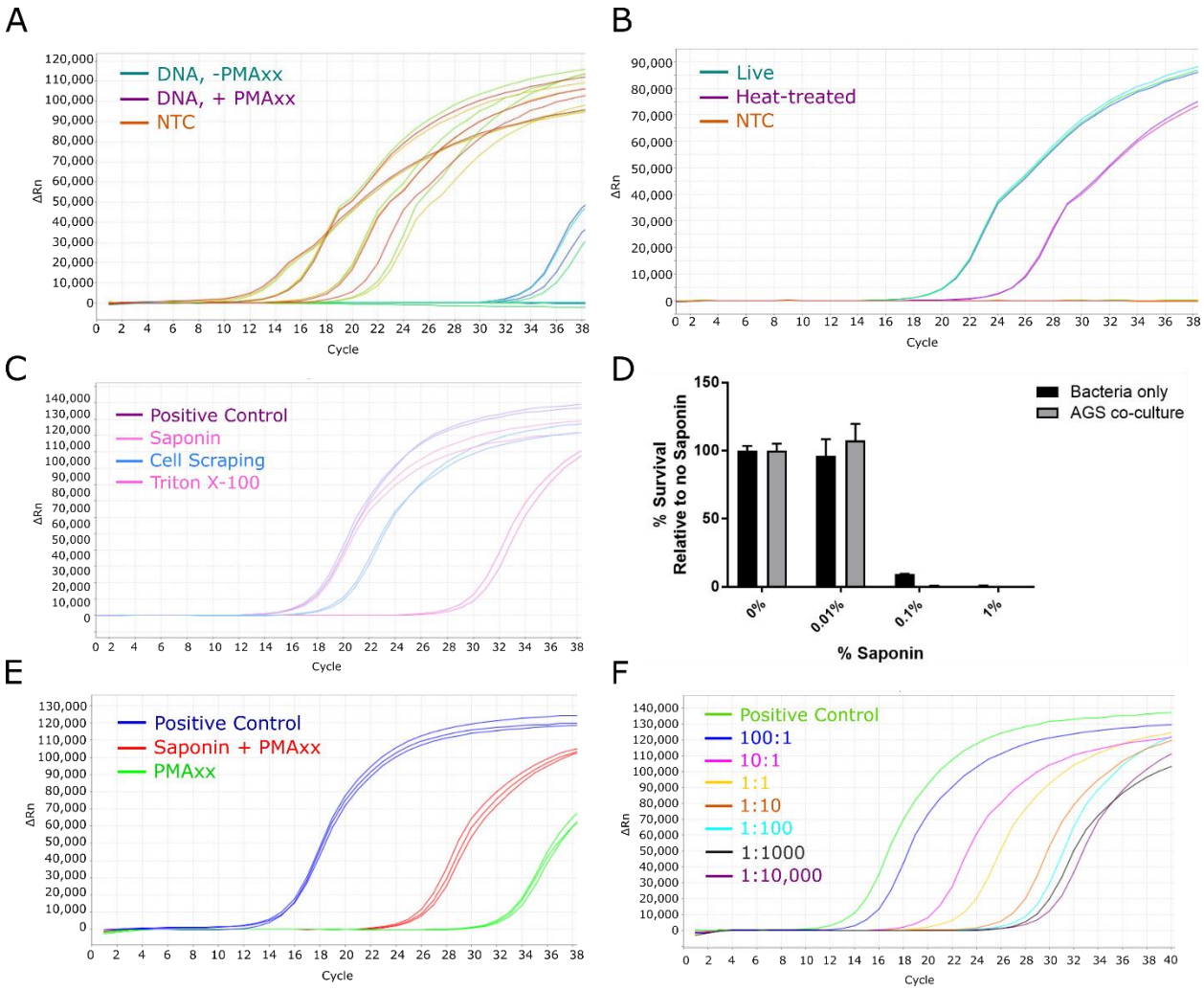

Figure 5.4 Optimization of AGS co-culture methodology.
Amplification curves of *H. pylori ureA* gene are shown for (A) *H. pylori* DNA with or without PMAxx dye treatment. The sequential curves represent 10-fold dilution curves for each sample. (B) Live and heat-treated *H. pylori* with PMAxx dye (+PMA), NTC=no template control. (C) *H. pylori* DNA from AGS co-cultures treated with saponin, cell scraping, or Triton X-100 compared to a positive control (DNA only). Panel D shows *H. pylori* survival in various concentrations of saponin (bacteria only or AGS co-cultures). (E) Amplification curves of *H. pylori* DNA from AGS co-cultures treated with saponin and PMAxx or PMAxx only. (F) Amplification of *H. pylori* DNA from AGS co-cultures at various multiplicities of infection (MOI).

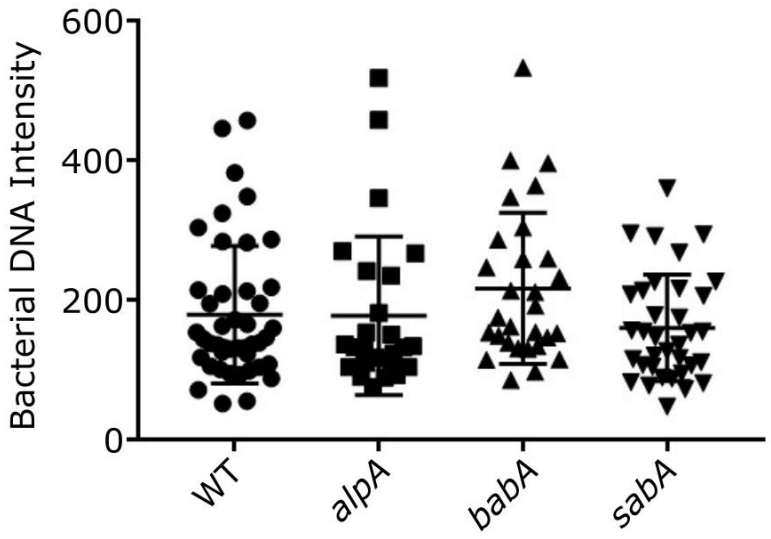

Figure 5.5 Immunofluorescence of AGS-adherent *H. pylori.*

AGS cells were co-cultured with an *H. pylori* wild-type strain or various *hop* mutants, including *alpA, babA,* and *sabA* mutants. DNA was stained with DAPI. Each point represents the fluorescent DAPI signal of bacterial cells in the peripheral regions of AGS cells (excluding the signal from AGS cell nuclei). The mutants did not show a significant difference from *H. pylori* WT. Statistical analysis was performed using a one-way ANOVA with Dunnett's multiple comparisons test.

CHAPTER VI

FUTURE DIRECTIONS

In summary, this work provides new insights into the existence of non-selective pressures that shape the *H. pylori* community during gastric colonization and the role of BabA as a determinant of *H. pylori* fitness. In addition, the quantitative monitoring of nucleotide barcodes described in this study offers numerous advantages compared to comparative analysis of strains labeled with different antibiotic markers. These methodologic approaches will be useful in future studies of *H. pylori* population dynamics, both *in vitro* and *in vivo.* For example, these methods can be used in future studies to evaluate *H. pylori* fitness in various *in vitro* environments to study effects of pH, medium composition, and other variables, as well as *H. pylori* fitness and population dynamics in varying in vivo environments. We anticipate that this approach will be generally useful for studying the roles of *H. pylori* genes in various phenotypes, as well as for analysis of the selective pressures that shape the evolution of *H. pylori* populations.

Development of an intragastric *H. pylori* community

One area of interest is to investigate the dynamics of *H. pylori* gastric colonization. Some individuals may become colonized with multiple unrelated strains of *H. pylori* over the course of a lifetime. A single *H. pylori* strain can also mutate within the stomach, and consequently, individuals are commonly colonized with a population of closely related strains. Work presented in Chapter IV, analyzing mice infected with multiple barcoded control strains, suggests the existence of a non-selective bottleneck phenomenon during early stages of *H. pylori* colonization (Figure 6.1). At present, it is unclear if this bottleneck

is attributable to effects of the host immune response or if the bottleneck corresponds simply to a stochastic process in which only a small proportion of an initial inoculum successfully colonizes the gastric mucosa (a phenomenon known as a founder effect) (Figure 6.1). In either case, we propose that the small number of *H. pylori* organisms initially colonizing the stomach subsequently expands until an equilibrium is achieved.

An important goal of future experiments is to further characterize the bottlenecks that occur during *H. pylori* colonization of the stomach. Most experiments presented in Chapter IV were conducted for 21 days. However, it is possible that *H. pylori* experiences multiple bottlenecks at different time points. A previous study reported various timepoints that are relevant to the murine host response to *H. pylori* infection (strain SS1) in FVB/N mice, including macrophage and neutrophil responses and lymphocyte infiltration in the stomachs of *H. pylori*-infected mice (118). These time points were reported to correlate with *H. pylori* colonization density in the same study. Specifically, a marked reduction in numbers of viable bacteria was detected at 3 days post-infection (compared to the initial inoculum), which suggests the existence of a bottleneck. This was followed by an increase in colonization density at later time points (118). To further evaluate bottlenecks in vivo, C57BL/6J mice would be infected with a barcoded control library and CFUs would be quantified at 1-, 2-, 3-, 10-, 21-, and 34-days post-infection (Figure 6.2) (118). Based on results of the previous study, we expect to see a reduction in colonization density at 3 days post-infection compared to earlier time points, followed by an increase in colonization density at later time points (118). In addition to assessing *H. pylori* colonization density, the composition of the output barcoded control libraries from mice infected with the barcoded control library (described above) would be quantified via next-

generation sequencing and compared to the input barcoded control library (Figure 6.2). We expect that the composition of the population will be most similar to the input population at early time points and will become increasingly varied as each bottleneck is encountered. A complementary approach would be to monitor changes in the composition of the colonizing control library population over time. Overall, these studies would allow us to define the role of bottlenecks in *H. pylori* colonization of the stomach.

Another area of interest is to further characterize the ability of a *H. pylori* founder population to block subsequent colonization by distinct strains. One study suggested that colonization resistance can be detected early in infection (124). In this study, they observed that a founder isogenic strain (expressing GFP) exhibited colonization resistance against a challenging isogenic strain (expressing td-Tomato) as early as 5 days post-infection. The approach used in that study was optimally suited for evaluating *H. pylori* localization in the stomach via imaging, but the method is not well suited for analyzing large numbers of animals. The methods described in this dissertation would allow for an expanded and more quantitative investigation of *H. pylori* founder effects. To characterize *H. pylori* founder effects, we would infect mice with 1:1 mixtures of two barcoded control strains, as well as infect mice with one barcoded control, and then challenge with a second barcoded control strain at different time points, such as those identified previously to be correlated with bottleneck events (Figure 6.3). As a control, we would assess the colonization density of individual barcoded control mutants in murine mono-infection studies (Figure 6.3). If founder effects play a role in shaping the output population of mice infected with various strains of *H. pylori,* we would expect to see colonization resistance of the founder strain to a challenging strain. Similarly, we expect

that the longer that mice are infected with the founding strain, the greater the degree of colonization resistance will be over time. Overall, these studies would allow us to characterize the founder effects that govern *H. pylori* infection.

Analysis of dominant output strains

In experiments in which mice were infected with a library of control barcoded strains, we observed marked variation among the mice and frequent emergence of dominant strains. The acquisition of spontaneous mutations could possibly account for the emergence of dominant strains. Previous literature has shown that *H. pylori* may undergo genetic changes in the stomach to adapt to the gastric environment. Potentially, these spontaneously occurring mutations could confer a fitness advantage over other strains *in vivo.* Work presented in Chapter IV shows that the composition of the output *H. pylori* population recovered from individual mice was markedly different from the input population. Furthermore, the output populations from individual mice were markedly different from output populations of other mice. We hypothesize that the control strain that dominates the output population acquired mutation(s) that give it a fitness advantage over other control strains. An alternative hypothesis is that the control strain that dominates the output population is established by chance. To test this, we would first evaluate if dominant output control strains are more fit than the input control strain *in vivo.* Mice would be infected with a 1:1 mixture of the input and output control strains (Figure 6.4). If the output control strain outcompetes the input control strain, we would then undertake genome sequencing studies (and maybe also RNA-seq), comparing the dominant strains in each output population from mice infected with the barcoded control library to another barcoded control strain in the input population. The output pools of mice infected with the

barcoded control library in Chapter IV were frozen down and can be used for this analysis. Follow-up studies would seek to define the role of individual mutations on *H. pylori* fitness in the stomach. Overall, these studies would determine whether the acquisition of spontaneous mutations results in a fitness advantage *in vivo* among members of the barcoded control mutant library.

Studies of BabA in *H. pylori* fitness and disease outcome

Temporal features of BabA-mediated *H. pylori* Fitness

Previous studies have shown that BabA is selected against in various animal models (51, 61). Our data suggest that a *babA* mutant outcompetes other *H. pylori* strains when the strains pooled together and introduced together into a mouse stomach. We also show that the J166 control mutant (LH-7) and the *babA* mutant colonize at similar densities during mono-infection studies, suggesting that the fitness advantage rendered by mutagenesis of *babA* is most evident in a competitive environment.

To determine if the fitness advantage of a *babA* mutant is dependent on timing of infection, one approach would be to first infect mice with a control mutant, and then subsequently infect the mice with a *babA* mutant (Figure 6.5). In parallel, mice would be infected with the *babA* mutant first and then subsequently infected with a control mutant (Figure 6.5). One challenge is to determine murine infection status prior to infection with the second strain. To assess whether a mouse is successfully infected with the first *H. pylori* strain, we would perform PCR to amplify a highly conserved *H. pylori* gene such as *ureA* from mouse feces. We hypothesize that the *babA* mutant might be able to displace

a founder control strain, but that a control strain would not be able to displace a founder *babA* mutant strain.

Another approach would be to use a TetR regulation system to turn *babA* expression on and off within the murine stomach (Figure 6.6). This system has been optimized for use in *H. pylori* studies in our lab (125). This system is based on the binding of the *Escherichia coli* Tet repressor (TetR) to the tet operator (*tetO*) DNA sequence introduced upstream of the gene of interest, thereby repressing the expression of the target gene (Figure 6.6 A). Anhydrotetracycline (ATc) or doxycycline (dox) binding to TetR triggers conformational changes that prevent TetR from binding *tetO*, therefore de-repressing the target gene (Figure 6.6 A). A TetR variant exhibits the reverse phenotype, where TetR binds *tetO* in the presence of ATc or dox. Both systems could be applied to these studies. With this approach, we would infect 4 groups of mice with 1:1 mixtures of two isogenic strains, one of which contains a Tet-regulated *babA* gene that could be turned on/off *in vivo* and the other containing a *babA* gene that is not regulated by Tet (Figure 6.6 B). Each strain would be engineered to contain a unique nucleotide barcode. In the first group of mice, *babA* would be expressed in both strains throughout the experiment via administration of ATc in the water supply. In the second group, the *babA* gene in the *tet*-regulated strain would be turned off for the entire experiment (no ATc in the water supply). In the third group, we would start the experiment with the *babA* gene on, but then turn it off 2 weeks post-infection. Finally, in the fourth group, the *babA* gene would be turned off at the beginning of the experiment, then turned on at 2 weeks post-infection. Based on the findings from Chapter IV, we would expect to see variability in the dominant strain colonizing the first group of mice. In the second group, we anticipate that

the control mutant that does not express *babA* would overtake the other strain, an example of competitive displacement. If the absence of *babA* leads to a fitness advantage irrespective of order of infection, we anticipate that the *babA* mutant will overtake the other strain in the third group. If order of infection does dictate the fitness advantage of a *babA* mutant, we anticipate that the *babA* mutant will not overtake the other strain in this control group. Finally, in the fourth group, we expect that the strain that does not express *babA* would outcompete the other control mutant and that this fitness advantage early in infection would have a lasting effect even after *babA* has been turned off. This experimental set-up would allow us to test whether the fitness advantage of a *babA* mutant is dependent on timing of infection. One possible problem with the proposed experiment is that the wild-type *H. pylori* strain used in these studies could undergo mutations in *babA*, resulting in loss of BabA production. To test for this, we would evaluate the wild-type strain's ability to produce BabA and ability to bind to Le^b antigen before and after co-infections. If we see that the wild-type *H. pylori* strain consistently loses BabA, we would generate a mutant that favors retention of BabA. Previous studies have shown that BabA is regulated by slipped strand mispairing within the CT repeats upstream of *babA.* To prevent slipped strand mispairing, we would delete these repeats upstream of *babA.*

Analysis of the fitness advantage of a *babA* mutant

Our data suggest that a *babA* mutant outcompetes other *H. pylori* strains when introduced into a murine stomach. Other groups have shown that BabA is selected against in numerous animal models. An important goal is to determine why a *babA* mutant exhibits a fitness advantage compared to strains containing a wild-type *babA* in mice.

One possibility is that the production of BabA is unfavorable. Another possibility is that the mutagenesis of the *babA* locus has indirect consequences that are favorable to *H. pylori* fitness *in vivo.*

To evaluate if the production of BabA is unfavorable in mice, we could generate an *H. pylori* mutant that overexpresses BabA by driving the expression of *babA* with the promoter of a constitutive gene, such as *ureA.* If the production of BabA is unfavorable in mice, we expect that the *H. pylori* mutant overexpressing BabA will be less fit than a wild-type strain. Another approach (described previously) would evaluate how early during infection a *babA* shows a fitness advantage compared to a control strain. If this timepoint correlates to the development of an adaptive immune response, experiments could be undertaken using various knockout mice (i.e. $RAG^{-/-}$ mice). Localization of *H. pylori* in close proximity to the gastric epithelium, as a consequence of BabA's adhesive functions, could also influence *H. pylori* fitness in the stomach. The use of various imaging approaches could be used to determine if there is a difference in localization of an *H. pylori* wild-type strain, compared to a *babA* mutant strain. Further characterization of the surface-exposed regions of BabA and analysis of their potential roles in *H. pylori* fitness could also be used to determine if the production of BabA is unfavorable to *H. pylori* fitness in mice. To test this, we would generate strains producing BabA variants via mutagenesis of regions encoding surface-exposed loops and test the fitness of these strains compared to a barcoded control strain in mice. Another approach would be to use naturally occurring *babA* variants, such as those that can and cannot bind Lewis b, to test the hypothesis that specific regions of BabA are unfavorable in mice. To do so, we would

alter the *babA* sequences in the strain(s) of interest by substituting the *babA* sequences from another *H. pylori* strain and then test for differences in *H. pylori* fitness *in vivo.*

To assess whether the mutagenesis of *babA* locus has indirect consequences that are favorable to *H. pylori* fitness in mice, we could analyze the outer membrane characteristics of a *babA* mutant compared to wild-type strain. It is also possible that the mutagenesis of *babA* leads to transcriptomic changes. We hypothesize that the fitness advantage rendered by the mutagenesis of *babA* might be the consequence of genes that become overexpressed (or underexpressed) in the *babA* mutant compared to the wild-type strain to compensate for the loss of *babA. H. pylori* encodes multiple OMPs that are thought to have redundant functions, such as those that functionally acting as adhesins. Therefore, we hypothesize that the absence of one OMP (such as BabA) may lead to overexpression of another OMP, or perhaps changes in other classes of genes or proteins. We propose to use an RNA-seq approach to evaluate differences in the transcriptome of the *babA* mutant compared to that of a wild-type control strain. Follow-up studies could potentially include the mutagenesis of target genes that are differentially expressed in the mutant compared to the wild-type strain to further investigate the role of these genes in *H. pylori* fitness. A complementary approach would be to perform comparative analysis of protein abundance in membrane fractions versus non-membrane fractions of a *babA* mutant compared to a wild-type control.

Mutagenesis of the *babA* locus might also have indirect effects on the activity of other *H. pylori* virulence factors. Multiple *H. pylori* virulence factors that have been associated with adverse disease outcome, including the presence of CagA (an effector protein secreted by the Cag T4SS) and active forms of VacA (a secreted pore-forming

toxin). Previous literature suggests that the production of CagA and/or Cag T4SS function are commonly lost during *H. pylori* colonization of mice, implying that CagA-positive strains or strains that have an active Cag T4SS have a selective disadvantage in mice. Additionally, BabA-mediated adherence if thought to act as a potentiator of *H. pylori* Type IV secretion system activity (59). It is possible that the mutagenesis of *babA* may lead to impaired CagA delivery into host cells or impaired T4SS activity, thereby conferring a selective advantage. To test this, we would evaluate *H. pylori* T4SS activity using assays with cultured gastric epithelial cell lines, such as NF-κB reporter assays, IL-8 ELISAs, and CagA phosphorylation assays.

Colonization of mice versus gerbils

Data presented in Chapter IV and V suggest that mutagenesis of *babA* leads to a fitness advantage in mice. We do not, however, know whether this fitness advantage occurs in other animal models. One study reported that BabA is selected against in the Mongolian gerbil model, similar to what is observed in mice (61). Therefore, we expect that the fitness advantage rendered by the mutagenesis of the *babA* gene in the murine model will also be observed in the Mongolian gerbil model. To test this, we would infect gerbils with 1:1 mixtures of a barcoded control mutant and a *babA* mutant. After 21 days, we would quantify the proportional abundance of each barcoded strain in the output compared to the input pool.

Strain-specific differences in colonization

Another approach would be to test the generality of the fitness advantage rendered by the mutagenesis of *babA* in strain J166 across other *H. pylori* strain backgrounds.

Preliminary work presented in Chapter V showed that mutagenesis of *babA* does not affect colonization density in strain J166 (Fig. 5.1) but leads to increased colonization density in the strain G27 background (Fig. 5.2). Follow-up studies would include the analysis of mice infected with 1:1 infections of a control strain with a *babA* mutant strain to determine if the fitness advantage in conserved across both strain backgrounds.

Disease outcome

Another unknown is whether BabA contributes to disease outcome. Mongolian gerbils are an animal model commonly used to study *H. pylori*-mediated disease. One previous study suggested that BabA might contribute to gastric damage and inflammation in the gerbil model, but further studies of this topic are needed (63). To investigate whether the mutagenesis of *babA* affects disease outcome, we would infect Mongolian gerbils with monocultures of an *H. pylori* wild-type strain or a *babA* mutant. After 4 months of infection, we would compare bacterial colonization density and disease outcome of gerbils infected with the *H. pylori* wild-type strain to the corresponding parameters in gerbils infected with the *babA* mutant. Gastric histology would be evaluated by a collaborating pathologist to define disease outcomes, including severity of inflammation, atrophic gastritis, ulceration, dysplasia, and invasive carcinoma. Alternatively, we could evaluate colonization density and disease outcome in an INS-GAS murine model, which is susceptible to *H. pylori*-mediated disease.

Characterization of the barcoded OMP library in the absence of a *babA* mutant

Chapter V shows that the mutagenesis of *babA* leads to a fitness advantage in the context of gastric colonization but not *in vitro*. Interestingly, we also observed that several

other mutants, such as *alpA, hopE,* and *oipA* mutants, exhibited a trend towards a fitness defect *in vivo* compared to their fitness *in vitro.* We wonder if potential fitness disadvantages of these mutants *in vivo* might have been overshadowed by the presence of a *babA* mutant. We hypothesize that mutagenesis of OMPs such as *alpA, hopE,* and *oipA* lead to a fitness disadvantage *in vivo,* which could be more easily detected in the absence of a *babA* mutant. Previous studies have shown that *alpA* is highly conserved across hundreds of clinical isolates. Similarly, expression of *alpA* is conserved *in vivo,* and studies have reported that the mutagenesis of *alpA* leads to a colonization defect *in vivo.* To test the role of these OMPs *in vivo,* C57Bl/6J mice could be infected with the barcoded OMP mutant library described in Chapter V, without the *babA* mutant (Figure 6.7). In parallel, we would infect mice with the control library and use the variability of the corresponding output population to facilitate statistical analysis of results obtained with the OMP mutant library.

OMP mutants that lead to a significant fitness disadvantage *in vivo,* would be followed-up with studies using independent mutants and complemented mutants. One limitation from Chapter III was that fitness was not restored via complementation into the *rdxA* locus *in vitro*. One possibility is that mutagenesis of the *rdxA* locus itself confers a fitness disadvantage *in vitro* (126). On the other hand, a previous study did not detect any fitness disadvantages associated with *rdxA* mutagenesis (126). An important step will be to develop improved methods for complementation to restore fitness phenotypes of *omp* mutants. One approach would be to complement the *omp* genes in trans at the *mdaB-hydA* locus by insertional mutagenesis with a barcoded antibiotic cassette. Several attempts have been done to generate a complemented mutant using a barcoded

kanamycin cassette, but thus far we have not been able to isolate transformants. This problem can potentially be resolved by using a different antibiotic marker, such as a gentamicin resistance cassette. As another approach, we could restore the target *omp* into its endogenous locus. To do so, we could use a negative selection method (for example, using a *cat-rdxA* or *cat-rpsL* cassette) (127, 128). Restoration of the target *omp* would be generated in an *rdxA* mutant background (or *rpsL* mutant background), which is resistant to metronidazole (Mtz^R) or streptomycin. We would first replace the target *omp* with a *cat-rdxA* cassette. This would make the strain resistant to chloramphenicol and sensitive to metronidazole (Chl^R/ Mtz^S). Next, we would transform with a cassette that carries the intact version of the target *omp.* This vector, in turn, would replace the *cat-rdxA* cassette, making this strain susceptible to chloramphenicol but resistant to metronidazole (Chl^R/ Mtz^R). In parallel, we would perform mono-infection studies to determine the colonization density of each mutant in animals experimentally infected with single strains.

Studies of *H. pylori* biofilms

Another area of interest is to investigate the dynamics of *H. pylori* biofilm formation. As mentioned previously, some individuals may encounter multiple strains of *H. pylori* over the course of a lifetime. One study suggested that *H. pylori* is capable of competitive exclusion and displacement in the gastric environment (124). Biofilm formation has been shown to be critical for the survival and persistence of numerous bacteria. Previous studies have shown that *H. pylori* is capable of biofilm formation, but not much is known about the dynamics of *H. pylori* biofilm formation. As such, it would be interesting to study whether multiple strains of *H. pylori* can integrate into the same biofilm community. To

test this, we could start by combining two barcoded isogenic control strains and quantifying the relative abundance of these mutants within the surface-attached biofilm compared to their relative abundance in a planktonic population (in the supernatant) (Figure 6.8). As a complementary approach, two fluorescent strains could be mixed, and the biofilm could be imaged via fluorescence microscopy (Figure 6.8). We expect that both isogenic control strains will be able to integrate into the biofilm. As a next step, we could test whether a challenging isogenic strain would be able to integrate into an established biofilm. We expect that the challenging isogenic strain would be able to integrate but would reside at the uppermost layer of the biofilm community.

Highlights

Overall, the work presented here provides a foundation for future studies of *H. pylori* population dynamics and the forces that govern the development of an intragastric *H. pylori* community. This work also provides a framework for studies of the fitness advantage conferred by the mutagenesis of *babA* and further characterization of the role of BabA in colonization and disease. Additionally, the methodology of quantitative monitoring of nucleotide barcodes developed during my thesis work provides a novel tool for our lab to study many aspects of *H. pylori* fitness in various environments.

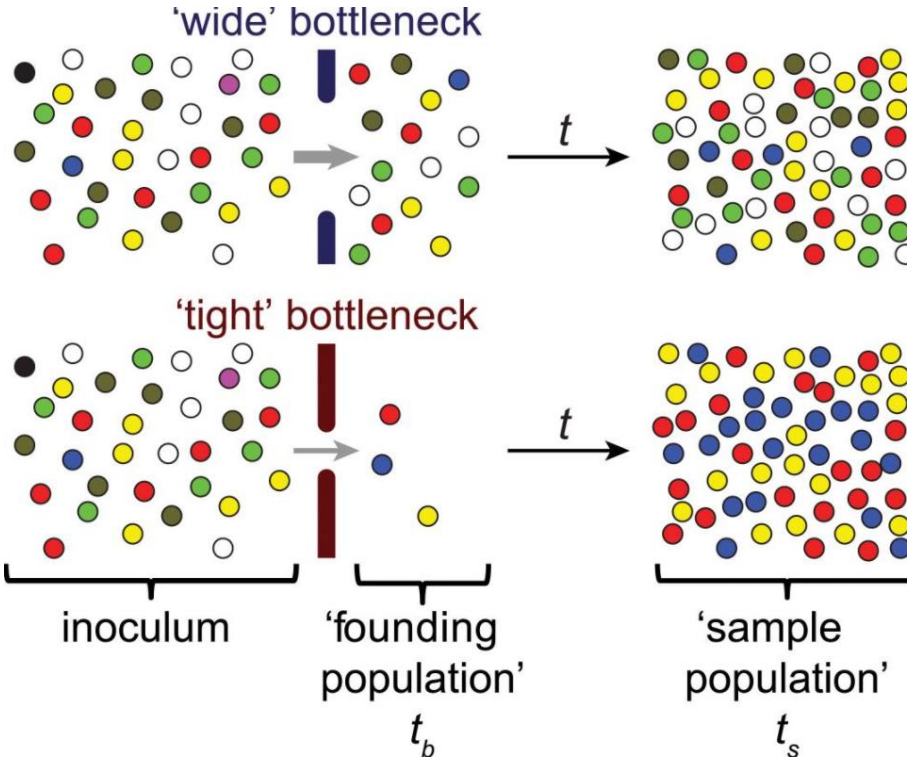

Figure 6.1: Schematic representation of the effect of bottlenecks on genetic diversity.

Individual organisms are shown as colored spheres; the colors represent distinguishable markers. The barriers to infection that constitute the bottleneck are shown by the solid bars and the size of the bottleneck is represented by the size of the gap between these bars. Bottlenecks are events that dramatically reduce the size of a population (for example, the inoculum of an infectious agent). In the context of infection, the founding population consists of the organisms that survive passage through the bottleneck and give rise to a population in a new environment, e.g., a new host or anatomical site. Often it is not feasible to sample directly after the bottleneck event (t_b); instead, populations are sampled (at time t_s) after the passage of time (t), represented by the black arrow. During this time, the founding population often replicates. Wide bottlenecks lead to limited loss of markers (e.g., the magenta and black spheres) and limited changes in the marker frequencies (e.g., over-representation of the blue and under-representation of the olive marker). In contrast, tight bottlenecks lead to stochastic loss of many markers and substantial changes in marker frequencies. These changes can be used to determine the magnitude of bottleneck events and the size of the founding population, even after the population size has increased, provided that the expansion has limited effect on the marker composition (i.e., markers are fitness neutral, and no additional genetic drift occurs) (129).

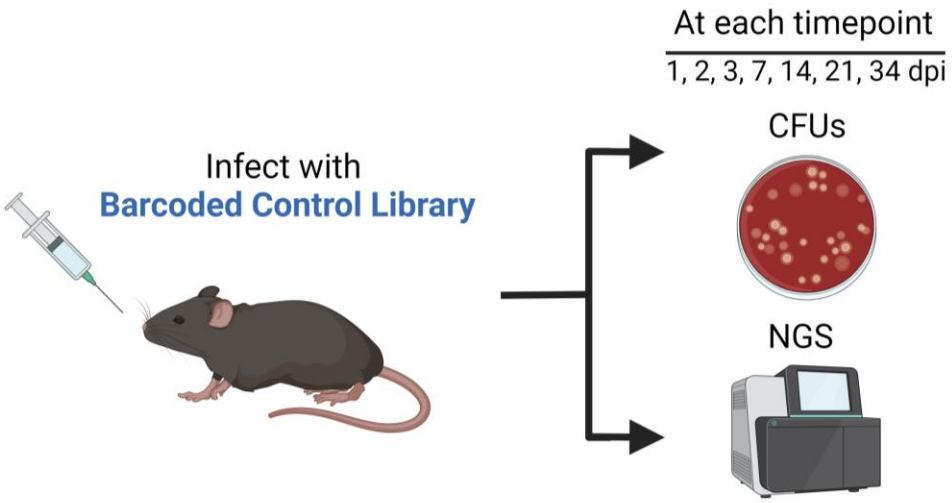

Figure 6.2: Experimental design to characterize *in vivo* bottlenecks and founder effects. C57BL/6J mice would be infected with the barcoded control library. At each timepoint (1, 2, 3, 7, 14, 21, 34 days post-infection, the colonization density would be quantified via CFU counts and the composition of the library would be analyzed via next-generation sequencing. Schematic created with BioRender.com

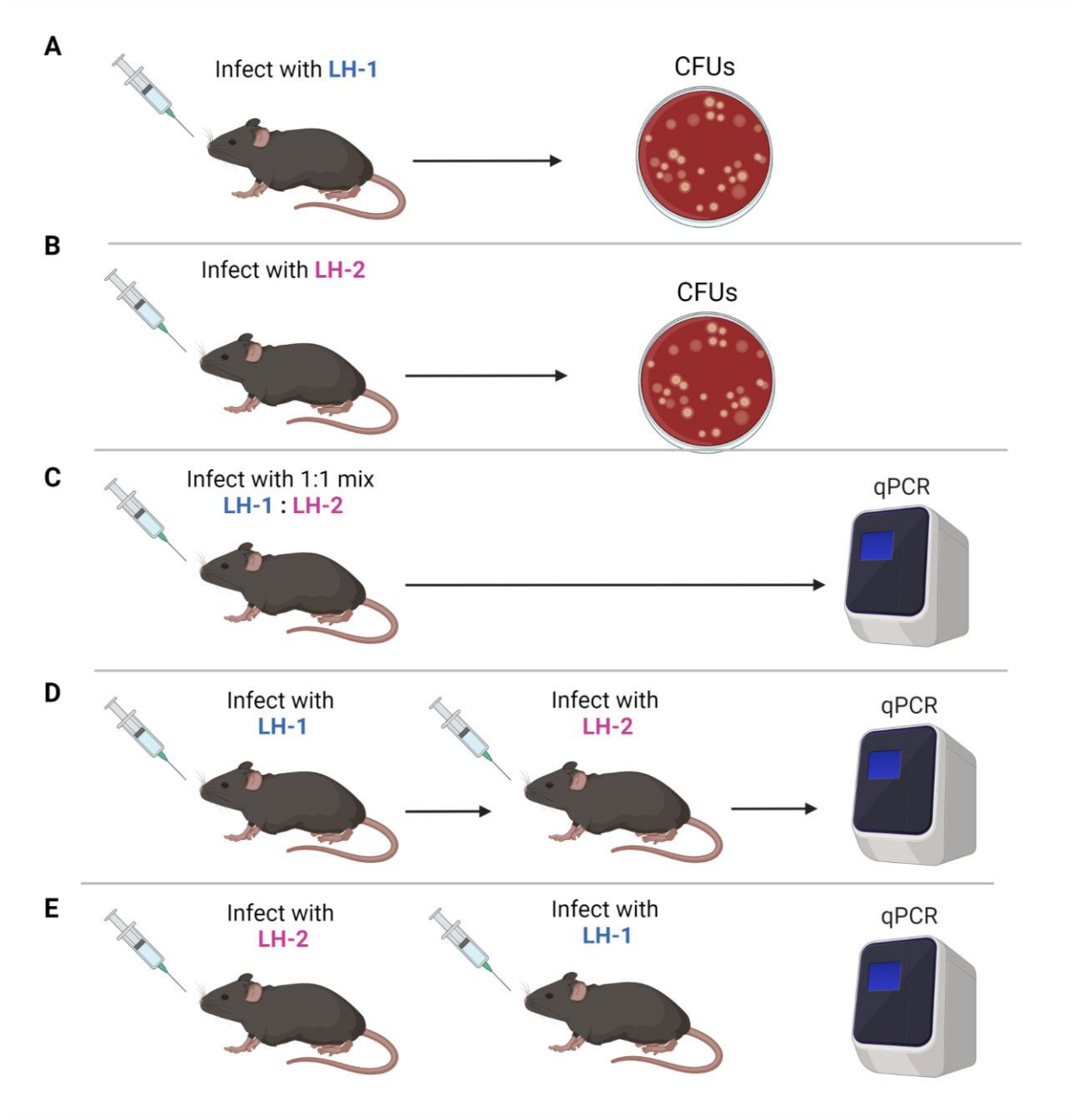

Figure 6.3: Experimental design to evaluate if an *H. pylori* founder population can block colonization by distinct strains.
C57BL/6J mice would be infected with mono-infections of J166 barcoded control strains A) LH-1 and B) LH-2 and colonization density would be quantified via CFU counts. C57BL/6J mice would be infected with C) a 1:1 mixture of barcoded control strain LH-1 and LH-2, D) the LH-1 strain first, followed by subsequent infection with LH-2, and E) the LH-2 strain first, followed by subsequent infection with LH-1. The composition of the input and/or outputs would be quantified via SYBR qPCR with mutant-specific primers. Schematic created with BioRender.com

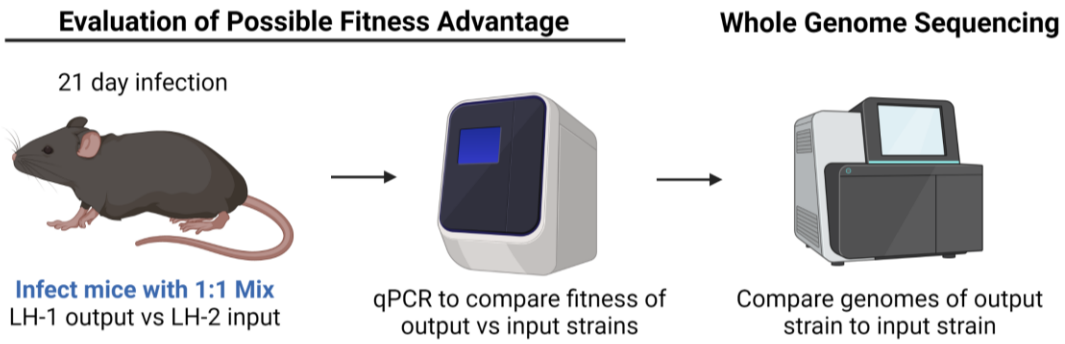

Figure 6.4: Experimental design to evaluate the characteristics of dominant founder *H. pylori* strains compared to non-founders in mice infected with the barcoded control library. Dominant strains from mice infected with the barcoded control library would be isolated from output pools. These dominant strains would be competed against a barcoded control input strain in mice. The composition of the input and output pools would be quantified via qPCR using mutant specific primers. If the output strain retains the fitness advantage over the input strain, the genomes of these strains would be sequenced via whole genome sequencing. Schematic created with BioRender.com

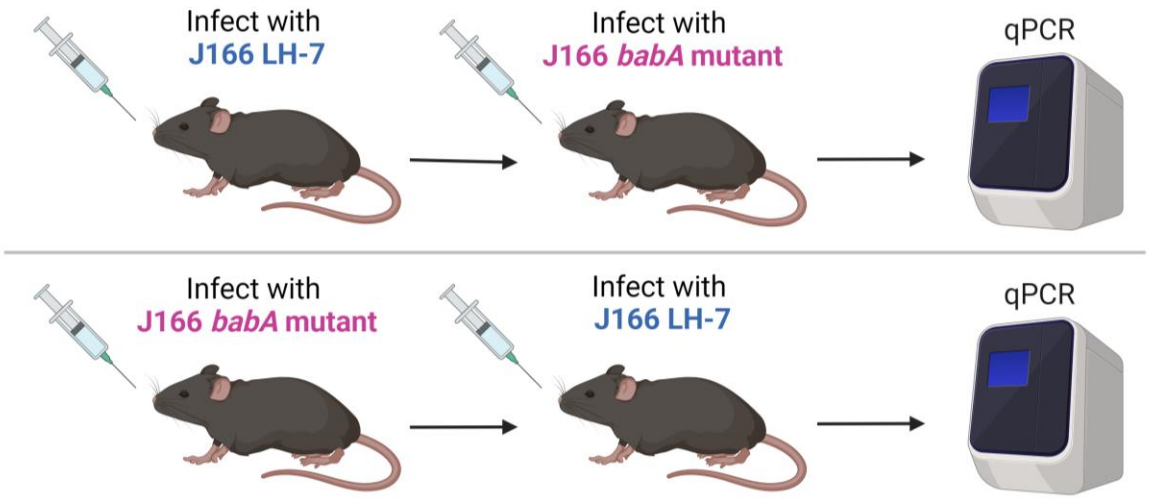

Figure 6.5: Experimental design to determine if the fitness advantage of a *babA* mutant is dependent on order of infection.
C57BL/6J mice would first be infected with A) LH-7 control strain, followed by subsequent infection with a *babA* mutant strain, or B) the *babA* mutant strain, followed by subsequent infection with LH-7 control strain. The composition of the outputs would be quantified via SYBR qPCR with mutant-specific primers. Schematic created with BioRender.com

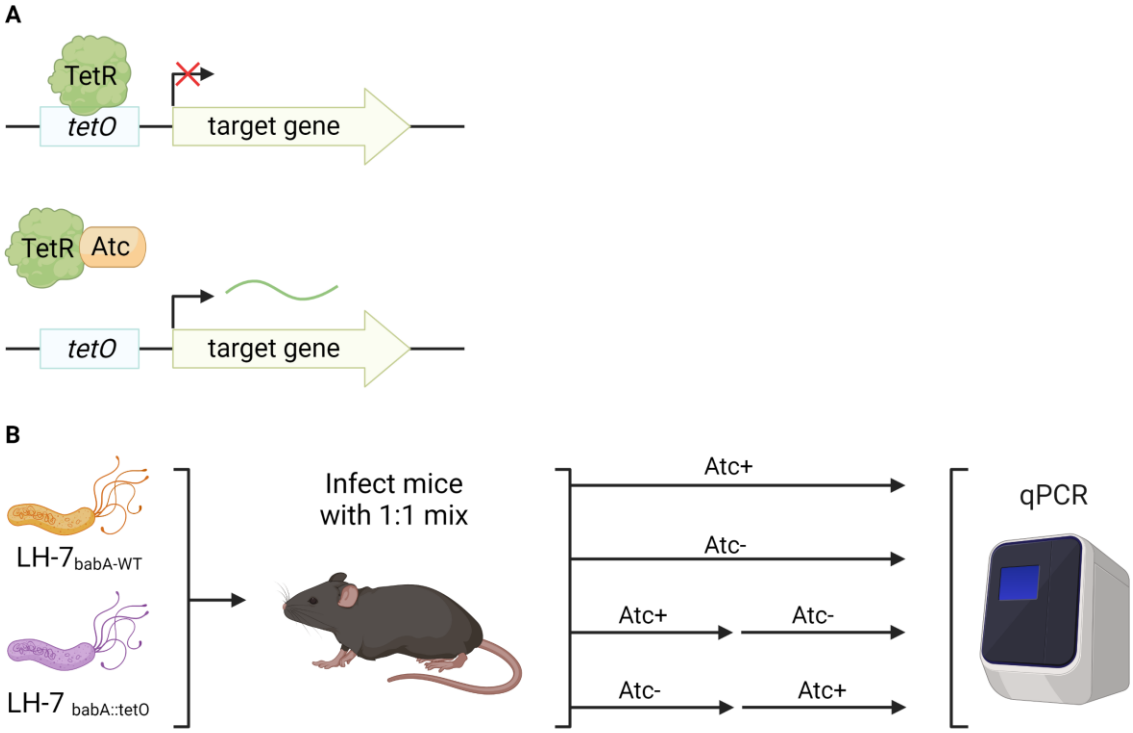

Figure 6.6: Alternative approach to determine if the fitness advantage of a *babA* mutant is dependent on order of infection.
A) Illustration of the Tet methodology. Under normal physiological conditions, the Tet repressor (TetR) binds to the *tet* operator site (*tetO*), which represses the expression of a target gene. In the presence of anhydrotetracycline (ATc), TetR does not bind to *tetO,* which de-represses the expression of the target gene. B) C57BL/6J mice would be infected with a 1:1 mixture of two barcoded control strains (LH-7$_{\text{babA-WT}}$ and LH-7$_{\textit{babA::tetO}}$). LH-7$_{\text{babA-WT}}$ encodes a wild-type version of the *babA* gene, whereas LH-7$_{\textit{babA::tetO}}$ encodes a *tet*-regulated *babA* gene. The first group of mice would receive ATc for the duration of the experiment. The second group would not receive ATc. The third group of mice would receive ATc for the first half of the experiment, whereas the fourth group would receive ATc during the second half of the experiment. The composition of the outputs would be quantified via SYBR qPCR with mutant-specific primers. Schematic created with BioRender.com

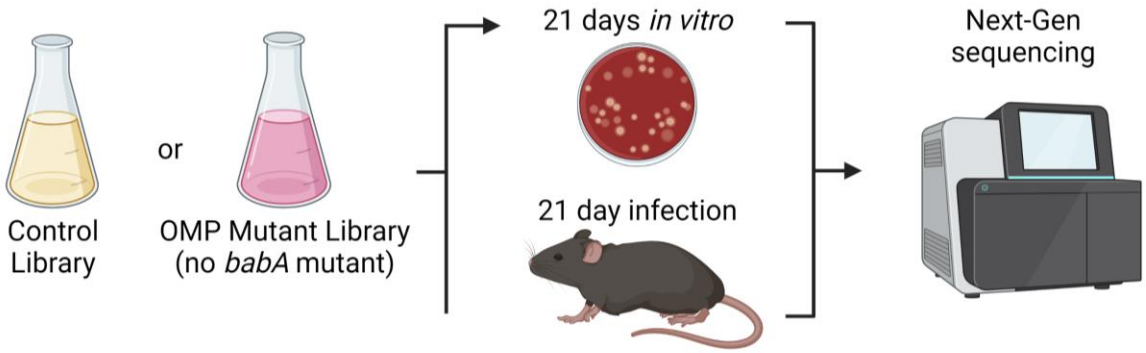

Figure 6.7: Experimental design to determine if Hops play a role in *H. pylori* fitness *in vivo,* but not *in vitro,* in the absence of a *babA* mutant.
C57BL/6J mice would first be infected either the control library or the OMP mutant library (without the *babA* mutant). In parallel, the libraries would be passaged *in vitro.* Subsequently, the composition of the input and output libraries would be quantified via next-generation sequencing. Schematic created with BioRender.com

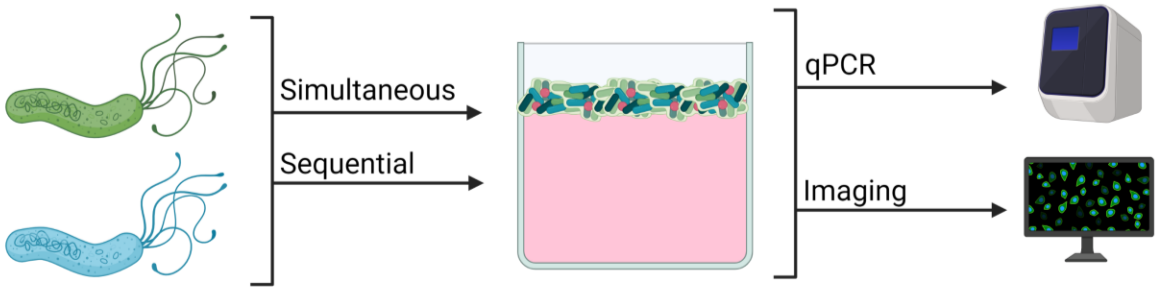

Figure 6.8: Experimental design to investigate the dynamics of *H. pylori* biofilm formation. Biofilm assays would be set up with either simultaneous or sequential inoculations, using multiple fluorescently-tagged barcoded *H. pylori* strains. Biofilm composition would be quantified by SYBR qPCR or next-generation sequencing, and the localization of individual strains within the biofilm would be imaged via fluorescence microscopy. Schematic created with BioRender.com

PUBLICATIONS

Harvey ML, Lin AS, Sun L, Koyama T, Shuman JHB, Loh JT, Algood HMS, Scholz MB, McClain MS, Cover TL. Enhanced fitness of a *Helicobacter pylori babA* mutant in a murine model. Infect Immun. Sep 2021, 89(10):e0072520. DOI: 10.1128/IAI.00725-20.

Lin AS*, McClain MS*, Beckett AC, Caston RR, **Harvey ML**, Dixon BREA, Campbell AM, Shuman JH, Sawhney N, Delgado AG, Loh JT, Piazuelo MB, Algood HM, Cover TL. Temporal control of the *Helicobacter pylori* Cag type IV secretion system *in vivo*. mBio. Jun 2020, 11 (3) e01296-20; DOI: 10.1128/mBio.01296-20.

Lin AS, Dooyema SDR, Frick-Cheng AE, **Harvey ML**, Suarez G, Loh JT, McDonald WH, McClain MS, Peek Jr. RM, Cover TL. Bacterial energetic requirements for *Helicobacter pylori* Cag type IV secretion system-dependent alterations in gastric epithelial cells. Infect. Immun. Jan 2020, 88(2); DOI: 10.1128/IAI.00790-19.

Hu, B., Khara, P., Song, L., Lin, A. S., Frick-Cheng, A., **Harvey, ML**, Cover, T. L., Christie, P. J. *In situ* molecular architecture of the *Helicobacter pylori* Cag Type IV secretion system. mBio. May 2019, 10(3); DOI: 10.1128/mBio.00849-19.

REFERENCES

1. Oleastro M, Ménard A. 2013. The role of Helicobacter pylori outer membrane proteins in adherence and pathogenesis doi:10.3390/biology2031110.
2. Bastos J, Peleteiro B, Barros R, Alves L, Severo M, de Fatima Pina M, Pinto H, Carvalho S, Marinho A, Guimaraes JT, Azevedo A, La Vecchia C, Barros H, Lunet N. 2013. Sociodemographic determinants of prevalence and incidence of Helicobacter pylori infection in Portuguese adults. Helicobacter 18:413-22.
3. Moayyedi P, Axon AT, Feltbower R, Duffett S, Crocombe W, Braunholtz D, Richards ID, Dowell AC, Forman D, Leeds HSG. 2002. Relation of adult lifestyle and socioeconomic factors to the prevalence of Helicobacter pylori infection. Int J Epidemiol 31:624-31.
4. van Blankenstein M, van Vuuren AJ, Looman CW, Ouwendijk M, Kuipers EJ. 2013. The prevalence of Helicobacter pylori infection in the Netherlands. Scand J Gastroenterol 48:794-800.
5. Burucoa C, Axon A. 2017. Epidemiology of Helicobacter pylori infection. Helicobacter 22 Suppl 1.
6. Puculek M, Machlowska J, Wierzbicki R, Baj J, Maciejewski R, Sitarz R. 2018. Helicobacter pylori associated factors in the development of gastric cancer with special reference to the early-onset subtype. Oncotarget 9:31146-31162.
7. Gonzalez CA, Megraud F, Buissonniere A, Lujan Barroso L, Agudo A, Duell EJ, Boutron-Ruault MC, Clavel-Chapelon F, Palli D, Krogh V, Mattiello A, Tumino R, Sacerdote C, Quiros JR, Sanchez-Cantalejo E, Navarro C, Barricarte A, Dorronsoro M, Khaw KT, Wareham N, Allen NE, Tsilidis KK, Bas Bueno-de-Mesquita H, Jeurnink SM, Numans ME, Peeters PHM, Lagiou P, Valanou E, Trichopoulou A, Kaaks R, Lukanova-McGregor A, Bergman MM, Boeing H, Manjer J, Lindkvist B, Stenling R, Hallmans G, Mortensen LM, Overvad K, Olsen A, Tjonneland A, Bakken K, Dumeaux V, Lund E, Jenab M, Romieu I, Michaud D, Mouw T, Carneiro F, Fenge C, et al. 2012. Helicobacter pylori infection assessed by ELISA and by immunoblot and noncardia gastric cancer risk in a prospective study: the Eurgast-EPIC project. Ann Oncol 23:1320-1324.
8. Machlowska J, Kapusta P, Baj J, Morsink FHM, Wolkow P, Maciejewski R, Offerhaus GJA, Sitarz R. 2020. High-Throughput Sequencing of Gastric Cancer Patients: Unravelling Genetic Predispositions Towards an Early-Onset Subtype. Cancers (Basel) 12.
9. Plummer M, Franceschi S, Vignat J, Forman D, de Martel C. 2015. Global burden of gastric cancer attributable to Helicobacter pylori. Int J Cancer 136:487-90.
10. Holcombe C, Omotara BA, Eldridge J, Jones DM. 1992. H. pylori, the most common bacterial infection in Africa: a random serological study. Am J Gastroenterol 87:28-30.
11. Agha A, Graham DY. 2005. Evidence-based examination of the African enigma in relation to Helicobacter pylori infection. Scand J Gastroenterol 40:523-9.
12. Graham DY, Lu H, Yamaoka Y. 2009. African, Asian or Indian enigma, the East Asian Helicobacter pylori: facts or medical myths. J Dig Dis 10:77-84.
13. Nikaido H. 2003. Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev 67:593-656.
14. Cullen TW, Giles DK, Wolf LN, Ecobichon C, Boneca IG, Trent MS. 2011. Helicobacter pylori versus the host: remodeling of the bacterial outer membrane is required for survival in the gastric mucosa. PLoS Pathog 7:e1002454.
15. Teghanemt A, Zhang D, Levis EN, Weiss JP, Gioannini TL. 2005. Molecular basis of reduced potency of underacylated endotoxins. J Immunol 175:4669-76.
16. Schleiff E, Soll J. 2005. Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. EMBO Rep 6:1023-7.

17. Tamm LK, Hong H, Liang B. 2004. Folding and assembly of beta-barrel membrane proteins. Biochim Biophys Acta 1666:250-63.
18. Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J. 2003. Role of a highly conserved bacterial protein in outer membrane protein assembly. Science 299:262-5.
19. Walther DM, Rapaport D, Tommassen J. 2009. Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. Cell Mol Life Sci 66:2789-804.
20. Fairman JW, Noinaj N, Buchanan SK. 2011. The structural biology of beta-barrel membrane proteins: a summary of recent reports. Curr Opin Struct Biol 21:523-31.
21. Wimley WC. 2003. The versatile beta-barrel membrane protein. Curr Opin Struct Biol 13:404-11.
22. Tomasek D, Kahne D. 2021. The assembly of beta-barrel outer membrane proteins. Curr Opin Microbiol 60:16-23.
23. Koebnik R, Locher KP, Van Gelder P. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol Microbiol 37:239-53.
24. Oleastro M, Cordeiro R, Menard A, Yamaoka Y, Queiroz D, Megraud F, Monteiro L. 2009. Allelic diversity and phylogeny of homB, a novel co-virulence marker of Helicobacter pylori. BMC Microbiol 9:248.
25. Harvey VC, Acio CR, Bredehoft AK, Zhu L, Hallinger DR, Quinlivan-Repasi V, Harvey SE, Forsyth MH. 2014. Repetitive sequence variations in the promoter region of the adhesin-encoding gene sabA of Helicobacter pylori affect transcription. J Bacteriol 196:3421-9.
26. Coppens F, Castaldo G, Debraekeleer A, Subedi S, Moonens K, Lo A, Remaut H. 2018. Hop-family Helicobacter outer membrane adhesins form a novel class of Type 5-like secretion proteins with an interrupted β-barrel domain. Molecular Microbiology doi:10.1111/mmi.14075.
27. Colbeck JC, Hansen LM, Fong JM, Solnick JV. 2006. Genotypic profile of the outer membrane proteins BabA and BabB in clinical isolates of Helicobacter pylori. Infect Immun 74:4375-8.
28. Backstrom A, Lundberg C, Kersulyte D, Berg DE, Boren T, Arnqvist A. 2004. Metastability of Helicobacter pylori bab adhesin genes and dynamics in Lewis b antigen binding. Proc Natl Acad Sci U S A 101:16923-8.
29. Alm RA, Bina J, Andrews BM, Doig P, Hancock REW, Trust TJ. 2000. Comparative genomics of Helicobacter pylori: Analysis of the outer membrane protein families. Infection and Immunity doi:10.1128/IAI.68.7.4155-4168.2000.
30. Pang SS, Nguyen STS, Perry AJ, Day CJ, Panjikar S, Tiralongo J, Whisstock JC, Kwok T. 2014. The three-dimensional structure of the extracellular adhesion domain of the sialic acid-binding adhesin SabA from Helicobacter pylori. J Biol Chem 289:6332-6340.
31. Hage N, Howard T, Phillips C, Brassington C, Overman R, Debreczeni J, Gellert P, Stolnik S, Winkler GS, Falcone FH. 2015. Structural basis of Lewis(b) antigen binding by the Helicobacter pylori adhesin BabA. Sci Adv 1:e1500315.
32. Moonens K, Gideonsson P, Subedi S, Bugaytsova J, Romao E, Mendez M, Norden J, Fallah M, Rakhimova L, Shevtsova A, Lahmann M, Castaldo G, Brannstrom K, Coppens F, Lo AW, Ny T, Solnick JV, Vandenbussche G, Oscarson S, Hammarstrom L, Arnqvist A, Berg DE, Muyldermans S, Boren T, Remaut H. 2016. Structural Insights into Polymorphic ABO Glycan Binding by Helicobacter pylori. Cell Host Microbe 19:55-66.
33. Javaheri A, Kruse T, Moonens K, Mejías-Luque R, Debraekeleer A, Asche CI, Tegtmeyer N, Kalali B, Bach NC, Sieber SA, Hill DJ, Königer V, Hauck CR, Moskalenko R, Haas R, Busch DH, Klaile E, Slevogt H, Schmidt A, Backert S, Remaut H, Singer BB, Gerhard M. 2016. Helicobacter pylori adhesin HopQ engages in a virulence-enhancing interaction with human CEACAMs. Nature Microbiology doi:10.1038/nmicrobiol.2016.189.

34. Bugaytsova JA, Bjornham O, Chernov YA, Gideonsson P, Henriksson S, Mendez M, Sjostrom R, Mahdavi J, Shevtsova A, Ilver D, Moonens K, Quintana-Hayashi MP, Moskalenko R, Aisenbrey C, Bylund G, Schmidt A, Aberg A, Brannstrom K, Koniger V, Vikstrom S, Rakhimova L, Hofer A, Ogren J, Liu H, Goldman MD, Whitmire JM, Aden J, Younson J, Kelly CG, Gilman RH, Chowdhury A, Mukhopadhyay AK, Nair GB, Papadakos KS, Martinez-Gonzalez B, Sgouras DN, Engstrand L, Unemo M, Danielsson D, Suerbaum S, Oscarson S, Morozova-Roche LA, Olofsson A, Grobner G, Holgersson J, Esberg A, Stromberg N, Landstrom M, Eldridge AM, Chromy BA, et al. 2017. Helicobacter pylori Adapts to Chronic Infection and Gastric Disease via pH-Responsive BabA-Mediated Adherence. Cell Host Microbe 21:376-389.
35. Sakamoto S, Watanabe T, Tokumaru T, Takagi H, Nakazato H, Lloyd KO. 1989. Expression of Lewisa, Lewisb, Lewisx, Lewisy, siayl-Lewisa, and sialyl-Lewisx blood group antigens in human gastric carcinoma and in normal gastric tissue. Cancer Res 49:745-52.
36. Mahdavi J, Sondén B, Hurtig M, Olfat FO, Forsberg L, Roche N, Ångström J, Larsson T, Teneberg S, Karlsson KA, Altraja S, Wadström T, Kersulyte D, Berg DE, Dubois A, Petersson C, Magnusson KE, Norberg T, Lindh F, Lundskog BB, Arnqvist A, Hammarström L, Borén T. 2002. Helicobacter pylori sabA adhesin in persistent infection and chronic inflammation. Science doi:10.1126/science.1069076.
37. Magalhaes A, Marcos-Pinto R, Nairn AV, Dela Rosa M, Ferreira RM, Junqueira-Neto S, Freitas D, Gomes J, Oliveira P, Santos MR, Marcos NT, Xiaogang W, Figueiredo C, Oliveira C, Dinis-Ribeiro M, Carneiro F, Moremen KW, David L, Reis CA. 2015. Helicobacter pylori chronic infection and mucosal inflammation switches the human gastric glycosylation pathways. Biochim Biophys Acta 1852:1928-39.
38. Kavermann H, Burns BP, Angermuller K, Odenbreit S, Fischer W, Melchers K, Haas R. 2003. Identification and characterization of Helicobacter pylori genes essential for gastric colonization. J Exp Med 197:813-22.
39. Baldwin DN, Shepherd B, Kraemer P, Hall MK, Sycuro LK, Pinto-Santini DM, Salama NR. 2007. Identification of Helicobacter pylori genes that contribute to stomach colonization. Infect Immun 75:1005-16.
40. Eaton KA, Krakowka S. 1994. Effect of gastric pH on urease-dependent colonization of gnotobiotic piglets by Helicobacter pylori. Infect Immun 62:3604-7.
41. Eaton KA, Suerbaum S, Josenhans C, Krakowka S. 1996. Colonization of gnotobiotic piglets by Helicobacter pylori deficient in two flagellin genes. Infection and Immunity doi:10.1128/iai.64.7.2445-2448.1996.
42. Tsuda M, Karita M, Morshed MG, Okita K, Nakazawa T. 1994. A urease-negative mutant of Helicobacter pylori constructed by allelic exchange mutagenesis lacks the ability to colonize the nude mouse stomach. Infect Immun 62:3586-9.
43. Salama NR, Shepherd B, Falkow S. 2004. Global transposon mutagenesis and essential gene analysis of Helicobacter pylori. Journal of Bacteriology doi:10.1128/JB.186.23.7926-7935.2004.
44. Kavermann H, Burns BP, Angermüller K, Odenbreit S, Fischer W, Melchers K, Haas R. 2003. Identification and characterization of Helicobacter pylori genes essential for gastric colonization. Journal of Experimental Medicine doi:10.1084/jem.20021531.
45. Linden S, Nordman H, Hedenbro J, Hurtig M, Boren T, Carlstedt I. 2002. Strain- and blood group-dependent binding of Helicobacter pylori to human gastric MUC5AC glycoforms. Gastroenterology 123:1923-30.
46. Jaff MS. 2010. Higher frequency of secretor phenotype in O blood group - its benefits in prevention and/or treatment of some diseases. Int J Nanomedicine 5:901-5.
47. Sherburne R, Taylor DE. 1995. Helicobacter pylori expresses a complex surface carbohydrate, Lewis X. Infect Immun 63:4564-8.

48. Moran AP. 1996. Helicobacter pylori expresses Lewis X. Helicobacter 1:190-1.
49. Negrini R, Savio A, Poiesi C, Appelmelk BJ, Buffoli F, Paterlini A, Cesari P, Graffeo M, Vaira D, Franzin G. 1996. Antigenic mimicry between Helicobacter pylori and gastric mucosa in the pathogenesis of body atrophic gastritis. Gastroenterology 111:655-65.
50. Hennig EE, Allen JM, Cover TL. 2006. Multiple chromosomal loci for the babA gene in Helicobacter pylori. Infect Immun 74:3046-51.
51. Hansen LM, Gideonsson P, Canfield DR, Boren T, Solnick JV. 2017. Dynamic Expression of the BabA Adhesin and Its BabB Paralog during Helicobacter pylori Infection in Rhesus Macaques. Infect Immun 85.
52. Kable ME, Hansen LM, Styer CM, Deck SL, Rakhimova O, Shevtsova A, Eaton KA, Martin ME, Gideonsson P, Boren T, Solnick JV. 2017. Host Determinants of Expression of the Helicobacter pylori BabA Adhesin. Sci Rep 7:46499.
53. Benktander J, Angstrom J, Breimer ME, Teneberg S. 2012. Redefinition of the carbohydrate binding specificity of Helicobacter pylori BabA adhesin. J Biol Chem 287:31712-24.
54. Ansari S, Yamaoka Y. 2017. Helicobacter pylori BabA in adaptation for gastric colonization. World J Gastroenterol 23:4158-4169.
55. Morozov V, Borkowski J, Hanisch FG. 2018. The Double Face of Mucin-Type O-Glycans in Lectin-Mediated Infection and Immunity. Molecules 23.
56. Magalhaes A, Reis CA. 2010. Helicobacter pylori adhesion to gastric epithelial cells is mediated by glycan receptors. Braz J Med Biol Res 43:611-8.
57. Odenbreit S, Swoboda K, Barwig I, Ruhl S, Borén T, Koletzko S, Haas R. 2009. Outer membrane protein expression profile in Helicobacter pylori clinical isolates. Infection and Immunity doi:10.1128/IAI.00364-09.
58. Chang WL, Yeh YC, Sheu BS. 2018. The impacts of H. pylori virulence factors on the development of gastroduodenal diseases. J Biomed Sci 25:68.
59. Ishijima N, Suzuki M, Ashida H, Ichikawa Y, Kanegae Y, Saito I, Borén T, Haas R, Sasakawa C, Mimuro H. 2011. BabA-mediated adherence is a potentiator of the helicobacter pylori type IV secretion system activity. Journal of Biological Chemistry doi:10.1074/jbc.M111.233601.
60. Yamaoka Y. 2008. Roles of Helicobacter pylori BabA in gastroduodenal pathogenesis. World J Gastroenterol 14:4265-72.
61. Styer CM, Hansen LM, Cooke CL, Gundersen AM, Choi SS, Berg DE, Benghezal M, Marshall BJ, Peek RM, Borén T, Solnick JV. 2010. Expression of the BabA adhesin during experimental infection with Helicobacter pylori. Infection and Immunity doi:10.1128/IAI.01297-09.
62. Solnick JV, Hansen LM, Salama NR, Boonjakuakul JK, Syvanen M. 2004. Modification of Helicobacter pylori outer membrane protein expression during experimental infection of rhesus macaques. Proceedings of the National Academy of Sciences of the United States of America doi:10.1073/pnas.0308573100.
63. Ohno T, Vallström A, Rugge M, Ota H, Graham DY, Arnqvist A, Yamaoka Y. 2011. Effects of blood group antigen-binding adhesin expression during Helicobacter pylori infection of Mongolian gerbils. Journal of Infectious Diseases doi:10.1093/infdis/jiq090.
64. Unemo M, Aspholm-Hurtig M, Ilver D, Bergstrom J, Boren T, Danielsson D, Teneberg S. 2005. The sialic acid binding SabA adhesin of Helicobacter pylori is essential for nonopsonic activation of human neutrophils. J Biol Chem 280:15390-7.
65. Dossumbekova A, Prinz C, Mages J, Lang R, Kusters JG, Van Vliet AH, Reindl W, Backert S, Saur D, Schmid RM, Rad R. 2006. Helicobacter pylori HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of hopH gene polymorphisms. J Infect Dis 194:1346-55.

66. Miftahussurur M, Yamaoka Y. 2015. Helicobacter pylori virulence genes and host genetic polymorphisms as risk factors for peptic ulcer disease. Expert Rev Gastroenterol Hepatol 9:1535-47.
67. Yamaoka Y, Kita M, Kodama T, Imamura S, Ohno T, Sawai N, Ishimaru A, Imanishi J, Graham DY. 2002. Helicobacter pylori infection in mice: Role of outer membrane proteins in colonization and inflammation. Gastroenterology 123:1992-2004.
68. Dossumbekova A, Prinz C, Mages J, Lang R, Kusters Johannes G, Van Vliet Arnoud HM, Reindl W, Backert S, Saur D, Schmid Roland M, Rad R. 2006. Helicobacter pylori HopH (OipA) and Bacterial Pathogenicity: Genetic and Functional Genomic Analysis of hopH Gene Polymorphisms. The Journal of Infectious Diseases doi:10.1086/508426.
69. Matsuo Y, Kido Y, Yamaoka Y. 2017. Helicobacter pylori outer membrane protein-related pathogenesis doi:10.3390/toxins9030101.
70. Liu J, He C, Chen M, Wang Z, Xing C, Yuan Y. 2013. Association of presence/absence and on/off patterns of Helicobacter pylori oipA gene with peptic ulcer disease and gastric cancer risks: a meta-analysis. BMC Infect Dis 13:555.
71. Alm RA, Trust TJ. 1999. Analysis of the genetic diversity of Helicobacter pylori: the tale of two genomes. J Mol Med (Berl) 77:834-46.
72. Alm RA, Bina J, Andrews BM, Doig P, Hancock RE, Trust TJ. 2000. Comparative genomics of Helicobacter pylori: analysis of the outer membrane protein families. Infect Immun 68:4155-68.
73. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R, Stadler PF, Vogel J. 2010. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature 464:250-5.
74. Rokbi B, Seguin D, Guy B, Mazarin V, Vidor E, Mion F, Cadoz M, Quentin-Millet MJ. 2001. Assessment of Helicobacter pylori gene expression within mouse and human gastric mucosae by real-time reverse transcriptase PCR. Infection and Immunity doi:10.1128/IAI.69.8.4759-4766.2001.
75. Odenbreit S, Till M, Hofreuter D, Faller G, Haas R. 1999. Genetic and functional characterization of the alpAB gene locus essential for the adhesion of Helicobacter pylori to human gastric tissue. Molecular Microbiology doi:10.1046/j.1365-2958.1999.01300.x.
76. Odenbreit S, Faller G, Haas R. 2002. Role of the alpAB proteins and lipopolysaccharide in adhesion of Helicobacter pylori to human gastric tissue. Int J Med Microbiol 292:247-56.
77. Senkovich OA, Yin J, Ekshyyan V, Conant C, Traylor J, Adegboyega P, McGee DJ, Rhoads RE, Slepenkov S, Testerman TL. 2011. Helicobacter pylori AlpA and AlpB Bind host laminin and influence gastric inflammation in gerbils. Infection and Immunity doi:10.1128/IAI.01275-10.
78. Lu H, Jeng YW, Beswick EJ, Ohno T, Odenbreit S, Haas R, Reyes VE, Kita M, Graham DY, Yamaoka Y. 2007. Functional and intracellular signaling differences associated with the Helicobacter pylori AlpAB adhesin from Western and East Asian strains. Journal of Biological Chemistry doi:10.1074/jbc.M611178200.
79. Yonezawa H, Osaki T, Fukutomi T, Hanawa T, Kurata S, Zaman C, Hojo F, Kamiya S. 2017. Diversification of the AlpB Outer Membrane Protein of Helicobacter pylori Affects Biofilm Formation and Cellular Adhesion. J Bacteriol 199.
80. De Jonge R, Durrani Z, Rijpkema SG, Kuipers EJ, Van Vliet AHM, Kusters JG. 2004. Role of the Helicobacter pylori outer-membrane proteins AlpA and AlpB in colonization of the guinea pig stomach. Journal of Medical Microbiology doi:10.1099/jmm.0.45551-0.
81. Sugimoto M, Ohno T, Graham DY, Yamaoka Y. 2011. Helicobacter pylori outer membrane proteins on gastric mucosal interleukin 6 and 11 expression in Mongolian gerbils. Journal of Gastroenterology and Hepatology (Australia) doi:10.1111/j.1440-1746.2011.06817.x.

82. Königer V, Holsten L, Harrison U, Busch B, Loell E, Zhao Q, Bonsor DA, Roth A, Kengmo-Tchoupa A, Smith SI, Mueller S, Sundberg EJ, Zimmermann W, Fischer W, Hauck CR, Haas R. 2016. Helicobacter pylori exploits human CEACAMs via HopQ for adherence and translocation of CagA. Nature Microbiology doi:10.1038/nmicrobiol.2016.188.
83. Cao P, Lee KJ, Blaser MJ, Cover TL. 2005. Analysis of hopQ alleles in East Asian and Western strains of Helicobacter pylori. FEMS Microbiol Lett 251:37-43.
84. Abadi AT, Mobarez AM, Bonten MJ, Wagenaar JA, Kusters JG. 2014. Clinical relevance of the cagA, tnpA and tnpB genes in Helicobacter pylori. BMC Gastroenterol 14:33.
85. Leylabadlo HE, Yekani M, Ghotaslou R. 2016. Helicobacter pylori hopQ alleles (type I and II) in gastric cancer. Biomed Rep 4:601-604.
86. Yakoob J, Abbas Z, Khan R, Salim SA, Awan S, Abrar A, Jafri W. 2016. Helicobacter pylori outer membrane protein Q allele distribution is associated with distinct pathologies in Pakistan. Infect Genet Evol 37:57-62.
87. Javaheri A, Kruse T, Moonens K, Mejias-Luque R, Debraekeleer A, Asche CI, Tegtmeyer N, Kalali B, Bach NC, Sieber SA, Hill DJ, Koniger V, Hauck CR, Moskalenko R, Haas R, Busch DH, Klaile E, Slevogt H, Schmidt A, Backert S, Remaut H, Singer BB, Gerhard M. 2016. Helicobacter pylori adhesin HopQ engages in a virulence-enhancing interaction with human CEACAMs. Nat Microbiol 2:16189.
88. Bonsor DA, Zhao Q, Schmidinger B, Weiss E, Wang J, Deredge D, Beadenkopf R, Dow B, Fischer W, Beckett D, Wintrode PL, Haas R, Sundberg EJ. 2018. The Helicobacter pylori adhesin protein HopQ exploits the dimer interface of human CEACAMs to facilitate translocation of the oncoprotein CagA. EMBO J 37.
89. Gur C, Maalouf N, Gerhard M, Singer BB, Emgard J, Temper V, Neuman T, Mandelboim O, Bachrach G. 2019. The Helicobacter pylori HopQ outermembrane protein inhibits immune cell activities. Oncoimmunology 8:e1553487.
90. Rossez Y, Gosset P, Boneca IG, Magalhães A, Ecobichon C, Reis CA, Cieniewski-Bernard C, Curt MJC, Léonard R, Maes E, Sperandio B, Slomianny C, Sansonetti PJ, Michalski JC, Robbe-Masselot C. 2014. The lacdiNAc-specific adhesin LabA mediates adhesion of Helicobacter pylori to human gastric mucosa. Journal of Infectious Diseases doi:10.1093/infdis/jiu239.
91. Paraskevopoulou V, Schimpl M, Overman RC, Stolnik S, Chen Y, Nguyen L, Winkler GS, Gellert P, Klassen JS, Falcone FH. 2021. Structural and binding characterization of the LacdiNAc-specific adhesin (LabA; HopD) exodomain from Helicobacter pylori. Curr Res Struct Biol 3:19-29.
92. Exner MM, Doig P, Trust TJ, Hancock REW. 1995. Isolation and characterization of a family of porin proteins from Helicobacter pylori. Infection and Immunity doi:10.1128/iai.63.4.1567-1572.1995.
93. Mthembu YH, Jin C, Padra M, Liu J, Edlund JO, Ma H, Padra J, Oscarson S, Boren T, Karlsson NG, Linden SK, Holgersson J. 2020. Recombinant mucin-type proteins carrying LacdiNAc on different O-glycan core chains fail to support H. pylori binding. Mol Omics 16:243-257.
94. Peck B, Ortkamp M, Diehl KD, Hundt E, Knapp B. 1999. Conservation, localization and expression of HopZ, a protein involved in adhesion of Helicobacter pylori. Nucleic Acids Research doi:10.1093/nar/27.16.3325.
95. Kennemann L, Brenneke B, Andres S, Engstrand L, Meyer TF, Aebischer T, Josenhans C, Suerbaum S. 2012. In vivo sequence variation in HopZ, a phase-variable outer membrane protein of Helicobacter pylori. Infect Immun 80:4364-73.
96. Giannakis M, Backhed HK, Chen SL, Faith JJ, Wu M, Guruge JL, Engstrand L, Gordon JI. 2009. Response of gastric epithelial progenitors to Helicobacter pylori Isolates obtained from Swedish patients with chronic atrophic gastritis. J Biol Chem 284:30383-94.

97. Scott DR, Weeks D, Hong C, Postius S, Melchers K, Sachs G. 1998. The role of internal urease in acid resistance of Helicobacter pylori. Gastroenterology 114:58-70.
98. Eaton KA, Morgan DR, Krakowka S. 1992. Motility as a factor in the colonisation of gnotobiotic piglets by Helicobacter pylori. J Med Microbiol 37:123-7.
99. Lertsethtakarn P, Ottemann KM, Hendrixson DR. 2011. Motility and chemotaxis in Campylobacter and Helicobacter. Annu Rev Microbiol 65:389-410.
100. Paul K, Gonzalez-Bonet G, Bilwes AM, Crane BR, Blair D. 2011. Architecture of the flagellar rotor. EMBO J 30:2962-71.
101. Loconte V, Kekez I, Matkovic-Calogovic D, Zanotti G. 2017. Structural characterization of FlgE2 protein from Helicobacter pylori hook. FEBS J 284:4328-4342.
102. Leying H, Suerbaum S, Geis G, Haas R. 1992. Cloning and genetic characterization of a Helicobacter pylori flagellin gene. Mol Microbiol 6:2863-74.
103. Ottemann KM, Lowenthal AC. 2002. Helicobacter pylori uses motility for initial colonization and to attain robust infection. Infect Immun 70:1984-90.
104. Kim JS, Chang JH, Chung SI, Yum JS. 1999. Molecular cloning and characterization of the Helicobacter pylori fliD gene, an essential factor in flagellar structure and motility. J Bacteriol 181:6969-76.
105. Nakajima K, Inatsu S, Mizote T, Nagata Y, Aoyama K, Fukuda Y, Nagata K. 2008. Possible involvement of put A gene in Helicobacter pylori colonization in the stomach and motility. Biomed Res 29:9-18.
106. Eaton KA, Suerbaum S, Josenhans C, Krakowka S. 1996. Colonization of gnotobiotic piglets by Helicobacter pylori deficient in two flagellin genes. Infect Immun 64:2445-8.
107. Josenhans C, Labigne A, Suerbaum S. 1995. Comparative ultrastructural and functional studies of Helicobacter pylori and Helicobacter mustelae flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in Helicobacter species. J Bacteriol 177:3010-20.
108. Linz B, Windsor HM, McGraw JJ, Hansen LM, Gajewski JP, Tomsho LP, Hake CM, Solnick JV, Schuster SC, Marshall BJ. 2014. A mutation burst during the acute phase of Helicobacter pylori infection in humans and rhesus macaques. Nat Commun 5:4165.
109. Dubois A, Berg DE, Incecik ET, Fiala N, Heman-Ackah LM, Del Valle J, Yang M, Wirth HP, Perez-Perez GI, Blaser MJ. 1999. Host specificity of Helicobacter pylori strains and host responses in experimentally challenged nonhuman primates. Gastroenterology 116:90-6.
110. Audano P, Vannberg F. 2014. KAnalyze: a fast versatile pipelined k-mer toolkit. Bioinformatics 30:2070-2.
111. Guo BP, Mekalanos JJ. 2002. Rapid genetic analysis of Helicobacter pylori gastric mucosal colonization in suckling mice. Proc Natl Acad Sci U S A 99:8354-9.
112. Jimenez-Soto LF, Rohrer S, Jain U, Ertl C, Sewald X, Haas R. 2012. Effects of cholesterol on Helicobacter pylori growth and virulence properties in vitro. Helicobacter 17:133-9.
113. Hawrylik SJ, Wasilko DJ, Haskell SL, Gootz TD, Lee SE. 1994. Bisulfite or sulfite inhibits growth of Helicobacter pylori. J Clin Microbiol 32:790-2.
114. Lin AS, McClain MS, Beckett AC, Caston RR, Harvey ML, Dixon B, Campbell AM, Shuman JHB, Sawhney N, Delgado AG, Loh JT, Piazuelo MB, Algood HMS, Cover TL. 2020. Temporal Control of the Helicobacter pylori Cag Type IV Secretion System in a Mongolian Gerbil Model of Gastric Carcinogenesis. mBio 11:e01296-20.
115. McClain MS, Voss BJ, Cover TL. 2020. Lipoprotein Processing and Sorting in Helicobacter pylori. mBio 11:e00911-20.
116. Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol Chapter 2:Unit 2 4.
117. Goodwin A, Kersulyte D, Sisson G, Veldhuyzen van Zanten SJ, Berg DE, Hoffman PS. 1998. Metronidazole resistance in Helicobacter pylori is due to null mutations in a gene

(rdxA) that encodes an oxygen-insensitive NADPH nitroreductase. Mol Microbiol 28:383-93.
118. Scott Algood HM, Gallo-Romero J, Wilson KT, Peek RM, Cover TL. 2007. Host response to Helicobacter pylori infection before initiation of the adaptive immune response. FEMS Immunology and Medical Microbiology doi:10.1111/j.1574-695X.2007.00338.x.
119. Servetas SL, Doster RS, Kim A, Windham IH, Cha JH, Gaddy JA, Merrell DS. 2018. ArsRS-Dependent Regulation of homB Contributes to Helicobacter pylori Biofilm Formation. Front Microbiol 9:1497.
120. Hathroubi S, Zerebinski J, Ottemann KM. 2018. Helicobacter pylori Biofilm Involves a Multigene Stress-Biased Response, Including a Structural Role for Flagella. mBio 9.
121. Wong EH, Ng CG, Chua EG, Tay AC, Peters F, Marshall BJ, Ho B, Goh KL, Vadivelu J, Loke MF. 2016. Comparative Genomics Revealed Multiple Helicobacter pylori Genes Associated with Biofilm Formation In Vitro. PLoS One 11:e0166835.
122. Shao C, Sun Y, Wang N, Yu H, Zhou Y, Chen C, Jia J. 2013. Changes of proteome components of Helicobacter pylori biofilms induced by serum starvation. Mol Med Rep 8:1761-6.
123. Voss BJ, Gaddy JA, McDonald WH, Cover TL. 2014. Analysis of surface-exposed outer membrane proteins in Helicobacter pylori. Journal of Bacteriology doi:10.1128/JB.01768-14.
124. Fung C, Tan S, Nakajima M, Skoog EC, Camarillo-Guerrero LF, Klein JA, Lawley TD, Solnick JV, Fukami T, Amieva MR. 2019. High-resolution mapping reveals that microniches in the gastric glands control Helicobacter pylori colonization of the stomach. PLoS Biol 17:e3000231.
125. McClain MS, Duncan SS, Gaddy JA, Cover TL. 2013. Control of gene expression in Helicobacter pylori using the Tet repressor. J Microbiol Methods 95:336-41.
126. Windham IH, Merrell DS. 2020. Analysis of fitness costs associated with metronidazole and amoxicillin resistance in Helicobacter pylori. Helicobacter 25:e12724.
127. Dailidiene D, Dailide G, Kersulyte D, Berg DE. 2006. Contraselectable streptomycin susceptibility determinant for genetic manipulation and analysis of Helicobacter pylori. Appl Environ Microbiol 72:5908-14.
128. Jeong JY, Mukhopadhyay AK, Dailidiene D, Wang Y, Velapatino B, Gilman RH, Parkinson AJ, Nair GB, Wong BC, Lam SK, Mistry R, Segal I, Yuan Y, Gao H, Alarcon T, Brea ML, Ito Y, Kersulyte D, Lee HK, Gong Y, Goodwin A, Hoffman PS, Berg DE. 2000. Sequential inactivation of rdxA (HP0954) and frxA (HP0642) nitroreductase genes causes moderate and high-level metronidazole resistance in Helicobacter pylori. J Bacteriol 182:5082-90.
129. Abel S, Abel zur Wiesch P, Davis BM, Waldor MK. 2015. Analysis of Bottlenecks in Experimental Models of Infection. PLoS Pathog 11:e1004823.