UNIFORMED SERVICES UNIVERSITY, SCHOOL OF MEDICINE GRADUATE PROGRAMS
Graduate Education Office (A 1045), 4301 Jones Bridge Road, Bethesda, MD 20814

DISSERTATION APPROVAL FOR THE DOCTORAL DISSERTATION IN THE EMERGING INFECTIOUS DISEASES GRADUATE PROGRAM

Title of Dissertation: "*Heliobactor pylori* Virulence Factors and Their Role in Pathogenesis"

Name of Candidate: Kathy R. Jones
Doctor of Philosophy Degree
March 25, 2011

DISSERTATION AND ABSTRACT APPROVED:

DATE:

Ann Jerse — 3/25/11
Dr. Ann Jerse
DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY
Committee Chairperson

D S Merrell — 4/1/11
Dr. D. Scott Merrell
DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY
Dissertation Advisor

3/25/2011
Dr. Stephen J. Davies
DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY
Committee Member

03-25-11
Dr. Anthony T. Maurelli
DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY
Committee Member

3-25-11
Dr. Tod J. Merkel
CBER, FDA
Committee Member

Copyright Statement

The author hereby certifies that the use of any copyrighted material in the thesis manuscript entitled:

"*Helicobacter pylori* Virulence Factors and Their Role in Pathogenesis"

is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.

Kathleen R. Jones

Kathleen R. Jones

Emerging Infectious Diseases Program

Department of Microbiology and Immunology

Uniformed Services University of the Health Sciences

Uniformed Services University of the Health Sciences

Abstract

Title of Dissertation:

Helicobacter pylori Virulence Factors and Their Role in Pathogenesis

Kathleen R. Jones, Doctor of Philosophy, 2011

Thesis directed by:

D. Scott Merrell, Ph.D.
Associate Professor, Department of Microbiology and Immunology

Helicobacter pylori is a Gram negative, microaerophilic, spiral shaped bacterium that is the causative agent of a variety of gastric maladies: gastritis, peptic ulcers (both duodenal and gastric), and two forms of gastric cancer. Currently, several bacterial virulence factors have been associated with more severe gastric disease, including specific polymorphismic forms of two bacterial toxins, CagA and VacA. These toxins have been shown to have numerous effects on host cells, including modulation of multiple cellular pathways that appear to ultimately lead to disease. Through the process of completing this thesis, we were the first group to show an association between East Asian CagA (EPIYA-ABD) and progression to gastric cancer in a South Korean population. We also examined the role of VacA polymorphisms within that population, and found that while the distribution of *vacA* alleles was not associated with disease state, it was associated with the distribution of *cagA* alleles and was integral in a three way interaction with the distribution of *cagA* alleles and disease state. Next, we analyzed the contribution of the newly described i region of VacA to disease development, and

identified an amino acid (196) that was important for the development of gastric cancer. Additionally, we identified some associations that were CagA-dependent, such as the association of VacA and disease state in the EPIYA-ABD population and the association of the distribution of amino acids at position 231 and disease state in the non EPIYA-ABD population. Moreover, we were able to optimize techniques that will ultimately be used to characterize CagA isogenic strains. Those future studies will serve to define the role of specific EPIYA motifs in *H. pylori*-induced host cellular damages both *in vitro* and *in vivo*. *En masse*, these data add to what we know about the complexity of *H. pylori*-induced pathogenesis, and it is becoming increasingly more evident that polymorphisms within CagA and VacA, alone and in concert, affect *H. pylori*- induced disease. However, the reason why only a portion of the population develops gastric cancer still remains unclear.

Helicobacter pylori Virulence Factors and Their Role in Pathogenesis

By

Kathleen R. Jones

Dissertation submitted to the Faculty of the

Emerging Infectious Diseases Interdisciplinary Graduate Program

of the Uniformed Services University of the Health Sciences

F. Edward Hébert School of Medicine

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy 2011

Acknowledgements

To my advisor, Scotty – thank you for this opportunity and your patience.

To my thesis committee – thank you for your support throughout this process.

To the Helicobacterites of the Merrell Lab – You are truly wonderful! Thank you for everything including: your extra hands, ideas, friendship, and laughter through good times and bad.

To Dr. Cha – thank you for your friendship, support, and collaboration.

To the faculty and staff – thank you for your support and wisdom.

To Cara and the staffs of BIC and LAM – thank you for being patient and teaching me so much.

To my fellow students – thank you! The laughter you introduced through this process has been invaluable.

To my outstanding friends – I cannot thank you enough for being there through EVERYTHING!!!!

To Michele and Kaleb Captain – thank you for always being there.

To Deb – thank you putting up with me for the last three plus years, for grounding me, for making me smile and laugh, for being my biggest advocate on bad days, and for being a partner in crime!

To Reyda and Esteban – Muchas gracias por dejarme ser parte de vuestra familia, por haber estado siempre alli cuando os necesitaba y por todos los momentos inolvidables que hemos compartido.

To Linda and Mike – thank you for adopting me and always being there.

To Jeff and Randi – thanks for giving me a family home closer to DC.

To Jeannette – there are not words to thank you enough. You are absolutely amazing, and without you, this would NOT have been possible. Thank you for being my "sister" and for adopting me into your family.

To the Roommates – thank you is SO inadequate, but thanks anyway.

To Gretchen and Kat – thank you for listening and being there. I would not be sane without your friendship.

To Sam, Greg, Jeff, and Justin – thank you for the laughter, smiles, sarcasm, and friendship. Your support has been invaluable!!!

To my kids, Brenna and Kevin – thank you for the hugs, love, and undying support that only true love can offer. I am so proud to be your aunt. I love you!!!

To Michelle- thank you for being there for the phone calls when things were great and rough, especially for the laughter on the long drives home. I love you!!!

To my best friend and partner in crime, Matt – thank you for everything. You have been there for every aspect my life (of course you are responsible for some of the trouble along the way), but I can't thank you enough. I love you big brother!!!

To my Mom – I am so proud and honored to be your daughter. Thank you for teaching me that there was nothing I couldn't do if I believed in it, teaching me to laugh, and for constantly teaching that no one can take my smile away. I love you!!!

To my Dad - I am so proud and honored to be your daughter. Thank you for teaching me that if I tried hard enough and stuck it out, that the prize would be worth it. You have been a constant source of support in this dream, and I could not have done this without you. I love you!!!

Dedication

To all of my friends who have stood beside me. To my family who constantly give me strength, faith, and love, especially for my dad.

Table of Contents

Approval Sheet i

Copyright Statement ii

Abstract iii

Title Page v

Acknowledgements vi

Dedication viii

Table of Contents ix

Tables of Tables xv

Table of Figures xvi

Chapter One: Introduction 1

Helicobacter pylori 1

Epidemiology and Disease 4

Treatment and Vaccines 10

Virulence Factors that Impact Host Cell Pathways 11

Cytotoxin-associated gene A - CagA 14

CagA and Disease 18

CagA Phosphorylation Independent Events 19

Physical Effects on Host Cells 19

Cellular Differentiation 20

Cell Cycle, Survival and Proliferation 21

Cell Scattering 23

Inflammation 24

CagA Phosphorylation Dependent Events 26

Targeting of SHP-2 26

Activation of ERK 26

Non-ERK Mediated Cytoskeletal Rearrangement and Scattering 27

CagA Feedback Loop, Src vs. Csk 28

Interactions with Unknown Function 29

CagA Independent/Redundancy 29

Vacuolating Cytotoxin A - VacA 33

VacA Functions 35

Anion Channel Formation and Vacuolation 35

Induction of Apoptosis 36

Disruption of Cellular Pathways 37

Immune Modulation 39

Interactions between CagA and VacA 41

Conclusion 42

Goals and Specific Aims 44

References 45

Chapter Two: Polymorphism in the CagA EPIYA Motif Impacts Development of Gastric Cancer 125

Abstract 125

Introduction 126

Materials and Methods 129

Bacterial Strains 129

Clinical Isolate Acquisition 129

cagA Genotyping 130

CagA Protein Expression 132

CagA Phosphorylation Assays 132

Cell Elongation Studies 134

IL-8 Induction Assay 134

Statistical Analysis 135

Nucleotide Sequence Accession Numbers 136

Results 136

Sample Acquisition/*cagA* Genotyping 136

CagA Protein Expression 146

Delivery and Phosphorylation of CagA 146

Cell Elongation Assay 149

Induction of IL-8 155

Discussion 155

Acknowledgements 160

References 161

Chapter Three: An Epidemiological Link Between Gastric Disease and Polymorphisms in VacA and CagA 184

Abstract 184

Introduction 185

Material and Methods 187

Bacterial Strains and Culture Conditions 187

vacA Genotyping 187

Statistical Analysis 191

Nucleotide Sequence Accession Numbers 192

Results 192

Sample Acquisition/*vacA* Genotyping 192

Distribution of *vacA* Allele and Gender 196

Associations among *vacA*, *cagA*, and Disease 196

Sequence Analysis of the Intermediate 1 Region 206

Discussion 209

Acknowledgements 213

References 214

Chapter Four: Polymorphisms in the Intermediate Region of VacA Impact *Helicobacter pylori*-Induced Disease Development 237

Abstract 237

Introduction 238

Material and Methods 242

Bacterial Strains and Culture Conditions 242

vacA i Region Sequencing 242

Statistical Analysis 243

Nucleotide Sequence Accession Numbers 243

Results 244

Sample Acquisition/*vacA* i Region Sequencing 244

Vac i Region, i1, i2, and i3 252

Amino Acid 151 256

Amino Acid 231 262

Amino Acid 196 263

Higher Order Associations 267

Discussion 267

Acknowledgements 272

References 273

Chapter Five: Construction and Analysis of Isogenic Strains of *Helicobacter pylori* Differing Only in the CagA EPIYA Motif 296

Introduction 296

Materials and Methods and Results 299

Bacterial Strain Growth Conditions 299

Creation of Isogenic Strains 303

ΔEPIYA Strain 303

Western Strains 312

EPIYA-AB^{t} Strain 313

EPIYA-AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC Strains 314

East Asian Strains 316

EPIYA-ABD Strain 316

EPIYA-AB^{t}D Strain 317

Restorant Strain 318

ΔcagA Strain 319

Growth Dynamics 321

Expression of CagA 324

Localization of Bacteria 326

Adherence and Internalization Assays 334

CagA Phosphorylation Assays 339

Morphological Changes 344

In vivo analysis of isogenic strains 348

Conclusions 355

References 360

Chapter Six: Discussion 366

Unanswered Questions Stemming from the Epidemiological Studies 367

CagA 367

VacA 371

Cag/Vac Interaction 375

A Hierarchy of Virulence Factors 377

Lessons Learned from the Current Project 378

Looking Forward 384

Importance and the Impact of Future Studies 387

Conclusions 390

References 392

Tables of Tables

Table 1: Primer Sequences 131

Table 2: Epidemiological Breakdown of the Korean Collection..... 137

Table 3: Analysis of CagA Expression and Function..... 151

Table S1: Complete Korean Collection……...171

Table 4: Primer sequences 188

Table 5: *vacA* genotyped isolates and disease state 193

Table 6: *vacA* genotype and *cagA* allele. 200

Table 7: *vacA* allele, *cagA* allele and disease state 203

Table S2: Complete Korean Collection……..225

Table 8: Distribution of *cagA* and *vacA* allele across the different disease states. 245

Table 9: Distribution of amino acids at polymorphic positions within the intermediate region 257

Table 10: P values of the distribution of amino acids at polymorphic positions across various disease states. 265

Table S3: Complete Korean Collection……..285

Table 11: Bacterial strains and plasmids 300

Table 12: Primer Sequences..... 302

Table of Figures

Figure 1: Cellular components targeted by commonly used antibiotics and mechanisms of resistance utilized by *H. pylori* 13

Figure 2: CagA and known host cell targets 17

Figure 3: VacA and known host cell targets 32

Figure 4: Genotyping of the *cagA* variable-EPIYA motif 142

Figure 5: Schematic depiction of the distribution of the *cagA* genotypes stratified by disease state within this Korean population 145

Figure 6: Expression, Delivery, and Phosphorylation of CagA 148

Figure 7: Morphological Changes and Induction of IL-8 157

Figure S1: Genotyping of the complex cagA variable EPIYA motifs……183

Figure 8: Genotyping of the *vacA* polymorphic regions 190

Figure 9: Schematic depiction of the distribution of the *vacA* genotypes stratified by disease state and *cagA* allele within this South Korean population 199

Figure 10: Amino acid alignment of the i1 type of VacA 208

Figure 11: Weblogo showing the VacA intermediate region's major polymorphic domains within the South Korean population 251

Figure 12: Consensus Sequences for the *vacA* intermediate region 255

Figure 13: Creation of isogenic strains 305

Figure 14: Creation of the various EPIYA motifs 309

Figure 14b: Creation of the various EPIYA motis…….3098

Figure 15: Growth dynamics of the isogenic strains 323

Figure 16: CagA Expression 328

Figure 17: Localization of bacteria on host cells 333

Figure 18: Adherence and Internalization 338

Figure 19: Phosphorylation of CagA 343

Figure 20: Morphology of infected host cells. 347

Figure 21: *In vivo* colonization 351

Figure 22: Isogenic strain-induced disease states 354

Figure 23: Histopathology of the new restorant 358

Chapter One

Introduction

Helicobacter pylori

Helicobacter pylori is a bacterium whose discovery has radically impacted the medical field. While reports of spiral shaped bacteria associated with the gastric mucosa were made multiple times since the 1870s, the presence of these bacteria was considered a contaminate of the process (as reviewed in 126). The idea that there was a pathogenic species of bacteria that could live within the hostile, and at the time presumed sterile, environment of the stomach was inconceivable. However, *H. pylori* was cultured in the early 1980s by the accidental incubation of plates for five days instead of the usual three days typically used to culture *Campylobacter* species (219). Originally, the organism was classified as *Campylobacter pyloridis,* but after subsequent study was reclassified to *Helicobacter pylori* in 1989 (1). Currently, there are 18 identified species within the *Helicobacter* genus. These species infect the gastric mucosa, intestinal tract, or hepatobilliary tract of various mammals, which range from rodents, to domestic dogs and cats, to farm animals (cattle and swine), to some non-human primates and humans (81, 85, 264, 330). *H. pylori* itself is a human and non-human primate specific organism (81, 85, 150), which is believed to have co-evolved with humans for at least the last 50,000 years (23, 66, 393).

H. pylori is a small, Gram negative, spiral shaped, microaerobe that has multiple (four to six) polar, sheathed flagellae that are responsible for its corkscrew motility (1,

120, 127, 219). The *H. pylori* genome is approximately 1.6 Mb, contains approximately 1,600 open reading frames and is A/T rich (60%; 12, 32, 124, 257, 336, 360). Approximately half of the characterized strains contain plasmids (281). *H. pylori* strains are naturally competent, which allows for the constant exchange of DNA. Given this and the fact that multiple strains can infect a single host (110, 122, 158, 164), it appears that *H. pylori*'s ability to take up new DNA *in vivo* acts as a mechanism of evolution. Indeed, analysis of an archived reference strain (J99) and isolates from the same patient taken six years apart shows high levels of genetic diversity (158). Collectively, the thirty new isolates from the original patient infected with J99 had lost up to 2.3% of the open reading frames compared to the archived J99 strain (158). Furthermore, the new isolates also contained additional DNA, which showed homology to genes from other *H. pylori* strains, that were not found in the original J99 strain. This DNA may have been acquired from a transient co-infection with a different *H. pylori* strain or another closely related bacterium (158). Overall, natural competence has been proposed to contribute to the vast allelic diversity of the organism, and to help account for the considerable genetic variability (6-7%) between strains (13, 124, 205, 352).

H. pylori is extremely well adapted for colonization of the human gastric mucosa thanks to a variety of colonization/virulence factors. For example, given that it is a neutrophile, one of the most daunting innate host defenses that this organism has to overcome is the extremely low pH of the stomach. To this end, *H. pylori* encodes a urease that appears to be the key factor in this process *in vivo*. Urease hydrolyzes urea to create ammonia, and the basic ammonia molecule in turn buffers the bacterial cytoplasm, as well as the surrounding micro-environment (86, 87, 231). The importance of this

process is evidenced by two different studies that identified genes necessary for colonization; multiple genes within the urease operon, as well as genes that encode components required to obtain and integrate nickel into the urease complex were found (31, 166). Since *H. pylori* colonizes within the mucus layer overlaying the gastric epithelium, the bacteria require flagellar motility in order to move through the gastric lumen to the proper site of colonization (88). Indeed, the largest number of characterized genes found to be important for animal colonization are genes involved in motility or chemotaxis (31, 166). Additionally, *H. pylori* also produces a mucinase that helps the bacteria penetrate the mucus layer to reach the epithelium (328). While a majority of the bacteria do live within the mucus layer, a small percentage actually adhere to the host epithelium. Adherence occurs through the use of adhesins such as the Lewis b-binding adhesion (BabA; 46, 155, 288) and the sialic acid-binding adhesion (SabA; 203). Other genes identified as essential for animal colonization encode outer membrane proteins, transporters, and factors important for cellular maintenance (such as those that control energy production, DNA modification, transcriptional regulation, iron metabolism, and cell division; 31, 166).

In addition to these colonization factors, *H. pylori* also has numerous virulence factors that affect the host cell directly. Two of those virulence factors, CagA and VacA, affect a multitude of host cellular pathways and are discussed in detail later in this thesis (29, 141-143, 215, 236, 267, 283, 308, 354, 355, 366). In addition to these two toxins, *H. pylori* encodes factors that affect inflammation. These include the N-terminus of the large subunit of the urease protein (204) and the neutrophil activating protein (NapA; 96). NapA was first identified as a bioferritin through amino acid sequence analysis (95);

however, it was later shown to be chemotactic for neutrophils and increases neutrophil adhesion to endothelial cells (96). Additionally, expression of the proinflammatory outer membrane protein (OipA) has been shown to be associated with severe neutrophil infiltration, gastric inflammation, and high colonization loads (395, 396). OipA seems to co-vary with other virulence factors, including *cagA* and *vacA* (395, 396). Other virulence factors that contribute to inflammation are the superoxide dismutase (SodB; 317) and the type IV secretion system encoded on the *cag* pathogencity island (71). *En masse* all of these virulence factors, or particular combinations of these virulence factors, allow *H. pylori* to be an effective pathogen despite its small genome and chosen niche.

Epidemiology and Disease

Worldwide, *H. pylori* infection is acquired by more than 50% of children by the age of 10 (285). Conversely, the acquisition of new infections in adults is significantly lower (174). Overall prevalence is approximately 50% of the world's population, though infection rates vary from 7-87% (3, 84, 104, 210, 225, 273, 345, 358, 403). The highest rates of infection are among developing or more recently developed countries, as well as countries in East Asia, such as South Korea (3, 225, 273). Interestingly, not only do prevalence rates vary between countries, but also within a country (182, 269). There appears to be an inverse relationship between socioeconomic status and infection with *H. pylori* (129, 208, 210-212, 326). Furthermore, in the U.S. race was also found to influence infection rates. Minorities, especially African Americans, are more likely to be infected than Caucasians and are more likely to maintain the infection (101, 129, 209). While it is generally believed that *H. pylori* infection lasts the lifetime of the host without

treatment, spontaneous clearance and transient infections have been reported (22, 130, 176, 209, 298). Even in treated individuals, recurrence of infection, particularly with new strains, is frequent in countries where overall *H. pylori* prevalence is high (123, 253). Reinfection is likely due to the fact that infection fails to induce a protective immunity. Recently, however, the global infection rate for *H. pylori* is slowly declining (33).

Transmission of *H. pylori* appears to occur via person to person spread, and many studies have correlated higher rates of *H. pylori* infection to living in close quarters: for instance, an institutionalized setting (44, 135, 171, 182, 190). In fact, infection rates were increased by as much as 52% in an institutionalized population as compared to gender and age matched controls (182). In support of increased person to person contact being a risk factor for infection, another study found increased rates of *H. pylori* infection in children living in communal apartments as compared to children with traditional families (212). Transmission is also thought to be familial; children born to an infected parent are more likely to be colonized and the colonizing strain is the same strain carried by the infected parent (173, 206, 300, 357). Other studies have looked at this trend from the other side and found that in families with *H. pylori* infected children, there were increased numbers of family members infected as compared to families with uninfected children (82, 228). Transmission has also been documented between spouses, and the risk of infection increased the longer the spouse lived with their infected partner (51, 325). These studies suggest that *H. pylori* is in fact spread person to person.

H. pylori infection is specific to humans and non-human primates, and although we do not know the exact vehicle of human transmission, it is believed to be spread through either gastro-oral, oral-oral, or fecal-oral routes of infection. Although rare, there

is some evidence that *H. pylori* can survive in a culturable form in milk and tap water for a few days (98, 153), and longer in a nonculturable coccoid form (373). Gastro-oral transmission of *H. pylori* has been postulated and is a risk factor among institutionalized adults in the Netherlands (44). In fact, *H. pylori* was detected in the vomitus of 100% of adult subjects and could be detected and cultured approximately 40% of the time from up to 0.3 meters away (277). Vomiting was also shown to be a risk factor for infection in children (201, 282). Another possible method of transmission is oral-oral spread. This method has been proposed due to studies in which identical *H. pylori* strains were isolated from the stomach and dental plaque (56, 99, 318). However, the rate of isolation of *H. pylori* from dental plaque is extremely low (39, 54, 56, 99, 318), and PCR detection rates in dental plaque vary (224). Interestingly, dental workers do not have an increased rate of infection (194, 207, 249). Another possible means of transmission is the fecal-oral route. Studies have been able to identify *H. pylori* DNA in feces, but the rates of detection vary from less than 10% (245) up to 90% (191, 216). While there have been studies that successfully cultured the organism from feces (168, 356), these studies looked at a population that was malnourished, where the bacteria would have experienced a shorter transient time in the intestine. In fact, another study verified that, by inducing a shorter time in the intestine via use of a diuretic, the ability to culture *H. pylori* from feces was increased (277). However, controversy still surrounds this method of transmission because *H. pylori* is sensitive to bile (134). *H. pylori* was also found not to be associated with other diseases that are transmitted fecal-orally, such as hepatitis A (113, 140, 200, 303, 380). Thus, overall scientists still do not know how *H. pylori* is

transmitted. Answering this question would go a long way to understanding *H. pylori* epidemiology and induced disease etiology.

H. pylori was first identified within areas of inflammation and was suggested to be responsible for the inflamed tissue. (219). This correlation was corroborated with Marshall's famous experiment, where upon consumption of an *H. pylori* culture, he developed gastritis (217). As mentioned above, the discovery of *H. pylori* was a controversial one, because people believed that bacteria could not survive the harsh environment of the stomach, let alone be the causative agent of gastric disease (218). Presently, it is now readily accepted that *H. pylori* is the etiologic agent of acute and chronic gastritis, peptic ulcer disease (75% of gastric and 90% of duodenal ulcers), and two forms of gastric cancer (MALT lymphoma and gastric adenocarcinoma; 40, 94, 275, 276, 349). Disease progression can include asymptomatic infection or can advance from acute to chronic gastritis and then to more severe forms of disease (335). If the inflammation is predominant in the antral end of the stomach, then gastritis may progress to duodenal ulcers, whereas, nonatrophic pangastritis can progress to MALT lymphoma, and corpus atrophic gastritis can lead to gastric ulcers, and eventually cancer via the following stages: intestinal metaplasia, dysplasia, and gastric adenocarcinoma (335). Deaths due to gastric cancer are still the second leading cause of cancer-related deaths worldwide (74, 248, 272, 392). For all of these reasons, the World Health Organization has classified *H. pylori* as a Class I carcinogen, and it is currently the only bacterium with this designation (4).

While there is not a gender bias among *H. pylori* infection rates, there is a bias among disease distributions. Females are more likely to suffer from gastritis, while

males are more likely to be afflicted with more severe disease: both ulcers and cancer (302). In fact, men are 1.5 to 2.5 times more likely than women to suffer from gastric cancer (reviewed in 297). While, overall gastric cancer rates vary drastically worldwide, it may not be surprising that gastric cancer rates are highest where *H. pylori* colonization rates are highest (74). For instance, South Korea, which has one of the highest rates of colonization, also has one of the highest rates of gastric cancer (42, 60, 133, 210, 323, 358, 403). In fact, 56.2% of the infection-related cancers in South Korea are due to *H. pylori*, and 45.1% of the cancer deaths due to infectious diseases are associated with *H. pylori* infection (322). This amazing statistic exemplifies the burden of *H. pylori*-induced disease.

While many insights into *H. pylori*-induced disease have been gained, the reason why some individuals develop more severe disease remains elusive. The epidemiologic triangle looks at disease as a contribution of three attributes: agent (in this case *H. pylori*), environment, and host. The first arm of the triangle is the agent, specifically bacterial factors that affect disease. Recent studies have focused on the role of the *H. pylori* toxins, CagA and VacA in disease development. Both of these toxins are polymorphic within different *H. pylori* strains and particular polymorphisms have been correlated to the development of particular disease states. The more virulent forms of these factors also correlate with the geographical areas with the highest rates of the most severe forms of disease, specifically gastric cancer. In fact, over 90% of South Korean isolates contain the most virulent form of CagA, which has been shown to impact gastric cancer risk (42, 60). Additionally, OipA expression was proposed to be a factor that affects disease progression (395). Although its role as a novel marker for certain disease types is

uncertain (17, 18, 125, 196), the duodenal ulcer-promoting gene (*dupA*) is associated with increased inflammation, but has an inverse relationship for gastric atrophy and cancer (17, 196, 287). Furthermore, the *Helicobacter* outer membrane protein B (HomB) impacts disease development (71, 163, 261). *homB* presence is associated with peptic ulcers and is a discriminative factor between gastric cancer and duodenal ulcers (163, 260, 261). The second arm of the epidemiological triangle focuses on the environment in which afflicted people live. We have already discussed that low socio-economic status, as well as proximity to other people, impact infection rates. However, other environmental factors play a role in disease development. These include co-infection with parasites and cigarette smoking. It has been shown that infection with helminths can actually diminish disease severity (83, 106, 382), whereas cigarette smoking increases gastric cancer rates (321, 324, 332). The final arm of the epidemiological triangle is the host. Host factors include both host genetics and diet. Different host genetic mutations can increase the likelihood of development of gastric cancer: IL-1β (91, 92, 112, 202) and PTPN11, the gene that encodes for SHP-2 (37, 149). Also included in the host arm are dietary factors, which impact disease development by affecting bacterial factors or changing the host environment. In fact, salt intake is implicated as important for *H. pylori*-induced gastric cancer development (107, 146, 364). Conversely, consumption of fruits and vegetables, which contain vitamin C, β carotene, and antioxidants can actually decrease the risk of gastric cancer (89, 93, 331, 387). As scientists and doctors are well aware, the process of disease development is complicated and involves not only the bacteria, but also environmental and host factors.

Treatment and Vaccines[1]

During the relatively short period of time that we have known about *H. pylori,* there have been many different treatment regimens developed (reviewed in 55). In fact, in 1994 there was a consensus from the National Institutes of Health (USA; 5), followed two years later by the Maastricht Consensus from the European *Helicobacter* Study Group (Netherlands; 2), which established treatment recommendations to treat *H. pylori* infection. Given the increase in incidence of antibiotic resistance, the Maastricht Consensus report was updated in 2000 and again in 2005 to increase the effectiveness of treatment regimens against *H. pylori* (213, 214).

The current recommendation for first line therapy in locations where clarithromycin resistance is low, is a protein pump inhibitor, clarithromycin, and either metronidazole (first choice) or amoxicillin (second choice) for 14 days (213). Additionally, these triple therapy regimens can be supplemented by the addition of bismuth in geographical areas where antibiotic resistance is high, though this combination is typically recommended as a second line therapy (213). Moreover, since bismuth is not available in many countries, a combination of a proton pump inhibitor, metronidazole, and either amoxicillin or tetracycline is sometimes recommended (213).

Primary and secondary therapies are not always successful at eradicating *H. pylori*; therefore, many alternative drugs are proposed for rescue therapy. These include fluoroquinolones (such as levofloxacin), rifamycins (such as rifabutin and rifampicin), nitrofurans (such as furazolidone) and other members (such as doxycycline) within

[1] Excerpts taken from the review article: K.R. Jones, J.H. Cha, and D.S. Merrell. 2008. Who's Winning the War? Molecular Mechanisms of Antibiotic Resistance in *Helicobacter pylori. Current Drug Therapies*. 3: 190-203.

families that are already used to treat *H. pylori* infection (reviewed in 55, 64). Of note, resistance has been found to all utilized primary and secondary antibiotics, as well as, to many of the antimicrobials used for rescue therapy (Fig. 1). This fact suggests that therapy success rates will continue to decline, and indicates that a detailed understanding of antibiotic resistance mechanisms may facilitate development of novel therapeutics.

Arguably the best option would be the production of a *H. pylori* vaccine, and several possible vaccine candidates are being researched (238, 327, 407). Vaccine components vary and include killed *H. pylori* whole cell extracts (289), heat shock proteins (100), flagellar antigens (327), adhesion antigens (340), lipopolysaccaride antigens (105), neutrophil activating protein (304), and urease (35). Unfortunately many of these vaccines are a long way from human trials (238, 327, 407), and the inactivated whole cell extract was proven ineffective in a human volunteer study (177).

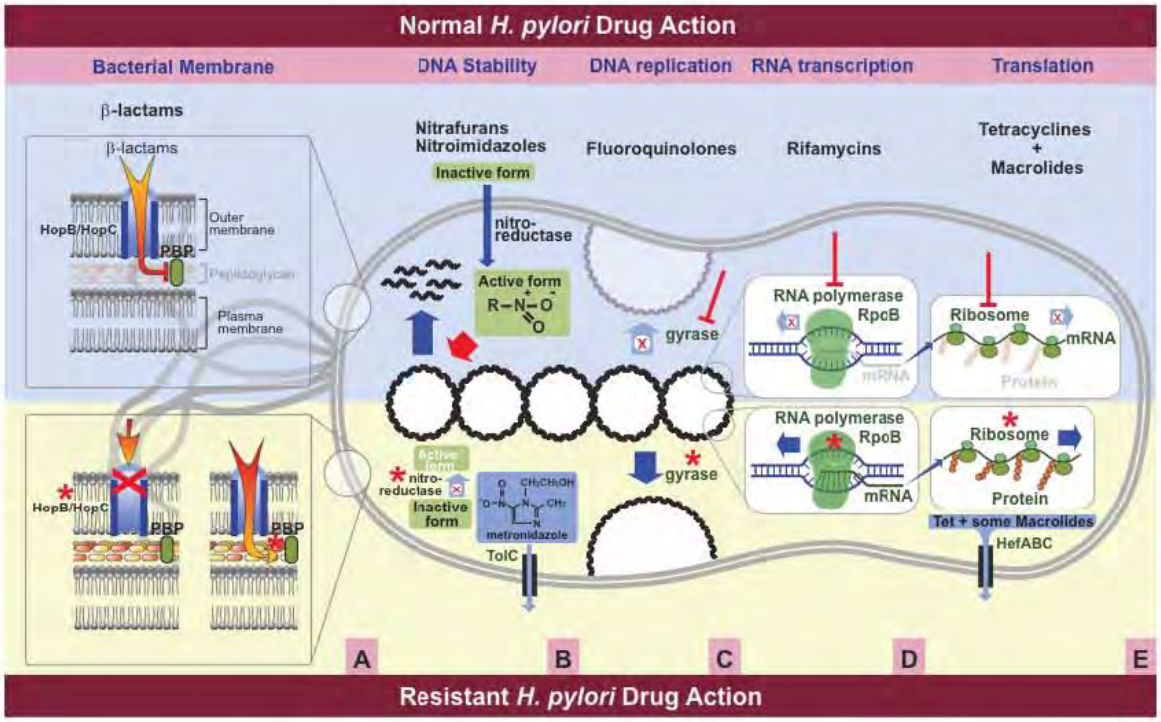
Virulence Factors that Impact Host Cell Pathways[2]

Due to *H. pylori*'s association with a variety of severe gastric diseases, many studies have been conducted to elucidate the bacterial, host, and environmental factors that impact disease progression. To date, several bacterial virulence factors have been associated with gastric cancer. For instance, the outer membrane proteins, HomB (163, 260) and BabA in Western type CagA containing strains (121), but not in East Asian type CagA containing strains (230) have been associated with progression to gastric cancer, while OipA (36, 254) and DupA (80, 154, 196, 251, 306) have more controversial roles

[2]Excerpts taken from the review article: K.R. Jones, Jeannette Whitmire, D.S. Merrell. A tale of two toxins: *Helicobacter pylori* CagA and VacA modulate host pathways that impact disease. *Frontiers in Cellular and Infection Microbiology.* doi: 10.3389/fmicb.2010.00115

Figure 1: Cellular components targeted by commonly used antibiotics and mechanisms of resistance utilized by H. pylori. The upper portion of the figure depicts normal drug interactions, while the lower portion depicts resistance mechanisms currently identified in *H. pylori*. * Denotes a mutation in the specified gene. A. β-lactams prevent the completion of the peptidoglycan layer of *H. pylori* through their interaction with penicillin binding proteins (PBP). Resistant bacteria contain either a mutated PBP, which prevents interaction with the β-lactams, or mutations in *hopB* or *hopC*, which decrease accumulation of the β-lactam within the bacterial cell. B. Nitrofurans and nitroimidizoles act in a similar manner against bacterial DNA. Both pro-drugs enter the cell and must be reduced by nitroreductases to become active. The activated form leads to formation of radicals that damage DNA. In resistant bacteria one or more of these nitroreductases are inactivated. The existence of a TolC efflux pump has also been identified as a mechanism of resistance to nitroimidizoles. C. Fluoroquinolones act upon gyrases, which are enzymes responsible for the conversion of DNA into a relaxed state required for DNA replication. Bacteria containing mutations in these gyrases are resistant to fluoroquinolones. D. Rifamycins act by blocking a subunit of the DNA-dependent RNA polymerase thereby terminating the production of mRNA. Resistant bacteria contain mutations within this subunit, which is encoded by the *rpoB* gene. E. Tetracyclines and macrolides both prevent the completion of translation, thereby preventing protein production. Mutations within specific ribosomal subunits cause *H. pylori* to be resistant to these drugs. Also, the HefABC efflux pump has been shown to remove tetracycline and some macrolides from the bacterial cell.

Figure 1: Cellular components targeted by commonly used antibiotics and mechanisms of resistance utilized by H. pylori



in disease development. The effectors, CagA and VacA, have also been shown to influence disease state, and are probably the most well studied virulence factors of *H. pylori*. These toxins have been shown to have multiple effects on host cells, as well as to modulate multiple cellular pathways in what appears to be a complex orchestration that ultimately leads to disease. To begin to shed some light on these pathways, as well as on the etiology of disease, this section will highlight some major findings regarding CagA, VacA, and their specific effects on host cells. Due to the large amount of literature on this subject and space limitations, an exhaustive review is not provided. However, we encourage readers to explore the excellent reviews by Cover and Blanke (67), Rieder *et al.* (294), and Hatakeyama and Higashi (139).

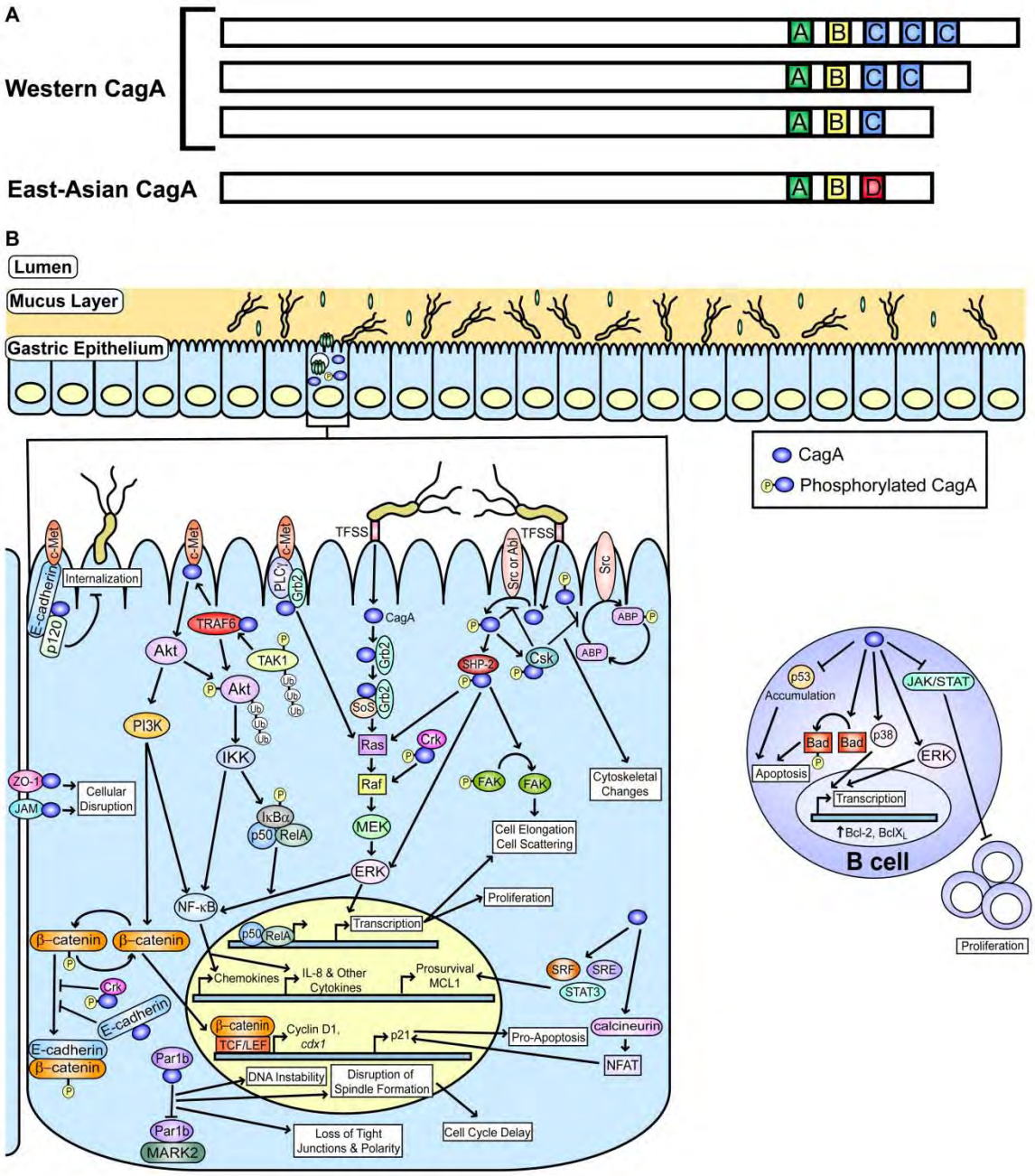
Cytotoxin-associated gene A - CagA

CagA is arguably the most well studied virulence factor of *H. pylori.* It is encoded on the *cag* pathogenicity island, which is a horizontally acquired 40 Kb DNA segment that encodes for a type IV secretion system, and is the only known effector protein to be injected into host cells (9, 57). *cagA* is the last gene on the *cag* pathogenicity island, and encodes for the 120-145 kDa immunodominant CagA protein (65, 367). Since its discovery, CagA has been shown to impact disease, especially more severe disease states like gastric cancer (42, 152, 178, 274, 388). *cagA* is present in ~70% of strains worldwide, but this rate varies geographically from between 90-95% in East Asian countries (South Korea, China, Japan) to only about 40% in Western countries (Australia, United States of America, England; 137).

Once injected into host cells, CagA can act directly in an unphosphorylated state to influence cellular tight junction (16, 30, 240, 262), cellular polarity (301, 406), cell proliferation and differentiation (58, 184, 227, 240), cell scattering (227), induction of the inflammatory response (49), and perhaps cellular elongation (Fig. 2; 301, 370). Moreover, upon entering the eukaryotic cell, CagA localizes to the plasma membrane where it can be phosphorylated by either Abl kinase or Src family kinases (284, 311, 333, 334, 350). These kinases phosphorylate tyrosine residues found in a five amino acid repeat, Glu-Pro-Ile-Tyr-Ala (EPIYA), within the carboxy-terminus of CagA (142, 143). These repeats can be categorized based on the amino acid sequences found within the regions flanking the EPIYA sequence to yield four distinct EPIYA motifs, which are known as EPIYA-A, -B, -C, and -D. Two combinations of these motifs predominate: Western CagA, which contains EPIYA-A, -B, and -C motifs (strains have been genotyped that contain up to five -C motifs) and East Asian CagA, which contains EPIYA-A, -B, and -D motifs (Fig. 2; 19, 65, 142, 143, 242, 250, 333). Additionally, there is a multimerization motif that consists of a 16 amino acid sequence present within the EPIYA repeat region (291). Once phosphorylated, CagA can form a complex with the CT10 regulator of kinase (Crk) adaptor protein (50, 344), Abl kinase and a splice variant of Crk, CrkII (350), or the Src homology 2 phosphatase (SHP-2; 143). Each of these interactions influences cellular shape and motility (50, 143, 344, 350). CagA that is phosphorylated at the primary phosphorylation sites (EPIYA-C and -D) shows varying affinities for SHP-2 based on the particular EPIYA variant as well as subsequent differential effects on the pathways influenced by the phosphorylated CagA/SHP-2 complex (142, 143).

Figure 2: CagA and known host cell targets. A. A schematic representation of CagA with the polymorphic region containing different EPIYA motif (A, B, C, and D) combinations is shown and is adapted from Hatakeyama and Higashi (139). B. A graphic depiction of the gastric mucosa and known host pathways impacted by phosphorylated and nonphosphorylated CagA is shown. Pathways targeted in epithelial cells and B cells are indicated. The actin binding proteins (ABP) affected by CagA include vinculin, cortactin, and ezrin. This figure was adapted from an earlier version by Rieder, *et al.* (294).

Figure 2: CagA and known host cell targets



CagA and Disease

The mere presence of CagA is associated with more severe disease forms (42, 65, 72, 133). In fact, cancer patients are at least twice as likely to be infected with an *H. pylori* strain that is *cagA* positive than one that is *cagA* negative (42, 133). It has additionally been demonstrated *in vivo* that *cagA* plays an important role in disease progression in a Mongolian gerbil model where gastric cancer develops within 12 weeks (108, 109, 255, 280, 295, 379, 386). The differences in affinity of various EPIYA motifs for SHP-2 and subsequent differences in induction levels of downstream pathways has been speculated to impact the differences in disease rates, especially gastric cancer (138, 142, 143). In fact, increasing numbers of EPIYA-C motifs have been suggested to be associated with cancer development (26, 305), and epidemiological studies have identified a correlation between increased number of -C motifs and heightened disease severity (20, 142, 394). Since EPIYA-D has the strongest affinity for SHP-2 (142), it is not surprising that East Asian CagA produces more inflammation and atrophy (27) as well as greater morphological changes in infected cells (142). Moreover, the variability in CagA is important when analyzing the geographic areas with the highest gastric cancer rates; these areas not only have the highest colonization rates, but also contain the highest percentage of *H. pylori* strains that carry the *cagA* gene, in particular the EPIYA-ABD allele (7, 74, 137, 392). Indeed, we identified an association between gastric cancer development and EPIYA-ABD CagA through a large scale molecular epidemiological study of strains from South Korea (162).

CagA Phosphorylation Independent Events

Physical Effects on Host Cells

Despite the importance of CagA phosphorylation, CagA has many effects on the host cell, and some of these effects are accomplished in a phosphorylation independent manner. One of the most noticeable CagA-dependent effects on host cells is the disruption of tight junctions and induction of changes in cell morphology. CagA has been shown to affect cellular tight junctions in a phosphorylation-independent manner (16, 30, 240, 262), and has been shown to be important for the recruitment of the junctional adhesion molecule (JAM) and the tight junction protein, zona occludens-1 (ZO-1) to points of bacterial contact (16, 30). Murata-Kamiya, *et al.* showed by immunoprecipitation that E-cadherin physically interacts with both wild type and phosphorylation resistant variants of CagA, and that this interaction inhibits the association of E-cadherin with β-catenin, which subsequently results in the accumulation of nuclear and cytoplasmic β-catenin (240).

Additionally, it has been demonstrated that CagA binds to and prevents the kinase activity of the partitioning-defective 1/microtubule affinity-regulating kinase (Par1b/MARK2), thereby escalating the loss of tight junctions and polarity (301, 406). CagA binds to Par1b, as well as other members of this kinase family, through the multimerization sequence (198, 301), specifically 14 of the 16 amino acids of the multimerization motif are required (FPLKRHDKVDDLSK; 247). This interaction has been suggested to contribute to host cell elongation (301, 370). East Asian CagA binds Par1b with a stronger affinity than CagA from Western strains, and the efficiency and strength of binding to Par1b among Western strains appears proportional to the number

of Western multimerization sequences (198). In addition to affects on cell elongation and disruption of cellular junctions, interaction of CagA and Par1b also causes spindle dysfunction, which delays progression from prophase to metaphase and is hypothesized to result in DNA instability (197).

Moreover, recent evidence has shown that CagA has an effect on how invasive *H. pylori* can be. Though considered an extracellular pathogen, numerous studies have shown that *H. pylori* is able to invade and survive inside host cells (15, 256, 314). Upon infection with *cagA* positive *H. pylori* strains, a multiprotein complex is formed through the association of CagA, c-Met, E-cadherin and p120-catenin, and this complex influences whether the bacteria can become intracellular (262). When this complex is present in a cell line that *H. pylori* can normally invade, it suppresses the ability of *H. pylori* to be internalized (262).

Cellular Differentiation

Since there is a causal link between gastric cancer, *H. pylori* infection, and the presence of *cagA*, there is no doubt that effects on host cell differentiation and proliferation are important for ultimate disease progression. In keeping with this, *cagA* positive strains of *H. pylori* influence a factor with known oncogenic potential, β-catenin (90, 108). β-catenin has two distinct functions, it links cadherins with the actin cytoskeleton and is part of the WNT signaling pathway (359). When unphosphorylated CagA binds E-cadherin, it prevents the formation of the E-cadherin/β-catenin complex, which ultimately leads to accumulation of β-catenin in both the nucleus and cytoplasm (240). Kurashima, *et al.* showed that while phosphorylation of the EPIYA motifs was not

necessary for deregulation of β-catenin, the CagA multimerization sequence was necessary (179). However, this process is likely multifactorial and complex since some evidence indicates that E-cadherin dissociation is independent of CagA and that the E-cadherin/β-catenin/p120^{ctn} complex is not affected to the same degree in all studies (381). When the E-cadherin/β-catenin complex is disrupted, cytoplasmic β-catenin is dephosphorylated and then translocates to the nucleus, where it forms heteromers with other transcription factors and transcribes a number of genes with oncogenic potential (193, 240). In fact CagA was found to upregulate the β-catenin-dependent *cdx1* gene (193, 240), which encodes a transcription factor important for transdifferentiation of intestinal cells (229), as well as to affect the expression of goblet-cell mucin MUC2, an intestinal-differentiation marker (240); both are indicators of gastric intestinal metaplasia.

Cell Cycle, Survival and Proliferation

Increased cellular proliferation is one indicator of cancer that has been demonstrated to result from infection with *cagA* positive strains of *H. pylori* (278, 279). This increased proliferation can happen through CagA-mediated activation of the ERK/ MAPK pathway. CagA activates ERK through interaction with growth factor receptor bound 2 (Grb2), and appears to interact with both phosphorylated and nonphosphorylated CagA (227, 313). However, it should be noted that the phosphorylation sequences themselves are essential for Grb2 binding to non-phosphorylated CagA (227), which similar to the strategy for binding of CagA to SHP-2 (192), likely occurs through CagA binding to the Grb2 SH2 domains. In a normal cell, upon receiving an extracellular signal, Grb2 binds son of sevenless (SoS), which increases the formation of the Ras-GTP

complex and activates the Raf→MEK→ERK pathway. This pathway increases activation of transcription factors involved in cell proliferation (114). In CagA intoxicated cells, CagA can bind Grb2, and then the CagA/Grb2/SoS complex activates the Ras→ERK pathway as described above (227).

Activity of the serum response factor and serum response element (SRF and SRE, respectively) transcription factors are also increased in CagA transfected cells in a CagA phosphorylation independent manner (145). Activation of SRE appears to be mediated by increased DNA-binding by the E-26 like protein-1 (Elk1; 145). Evidence of the role of CagA in SRE/SRF activation can be found through CagA-mediated increases in levels of the anti-apoptotic protein myeloid cell leukemia sequence-1 (MCL1), which acts as a pro-survival factor (226). Furthermore, the tendency of CagA to affect transcription factor activity seems to be a common theme since the signal transducer and activator of transcription 3 (STAT3) pathway, which induces cellular proliferation (175), has been shown to be induced *in vitro* and *in vivo* in a CagA-dependent but phosphorylation-independent manner (52). Moreover, it was recently confirmed that nonphosphorylated CagA preferentially activates STAT3 (184). The activation of so many transcription factors by CagA is evidence of the broad influence of CagA on a wide variety of cellular functions.

CagA appears to influence the development of MALT lymphoma in a phosphorylation independent manner; CagA can inhibit apoptosis of B-cells through inhibition of the accumulation of p53 due to decreased *p53* transcription (371). Interestingly, B-cell survival is also likely due in part to an increase in phosphorylated ERK1/2 (409), which when moderately increased can inhibit apoptosis and promote

proliferation (271). However, it should be noted that transfection with CagA also leads to phosphorylation of the pro-apoptotic protein, Bad (409). Conversely, translocation of CagA by *H. pylori* results in upregulation of both the ERK and p38 pathways, which lead to upregulation of the pro-survival proteins, Bcl-2 and Bcl-X_L (195). Clearly, CagA has an effect on cell survival, and most of the literature suggests that CagA can inhibit apoptosis of B cells, which likely promotes the development of MALT lymphoma.

Conversely, ectopically expressed CagA can also suppress cellular proliferation in IL-3-dependent B-lymphoid cells through suppression of JAK-STAT signaling (371). Moreover, transfection of AGS cells with CagA results in increased expression of the pro-apoptotic factor $p21^{WAF1/Cip1}$ due to CagA-mediated nuclear translocation of the nuclear factor of activated T cells family transcription factor (NFATc3; 404). Interestingly, $p21^{WAF1}$ expression can also occur as a direct result of excessive ERK1/2 activation (38, 270, 405). Finally, CagA-induced deregulation of β-catenin, increases the expression of Cyclin D1 (369), which influences progression of cells from G1 to S phase; thereby promoting cell survival in a CagA phosphorylation-independent manner (58, 240). CagA obviously affects cell survival, proliferation, and differentiation, all of which can affect the progression of disease, including development of gastric cancer.

Cell Scattering

In order for cancer cells to spread or metastasize, they must detach and scatter to a new area. CagA increases cell scattering by targeting the hepatocyte growth factor receptor (c-Met), which acts as an adaptor molecule for proteins like Grb2, phospholipase Cγ (PLCγ), and STAT3 (63, 227). CagA/Grb2/SoS→Ras-GTP complex→

Raf→MEK→ERK signaling leads not only to an increase in transcription factors that promote cellular proliferation, but also to an increase in cell scattering (227). In support of this, cell scattering due to *H. pylori* infection is suppressed by blocking PLCγ activity (63). Though once again, the role of CagA in this process is complex; some work suggests no association between enzymatic activity of PLCγ and the *cagA* status of *H. pylori* strains (43). However, it is clear that nonphosphorylated CagA interacts with c-Met through interaction with the multimerization domain (343). The consequence of this interaction is activation of phosphatidylinositol 3-kinase (PI3K) signaling through Akt, which subsequently activates NF-κB and β-catenin (343). It has been suggested that CagA binds to c-Met and then recruits TNF receptor associated factor 6 (TRAF6; 337) and poly-ubiquitinated transforming growth factor-β-activating kinase 1(TAK1; 180, 181). TRAF6 then activates Akt (399), which potentially activates NF-κB and RelA through activation of the IκB kinase (IKK) complex (343, 348). CagA modulates multiple pathways that impact cell scattering, and these pathways that are activated appear to have multiple downstream targets that can affect numerous cellular processes in addition to cell scattering.

Inflammation

A hallmark of *Helicobacter pylori* infection is increased and chronic inflammation. This appears to occur due to activation of NF-κB and persistent induction of IL-8. While it remains controversial (73, 181, 307), this IL-8 induction has been shown to be CagA dependent through studies that ectopically expressed CagA or various CagA EPIYA motifs (49, 170), through IL-8 promoter reporter assays (319), and through

analysis of inflammation and NF-κB activation in Mongolian gerbils infected with *cagA* positive and *cagA* negative *H. pylori* strains (320). Moreover, Keates, *et al.* showed that IL-8 secretion is affected by the activation of MAPKs by *cagA* positive *H. pylori* strains (167). It has been demonstrated that activation of NF-κB and induction of IL-8 occurs through the activation of the Ras→Raf→Mek→ERK→ NF-κB pathway and is independent of SHP-2 or c-Met (49). Indeed, Brandt, *et al.* demonstrated that IL-8 induction was CagA phosphorylation-independent (49). Kim, *et al.* confirmed that NF-κB activation and subsequent induction of IL-8 were due to activation of the MAPK pathways and also analyzed the role of the different EPIYA motifs by analysis of transfected cells with CagA constructs that differed only in the EPIYA region. Analysis of Western CagA-specific sequences and East Asian CagA-specific sequences revealed that the levels of IL-8 induction are not significantly different between the CagA variants (170). However, it should be noted that Argent, *et al.* demonstrated that CagA-related differences in IL-8 induction were dependent on the EPIYA motifs and that strains containing East Asian CagA induce the greatest levels of IL-8 (19). However, they did not investigate the influence of phosphorylation status on these differences. Since persistent inflammation is a hallmark of *H. pylori* infection and is linked to more severe diseases, and since CagA affects the inflammatory process, it is easy to envision what an important role CagA plays in persistent infection and disease development.

CagA Phosphorylation Dependent Events

Targeting of SHP-2

The most striking *H. pylori* induced morphological change to host cells is the induction of the "hummingbird phenotype," which occurs as a direct result of phosphorylated CagA complexing with SHP-2 and subsequent increased ERK1/2 activation. Normally, SHP-2 functions to increase cellular proliferation and motility and is activated by interacting with a phosphorylated Gab protein (246). CagA has been shown to be able to mimic the function of the eukaryotic Gab protein (47, 136). Both *in vitro* and *in vivo* (398), CagA forms a complex with SHP-2 after phosphorylation of an EPIYA-C or -D motif (142). The formation of this complex, as well as the subsequent deregulation of SHP-2 as a means of CagA-mediated effects on gastric cancer, is of relevance since mutations within the gene encoding for SHP-2 (*PTPN11*) have been identified in multiple forms of cancer (37, 351). Additionally, there is an increase in the risk of gastric cancer development in *H. pylori* infected patients with certain *PTPN11* polymorphisms (128). Again, this demonstrates the potential of CagA to impact disease progression.

Activation of ERK

The ERK MAP kinases are activated in a CagA phosphorylation-dependent manner upon infection with *H. pylori cagA* positive strains, leading to a SHP-2 dependent change in cell motility and morphology (246). Indeed, inhibition of the phosphorylation of CagA, knockdown of SHP-2 expression, or the disruption of the CagA/SHP-2 complex abrogates cell elongation, thus, indicating that the "hummingbird phenotype" is a product

of the SHP-2/CagA complex (141, 143, 144, 308). In fact, this complex activates the ERK pathway by activating Rap1→B-Raf→ERK and has been proven to activate ERK in both a Ras-independent and dependent manner (141). In addition, CagA promotes cell proliferation through activation of ERK, which subsequently promotes progression through the cell cycle (299, 365).

Recent data indicate that the phosphorylation status of CagA may act as a signaling switch between the JAK/STAT3 and SHP-2/ERK pathways. This process is mediated through gp130 (185). Unphosphorylated Cag activates STAT3, while phosphorylated CagA preferentially activates ERK1/2 phosphorylation (185). This differential activation based on phosphorylation status illustrates the complexity of the effects that CagA has on the host cell.

Non-ERK Mediated Cytoskeletal Rearrangement and Scattering

Important steps in the creation of elongated cells include the decrease in cellular adhesions and the deregulation of the actin-binding proteins that maintain proper cellular shape (48, 233). Phosphorylated CagA binding and activation of SHP-2 leads to increased tyrosine dephosphorylation of the focal adhesion kinase (FAK; 366), which is important for elongation of host cells; when dominant-negative FAK is expressed, host cells change morphology, while constitutively active FAK abrogates the formation of this morphological change (366). Additionally, studies have identified multiple actin-binding proteins that, when tyrosine-dephosphorylated, promote CagA phosphorylation dependent cell elongation. These include vinculin (232), cortactin (309, 312), and ezrin (310). Since the SHP-2 phosphatase is not required for the dephosphorylation of

cortacin, the dephosphorylation of these actin-binding proteins is likely a result of blocked activity of a kinase, and therefore a product of the phosphorylated CagA negative feedback loop that inhibits the Src kinase (312). In fact, this is the case for ezrin as inhibition of Src family kinases increases dephosphorylation of ezrin (310), and host elongation is achieved simply through inactivation of Src, which results in dephosphorylation of all Src subtrates, including vinculin, ezrin, and cortactin (28). This inhibition of Src could occur either directly or through the recruitment of C-terminal Src kinase (Csk), as described in the next section (28, 312, 365). Finally, phosphorylated CagA can bind Crk adaptor proteins (Crk-I, Crk-II, and Crk-L; 344). This interaction is important for cell scattering, disruption of E-cadherin/catenin, and activation of Raf (344). Furthermore, it was recently shown that CagA phosphorylation could occur via Abl instead of Src, thereby activating downstream effects, specifically cell scattering and motility (284, 350). Moreover, Abl could also form a complex with CrkII and CagA (350). Taken together, these findings show that phosphorylation of CagA is very important for host cell shape and adhesion. This fact implicates the degree of phosphorylation as a consequence of the *cagA* allele carried by a strain as being important for development of gastric carcinomas.

CagA Feedback Loop, Src vs. Csk

As mentioned above, CagA participates in a negative feedback loop that allows regulation of the amount of phosphorylated CagA. CagA can bind Csk via direct interaction with the EPIYA-A and -B motifs (365, 366). Formation of this complex leads to inhibition of the Src family kinases (SFKs) through Csk tyrosine phosphorylation of an

inhibitory C-terminal residue on Src (138, 365, 366). CagA can also directly inhibit SFK activity (312). While the purpose of this negative feedback system is not completely clear, it appears that in its absence, CagA is excessively toxic to cells (137, 138, 365). Thus, this loop has been hypothesized to promote long-term colonization of *cagA* positive *H. pylori* strains (137).

Interactions with Unknown Function

Recent proteomic screens identified a number of proteins that appear to interact with phosphorylated CagA (313). These include PI3K, Grb2, Ras-GAP, Grb7, and Shp1. The consequences of these interactions with phosphorylated CagA is still unknown (313). However, it is clear that both Grb2 and PI3K are actively involved in *H. pylori* pathogenesis when CagA is not phosphorylated (227, 343). Indeed, unphosphorylated CagA is known to bind to Grb2, which activates ERK signaling, and leads to increased cellular proliferation, transcription factor activation, cell scattering (227), and activation of Akt and PI3K signaling. These activation events subsequently stimulate both the β-catenin and NF-κB pathways (343). So the interaction of some of these proteins with phosphorylated CagA may represent redundant mechanisms of action.

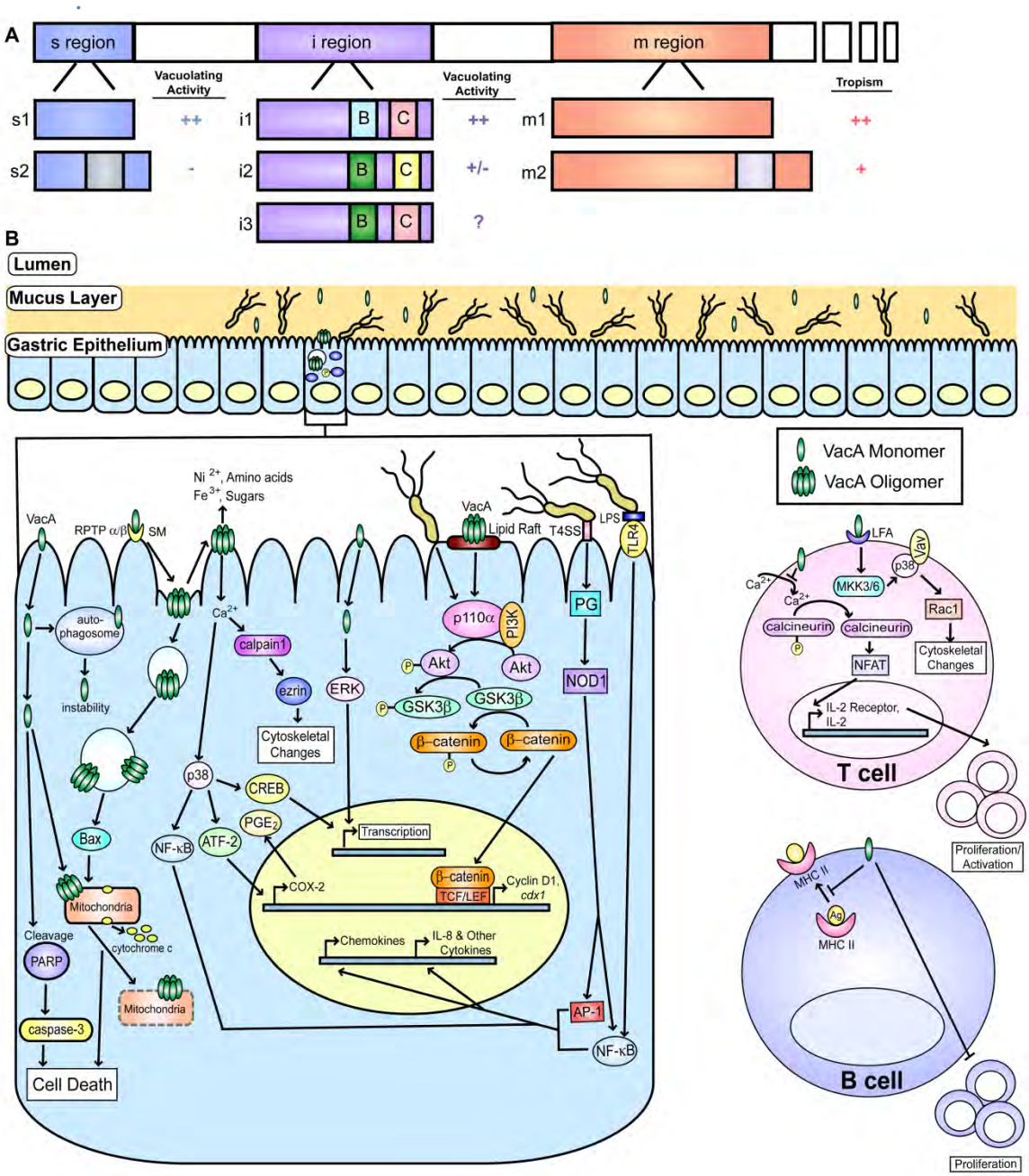
CagA Independent/Redundancy

Inflammation appears important for *H. pylori* growth *in vivo*. For instance, *H. pylori* induced inflammation results in a decrease in the inhibitor of gastrin; gastrin has been proven to be a *H. pylori* growth factor (41, 61, 189, 239). Given the importance of inflammation for *H. pylori* colonization and persistence, there are redundant mechanisms

by which *H. pylori* induces inflammation. For instance, in addition to CagA effects, NF-κB activation can also be achieved through TLR4 recognition of LPS (180, 258) or through type IV secretion system (T4SS) delivered peptidoglycan (PG) binding to nucleotide-binding oligomerization domain 1 (NOD1; Fig. 3; 165, 372). The inflammatory response caused by the interaction of peptidoglycan-NOD1 may be a result of the activation of AP-1 (11), which functions as a transcription factor for cytokines and chemokines such as IL-8 (59, 97, 159, 290). Recently, it has also been demonstrated that the binding of NOD1 to its ligand, in this case peptidoglycan, activates RICK, which allows RICK to interact directly with TRAF3 followed by TRAF3 interaction with TBK1 (378). TBK1, as well as IKKε, leads to the production of cytokines, such as the type 1 interferons (IFN) like IFN-β. The production of IFN-β is responsible for NOD1's ability to increase the level of the chemokine IP-10 as well as the induction of and nuclear translocation of the transcription factor interferon-stimulated gene factor 3 (ISGF3; 378). The presence of the *cag* pathogenicity island has also been shown to lead to increased inflammation (57); 14 of the 27 genes within the island are essential for IL-8 induction (102). With multiple bacterial factors that induce NF-κB, this begs the question of why the bacterium needs mechanisms for such redundancy. Lamb, *et al.* postulate that since CagA and peptidoglycan target different cellular signaling molecules, they may synergistically activate NF-κB, and that this synergy may be important. Alternately, in strains where CagA is not present or is not a potent inducer of IL-8 and NF-κB, peptidoglycan may serve as the major inducer of the inflammatory response (180). Both of these hypotheses highlight the importance of the induction of an inflammatory response for *H. pylori*, probably due to the requirement for gastrin or other nutrients.

Figure 3: VacA and known host cell targets. A. A schematic representation of VacA with the three major regions of polymorphisms (s, i, and m) is shown. Additionally, schematics of the known alleles of each region are shown. The i region contains two important polymorphic regions known as Cluster B and Cluster C, which are designated by a B and C, respectively on the diagram. The activity attributed to each of the regions of the toxin (vacuolating activity or cellular tropism) are indicated, and the impact of each allele on these effects is shown. The highest level of activity or the broadest tropism is defined as ++, intermediate tropism is indicated by a +, low activity is indicated as a +/-, no activity is designated by a -, and incomplete information is indicated by a ?. B. A depiction of the gastric mucosa and known host pathways targeted by VacA is shown. One of the receptors, sphingomylein is designated by SM. Pathways targeted in epithelial cells and B and T cells are indicated. Additionally, activation of several pathways by peptidoglycan (PG) and LPS are shown. This figure was adapted from an earlier version by Rieder, *et al.* (294).

Figure 3: VacA and known host cell targets



Vacuolating Cytotoxin A - VacA

VacA is another important factor that has been indicated to have effects on *H. pylori* virulence and to target numerous host cell pathways (Fig. 3). Activity of this protein was found when *H. pylori* filtrates were shown to induce large host cell vacuoles (188). The VacA cytotoxin appears to be produced and secreted by most, if not all, *H. pylori* strains, but possesses no similarity to any other known bacterial or eukaryotic protein (25, 67). Once produced, VacA can remain on the bacterial surface (156) or be secreted as an approximately 88 KDa toxin (68). Secreted VacA monomers oligomerize (6, 69, 183, 199) but dissociate upon exposure to a non-neutral environment. In fact, exposure to alkaline or acidic conditions actually amplifies the activity of VacA (69, 77, 234, 389). Once secreted, VacA undergoes proteolytic cleavage to yield two smaller products, p33 and p55. However, to date the consequence of this cleavage is not understood (252, 354, 362, 385, 400). The smaller p33 product and about 100 amino acids of p55 are responsible for the vacuolating activity of VacA (75, 76, 402). The p33 domain is strongly hydrophobic and contains characteristic transmembrane dimerization motifs that are responsible for insertion into the host cellular membrane and vacuolating activity (169, 221, 223, 374, 400, 401), whereas the p55 domain has a crucial role in binding to host cells (116, 266, 292, 376, 377).

Like CagA, VacA is polymorphic. However, unlike CagA, this variation begins within the amino-terminus of VacA. Three regions of variation have been defined and there are at least two primary variants in each region; the regions are designated as the signal (s), intermediate (i), and middle (m) regions (Fig. 3; 24, 62, 293). The s region of

VacA is found in the p33 portion of the toxin and influences vacuolating activity and efficiency of anion channel formation due to the hydrophobic nature of the amino acid residues found near the proteolytic cleavage site (222, 286). The s2 variant undergoes cleavage at an alternate site, thereby providing an extension of 12 hydrophilic amino acids (24). The s1 variant contains more hydrophobic amino acids near the cleavage site than the s2 variant; thus, the s1 sequence is more easily inserted into the host cell membrane (186, 222). The m region is found in the p55 portion of the toxin and influences host cell tropism; the m1 region is toxic to a wider range of host cells (14, 161, 266). The i region is located between the s and m regions and is the most recent region to be described. The i region has been suggested to be the best indicator of disease severity (293) and three primary variants have been identified (62). The i1 region is believed to be associated with stronger vacuolating activity and more severe disease states than the i2 region (293). Furthermore, strains carrying VacA s1, i1, m1 or any combinations of these alleles are overall associated with more severe disease (34, 160, 187, 293). This association could be due to increased anion channel formation, vacuolating activity, and cell tropism from having the s1, i1, and m1 regions, respectively.

In recent years, a number of studies have elucidated multiple receptors for VacA and shown that VacA uses different receptors based on different host cell types (315). On epithelial cells, several different receptors have been identified. Among these are RPTPα, which is a receptor-like protein tyrosine phosphatase that appears to be used by VacA on G401 cells (a human kidney tumor cell line) and in AGS cells (an adenocarcinenoma cell line; 353, 390). Another receptor which needs to be glycosylated for VacA to bind to it is RPTPβ, which can be used by VacA on AZ-521 (gastric

epithelial-derived cells; 389, 391). When RPTPβ is artificially increased in some cell lines, toxicity to VacA also increases (265). The importance of this receptor *in vivo* has been demonstrated in RPTPβ knock out mice, which become resistant to VacA mediated ulceration (111). Additionally, sphingomyelin was recently identified as a receptor that is important for VacA binding and vacuolating activity of the toxin (131, 132). Finally, VacA can also bind to T-cells using the lymphocyte function-associated antigen-1 (LFA-1; 316). The fact that VacA can use different receptors based on the cell type targeted may help explain this toxin's diverse functions. While much is known about the various functions of the toxin, relatively little is currently known about the exact host signaling pathways affected by the toxin. Thus, herein we discuss what is currently known about the major cellular processes affected by VacA.

VacA Functions

Anion Channel Formation and Vacuolation

VacA can oligomerize within the plasma membrane and can cause formation of anion-selective channels (346). These channels may be used to increase the efflux of complex molecules, such as bicarbonate and urea, out of the host cell (78, 346, 361), which may aid *H. pylori* growth (237). In addition to forming anion-selective channels *in vitro*, VacA can reduce the transepithelial electrical resistance of polarized cells by increasing paracellular epithelial permeability. This allows the release of some cations, such as Fe^{3+} and Ni^{2+}, as well as more complex molecules such as amino acids and sugars (268).

One of the most striking effects of VacA on host cells is the creation of large cytoplasmic vacuoles that contain the markers for late endosomes and lysosomes (235). Once VacA is internalized by the host cell, it is trafficked to the early endosome by F-actin containing structures (118). Subsequently, the CD2-associated protein is essential for transferring VacA from early endosomes to late endosomes (117, 118). The process of vacuolation is then dependent on syntaxin 7 and vesicle associated membrane protein 7 (VAMP7), both of which are integral to late endosomes and lysosomes (220, 341). Additionally, this process requires the vacuolar ATPase (V-ATPase) activity and dynamin, which are enzymes crucial for formation and stability of vesicles, (70, 342).

Induction of Apoptosis

Although the fact that VacA causes apoptosis has been known for a while, the exact mechanism or mechanisms by which this occurs are still not completely understood. Evidence shows that VacA-mediated apoptosis is dependent on interaction with the mitochondria (79, 103, 115, 172, 259, 383, 384). Indeed, VacA has been proven to reduce the membrane potential of the mitochondria, thereby allowing the release of cytochrome c (115, 172, 383, 384). The modulation of the mitochondrial membrane potential by VacA also results in impaired cell cycle progression and a drop in ATP concentration (172). Several early studies showed that VacA that is deficient in its ability to form channels inhibits cytochrome c release (383, 384), and blocks the ability to modulate the mitochondrial membrane potential (383), suggesting that channel formation is essential for these events.

However, some additional work has demonstrated that most VacA is localized to vacuoles inside host cells (397). This finding suggests that VacA-mediated cell death might not be a result of direct binding of VacA to the mitochondria, but perhaps suggests that VacA-mediated induction of the pro-apoptotic factors in the Bcl-2 family might be involved. In keeping with this, it has been suggested that these pro-apoptotic factors actually interact with the mitochondria to release cytochrome c, and VacA has been shown to increase the level of pro-apoptotic Bax in a manner that mirrors the release of cytochrome c from the mitochondria (397). Additionally, VacA can also induce the cleavage of poly (ADP-ribose) polymerase (PARP), by the activation of the death factor, caspase-3 in transfected cells. Furthermore, this cleavage can be inhibited by the overexpression of the pro-apoptotic factor Bcl2 (115, 397). Taken together, these data suggest that VacA has two potential mechanisms to induce apoptosis in intoxicated cells.

Disruption of Cellular Pathways

VacA deregulates multiple cellular pathways as well as inducing inflammation. VacA intoxication induces production of a variety of inflammatory cytokines that include TNFα, IL-1β, IL-6, IL-10, and IL-13 (339). Moreover, IL-8 is produced by several different cell lines in response to VacA-mediated activation of the p38 MAPK through an increase in intracellular calcium and the subsequent activation of ATF-2, CREB, and NF-κB (147).

VacA increases the activity of p38, ERK, and the activating transcription factor 2 (ATF-2; 244). Through the p38/ATF-2 cascade, COX-2 is upregulated, which leads to increased production of prostaglandin E_2 (PGE_2; 148). Conversely, in mice VacA

inhibits PGE $_2$-stimulated duodenal epithelial bicarbonate (HCO_3^-) secretion by inducing the release of mucosal histamine (368). While the reason for this discrepancy in VacA-mediated increase in PGE_2 effects is unclear, it should be noted that decreased duodenal epithelial HCO_3^- secretion is associated with duodenal ulcers and may leave the mucosal layer less able to repair itself (157). This could account for the role of VacA in gastric damage (157). Furthermore, VacA can also inhibit gastric acid secretion by increasing the mobilization of intracellular calcium, which in turn activates calpain 1 to hydrolyze the CagA targeted cytoskeletal protein, ezrin (310, 375). As well as the inflammatory pathway, VacA activates the p38 and the ERK pathways leading to deregulation of molecules that directly correlate with gastric damage.

Like CagA, VacA has also been shown to affect the β-catenin signaling pathway and therefore, perhaps the oncogenic potential of *H. pylori*. As stated earlier, deregulation of this molecule affects many cellular pathways involved in migration, cell cycle, polarity, and apoptosis, and numerous studies have demonstrated the effect that *H. pylori* has on the β-catenin pathway (179, 241, 243, 329, 347). Recently, Tabassam, *et al.* showed that VacA activates PI3K/p110α, which in turn activates Akt to phosphorylate GSK3β. This phosphorylation ultimately frees β-catenin to translocate to the nucleus to bind TCF/LEF and allows transcription of β-catenin dependent genes such as cyclin D1 and potentially other oncogenic genes such as *cdx1* (193, 229, 347). Thus this affect on the β-catenin pathway again lends credence to the fact that the more virulent alleles of VacA, which potentially could cause greater induction of this pathway, are associated with more severe disease manifestations.

Finally, VacA also affects the autophagy pathway; VacA induces the formation of autophagosomes and is associated with these structures (355). Moreover, the stability of intracellular VacA is impacted by the presence of autophagosomes, and VacA stability is increased when autophagy is inhibited (355).

Immune Modulation

VacA has many roles, but one important function that may directly impact *H. pylori* colonization and persistence is its ability to act as an immune modulator. This immune modulation occurs through several distinct mechanisms. For instance, VacA disrupts the process of phagosome maturation through recruitment and retention of coronin 1, which is also known as tryptophan-aspartate- containing coat protein (TACO; 408). However, despite this disruption in phagosome maturation, VacA does not seem to impact the intracellular survival of *H. pylori* within monocytes (296). Next, *H. pylori* infected macrophages form large vesicular compartments called megasomes and VacA supports this process by increasing homotypic phagosome fusion (10, 408). This allows *H. pylori* to persist in macrophages instead of being killed. VacA has also been proven to inhibit the invariant chain dependent pathway of antigen presentation by MHC class II molecules (236), and has been reported to interfere with the presentation of antigen in B cells (236). More recently, VacA has been reported to inhibit both PMA/anti-IgM and T cell induced B cell proliferation (363).

In addition to these pathways, T cells are also affected by VacA. VacA can enter activated, migrating primary T lymphocytes by binding to β2 integrin (CD18) and LFA-1 (316); LFA-1 is essential for this process since T cells deficient in LFA-1 are resistant to the effects of VacA (316). Intoxication by VacA can then inhibit the proliferation of

$CD8^+$ T cells (363) through down-regulation of the expression of the interleukin 2 (IL-2) surface receptor-α and inhibition of the production of IL-2, both of which are required for T-cell proliferation and survival. These IL-2 effects occur through the inhibition of NFAT (45, 119, 338). This disruption in normal NFAT signaling may be due to blocked dephosphorylation of NFAT, which could occur by blocking the influx of calcium that is required for dephosphorylation by the calcium-calmodulin- dependent phosphatase calcineurin and subsequent nuclear translocation of NFAT (45, 119). The down-regulation of IL-2 decreases cyclins D3 and E, which in turn decreases production of the retinoblastoma protein. This decrease induces cell cycle arrest in the G_1 phase (119).

With expression of over 100 genes altered in T cells upon intoxication with VacA (119), it is perhaps not surprising that some redundancy exists in this process. A recent study found that VacA inhibition of $CD4^+$ T cell proliferation is independent of NFAT induced IL-2 activation (338). Furthermore, VacA can induce the p38 MAPK pathway within T cells, neutrophils and macrophages (45). In T cells, p38 is activated through activation of serine-threonine kinases (MKK3/6), which are linked to signaling molecules through a Rho family GTPase exchange factor, Vav (45, 53, 263). Through its exchange activity on Rac, Vav is linked to the reorganization of the cytoskeleton (53), and VacA uses Rac1 to rearrange the host cell cytoskeleton (45, 151). Taken together, all of these immune modulations by VacA probably allow for persistent infection with *H. pylori*. Additionally, perhaps the importance of immune modulation has led to the selective pressure to maintain expression of VacA in most *H. pylori* strains.

Interactions between CagA and VacA

There is a growing amount of literature that suggests that the CagA and VacA toxins interact, and that this interaction has an effect on disease severity (160). Early on, Yokoyama, *et al.* showed an antagonistic effect between CagA and VacA on the NFAT pathway (404); CagA activates the NFAT pathway via activation of calcineurin through phospholipase Cγ, whereas VacA inhibits NFAT through prevention of calcineurin activation through decreased calcium influx due to VacA mediated pores (404). Moreover, recent transfection assays showed that CagA blocks the apoptotic activity of VacA by two different mechanisms (259); phosphorylated CagA blocks the ability of VacA to traffic to intracellular compartments, whereas unphosphorylated CagA blocks apoptosis in a manner that mimics Bcl2 (an anti-apoptotic factor) overexpression (259). However, Bcl2 expression was not shown to be increased by CagA (259). In fact, CagA not only blocks the cytotoxicity of VacA, but also blocks the ablity of VacA to enter host cells (8).

Additionally, VacA and CagA show antagonistic activities in regards to cellular morphology. In cells co-cultured with isogenic *H. pylori* mutant strains deficient in *cagA* or *vacA*, increased vacuolation was seen in cells infected with *cagA* mutants, whereas cells infected with *vacA* mutants showed greater elongation of cells (21). In other words, protrusion length was reduced in cells displaying vacuoles, and the number of vacuoles was decreased in elongated cells (21). At a mechanistic level, activation of ERK1/2 by CagA is important for cell scattering and morphological changes (246), and VacA inhibits activation of ERK1/2 through inhibition of the activation of the epidermal growth factor receptor (EGFR) and the human epidermal growth factor receptor 2 (HER2/Neu;

353). One explanation for this antagonism, which was suggested by Akada, *et al,* is that CagA is injected into the cells that the bacteria are attached to, which then protects those cells from the cytotoxic ativity of VacA. VacA then proceeds to attack distant cells, thereby freeing nutrients (8). Overall, these combined interactions may explain our observation of a link between the most active *vacA* allele (s1/i1/m1), the most pathogenic *cagA* allele (EPIYA-ABD), and more severe disease manifestations (160).

Conclusion

H. pylori is a medically important bacterium that possesses a wide variety of virulence factors that allow it to thrive in the hostile environment of the stomach. The causal link between *H. pylori* infection and gastric cancer development has led to numerous studies designed to ascertain the role of virulence factors in the establishment of disease (40, 275, 276, 349). Some virulence factors, such as HomB (163, 260) and BabA (121, 230) are just beginning to be linked via epidemiological evidence to support a role in the development of more severe disease. Conversely, CagA (42, 152, 178, 274, 388) and VacA (34, 160, 187, 293, 388) have been studied extensively.

CagA was the first *H. pylori* virulence factor to be associated with more severe disease (42, 65, 72, 133) and has been shown to affect cellular processes that include β-catenin (240), ERK (141, 143, 144, 227), and the inflammatory pathways (49, 170) to name a few (Figure1). This toxin functions in both a phosphorylation-dependent and -independent manner, and polymorphisms located in the carboxyl terminus lead to differential induction of several cellular pathways (138, 142, 143). Since these polymorphisms affect the phosphorylation sites, one might assume that these variations

would only affect pathways that are CagA phosphorylation dependent. However, even in its unphosphorylated state the sequence differences within this region of CagA affect the ability of the protein to multimerize, thereby leading to differential induction of CagA phosphorylation independent pathways as well (179, 198). Meanwhile, a multitude of studies have assessed the effects of VacA on host cells (Fig. 3). This toxin also has a vast array of functions that span induction of apoptosis (115, 172, 383, 384, 397) to modulation of the immune system (Fig. 3; 236, 363, 408). Again, there are polymorphisms within the VacA toxin that affect the range of cells it can intoxicate (161, 266), as well as its ability to integrate into membranes and cause downstream effects (186, 222, 293).

These two distinct toxins clearly have some overlap in their functions. Both are able to affect cell shape (141,143-144, 227, 310, 375), affect immune cells (235, 363, 371, 408-409), and activate oncogenic pathways such as β-catenin (240, 347). They also clearly have antagonistic effects on each other, such as dampening the phenotypic effects on the host cell (cellular elongation induced by CagA versus vacuolation caused by VacA; 21). CagA also has the ability to prevent VacA induced apoptosis, whereas VacA can prevent CagA induced nuclear translocation of NFAT (404). It is believed that this antagonistic relationship exists to increase the life of the host cell (21), and it has been shown that the more active form of VacA is often associated with the more active form of CagA and is thus, further linked to more severe gastric disease (160). In addition to investigating the impacts of these toxins individually on host cells, more knowledge is needed on the interaction of these toxins and their combined impact on gastric disease.

Goals and Specific Aims

The goal of the work reported within this dissertation was to explore the role of the different polymorphisms within CagA and VacA in the development of gastric disease (Chapters Two, Three, and Four). At the same time, different *in vitro* and *in vivo* assays were optimized to be able to detect biological differences among the different polymorphic forms of CagA (Chapter Five). In order to accomplish these goals, large epidemiological and molecular epidemiological studies were conducted. Furthermore, isogenic strains containing the different polymorphic forms of CagA were created and characterized. Although the parental strain was ultimately found to contain a secondary mutation, the optimized techniques are now a part of the lab protocol repertoire. Taken together, this work explores the role of different virulence factors in the development of severe gastric disease, and provides novel information on the hierarchy of interaction of these virulence factors. Ultimately, this information could be used to create more efficacious prevention strategies or treatment options for gastric cancer.

References

1. 1989. *Campylobacter pylori* becomes *Helicobacter pylori*. Lancet **2:**1019-20.

2. 1997. Current European concepts in the management of *Helicobacter pylori* infection. The Maastricht Consensus Report. European *Helicobacter Pylori* Study Group. Gut **41:**8-13.

3. 1993. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. The EUROGAST Study Group. Gut **34:**1672-6.

4. 1994. Infection with *Helicobacter pylori*. IARC Monogr Eval Carcinog Risks Hum **61:**177-240.

5. 1994. NIH Consensus Conference. *Helicobacter pylori* in peptic ulcer disease. NIH Consensus Development Panel on *Helicobacter pylori* in Peptic Ulcer Disease. JAMA **272:**65-9.

6. **Adrian, M., T. L. Cover, J. Dubochet, and J. E. Heuser.** 2002. Multiple oligomeric states of the *Helicobacter pylori* vacuolating toxin demonstrated by cryo-electron microscopy. J Mol Biol **318:**121-33.

7. **Ahn, Y. O., B. J. Park, K. Y. Yoo, N. K. Kim, D. S. Heo, J. K. Lee, H. S. Ahn, D. H. Kang, H. Kim, M. S. Lee, and T. S. Park.** 1991. Incidence estimation of stomach cancer among Koreans. J Korean Med Sci **6:**7-14.

8. **Akada, J. K., H. Aoki, Y. Torigoe, T. Kitagawa, H. Kurazono, H. Hoshida, J. Nishikawa, S. Terai, M. Matsuzaki, T. Hirayama, T. Nakazawa, R. Akada, and K. Nakamura.** 2010. *Helicobacter pylori* CagA inhibits endocytosis of cytotoxin VacA in host cells. Dis Model Mech **3:**605-17.

9. **Akopyants, N. S., S. W. Clifton, D. Kersulyte, J. E. Crabtree, B. E. Youree, C. A. Reece, N. O. Bukanov, E. S. Drazek, B. A. Roe, and D. E. Berg.** 1998. Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. Mol Microbiol **28:**37-53.

10. **Allen, L. A., L. S. Schlesinger, and B. Kang.** 2000. Virulent strains of *Helicobacter pylori* demonstrate delayed phagocytosis and stimulate homotypic phagosome fusion in macrophages. J Exp Med **191:**115-28.

11. **Allison, C. C., T. A. Kufer, E. Kremmer, M. Kaparakis, and R. L. Ferrero.** 2009. *Helicobacter pylor*i induces MAPK phosphorylation and AP-1 activation via a NOD1-dependent mechanism. J Immunol **183:**8099-109.

12. **Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176-80.

13. **Alm, R. A., and T. J. Trust.** 1999. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. J Mol Med **77:**834-46.

14. **Amieva, M. R., and E. M. El-Omar.** 2008. Host-bacterial interactions in *Helicobacter pylori* infection. Gastroenterology **134:**306-23.

15. **Amieva, M. R., N. R. Salama, L. S. Tompkins, and S. Falkow.** 2002. *Helicobacter pylori* enter and survive within multivesicular vacuoles of epithelial cells. Cell Microbiol **4:**677-90.

16. **Amieva, M. R., R. Vogelmann, A. Covacci, L. S. Tompkins, W. J. Nelson, and S. Falkow.** 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science **300:**1430-4.

17. **Arachchi, H. S., V. Kalra, B. Lal, V. Bhatia, C. S. Baba, S. Chakravarthy, S. Rohatgi, P. M. Sarma, V. Mishra, B. Das, and V. Ahuja.** 2007. Prevalence of

duodenal ulcer-promoting gene (*dupA*) of *Helicobacter pylori* in patients with duodenal ulcer in North Indian population. Helicobacter **12:**591-7.

18. **Argent, R. H., A. Burette, V. Y. Miendje Deyi, and J. C. Atherton.** 2007. The presence of *dupA* in *Helicobacter pylori* is not significantly associated with duodenal ulceration in Belgium, South Africa, China, or North America. Clin Infect Dis **45:**1204-6.

19. **Argent, R. H., J. L. Hale, E. M. El-Omar, and J. C. Atherton.** 2008. Differences in *Helicobacter pylori* CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. J Med Microbiol **57:**1062-7.

20. **Argent, R. H., M. Kidd, R. J. Owen, R. J. Thomas, M. C. Limb, and J. C. Atherton.** 2004. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of *Helicobacter pylori*. Gastroenterology **127:**514-23.

21. **Argent, R. H., R. J. Thomas, D. P. Letley, M. G. Rittig, K. R. Hardie, and J. C. Atherton.** 2008. Functional association between the *Helicobacter pylori* virulence factors VacA and CagA. J Med Microbiol **57:**145-50.

22. **Ashorn, M., M. Maki, M. Hallstrom, M. Uhari, H. K. Akerblom, J. Viikari, and A. Miettinen.** 1995. *Helicobacter pylori* infection in Finnish children and adolescents. A serologic cross-sectional and follow-up study. Scand J Gastroenterol **30:**876-9.

23. **Atherton, J. C., and M. J. Blaser.** 2009. Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. J Clin Invest **119:**2475-87.

24. **Atherton, J. C., P. Cao, R. M. Peek, Jr., M. K. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori.* Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem **270:**17771-7.

25. **Atherton, J. C., T. L. Cover, R. J. Twells, M. R. Morales, C. J. Hawkey, and M. J. Blaser.** 1999. Simple and accurate PCR-based system for typing vacuolating cytotoxin alleles of *Helicobacter pylori*. J Clin Microbiol **37:**2979-82.

26. **Azuma, T., A. Yamakawa, S. Yamazaki, K. Fukuta, M. Ohtani, Y. Ito, M. Dojo, Y. Yamazaki, and M. Kuriyama.** 2002. Correlation between variation of the 3' region of the *cagA* gene in *Helicobacter pylori* and disease outcome in Japan. J Infect Dis **186:**1621-30.

27. **Azuma, T., S. Yamazaki, A. Yamakawa, M. Ohtani, A. Muramatsu, H. Suto, Y. Ito, M. Dojo, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Higashi, and M. Hatakeyama.** 2004. Association between diversity in the Src homology 2 domain--containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis **189:**820-7.

28. **Backert, S., N. Tegtmeyer, and M. Selbach.** 2010. The versatility of *Helicobacter pylori* CagA effector protein functions: The master key hypothesis. Helicobacter **15:**163-76.

29. **Backhed, F., E. Torstensson, D. Seguin, A. Richter-Dahlfors, and B. Rokbi.** 2003. *Helicobacter pylori* infection induces interleukin-8 receptor expression in the human gastric epithelium. Infect Immun **71:**3357-60.

30. **Bagnoli, F., L. Buti, L. Tompkins, A. Covacci, and M. R. Amieva.** 2005. *Helicobacter pylori* CagA induces a transition from polarized to invasive phenotypes in MDCK cells. Proc Natl Acad Sci U S A **102:**16339-44.

31. **Baldwin, D. N., B. Shepherd, P. Kraemer, M. K. Hall, L. K. Sycuro, D. M. Pinto-Santini, and N. R. Salama.** 2007. Identification of *Helicobacter pylori* genes that contribute to stomach colonization. Infect Immun **75:**1005-16.

32. **Baltrus, D. A., M. R. Amieva, A. Covacci, T. M. Lowe, D. S. Merrell, K. M. Ottemann, M. Stein, N. R. Salama, and K. Guillemin.** 2009. The complete genome sequence of *Helicobacter pylori* strain G27. J Bacteriol **191:**447-8.

33. **Banatvala, N., K. Mayo, F. Megraud, R. Jennings, J. J. Deeks, and R. A. Feldman.** 1993. The cohort effect and *Helicobacter pylori*. J Infect Dis **168:**219-21.

34. **Basso, D., C. F. Zambon, D. P. Letley, A. Stranges, A. Marchet, J. L. Rhead, S. Schiavon, G. Guariso, M. Ceroti, D. Nitti, M. Rugge, M. Plebani, and J. C. Atherton.** 2008. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology **135:**91-9.

35. **Begue, R. E., A. R. Cruz, A. Ramgoolam, and M. B. Breslin.** 2007. Immunogenicity of *Helicobacter pylori* urease B protein and DNA vaccines in a mouse model. J Pediatr Gastroenterol Nutr **45:**493-6.

36. **Ben Mansour, K., C. Fendri, M. Zribi, A. Masmoudi, M. Labbene, A. Fillali, N. Ben Mami, T. Najjar, A. Meherzi, T. Sfar, and C. Burucoa.** 2010. Prevalence of *Helicobacter pylori vacA*, *cagA*, *iceA* and *oipA* genotypes in Tunisian patients. Ann Clin Microbiol Antimicrob **9:**10.

37. **Bentires-Alj, M., J. G. Paez, F. S. David, H. Keilhack, B. Halmos, K. Naoki, J. M. Maris, A. Richardson, A. Bardelli, D. J. Sugarbaker, W. G. Richards, J. Du, L. Girard, J. D. Minna, M. L. Loh, D. E. Fisher, V. E. Velculescu, B. Vogelstein, M. Meyerson, W. R. Sellers, and B. G. Neel.** 2004. Activating mutations of the noonan syndrome-associated *SHP2/PTPN11* gene in human solid tumors and adult acute myelogenous leukemia. Cancer Res **64:**8816-20.

38. **Bhat, N. R., and P. Zhang.** 1999. Hydrogen peroxide activation of multiple mitogen-activated protein kinases in an oligodendrocyte cell line: role of extracellular signal-regulated kinase in hydrogen peroxide-induced cell death. J Neurochem **72:**112-9.

39. **Bickley, J., R. J. Owen, A. G. Fraser, and R. E. Pounder.** 1993. Evaluation of the polymerase chain reaction for detecting the urease C gene of *Helicobacter pylori* in gastric biopsy samples and dental plaque. J Med Microbiol **39:**338-44.

40. **Blaser, M. J.** 1998. *Helicobacter pylori* and gastric diseases. Bmj **316:**1507-10.

41. **Blaser, M. J., and J. C. Atherton.** 2004. *Helicobacter pylori* persistence: biology and disease. J Clin Invest **113:**321-33.

42. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with

Helicobacter pylori strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

43. **Bode, G., R. Barth, Q. Song, and G. Adler.** 2001. Phospholipase C activity of *Helicobacter pylori* is not associated with the presence of the *cagA* gene. Eur J Clin Invest **31:**344-8.

44. **Bohmer, C. J., E. C. Klinkenberg-Knol, E. J. Kuipers, M. C. Niezen-de Boer, H. Schreuder, F. Schuckink-Kool, and S. G. Meuwissen.** 1997. The prevalence of *Helicobacter pylori* infection among inhabitants and healthy employees of institutes for the intellectually disabled. Am J Gastroenterol **92:**1000-4.

45. **Boncristiano, M., S. R. Paccani, S. Barone, C. Ulivieri, L. Patrussi, D. Ilver, A. Amedei, M. M. D'Elios, J. L. Telford, and C. T. Baldari.** 2003. The *Helicobacter pylori* vacuolating toxin inhibits T cell activation by two independent mechanisms. J Exp Med **198:**1887-97.

46. **Boren, T., P. Falk, K. A. Roth, G. Larson, and S. Normark.** 1993. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science **262:**1892-5.

47. **Botham, C. M., A. M. Wandler, and K. Guillemin.** 2008. A transgenic *Drosophila* model demonstrates that the *Helicobacter pylori* CagA protein functions as a eukaryotic Gab adaptor. PLoS Pathog **4:**e1000064.

48. **Bourzac, K. M., C. M. Botham, and K. Guillemin.** 2007. *Helicobacter pylori* CagA induces AGS cell elongation through a cell retraction defect that is independent of Cdc42, Rac1, and Arp2/3. Infect Immun **75:**1203-13.

49. **Brandt, S., T. Kwok, R. Hartig, W. Konig, and S. Backert.** 2005. NF-kappaB activation and potentiation of proinflammatory responses by the *Helicobacter pylori* CagA protein. Proc Natl Acad Sci U S A **102:**9300-5.

50. **Brandt, S., S. Shafikhani, P. Balachandran, S. Jin, R. Hartig, W. Konig, J. Engel, and S. Backert.** 2007. Use of a novel coinfection system reveals a role for Rac1, H-Ras, and CrkII phosphorylation in *Helicobacter pylori*-induced host cell actin cytoskeletal rearrangements. FEMS Immunol Med Microbiol **50:**190-205.

51. **Brenner, H., D. Rothenbacher, G. Bode, P. Dieudonne, and G. Adler.** 1999. Active infection with *Helicobacter pylori* in healthy couples. Epidemiol Infect **122:**91-5.

52. **Bronte-Tinkew, D. M., M. Terebiznik, A. Franco, M. Ang, D. Ahn, H. Mimuro, C. Sasakawa, M. J. Ropeleski, R. M. Peek, Jr., and N. L. Jones.**

2009. *Helicobacter pylori* cytotoxin-associated gene A activates the signal transducer and activator of transcription 3 pathway *in vitro* and *in vivo*. Cancer Res **69:**632-9.

53. **Bustelo, X. R.** 2000. Regulatory and signaling properties of the Vav family. Mol Cell Biol **20:**1461-77.

54. **Cammarota, G., A. Tursi, M. Montalto, A. Papa, G. Veneto, S. Bernardi, A. Boari, V. Colizzi, G. Fedeli, and G. Gasbarrini.** 1996. Role of dental plaque in the transmission of *Helicobacter pylori* infection. J Clin Gastroenterol **22:**174-7.

55. **Candelli, M., E. C. Nista, E. Carloni, G. Pignataro, M. A. Zocco, A. Cazzato, C. Di Campli, L. Fini, G. Gasbarrini, and A. Gasbarrini.** 2005. Treatment of *H. pylori* infection: a review. Curr Med Chem **12:**375-84.

56. **Cellini, L., N. Allocati, A. Piattelli, I. Petrelli, P. Fanci, and B. Dainelli.** 1995. Microbiological evidence of *Helicobacter pylori* from dental plaque in dyspeptic patients. New Microbiol **18:**187-92.

57. **Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiara, M. Borodovsky, R. Rappuoli, and A. Covacci.** 1996. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc Natl Acad Sci U S A **93:**14648-53.

58. **Chang, Y. J., M. S. Wu, J. T. Lin, R. G. Pestell, M. J. Blaser, and C. C. Chen.** 2006. Mechanisms for *Helicobacter pylori* CagA-induced cyclin D1 expression that affect cell cycle. Cell Microbiol **8:**1740-52.

59. **Chauhan, D., S. Kharbanda, H. Uchiyama, M. Urashima, R. Fragoso, J. Sen, D. W. Kufe, and K. C. Anderson.** 1994. Identification of upstream signals regulating interleukin-6 gene expression during *in vitro* treatment of human B cells with pokeweed mitogen. Blood **84:**2243-52.

60. **Choi, K. D., N. Kim, D. H. Lee, J. M. Kim, J. S. Kim, H. C. Jung, and I. S. Song.** 2007. Analysis of the 3' variable region of the *cagA* gene of *Helicobacter pylori* isolated in Koreans. Dig Dis Sci **52:**960-6.

61. **Chowers, M. Y., N. Keller, S. Bar-Meir, and Y. Chowers.** 2002. A defined human gastrin sequence stimulates the growth of *Helicobacter pylori*. FEMS Microbiol Lett **217:**231-6.

62. **Chung, C., A. Olivares, E. Torres, O. Yilmaz, H. Cohen, and G. Perez-Perez.** 2010. Diversity of VacA intermediate region among *Helicobacter pylori* strains from several regions of the world. J Clin Microbiol **48:**690-6.

63. **Churin, Y., L. Al-Ghoul, O. Kepp, T. F. Meyer, W. Birchmeier, and M. Naumann.** 2003. *Helicobacter pylori* CagA protein targets the c-Met receptor and enhances the motogenic response. J Cell Biol **161:**249-55.

64. **Cianci, R., M. Montalto, F. Pandolfi, G. B. Gasbarrini, and G. Cammarota.** 2006. Third-line rescue therapy for *Helicobacter pylori* infection. World J Gastroenterol **12:**2313-9.

65. **Covacci, A., S. Censini, M. Bugnoli, R. Petracca, D. Burroni, G. Macchia, A. Massone, E. Papini, Z. Xiang, N. Figura, and R. Rappuoli.** 1993. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci U S A **90:**5791-5.

66. **Covacci, A., J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328-33.

67. **Cover, T. L., and S. R. Blanke.** 2005. *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality. Nat Rev Microbiol **3:**320-32.

68. **Cover, T. L., and M. J. Blaser.** 1992. Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. J Biol Chem **267:**10570-5.

69. **Cover, T. L., P. I. Hanson, and J. E. Heuser.** 1997. Acid-induced dissociation of VacA, the *Helicobacter pylori* vacuolating cytotoxin, reveals its pattern of assembly. J Cell Biol **138:**759-69.

70. **Cover, T. L., L. Y. Reddy, and M. J. Blaser.** 1993. Effects of ATPase inhibitors on the response of HeLa cells to *Helicobacter pylori* vacuolating toxin. Infect Immun **61:**1427-31.

71. **Crabtree, J. E.** 1998. Role of cytokines in pathogenesis of *Helicobacter pylori*-induced mucosal damage. Dig Dis Sci **43:**46S-55S.

72. **Crabtree, J. E., J. D. Taylor, J. I. Wyatt, R. V. Heatley, T. M. Shallcross, D. S. Tompkins, and B. J. Rathbone.** 1991. Mucosal IgA recognition of *Helicobacter pylori* 120 kDa protein, peptic ulceration, and gastric pathology. Lancet **338:**332-5.

73. **Crabtree, J. E., Z. Xiang, I. J. Lindley, D. S. Tompkins, R. Rappuoli, and A. Covacci.** 1995. Induction of interleukin-8 secretion from gastric epithelial cells by a *cagA* negative isogenic mutant of *Helicobacter pylori*. J Clin Pathol **48:**967-9.

74. **Crew, K. D., and A. I. Neugut.** 2006. Epidemiology of gastric cancer. World J Gastroenterol **12:**354-62.

75. **de Bernard, M., B. Arico, E. Papini, R. Rizzuto, G. Grandi, R. Rappuoli, and C. Montecucco.** 1997. *Helicobacter pylori* toxin VacA induces vacuole formation by acting in the cell cytosol. Mol Microbiol **26:**665-74.

76. **de Bernard, M., D. Burroni, E. Papini, R. Rappuoli, J. Telford, and C. Montecucco.** 1998. Identification of the *Helicobacter pylori* VacA toxin domain active in the cell cytosol. Infect Immun **66:**6014-6.

77. **de Bernard, M., E. Papini, V. de Filippis, E. Gottardi, J. Telford, R. Manetti, A. Fontana, R. Rappuoli, and C. Montecucco.** 1995. Low pH activates the vacuolating toxin of *Helicobacter pylori*, which becomes acid and pepsin resistant. J Biol Chem **270:**23937-40.

78. **Debellis, L., E. Papini, R. Caroppo, C. Montecucco, and S. Curci.** 2001. *Helicobacter pylori* cytotoxin VacA increases alkaline secretion in gastric epithelial cells. Am J Physiol Gastrointest Liver Physiol **281:**G1440-8.

79. **Domanska, G., C. Motz, M. Meinecke, A. Harsman, P. Papatheodorou, B. Reljic, E. A. Dian-Lothrop, A. Galmiche, O. Kepp, L. Becker, K. Gunnewig, R. Wagner, and J. Rassow.** 2010. *Helicobacter pylori* VacA toxin/subunit p34: targeting of an anion channel to the inner mitochondrial membrane. PLoS Pathog **6:**e1000878.

80. **Douraghi, M., M. Mohammadi, A. Oghalaie, A. Abdirad, M. A. Mohagheghi, M. E. Hosseini, H. Zeraati, A. Ghasemi, M. Esmaieli, and N. Mohajerani.** 2008. *dupA* as a risk determinant in *Helicobacter pylori* infection. J Med Microbiol **57:**554-62.

81. **Drazek, E. S., A. Dubois, and R. K. Holmes.** 1994. Characterization and presumptive identification of *Helicobacter pylori* isolates from rhesus monkeys. J Clin Microbiol **32:**1799-804.

82. **Drumm, B., G. I. Perez-Perez, M. J. Blaser, and P. M. Sherman.** 1990. Intrafamilial clustering of *Helicobacter pylori* infection. N Engl J Med **322:**359-63.

83. **Du, Y., A. Agnew, X. P. Ye, P. A. Robinson, D. Forman, and J. E. Crabtree.** 2006. *Helicobacter pylori* and *Schistosoma japonicum* co-infection in a Chinese population: helminth infection alters humoral responses to *H. pylori* and serum pepsinogen I/II ratio. Microbes Infect **8:**52-60.

84. **Dube, C., T. C. Nkosi, A. M. Clarke, N. Mkwetshana, E. Green, and R. N. Ndip.** 2009. *Helicobacter pylori* antigenemia in an asymptomatic population of Eastern Cape Province, South Africa: public health implications. Rev Environ Health **24:**249-55.

85. **Dubois, A., N. Fiala, L. M. Heman-Ackah, E. S. Drazek, A. Tarnawski, W. N. Fishbein, G. I. Perez-Perez, and M. J. Blaser.** 1994. Natural gastric infection with *Helicobacter pylori* in monkeys: a model for spiral bacteria infection in humans. Gastroenterology **106:**1405-17.

86. **Dunn, B. E., H. Cohen, and M. J. Blaser.** 1997. *Helicobacter pylori*. Clin Microbiol Rev **10:**720-41.

87. **Eaton, K. A., and S. Krakowka.** 1994. Effect of gastric pH on urease-dependent colonization of gnotobiotic piglets by *Helicobacter pylori*. Infect Immun **62:**3604-7.

88. **Eaton, K. A., D. R. Morgan, and S. Krakowka.** 1992. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. J Med Microbiol **37:**123-7.

89. **Ekstrom, A. M., M. Serafini, O. Nyren, L. E. Hansson, W. Ye, and A. Wolk.** 2000. Dietary antioxidant intake and the risk of cardia cancer and noncardia cancer of the intestinal and diffuse types: a population-based case-control study in Sweden. Int J Cancer **87:**133-40.

90. **El-Etr, S. H., A. Mueller, L. S. Tompkins, S. Falkow, and D. S. Merrell.** 2004. Phosphorylation-independent effects of CagA during interaction between *Helicobacter pylori* and T84 polarized monolayers. J Infect Dis **190:**1516-23.

91. **El-Omar, E. M., M. Carrington, W. H. Chow, K. E. McColl, J. H. Bream, H. A. Young, J. Herrera, J. Lissowska, C. C. Yuan, N. Rothman, G. Lanyon, M. Martin, J. F. Fraumeni, Jr., and C. S. Rabkin.** 2000. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature **404:**398-402.

92. **El-Omar, E. M., M. Carrington, W. H. Chow, K. E. McColl, J. H. Bream, H. A. Young, J. Herrera, J. Lissowska, C. C. Yuan, N. Rothman, G. Lanyon, M. Martin, J. F. Fraumeni, Jr., and C. S. Rabkin.** 2001. The role of interleukin-1 polymorphisms in the pathogenesis of gastric cancer. Nature **412:**99.

93. **Epplein, M., A. M. Nomura, J. H. Hankin, M. J. Blaser, G. Perez-Perez, G. N. Stemmermann, L. R. Wilkens, and L. N. Kolonel.** 2008. Association of *Helicobacter pylori* infection and diet on the risk of gastric cancer: a case-control study in Hawaii. Cancer Causes Control **19:**869-77.

94. **Ernst, P. B., and B. D. Gold.** 2000. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu Rev Microbiol **54:**615-40.

95. **Evans, D. J., Jr., D. G. Evans, H. C. Lampert, and H. Nakano.** 1995. Identification of four new prokaryotic bacterioferritins, from *Helicobacter pylori, Anabaena variabilis, Bacillus subtilis* and *Treponema pallidum*, by analysis of gene sequences. Gene **153:**123-7.

96. **Evans, D. J., Jr., D. G. Evans, T. Takemura, H. Nakano, H. C. Lampert, D. Y. Graham, D. N. Granger, and P. R. Kvietys.** 1995. Characterization of a *Helicobacter pylori* neutrophil-activating protein. Infect Immun **63:**2213-20.

97. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res **20:**3-26.

98. **Fan, X. G., A. Chua, T. G. Li, and Q. S. Zeng.** 1998. Survival of *Helicobacter pylori* in milk and tap water. J Gastroenterol Hepatol **13:**1096-8.

99. **Ferguson, D. A., Jr., C. Li, N. R. Patel, W. R. Mayberry, D. S. Chi, and E. Thomas.** 1993. Isolation of *Helicobacter pylori* from saliva. J Clin Microbiol **31:**2802-4.

100. **Ferrero, R. L., J. M. Thiberge, I. Kansau, N. Wuscher, M. Huerre, and A. Labigne.** 1995. The GroES homolog of *Helicobacter pylori* confers protective

immunity against mucosal infection in mice. Proc Natl Acad Sci U S A **92:**6499-503.

101. **Fiedorek, S. C., H. M. Malaty, D. L. Evans, C. L. Pumphrey, H. B. Casteel, D. J. Evans, Jr., and D. Y. Graham.** 1991. Factors influencing the epidemiology of *Helicobacter pylori* infection in children. Pediatrics **88:**578-82.

102. **Fischer, W., J. Puls, R. Buhrdorf, B. Gebert, S. Odenbreit, and R. Haas.** 2001. Systematic mutagenesis of the *Helicobacter pylori cag* pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. Mol Microbiol **42:**1337-48.

103. **Foo, J. H., J. G. Culvenor, R. L. Ferrero, T. Kwok, T. Lithgow, and K. Gabriel.** 2010. Both p33 and p55 subunits of *Helicobacter pylori* VacA are targeted to mammalian mitochondria. J Mol Biol **401:**792-798.

104. **Ford, A. C., and A. T. Axon.** 2010. Epidemiology of *Helicobacter pylori* infection and public health implications. Helicobacter **15 Suppl 1:**1-6.

105. **Fowler, M., R. J. Thomas, J. Atherton, I. S. Roberts, and N. J. High.** 2006. Galectin-3 binds to *Helicobacter pylori* O-antigen: it is upregulated and rapidly secreted by gastric epithelial cells in response to *H. pylori* adhesion. Cell Microbiol **8:**44-54.

106. **Fox, J. G., P. Beck, C. A. Dangler, M. T. Whary, T. C. Wang, H. N. Shi, and C. Nagler-Anderson.** 2000. Concurrent enteric helminth infection modulates inflammation and gastric immune responses and reduces helicobacter-induced gastric atrophy. Nat Med **6:**536-42.

107. **Fox, J. G., C. A. Dangler, N. S. Taylor, A. King, T. J. Koh, and T. C. Wang.** 1999. High-salt diet induces gastric epithelial hyperplasia and parietal cell loss, and enhances *Helicobacter pylori* colonization in C57BL/6 mice. Cancer Res **59:**4823-8.

108. **Franco, A. T., D. A. Israel, M. K. Washington, U. Krishna, J. G. Fox, A. B. Rogers, A. S. Neish, L. Collier-Hyams, G. I. Perez-Perez, M. Hatakeyama, R. Whitehead, K. Gaus, D. P. O'Brien, J. Romero-Gallo, and R. M. Peek, Jr.** 2005. Activation of beta-catenin by carcinogenic *Helicobacter pylori*. Proc Natl Acad Sci U S A **102:**10646-51.

109. **Franco, A. T., E. Johnston, U. Krishna, Y. Yamaoka, D. A. Israel, T. A. Nagy, L. E. Wroblewski, M. B. Piazuelo, P. Correa, and R. M. Peek, Jr.** 2008. Regulation of gastric carcinogenesis by *Helicobacter pylori* virulence factors. Cancer Res **68:**379-87.

110. **Fritz, E. L., T. Slavik, W. Delport, B. Olivier, and S. W. van der Merwe.** 2006. Incidence of *Helicobacter felis* and the effect of coinfection with *Helicobacter pylori* on the gastric mucosa in the African population. J Clin Microbiol **44:**1692-6.

111. **Fujikawa, A., D. Shirasaka, S. Yamamoto, H. Ota, K. Yahiro, M. Fukada, T. Shintani, A. Wada, N. Aoyama, T. Hirayama, H. Fukamachi, and M. Noda.** 2003. Mice deficient in protein tyrosine phosphatase receptor type Z are resistant to gastric ulcer induction by VacA of *Helicobacter pylori*. Nat Genet **33:**375-81.

112. **Furuta, T., E. M. El-Omar, F. Xiao, N. Shirai, M. Takashima, and H. Sugimura.** 2002. Interleukin 1beta polymorphisms increase risk of hypochlorhydria and atrophic gastritis and reduce risk of duodenal ulcer recurrence in Japan. Gastroenterology **123:**92-105.

113. **Furuta, T., T. Kamata, M. Takashima, H. Futami, H. Arai, H. Hanai, and E. Kaneko.** 1997. Study of transmission routes of *Helicobacter pylori* in relation to seroprevalence of hepatitis A virus. J Clin Microbiol **35:**1891-3.

114. **Gale, N. W., S. Kaplan, E. J. Lowenstein, J. Schlessinger, and D. Bar-Sagi.** 1993. Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. Nature **363:**88-92.

115. **Galmiche, A., J. Rassow, A. Doye, S. Cagnol, J. C. Chambard, S. Contamin, V. de Thillot, I. Just, V. Ricci, E. Solcia, E. Van Obberghen, and P. Boquet.** 2000. The N-terminal 34 kDa fragment of *Helicobacter pylori* vacuolating cytotoxin targets mitochondria and induces cytochrome c release. Embo J **19:**6361-70.

116. **Garner, J. A., and T. L. Cover.** 1996. Binding and internalization of the *Helicobacter pylori* vacuolating cytotoxin by epithelial cells. Infect Immun **64:**4197-203.

117. **Gauthier, N. C., P. Monzo, T. Gonzalez, A. Doye, A. Oldani, P. Gounon, V. Ricci, M. Cormont, and P. Boquet.** 2007. Early endosomes associated with dynamic F-actin structures are required for late trafficking of *H. pylori* VacA toxin. J Cell Biol **177:**343-54.

118. **Gauthier, N. C., P. Monzo, V. Kaddai, A. Doye, V. Ricci, and P. Boquet.** 2005. *Helicobacter pylori* VacA cytotoxin: a probe for a clathrin-independent and Cdc42-dependent pinocytic pathway routed to late endosomes. Mol Biol Cell **16:**4852-66.

119. **Gebert, B., W. Fischer, E. Weiss, R. Hoffmann, and R. Haas.** 2003. *Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation. Science **301:**1099-102.

120. **Geis, G., S. Suerbaum, B. Forsthoff, H. Leying, and W. Opferkuch.** 1993. Ultrastructure and biochemical studies of the flagellar sheath of *Helicobacter pylori*. J Med Microbiol **38:**371-7.

121. **Gerhard, M., N. Lehn, N. Neumayer, T. Boren, R. Rad, W. Schepp, S. Miehlke, M. Classen, and C. Prinz.** 1999. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proc Natl Acad Sci U S A **96:**12778-83.

122. **Ghose, C., G. I. Perez-Perez, L. J. van Doorn, M. G. Dominguez-Bello, and M. J. Blaser.** 2005. High frequency of gastric colonization with multiple *Helicobacter pylori* strains in Venezuelan subjects. J Clin Microbiol **43:**2635-41.

123. **Gisbert, J. P.** 2005. The recurrence of *Helicobacter pylori* infection: incidence and variables influencing it. A critical review. Am J Gastroenterol **100:**2083-99.

124. **Go, M. F., V. Kapur, D. Y. Graham, and J. M. Musser.** 1996. Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. J Bacteriol **178:**3934-8.

125. **Gomes, L. I., G. A. Rocha, A. M. Rocha, T. F. Soares, C. A. Oliveira, P. F. Bittencourt, and D. M. Queiroz.** 2008. Lack of association between *Helicobacter pylori* infection with *dupA*-positive strains and gastroduodenal diseases in Brazilian patients. Int J Med Microbiol **298:**223-30.

126. **Goodwin, C. S., J. A. Armstrong, and B. J. Marshall.** 1986. *Campylobacter pyloridis*, gastritis, and peptic ulceration. J Clin Pathol **39:**353-65.

127. **Goodwin, C. S., R. K. McCulloch, J. A. Armstrong, and S. H. Wee.** 1985. Unusual cellular fatty acids and distinctive ultrastructure in a new spiral bacterium (*Campylobacter pyloridis*) from the human gastric mucosa. J Med Microbiol **19:**257-67.

128. **Goto, Y., T. Ando, K. Yamamoto, A. Tamakoshi, E. El-Omar, H. Goto, and N. Hamajima.** 2006. Association between serum pepsinogens and polymorphismof *PTPN11* encoding SHP-2 among *Helicobacter pylori* seropositive Japanese. Int J Cancer **118:**203-8.

129. **Graham, D. Y., H. M. Malaty, D. G. Evans, D. J. Evans, Jr., P. D. Klein, and E. Adam.** 1991. Epidemiology of *Helicobacter pylori* in an asymptomatic population in the United States. Effect of age, race, and socioeconomic status. Gastroenterology **100:**1495-501.

130. **Granstrom, M., Y. Tindberg, and M. Blennow.** 1997. Seroepidemiology of *Helicobacter pylori* infection in a cohort of children monitored from 6 months to 11 years of age. J Clin Microbiol **35:**468-70.

131. **Gupta, V. R., H. K. Patel, S. S. Kostolansky, R. A. Ballivian, J. Eichberg, and S. R. Blanke.** 2008. Sphingomyelin functions as a novel receptor for *Helicobacter pylori* VacA. PLoS Pathog **4:**e1000073.

132. **Gupta, V. R., B. A. Wilson, and S. R. Blanke.** 2010. Sphingomyelin is important for the cellular entry and intracellular localization of *Helicobacter pylori* VacA. Cell Microbiol **12:**1517-1533.

133. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

134. **Hanninen, M. L.** 1991. Sensitivity of *Helicobacter pylori* to different bile salts. Eur J Clin Microbiol Infect Dis **10:**515-8.

135. **Harris, A. W., A. Douds, E. V. Meurisse, M. Dennis, S. Chambers, and S. R. Gould.** 1995. Seroprevalence of *Helicobacter pylori* in residents of a hospital for people with severe learning difficulties. Eur J Gastroenterol Hepatol **7:**21-3.

136. **Hatakeyama, M.** 2003. *Helicobacter pylori* CagA--a potential bacterial oncoprotein that functionally mimics the mammalian Gab family of adaptor proteins. Microbes Infect **5:**143-50.

137. **Hatakeyama, M.** 2006. *Helicobacter pylori* CagA -- a bacterial intruder conspiring gastric carcinogenesis. Int J Cancer **119:**1217-23.

138. **Hatakeyama, M.** 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat Rev Cancer **4:**688-94.

139. **Hatakeyama, M., and H. Higashi.** 2005. *Helicobacter pylori* CagA: a new paradigm for bacterial carcinogenesis. Cancer Sci **96:**835-43.

140. **Hazell, S. L., H. M. Mitchell, M. Hedges, X. Shi, P. J. Hu, Y. Y. Li, A. Lee, and E. Reiss-Levy.** 1994. Hepatitis A and evidence against the community dissemination of *Helicobacter pylori* via feces. J Infect Dis **170:**686-9.

141. **Higashi, H., A. Nakaya, R. Tsutsumi, K. Yokoyama, Y. Fujii, S. Ishikawa, M. Higuchi, A. Takahashi, Y. Kurashima, Y. Teishikata, S. Tanaka, T. Azuma, and M. Hatakeyama.** 2004. *Helicobacter pylori* CagA induces Ras-independent morphogenetic response through SHP-2 recruitment and activation. J Biol Chem **279:**17205-16.

142. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

143. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **295:**683-6.

144. **Higuchi, M., R. Tsutsumi, H. Higashi, and M. Hatakeyama.** 2004. Conditional gene silencing utilizing the lac repressor reveals a role of SHP-2 in *cagA*-positive *Helicobacter pylori* pathogenicity. Cancer Sci **95:**442-7.

145. **Hirata, Y., S. Maeda, Y. Mitsuno, K. Tateishi, A. Yanai, M. Akanuma, H. Yoshida, T. Kawabe, Y. Shiratori, and M. Omata.** 2002. *Helicobacter pylori* CagA protein activates serum response element-driven transcription independently of tyrosine phosphorylation. Gastroenterology **123:**1962-71.

146. **Hirohata, T., and S. Kono.** 1997. Diet/nutrition and stomach cancer in Japan. Int J Cancer **Suppl 10:**34-6.

147. **Hisatsune, J., M. Nakayama, H. Isomoto, H. Kurazono, N. Mukaida, A. K. Mukhopadhyay, T. Azuma, Y. Yamaoka, J. Sap, E. Yamasaki, K. Yahiro, J. Moss, and T. Hirayama.** 2008. Molecular characterization of *Helicobacter pylori* VacA induction of IL-8 in U937 cells reveals a prominent role for p38MAPK in activating transcription factor-2, cAMP response element binding protein, and NF-kappaB activation. J Immunol **180:**5017-27.

148. **Hisatsune, J., E. Yamasaki, M. Nakayama, D. Shirasaka, H. Kurazono, Y. Katagata, H. Inoue, J. Han, J. Sap, K. Yahiro, J. Moss, and T. Hirayama.** 2007. *Helicobacter pylori* VacA enhances prostaglandin E2 production through induction of cyclooxygenase 2 expression via a p38 mitogen-activated protein kinase/activating transcription factor 2 cascade in AZ-521 cells. Infect Immun **75:**4472-81.

149. **Hishida, A., K. Matsuo, Y. Goto, and N. Hamajima.** Genetic predisposition to *Helicobacter pylori*-induced gastric precancerous conditions. World J Gastrointest Oncol **2:**369-79.

150. **Ho, S. A., J. A. Hoyle, F. A. Lewis, A. D. Secker, D. Cross, N. P. Mapstone, M. F. Dixon, J. I. Wyatt, D. S. Tompkins, G. R. Taylor, and et al.** 1991. Direct polymerase chain reaction test for detection of *Helicobacter pylori* in humans and animals. J Clin Microbiol **29:**2543-9.

151. **Hotchin, N. A., T. L. Cover, and N. Akhtar.** 2000. Cell vacuolation induced by the VacA cytotoxin of *Helicobacter pylori* is regulated by the Rac1 GTPase. J Biol Chem **275:**14009-12.

152. **Huang, J. Q., G. F. Zheng, K. Sumanac, E. J. Irvine, and R. H. Hunt.** 2003. Meta-analysis of the relationship between *cagA* seropositivity and gastric cancer. Gastroenterology **125:**1636-44.

153. **Hulten, K., S. W. Han, H. Enroth, P. D. Klein, A. R. Opekun, R. H. Gilman, D. G. Evans, L. Engstrand, D. Y. Graham, and F. A. El-Zaatari.** 1996. *Helicobacter pylori* in the drinking water in Peru. Gastroenterology **110:**1031-5.

154. **Hussein, N. R.** 2010. The association of *dupA* and *Helicobacter pylori*-related gastroduodenal diseases. Eur J Clin Microbiol Infect Dis **29:**817-21.

155. **Ilver, D., A. Arnqvist, J. Ogren, I. M. Frick, D. Kersulyte, E. T. Incecik, D. E. Berg, A. Covacci, L. Engstrand, and T. Boren.** 1998. *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. Science **279:**373-7.

156. **Ilver, D., S. Barone, D. Mercati, P. Lupetti, and J. L. Telford.** 2004. *Helicobacter pylori* toxin VacA is transferred to host cells via a novel contact-dependent mechanism. Cell Microbiol **6:**167-74.

157. **Isenberg, J. I., J. A. Selling, D. L. Hogan, and M. A. Koss.** 1987. Impaired proximal duodenal mucosal bicarbonate secretion in patients with duodenal ulcer. N Engl J Med **316:**374-9.

158. **Israel, D. A., N. Salama, U. Krishna, U. M. Rieger, J. C. Atherton, S. Falkow, and R. M. Peek, Jr.** 2001. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. Proc Natl Acad Sci U S A **98:**14625-30.

159. **Jain, J., V. E. Valge-Archer, and A. Rao.** 1992. Analysis of the AP-1 sites in the IL-2 promoter. J Immunol **148:**1240-50.

160. **Jang, S., K. R. Jones, C. H. Olsen, Y. M. Joo, Y. J. Yoo, I. S. Chung, J. H. Cha, and D. S. Merrell.** 2010. Epidemiological link between gastric disease and polymorphisms in VacA and CagA. J Clin Microbiol **48:**559-67.

161. **Ji, X., T. Fernandez, D. Burroni, C. Pagliaccia, J. C. Atherton, J. M. Reyrat, R. Rappuoli, and J. L. Telford.** 2000. Cell specificity of *Helicobacter pylori* cytotoxin is determined by a short region in the polymorphic midregion. Infect Immun **68:**3754-7.

162. **Jones, K. R., Y. M. Joo, S. Jang, Y. J. Yoo, H. S. Lee, I. S. Chung, C. H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in

the CagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol **47:**959-68.

163. **Jung, S. W., M. Sugimoto, D. Y. Graham, and Y. Yamaoka.** 2009. *homB* status of *Helicobacter pylori* as a novel marker to distinguish gastric cancer from duodenal ulcer. J Clin Microbiol **47:**3241-5.

164. **Kang, J., and M. J. Blaser.** 2006. Bacterial populations as perfect gases: genomic integrity and diversification tensions in *Helicobacter pylori*. Nat Rev Microbiol **4:**826-36.

165. **Kaparakis, M., L. Turnbull, L. Carneiro, S. Firth, H. A. Coleman, H. C. Parkington, L. Le Bourhis, A. Karrar, J. Viala, J. Mak, M. L. Hutton, J. K. Davies, P. J. Crack, P. J. Hertzog, D. J. Philpott, S. E. Girardin, C. B. Whitchurch, and R. L. Ferrero.** 2010. Bacterial membrane vesicles deliver peptidoglycan to NOD1 in epithelial cells. Cell Microbiol **12:**372-85.

166. **Kavermann, H., B. P. Burns, K. Angermuller, S. Odenbreit, W. Fischer, K. Melchers, and R. Haas.** 2003. Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. J Exp Med **197:**813-22.

167. **Keates, S., A. C. Keates, M. Warny, R. M. Peek, Jr., P. G. Murray, and C. P. Kelly.** 1999. Differential activation of mitogen-activated protein kinases in AGS

gastric epithelial cells by *cag*+ and *cag*- *Helicobacter pylori*. J Immunol **163:**5552-9.

168. **Kelly, S. M., M. C. Pitcher, S. M. Farmery, and G. R. Gibson.** 1994. Isolation of *Helicobacter pylori* from feces of patients with dyspepsia in the United Kingdom. Gastroenterology **107:**1671-4.

169. **Kim, S., A. K. Chamberlain, and J. U. Bowie.** 2004. Membrane channel structure of *Helicobacter pylori* vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. Proc Natl Acad Sci U S A **101:**5988-91.

170. **Kim, S. Y., Y. C. Lee, H. K. Kim, and M. J. Blaser.** 2006. *Helicobacter pylori* CagA transfection of gastric epithelial cells induces interleukin-8. Cell Microbiol **8:**97-106.

171. **Kimura, A., T. Matsubasa, H. Kinoshita, N. Kuriya, Y. Yamashita, T. Fujisawa, H. Terakura, and M. Shinohara.** 1999. *Helicobacter pylori* seropositivity in patients with severe neurologic impairment. Brain Dev **21:**113-7.

172. **Kimura, M., S. Goto, A. Wada, K. Yahiro, T. Niidome, T. Hatakeyama, H. Aoyagi, T. Hirayama, and T. Kondo.** 1999. Vacuolating cytotoxin purified from *Helicobacter pylori* causes mitochondrial damage in human gastric cells. Microb Pathog **26:**45-52.

173. **Kivi, M., A. L. Johansson, M. Reilly, and Y. Tindberg.** 2005. *Helicobacter pylor*i status in family members as risk factors for infection in children. Epidemiol Infect **133:**645-52.

174. **Kivi, M., and Y. Tindberg.** 2006. *Helicobacter pylori* occurrence and transmission: a family affair? Scand J Infect Dis **38:**407-17.

175. **Klampfer, L.** 2006. Signal transducers and activators of transcription (STATs): Novel targets of chemopreventive and chemotherapeutic drugs. Curr Cancer Drug Targets **6:**107-21.

176. **Klein, P. D., R. H. Gilman, R. Leon-Barua, F. Diaz, E. O. Smith, and D. Y. Graham.** 1994. The epidemiology of *Helicobacter pylori* in Peruvian children between 6 and 30 months of age. Am J Gastroenterol **89:**2196-200.

177. **Kotloff, K. L., M. B. Sztein, S. S. Wasserman, G. A. Losonsky, S. C. DiLorenzo, and R. I. Walker.** 2001. Safety and immunogenicity of oral inactivated whole-cell *Helicobacter pylori* vaccine with adjuvant among volunteers with or without subclinical infection. Infect Immun **69:**3581-90.

178. **Kuipers, E. J., G. I. Perez-Perez, S. G. Meuwissen, and M. J. Blaser.** 1995. *Helicobacter pylori* and atrophic gastritis: importance of the *cagA* status. J Natl Cancer Inst **87:**1777-80.

179. **Kurashima, Y., N. Murata-Kamiya, K. Kikuchi, H. Higashi, T. Azuma, S. Kondo, and M. Hatakeyama.** 2008. Deregulation of beta-catenin signal by *Helicobacter pylori* CagA requires the CagA-multimerization sequence. Int J Cancer **122:**823-31.

180. **Lamb, A., and L. F. Chen.** 2010. The many roads traveled by *Helicobacter pylori* to NFKappaB activation. Gut Microbes **1:**109-113.

181. **Lamb, A., X. D. Yang, Y. H. Tsang, J. D. Li, H. Higashi, M. Hatakeyama, R. M. Peek, S. R. Blanke, and L. F. Chen.** 2009. *Helicobacter pylori* CagA activates NF-kappaB by targeting TAK1 for TRAF6-mediated Lys 63 ubiquitination. EMBO Rep **10:**1242-9.

182. **Lambert, J. R., S. K. Lin, W. Sievert, L. Nicholson, M. Schembri, and C. Guest.** 1995. High prevalence of *Helicobacter pylori* antibodies in an institutionalized population: evidence for person-to-person transmission. Am J Gastroenterol **90:**2167-71.

183. **Lanzavecchia, S., P. L. Bellon, P. Lupetti, R. Dallai, R. Rappuoli, and J. L. Telford.** 1998. Three-dimensional reconstruction of metal replicas of the *Helicobacter pylori* vacuolating cytotoxin. J Struct Biol **121:**9-18.

184. **Lee, I. O., J. H. Kim, Y. J. Choi, M. H. Pillinger, S. Y. Kim, M. J. Blaser, and Y. C. Lee.** 2010. *Helicobacter pylori* CagA phosphorylation status determines the gp130-activated SHP2/ERK and JAK/STAT signal transduction pathways in gastric epithelial cells. J Biol Chem **285:**16042-50.

185. **Lee, I. O., J. H. Kim, Y. J. Choi, M. H. Pillinger, S. Y. Kim, M. J. Blaser, and Y. C. Lee.** *Helicobacter pylori* CagA phosphorylation status determines the gp130-activated SHP2/ERK and JAK/STAT signal transduction pathways in gastric epithelial cells. J Biol Chem **285:**16042-50.

186. **Letley, D. P., and J. C. Atherton.** 2000. Natural diversity in the N terminus of the mature vacuolating cytotoxin of *Helicobacter pylori* determines cytotoxin activity. J Bacteriol **182:**3278-80.

187. **Letley, D. P., J. L. Rhead, R. J. Twells, B. Dove, and J. C. Atherton.** 2003. Determinants of non-toxicity in the gastric pathogen *Helicobacter pylori*. J Biol Chem **278:**26734-41.

188. **Leunk, R. D., P. T. Johnson, B. C. David, W. G. Kraft, and D. R. Morgan.** 1988. Cytotoxic activity in broth-culture filtrates of *Campylobacter pylori*. J Med Microbiol **26:**93-9.

189. **Levi, S., K. Beardshall, G. Haddad, R. Playford, P. Ghosh, and J. Calam.** 1989. *Campylobacter pylori* and duodenal ulcers: the gastrin link. Lancet **1:**1167-8.

190. **Lewindon, P. J., D. Lau, A. Chan, P. Tse, and P. B. Sullivan.** 1997. *Helicobacter pylori* in an institution for disabled children in Hong Kong. Dev Med Child Neurol **39:**682-5.

191. **Li, C., T. Ha, D. A. Ferguson, Jr., D. S. Chi, R. Zhao, N. R. Patel, G. Krishnaswamy, and E. Thomas.** 1996. A newly developed PCR assay of *H. pylori* in gastric biopsy, saliva, and feces. Evidence of high prevalence of *H. pylori* in saliva supports oral transmission. Dig Dis Sci **41:**2142-9.

192. **Li, N., A. Batzer, R. Daly, V. Yajnik, E. Skolnik, P. Chardin, D. Bar-Sagi, B. Margolis, and J. Schlessinger.** 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature **363:**85-8.

193. **Lickert, H., C. Domon, G. Huls, C. Wehrle, I. Duluc, H. Clevers, B. I. Meyer, J. N. Freund, and R. Kemler.** 2000. Wnt/(beta)-catenin signaling regulates the expression of the homeobox gene *Cdx1* in embryonic intestine. Development **127:**3805-13.

194. **Lin, S. K., J. R. Lambert, M. A. Schembri, L. Nicholson, and I. H. Johnson.** 1998. The prevalence of *Helicobacter pylori* in practising dental staff and dental students. Aust Dent J **43:**35-9.

195. **Lin, W. C., H. F. Tsai, S. H. Kuo, M. S. Wu, C. W. Lin, P. I. Hsu, A. L. Cheng, and P. N. Hsu.** 2010. Translocation of Helicobacter pylori CagA into Human B lymphocytes, the origin of mucosa-associated lymphoid tissue lymphoma. Cancer Res **70:**5740-8.

196. **Lu, H., P. I. Hsu, D. Y. Graham, and Y. Yamaoka.** 2005. Duodenal ulcer promoting gene of *Helicobacter pylori*. Gastroenterology **128:**833-48.

197. **Lu, H., N. Murata-Kamiya, Y. Saito, and M. Hatakeyama.** 2009. Role of partitioning-defective 1/microtubule affinity-regulating kinases in the morphogenetic activity of *Helicobacter pylori* CagA. J Biol Chem **284:**23024-36.

198. **Lu, H. S., Y. Saito, M. Umeda, N. Murata-Kamiya, H. M. Zhang, H. Higashi, and M. Hatakeyama.** 2008. Structural and functional diversity in the

PAR1b/MARK2-binding region of *Helicobacter pylori* CagA. Cancer Sci **99:**2004-11.

199. **Lupetti, P., J. E. Heuser, R. Manetti, P. Massari, S. Lanzavecchia, P. L. Bellon, R. Dallai, R. Rappuoli, and J. L. Telford.** 1996. Oligomeric and subunit structure of the *Helicobacter pylori* vacuolating cytotoxin. J Cell Biol **133:**801-7.

200. **Luzza, F., M. Imeneo, M. Maletta, G. Paluccio, A. Giancotti, F. Perticone, A. Foca, and F. Pallone.** 1997. Seroepidemiology of *Helicobacter pylori* infection and hepatitis A in a rural area: evidence against a common mode of transmission. Gut **41:**164-8.

201. **Luzza, F., M. Mancuso, M. Imeneo, A. Contaldo, L. Giancotti, L. Pensabene, P. Doldo, M. C. Liberto, P. Strisciuglio, A. Foca, S. Guandalini, and F. Pallone.** 2000. Evidence favouring the gastro-oral route in the transmission of *Helicobacter pylori* infection in children. Eur J Gastroenterol Hepatol **12:**623-7.

202. **Machado, J. C., P. Pharoah, S. Sousa, R. Carvalho, C. Oliveira, C. Figueiredo, A. Amorim, R. Seruca, C. Caldas, F. Carneiro, and M. Sobrinho-Simoes.** 2001. Interleukin 1B and interleukin 1RN polymorphisms are associated with increased risk of gastric carcinoma. Gastroenterology **121:**823-9.

203. **Mahdavi, J., B. Sonden, M. Hurtig, F. O. Olfat, L. Forsberg, N. Roche, J. Angstrom, T. Larsson, S. Teneberg, K. A. Karlsson, S. Altraja, T. Wadstrom, D. Kersulyte, D. E. Berg, A. Dubois, C. Petersson, K. E. Magnusson, T. Norberg, F. Lindh, B. B. Lundskog, A. Arnqvist, L. Hammarstrom, and T. Boren.** 2002. *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. Science **297:**573-8.

204. **Mai, U. E., G. I. Perez-Perez, J. B. Allen, S. M. Wahl, M. J. Blaser, and P. D. Smith.** 1992. Surface proteins from *Helicobacter pylori* exhibit chemotactic activity for human leukocytes and are present in gastric mucosa. J Exp Med **175:**517-25.

205. **Majewski, S. I., and C. S. Goodwin.** 1988. Restriction endonuclease analysis of the genome of *Campylobacter pylori* with a rapid extraction method: evidence for considerable genomic variation. J Infect Dis **157:**465-71.

206. **Malaty, H. M.** 2007. Epidemiology of *Helicobacter pylori* infection. Best Pract Res Clin Gastroenterol **21:**205-14.

207. **Malaty, H. M., D. J. Evans, Jr., K. Abramovitch, D. G. Evans, and D. Y. Graham.** 1992. *Helicobacter pylori* infection in dental workers: a seroepidemiology study. Am J Gastroenterol **87:**1728-31.

208. **Malaty, H. M., and D. Y. Graham.** 1994. Importance of childhood socioeconomic status on the current prevalence of *Helicobacter pylori* infection. Gut **35:**742-5.

209. **Malaty, H. M., D. Y. Graham, W. A. Wattigney, S. R. Srinivasan, M. Osato, and G. S. Berenson.** 1999. Natural history of *Helicobacter pylori* infection in childhood: 12-year follow-up cohort study in a biracial community. Clin Infect Dis **28:**279-82.

210. **Malaty, H. M., J. G. Kim, S. D. Kim, and D. Y. Graham.** 1996. Prevalence of *Helicobacter pylori* infection in Korean children: inverse relation to socioeconomic status despite a uniformly high prevalence in adults. Am J Epidemiol **143:**257-62.

211. **Malaty, H. M., N. D. Logan, D. Y. Graham, and J. E. Ramchatesingh.** 2001. *Helicobacter pylori* infection in preschool and school-aged minority children: effect of socioeconomic indicators and breast-feeding practices. Clin Infect Dis **32:**1387-92.

212. **Malaty, H. M., V. Paykov, O. Bykova, A. Ross, D. P. Graham, J. F. Anneger, and D. Y. Graham.** 1996. *Helicobacter pylori* and socioeconomic factors in Russia. Helicobacter **1:**82-7.

213. **Malfertheiner, P., F. Megraud, C. O'Morain, F. Bazzoli, E. El-Omar, D. Graham, R. Hunt, T. Rokkas, N. Vakil, and E. J. Kuipers.** 2007. Current concepts in the management of *Helicobacter pylori* infection: the Maastricht III Consensus Report. Gut **56:**772-81.

214. **Malfertheiner, P., F. Megraud, C. O'Morain, A. P. Hungin, R. Jones, A. Axon, D. Y. Graham, and G. Tytgat.** 2002. Current concepts in the management of *Helicobacter pylori* infection--the Maastricht 2-2000 Consensus Report. Aliment Pharmacol Ther **16:**167-80.

215. **Manente, L., A. Perna, E. Buommino, L. Altucci, A. Lucariello, G. Citro, A. Baldi, G. Iaquinto, M. A. Tufano, and A. De Luca.** 2008. The *Helicobacter pylori's* protein VacA has direct effects on the regulation of cell cycle and apoptosis in gastric epithelial cells. J Cell Physiol **214:**582-7.

216. **Mapstone, N. P., D. A. Lynch, F. A. Lewis, A. T. Axon, D. S. Tompkins, M. F. Dixon, and P. Quirke.** 1993. PCR identification of *Helicobacter pylori* in faeces from gastritis patients. Lancet **341:**447.

217. **Marshall, B. J., J. A. Armstrong, D. B. McGechie, and R. J. Glancy.** 1985. Attempt to fulfil Koch's postulates for pyloric *Campylobacter*. Med J Aust **142:**436-9.

218. **Marshall, B. J., D. B. McGechie, P. A. Rogers, and R. J. Glancy.** 1985. Pyloric *Campylobacter* infection and gastroduodenal disease. Med J Aust **142:**439-44.

219. **Marshall, B. J., and J. R. Warren.** 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet **1:**1311-5.

220. **Mashima, H., J. Suzuki, T. Hirayama, Y. Yoshikumi, H. Ohno, H. Ohnishi, H. Yasuda, T. Fujita, and M. Omata.** 2008. Involvement of vesicle-associated membrane protein 7 in human gastric epithelial cell vacuolation induced by *Helicobacter pylori*-produced VacA. Infect Immun **76:**2296-303.

221. **McClain, M. S., P. Cao, and T. L. Cover.** 2001. Amino-terminal hydrophobic region of *Helicobacter pylori* vacuolating cytotoxin (VacA) mediates transmembrane protein dimerization. Infect Immun **69:**1181-4.

222. **McClain, M. S., P. Cao, H. Iwamoto, A. D. Vinion-Dubiel, G. Szabo, Z. Shao, and T. L. Cover.** 2001. A 12-amino-acid segment, present in type s2 but not type s1 *Helicobacter pylori* VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol **183:**6499-508.

223. **McClain, M. S., H. Iwamoto, P. Cao, A. D. Vinion-Dubiel, Y. Li, G. Szabo, Z. Shao, and T. L. Cover.** 2003. Essential role of a GXXXG motif for membrane

channel formation by *Helicobacter pylori* vacuolating toxin. J Biol Chem **278:**12101-8.

224. **Megraud, F.** 1995. Transmission of *Helicobacter pylori*: faecal-oral versus oral-oral route. Aliment Pharmacol Ther **9 Suppl 2:**85-91.

225. **Megraud, F., M. P. Brassens-Rabbe, F. Denis, A. Belbouri, and D. Q. Hoa.** 1989. Seroepidemiology of *Campylobacter pylori* infection in various populations. J Clin Microbiol **27:**1870-3.

226. **Mimuro, H., T. Suzuki, S. Nagai, G. Rieder, M. Suzuki, T. Nagai, Y. Fujita, K. Nagamatsu, N. Ishijima, S. Koyasu, R. Haas, and C. Sasakawa.** 2007. *Helicobacter pylori* dampens gut epithelial self-renewal by inhibiting apoptosis, a bacterial strategy to enhance colonization of the stomach. Cell Host Microbe **2:**250-63.

227. **Mimuro, H., T. Suzuki, J. Tanaka, M. Asahi, R. Haas, and C. Sasakawa.** 2002. Grb2 is a key mediator of *Helicobacter pylori* CagA protein activities. Mol Cell **10:**745-55.

228. **Mitchell, H. M., T. Bohane, R. A. Hawkes, and A. Lee.** 1993. *Helicobacter pylori* infection within families. Zentralbl Bakteriol **280:**128-36.

229. **Mizoshita, T., K. Inada, T. Tsukamoto, K. Nozaki, T. Joh, M. Itoh, Y. Yamamura, T. Ushijima, S. Nakamura, and M. Tatematsu.** 2004. Expression of the intestine-specific transcription factors, Cdx1 and Cdx2, correlates shift to an intestinal phenotype in gastric cancer cells. J Cancer Res Clin Oncol **130:**29-36.

230. **Mizushima, T., T. Sugiyama, Y. Komatsu, J. Ishizuka, M. Kato, and M. Asaka.** 2001. Clinical relevance of the *babA2* genotype of *Helicobacter pylori* in Japanese clinical isolates. J Clin Microbiol **39:**2463-5.

231. **Mobley, H. L., M. J. Cortesia, L. E. Rosenthal, and B. D. Jones.** 1988. Characterization of urease from *Campylobacter pylori*. J Clin Microbiol **26:**831-6.

232. **Moese, S., M. Selbach, V. Brinkmann, A. Karlas, B. Haimovich, S. Backert, and T. F. Meyer.** 2007. The *Helicobacter pylori* CagA protein disrupts matrix adhesion of gastric epithelial cells by dephosphorylation of vinculin. Cell Microbiol **9:**1148-61.

233. **Moese, S., M. Selbach, T. Kwok, V. Brinkmann, W. Konig, T. F. Meyer, and S. Backert.** 2004. *Helicobacter pylori* induces AGS cell motility and elongation via independent signaling pathways. Infect Immun **72:**3646-9.

234. **Molinari, M., C. Galli, M. de Bernard, N. Norais, J. M. Ruysschaert, R. Rappuoli, and C. Montecucco.** 1998. The acid activation of *Helicobacter pylori* toxin VacA: structural and membrane binding studies. Biochem Biophys Res Commun **248:**334-40.

235. **Molinari, M., C. Galli, N. Norais, J. L. Telford, R. Rappuoli, J. P. Luzio, and C. Montecucco.** 1997. Vacuoles induced by *Helicobacter pylori* toxin contain both late endosomal and lysosomal markers. J Biol Chem **272:**25339-44.

236. **Molinari, M., M. Salio, C. Galli, N. Norais, R. Rappuoli, A. Lanzavecchia, and C. Montecucco.** 1998. Selective inhibition of Ii-dependent antigen presentation by *Helicobacter pylori* toxin VacA. J Exp Med **187:**135-40.

237. **Montecucco, C., and R. Rappuoli.** 2001. Living dangerously: how *Helicobacter pylori* survives in the human stomach. Nat Rev Mol Cell Biol **2:**457-66.

238. **Morihara, F., R. Fujii, E. Hifumi, A. Nishizono, and T. Uda.** 2007. Effects of vaccination by a recombinant antigen ureB138 (a segment of the beta-subunit of urease) against *Helicobacter pylori* infection. J Med Microbiol **56:**847-53.

239. **Moss, S. F., S. Legon, A. E. Bishop, J. M. Polak, and J. Calam.** 1992. Effect of *Helicobacter pylori* on gastric somatostatin in duodenal ulcer disease. Lancet **340:**930-2.

240. **Murata-Kamiya, N., Y. Kurashima, Y. Teishikata, Y. Yamahashi, Y. Saito, H. Higashi, H. Aburatani, T. Akiyama, R. M. Peek, Jr., T. Azuma, and M. Hatakeyama.** 2007. *Helicobacter pylori* CagA interacts with E-cadherin and deregulates the beta-catenin signal that promotes intestinal transdifferentiation in gastric epithelial cells. Oncogene **26:**4617-26.

241. **Nagy, T. A., M. R. Frey, F. Yan, D. A. Israel, D. B. Polk, and R. M. Peek, Jr.** 2009. *Helicobacter pylori* regulates cellular migration and apoptosis by activation of phosphatidylinositol 3-kinase signaling. J Infect Dis **199:**641-51.

242. **Naito, M., T. Yamazaki, R. Tsutsumi, H. Higashi, K. Onoe, S. Yamazaki, T. Azuma, and M. Hatakeyama.** 2006. Influence of EPIYA-repeat polymorphism on the phosphorylation-dependent biological activity of *Helicobacter pylori* CagA. Gastroenterology **130:**1181-90.

243. **Nakayama, M., J. Hisatsune, E. Yamasaki, H. Isomoto, H. Kurazono, M. Hatakeyama, T. Azuma, Y. Yamaoka, K. Yahiro, J. Moss, and T. Hirayama.** 2009. *Helicobacter pylori* VacA-induced inhibition of GSK3 through the PI3K/Akt signaling pathway. J Biol Chem **284:**1612-9.

244. **Nakayama, M., M. Kimura, A. Wada, K. Yahiro, K. Ogushi, T. Niidome, A. Fujikawa, D. Shirasaka, N. Aoyama, H. Kurazono, M. Noda, J. Moss, and T.**

Hirayama. 2004. *Helicobacter pylori* VacA activates the p38/activating transcription factor 2-mediated signal pathway in AZ-521 cells. J Biol Chem **279:**7024-8.

245. **Namavar, F., R. Roosendaal, E. J. Kuipers, P. de Groot, M. W. van der Bijl, A. S. Pena, and J. de Graaff.** 1995. Presence of *Helicobacter pylori* in the oral cavity, oesophagus, stomach and faeces of patients with gastritis. Eur J Clin Microbiol Infect Dis **14:**234-7.

246. **Neel, B. G., H. Gu, and L. Pao.** 2003. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends Biochem Sci **28:**284-93.

247. **Nesic, D., M. C. Miller, Z. T. Quinkert, M. Stein, B. T. Chait, and C. E. Stebbins.** 2010. *Helicobacter pylori* CagA inhibits PAR1-MARK family kinases by mimicking host substrates. Nat Struct Mol Biol **17:**130-2.

248. **Neugut, A. I., M. Hayek, and G. Howe.** 1996. Epidemiology of gastric cancer. Semin Oncol **23:**281-91.

249. **Nguyen, A. M., F. A. el-Zaatari, and D. Y. Graham.** 1995. *Helicobacter pylori* in the oral cavity. A critical review of the literature. Oral Surg Oral Med Oral Pathol Oral Radiol Endod **79:**705-9.

250. **Nguyen, L. T., T. Uchida, K. Murakami, T. Fujioka, and M. Moriyama.** 2008. *Helicobacter pylori* virulence and the diversity of gastric cancer in Asia. J Med Microbiol **57:**1445-53.

251. **Nguyen, L. T., T. Uchida, Y. Tsukamoto, A. Kuroda, T. Okimoto, M. Kodama, K. Murakami, T. Fujioka, and M. Moriyama.** 2009. *Helicobacter pylori dupA* gene is not associated with clinical outcomes in the Japanese population. Clin Microbiol Infect **16:**1264-1269.

252. **Nguyen, V. Q., R. M. Caprioli, and T. L. Cover.** 2001. Carboxy-terminal proteolytic processing of *Helicobacter pylori* vacuolating toxin. Infect Immun **69:**543-6.

253. **Niv, Y., and R. Hazazi.** 2008. *Helicobacter pylori* recurrence in developed and developing countries: meta-analysis of 13C-urea breath test follow-up after eradication. Helicobacter **13:**56-61.

254. **Odenbreit, S., K. Swoboda, I. Barwig, S. Ruhl, T. Boren, S. Koletzko, and R. Haas.** 2009. Outer membrane protein expression profile in *Helicobacter pylori* clinical isolates. Infect Immun **77:**3782-90.

255. **Ogura, K., S. Maeda, M. Nakao, T. Watanabe, M. Tada, T. Kyutoku, H. Yoshida, Y. Shiratori, and M. Omata.** 2000. Virulence factors of *Helicobacter pylori* responsible for gastric diseases in Mongolian gerbil. J Exp Med **192:**1601-10.

256. **Oh, J. D., S. M. Karam, and J. I. Gordon.** 2005. Intracellular *Helicobacter pylori* in gastric epithelial progenitors. Proc Natl Acad Sci U S A **102:**5186-91.

257. **Oh, J. D., H. Kling-Backhed, M. Giannakis, J. Xu, R. S. Fulton, L. A. Fulton, H. S. Cordum, C. Wang, G. Elliott, J. Edwards, E. R. Mardis, L. G. Engstrand, and J. I. Gordon.** 2006. The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. Proc Natl Acad Sci U S A **103:**9999-10004.

258. **Ohmae, T., Y. Hirata, S. Maeda, W. Shibata, A. Yanai, K. Ogura, H. Yoshida, T. Kawabe, and M. Omata.** 2005. *Helicobacter pylori* activates NF-kappaB via the alternative pathway in B lymphocytes. J Immunol **175:**7162-9.

259. **Oldani, A., M. Cormont, V. Hofman, V. Chiozzi, O. Oregioni, A. Canonici, A. Sciullo, P. Sommi, A. Fabbri, V. Ricci, and P. Boquet.** 2009. *Helicobacter pylori* counteracts the apoptotic action of its VacA toxin by injecting the CagA protein into gastric epithelial cells. PLoS Pathog **5:**e1000603.

260. **Oleastro, M., R. Cordeiro, Y. Yamaoka, D. Queiroz, F. Megraud, L. Monteiro, and A. Menard.** 2009. Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. Gut Pathog **1:**12.

261. **Oleastro, M., L. Monteiro, P. Lehours, F. Megraud, and A. Menard.** 2006. Identification of markers for *Helicobacter pylori* strains isolated from children with peptic ulcer disease by suppressive subtractive hybridization. Infect Immun **74:**4064-74.

262. **Oliveira, M. J., A. M. Costa, A. C. Costa, R. M. Ferreira, P. Sampaio, J. C. Machado, R. Seruca, M. Mareel, and C. Figueiredo.** 2009. CagA associates with c-Met, E-cadherin, and p120-catenin in a multiproteic complex that suppresses *Helicobacter pylori*-induced cell-invasive phenotype. J Infect Dis **200:**745-55.

263. **Ono, K., and J. Han.** 2000. The p38 signal transduction pathway: activation and function. Cell Signal **12:**1-13.

264. **Owen, R. J.** 1998. *Helicobacter*--species classification and identification. Br Med Bull **54:**17-30.

265. **Padilla, P. I., A. Wada, K. Yahiro, M. Kimura, T. Niidome, H. Aoyagi, A. Kumatori, M. Anami, T. Hayashi, J. Fujisawa, H. Saito, J. Moss, and T. Hirayama.** 2000. Morphologic differentiation of HL-60 cells is associated with appearance of RPTPbeta and induction of *Helicobacter pylori* VacA sensitivity. J Biol Chem **275:**15200-6.

266. **Pagliaccia, C., M. de Bernard, P. Lupetti, X. Ji, D. Burroni, T. L. Cover, E. Papini, R. Rappuoli, J. L. Telford, and J. M. Reyrat.** 1998. The m2 form of the *Helicobacter pylori* cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A **95:**10212-7.

267. **Pai, R., T. L. Cover, and A. S. Tarnawski.** 1999. *Helicobacter pylori* vacuolating cytotoxin (VacA) disorganizes the cytoskeletal architecture of gastric epithelial cells. Biochem Biophys Res Commun **262:**245-50.

268. **Papini, E., B. Satin, N. Norais, M. de Bernard, J. L. Telford, R. Rappuoli, and C. Montecucco.** 1998. Selective increase of the permeability of polarized epithelial cell monolayers by *Helicobacter pylori* vacuolating toxin. J Clin Invest **102:**813-20.

269. **Parente, J. M., B. B. da Silva, M. P. Palha-Dias, S. Zaterka, N. F. Nishimura, and J. M. Zeitune.** 2006. *Helicobacter pylori* infection in children of low and

high socioeconomic status in northeastern Brazil. Am J Trop Med Hyg **75:**509-12.

270. **Park, B. G., C. I. Yoo, H. T. Kim, C. H. Kwon, and Y. K. Kim.** 2005. Role of mitogen-activated protein kinases in hydrogen peroxide-induced cell death in osteoblastic cells. Toxicology **215:**115-25.

271. **Park, K. S., S. H. Jeon, J. W. Oh, and K. Y. Choi.** 2004. p21Cip/WAF1 activation is an important factor for the ERK pathway dependent anti-proliferation of colorectal cancer cells. Exp Mol Med **36:**557-62.

272. **Parkin, D. M., F. Bray, J. Ferlay, and P. Pisani.** 2005. Global cancer statistics, 2002. CA Cancer J Clin **55:**74-108.

273. **Parsonnet, J.** 1995. The incidence of *Helicobacter pylori* infection. Aliment Pharmacol Ther **9 Suppl 2:**45-51.

274. **Parsonnet, J., G. D. Friedman, N. Orentreich, and H. Vogelman.** 1997. Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. Gut **40:**297-301.

275. **Parsonnet, J., G. D. Friedman, D. P. Vandersteen, Y. Chang, J. H. Vogelman, N. Orentreich, and R. K. Sibley.** 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med **325:**1127-31.

276. **Parsonnet, J., S. Hansen, L. Rodriguez, A. B. Gelb, R. A. Warnke, E. Jellum, N. Orentreich, J. H. Vogelman, and G. D. Friedman.** 1994. *Helicobacter pylori* infection and gastric lymphoma. N Engl J Med **330:**1267-71.

277. **Parsonnet, J., H. Shmuely, and T. Haggerty.** 1999. Fecal and oral shedding of *Helicobacter pylori* from healthy infected adults. Jama **282:**2240-5.

278. **Peek, R. M., Jr., M. J. Blaser, D. J. Mays, M. H. Forsyth, T. L. Cover, S. Y. Song, U. Krishna, and J. A. Pietenpol.** 1999. *Helicobacter pylori* strain-specific genotypes and modulation of the gastric epithelial cell cycle. Cancer Res **59:**6124-31.

279. **Peek, R. M., Jr., S. F. Moss, K. T. Tham, G. I. Perez-Perez, S. Wang, G. G. Miller, J. C. Atherton, P. R. Holt, and M. J. Blaser.** 1997. *Helicobacter pylori cagA*+ strains and dissociation of gastric epithelial cell proliferation from apoptosis. J Natl Cancer Inst **89:**863-8.

280. **Peek, R. M., Jr., H. P. Wirth, S. F. Moss, M. Yang, A. M. Abdalla, K. T. Tham, T. Zhang, L. H. Tang, I. M. Modlin, and M. J. Blaser.** 2000.

Helicobacter pylori alters gastric epithelial cell cycle events and gastrin secretion in Mongolian gerbils. Gastroenterology **118:**48-59.

281. **Penfold, S. S., A. J. Lastovica, and B. G. Elisha.** 1988. Demonstration of plasmids in *Campylobacter pylori*. J Infect Dis **157:**850-1.

282. **Perry, S., M. de la Luz Sanchez, S. Yang, T. D. Haggerty, P. Hurst, G. Perez-Perez, and J. Parsonnet.** 2006. Gastroenteritis and transmission of *Helicobacter pylori* infection in households. Emerg Infect Dis **12:**1701-8.

283. **Phadnis, S. H., D. Ilver, L. Janzon, S. Normark, and T. U. Westblom.** 1994. Pathological significance and molecular characterization of the vacuolating toxin gene of *Helicobacter pylori*. Infect Immun **62:**1557-65.

284. **Poppe, M., S. M. Feller, G. Romer, and S. Wessler.** 2007. Phosphorylation of *Helicobacter pylori* CagA by c-Abl leads to cell motility. Oncogene **26:**3462-72.

285. **Pounder, R. E., and D. Ng.** 1995. The prevalence of *Helicobacter pylori* infection in different countries. Aliment Pharmacol Ther **9 Suppl 2:**33-9.

286. **Pugsley, A. P.** 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiol Rev **57:**50-108.

287. **Queiroz, D. M., G. A. Rocha, A. M. Rocha, S. B. Moura, I. E. Saraiva, L. I. Gomes, T. F. Soares, F. F. Melo, M. M. Cabral, and C. A. Oliveira.** 2010. *dupA* polymorphisms and risk of *Helicobacter pylor*i-associated diseases. Int J Med Microbiol.

288. **Rad, R., M. Gerhard, R. Lang, M. Schoniger, T. Rosch, W. Schepp, I. Becker, H. Wagner, and C. Prinz.** 2002. The *Helicobacter pylori* blood group antigen-binding adhesin facilitates bacterial colonization and augments a nonspecific immune response. J Immunol **168:**3033-41.

289. **Raghavan, S., M. Hjulstrom, J. Holmgren, and A. M. Svennerholm.** 2002. Protection against experimental *Helicobacter pylori* infection after immunization with inactivated *H. pylori* whole-cell vaccines. Infect Immun **70:**6383-8.

290. **Rao, A.** 1994. NF-ATp: a transcription factor required for the co-ordinate induction of several cytokine genes. Immunol Today **15:**274-81.

291. **Ren, S., H. Higashi, H. Lu, T. Azuma, and M. Hatakeyama.** 2006. Structural basis and functional consequence of *Helicobacter pylori* CagA multimerization in cells. J Biol Chem **281:**32344-52.

292. **Reyrat, J. M., S. Lanzavecchia, P. Lupetti, M. de Bernard, C. Pagliaccia, V. Pelicic, M. Charrel, C. Ulivieri, N. Norais, X. Ji, V. Cabiaux, E. Papini, R.**

Rappuoli, and J. L. Telford. 1999. 3D imaging of the 58 kDa cell binding subunit of the *Helicobacter pylori* cytotoxin. J Mol Biol **290:**459-70.

293. **Rhead, J. L., D. P. Letley, M. Mohammadi, N. Hussein, M. A. Mohagheghi, M. Eshagh Hosseini, and J. C. Atherton.** 2007. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology **133:**926-36.

294. **Rieder, G., W. Fischer, and R. Haas.** 2005. Interaction of *Helicobacter pylori* with host cells: function of secreted and translocated molecules. Curr Opin Microbiol **8:**67-73.

295. **Rieder, G., J. L. Merchant, and R. Haas.** 2005. *Helicobacter pylori cag*-type IV secretion system facilitates corpus colonization to induce precancerous conditions in Mongolian gerbils. Gastroenterology **128:**1229-42.

296. **Rittig, M. G., B. Shaw, D. P. Letley, R. J. Thomas, R. H. Argent, and J. C. Atherton.** 2003. H*elicobacter pylori*-induced homotypic phagosome fusion in human monocytes is independent of the bacterial *vacA* and *cag* status. Cell Microbiol **5:**887-99.

297. **Roder, D. M.** 2002. The epidemiology of gastric cancer. Gastric Cancer **5 Suppl 1:**5-11.

298. **Roosendaal, R., E. J. Kuipers, J. Buitenwerf, C. van Uffelen, S. G. Meuwissen, G. J. van Kamp, and C. M. Vandenbroucke-Grauls.** 1997. *Helicobacter pylori* and the birth cohort effect: evidence of a continuous decrease of infection rates in childhood. Am J Gastroenterol **92:**1480-2.

299. **Roovers, K., and R. K. Assoian.** 2000. Integrating the MAP kinase signal into the G1 phase cell cycle machinery. Bioessays **22:**818-26.

300. **Rothenbacher, D., M. Winkler, T. Gonser, G. Adler, and H. Brenner.** 2002. Role of infected parents in transmission of *Helicobacter pylori* to their children. Pediatr Infect Dis J **21:**674-9.

301. **Saadat, I., H. Higashi, C. Obuse, M. Umeda, N. Murata-Kamiya, Y. Saito, H. Lu, N. Ohnishi, T. Azuma, A. Suzuki, S. Ohno, and M. Hatakeyama.** 2007. *Helicobacter pylori* CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. Nature **447:**330-3.

302. **Salih, B. A., M. F. Abasiyanik, N. Bayyurt, and E. Sander.** 2007. *H pylori* infection and other risk factors associated with peptic ulcers in Turkish patients: a retrospective study. World J Gastroenterol **13:**3245-8.

303. **Sathar, M. A., E. Gouws, A. E. Simjee, and A. M. Mayat.** 1997. Seroepidemiological study of *Helicobacter pylori* infection in South African children. Trans R Soc Trop Med Hyg **91:**393-5.

304. **Satin, B., G. Del Giudice, V. Della Bianca, S. Dusi, C. Laudanna, F. Tonello, D. Kelleher, R. Rappuoli, C. Montecucco, and F. Rossi.** 2000. The neutrophil-activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor. J Exp Med **191:**1467-76.

305. **Satomi, S., A. Yamakawa, S. Matsunaga, R. Masaki, T. Inagaki, T. Okuda, H. Suto, Y. Ito, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Kutsumi, and T. Azuma.** 2006. Relationship between the diversity of the *cagA* gene of *Helicobacter pylori* and gastric cancer in Okinawa, Japan. J Gastroenterol **41:**668-73.

306. **Schmidt, H. M., S. Andres, N. O. Kaakoush, L. Engstrand, L. Eriksson, K. L. Goh, K. M. Fock, I. Hilmi, S. Dhamodaran, D. Forman, and H. Mitchell.** 2009. The prevalence of the duodenal ulcer promoting gene (*dupA*) in *Helicobacter pylori* isolates varies by ethnic group and is not universally associated with disease development: a case-control study. Gut Pathog **1:**5.

307. **Schweitzer, K., O. Sokolova, P. M. Bozko, and M. Naumann.** 2010. *Helicobacter pylori* induces NF-kappaB independent of CagA. EMBO Rep **11:**10-1; author reply 11-2.

308. **Segal, E. D., J. Cha, J. Lo, S. Falkow, and L. S. Tompkins.** 1999. Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*. Proc Natl Acad Sci U S A **96:**14559-64.

309. **Selbach, M., and S. Backert.** 2005. Cortactin: an Achilles' heel of the actin cytoskeleton targeted by pathogens. Trends Microbiol **13:**181-9.

310. **Selbach, M., S. Moese, S. Backert, P. R. Jungblut, and T. F. Meyer.** 2004. The *Helicobacter pylori* CagA protein induces tyrosine dephosphorylation of ezrin. Proteomics **4:**2961-8.

311. **Selbach, M., S. Moese, C. R. Hauck, T. F. Meyer, and S. Backert.** 2002. Src is the kinase of the *Helicobacter pylori* CagA protein *in vitro* and *in vivo*. J Biol Chem **277:**6775-8.

312. **Selbach, M., S. Moese, R. Hurwitz, C. R. Hauck, T. F. Meyer, and S. Backert.** 2003. The *Helicobacter pylori* CagA protein induces cortactin dephosphorylation and actin rearrangement by c-Src inactivation. Embo J **22:**515-28.

313. **Selbach, M., F. E. Paul, S. Brandt, P. Guye, O. Daumke, S. Backert, C. Dehio, and M. Mann.** 2009. Host cell interactome of tyrosine-phosphorylated bacterial proteins. Cell Host Microbe **5:**397-403.

314. **Semino-Mora, C., S. Q. Doi, A. Marty, V. Simko, I. Carlstedt, and A. Dubois.** 2003. Intracellular and interstitial expression of *Helicobacter pylori* virulence genes in gastric precancerous intestinal metaplasia and adenocarcinoma. J Infect Dis **187:**1165-77.

315. **Sewald, X., W. Fischer, and R. Haas.** 2008. Sticky socks: *Helicobacter pylori* VacA takes shape. Trends Microbiol **16:**89-92.

316. **Sewald, X., B. Gebert-Vogl, S. Prassl, I. Barwig, E. Weiss, M. Fabbri, R. Osicka, M. Schiemann, D. H. Busch, M. Semmrich, B. Holzmann, P. Sebo, and R. Haas.** 2008. Integrin subunit CD18 Is the T-lymphocyte receptor for the *Helicobacter pylori* vacuolating cytotoxin. Cell Host Microbe **3:**20-9.

317. **Seyler, R. W., Jr., J. W. Olson, and R. J. Maier.** 2001. Superoxide dismutase-deficient mutants of *Helicobacter pylori* are hypersensitive to oxidative stress and defective in host colonization. Infect Immun **69:**4034-40.

318. **Shames, B., S. Krajden, M. Fuksa, C. Babida, and J. L. Penner.** 1989. Evidence for the occurrence of the same strain of *Campylobacter pylori* in the stomach and dental plaque. J Clin Microbiol **27:**2849-50.

319. **Sharma, S. A., M. K. Tummuru, M. J. Blaser, and L. D. Kerr.** 1998. Activation of IL-8 gene expression by *Helicobacter pylori* is regulated by transcription factor nuclear factor-kappa B in gastric epithelial cells. J Immunol **160:**2401-7.

320. **Shibata, W., Y. Hirata, S. Maeda, K. Ogura, T. Ohmae, A. Yanai, Y. Mitsuno, Y. Yamaji, M. Okamoto, H. Yoshida, T. Kawabe, and M. Omata.** 2006. CagA protein secreted by the intact type IV secretion system leads to gastric epithelial inflammation in the Mongolian gerbil model. J Pathol **210:**306-14.

321. **Shikata, K., Y. Doi, K. Yonemoto, H. Arima, T. Ninomiya, M. Kubo, Y. Tanizaki, T. Matsumoto, M. Iida, and Y. Kiyohara.** 2008. Population-based prospective study of the combined influence of cigarette smoking and *Helicobacter pylori* infection on gastric cancer incidence: the Hisayama Study. Am J Epidemiol **168:**1409-15.

322. **Shin, A., S. Park, H. R. Shin, E. H. Park, S. K. Park, J. K. Oh, M. K. Lim, B. Y. Choi, M. Boniol, and P. Boffetta.** Population attributable fraction of infection-related cancers in Korea. Ann Oncol.

323. **Shin, A., H. R. Shin, D. Kang, S. K. Park, C. S. Kim, and K. Y. Yoo.** 2005. A nested case-control study of the association of *Helicobacter pylori* infection with gastric adenocarcinoma in Korea. Br J Cancer **92:**1273-5.

324. **Siman, J. H., A. Forsgren, G. Berglund, and C. H. Floren.** 2001. Tobacco smoking increases the risk for gastric adenocarcinoma among *Helicobacter pylori*-infected individuals. Scand J Gastroenterol **36:**208-13.

325. **Singh, V., B. Trikha, K. Vaiphei, C. K. Nain, K. Thennarasu, and K. Singh.** 1999. *Helicobacter pylori*: evidence for spouse-to-spouse transmission. J Gastroenterol Hepatol **14:**519-22.

326. **Sitas, F., D. Forman, J. W. Yarnell, M. L. Burr, P. C. Elwood, S. Pedley, and K. J. Marks.** 1991. *Helicobacter pylori* infection rates in relation to age and social class in a population of Welsh men. Gut **32:**25-8.

327. **Skene, C., A. Young, A. Every, and P. Sutton.** 2007. *Helicobacter pylori* flagella: antigenic profile and protective immunity. FEMS Immunol Med Microbiol **50:**249-56.

328. **Smith, A. W., B. Chahal, and G. L. French.** 1994. The human gastric pathogen *Helicobacter pylori* has a gene encoding an enzyme first classified as a mucinase in *Vibrio cholerae*. Mol Microbiol **13:**153-60.

329. **Sokolova, O., P. M. Bozko, and M. Naumann.** 2008. *Helicobacter pylori* suppresses glycogen synthase kinase 3beta to promote beta-catenin activity. J Biol Chem **283:**29367-74.

330. **Solnick, J. V., and P. Vandamme.** 2001. Taxonomy of the *Helicobacter* genus. *In* H. L. T. Mobley, G. L. Mendz, and H. S.L. (ed.), *Helicobacter pylori* Physiology and Genetics. ASM Press, Washington D.C.

331. **Stanner, S. A., J. Hughes, C. N. Kelly, and J. Buttriss.** 2004. A review of the epidemiological evidence for the 'antioxidant hypothesis'. Public Health Nutr **7:**407-22.

332. **Steevens, J., L. J. Schouten, R. A. Goldbohm, and P. A. van den Brandt.** 2010. Alcohol consumption, cigarette smoking and risk of subtypes of oesophageal and gastric cancer: a prospective cohort study. Gut **59:**39-48.

333. **Stein, M., F. Bagnoli, R. Halenbeck, R. Rappuoli, W. J. Fantl, and A. Covacci.** 2002. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. Mol Microbiol **43:**971-80.

334. **Stein, M., R. Rappuoli, and A. Covacci.** 2000. Tyrosine phosphorylation of the *Helicobacter pylori* CagA antigen after cag-driven host cell translocation. Proc Natl Acad Sci U S A **97:**1263-8.

335. **Suerbaum, S., and C. Josenhans.** 2007. *Helicobacter pylori* evolution and phenotypic diversification in a changing host. Nat Rev Microbiol **5:**441-52.

336. **Suerbaum, S., J. M. Smith, K. Bapumia, G. Morelli, N. H. Smith, E. Kunstmann, I. Dyrek, and M. Achtman.** 1998. Free recombination within *Helicobacter pylori*. Proc Natl Acad Sci U S A **95:**12619-24.

337. **Sun, L., L. Deng, C. K. Ea, Z. P. Xia, and Z. J. Chen.** 2004. The TRAF6 ubiquitin ligase and TAK1 kinase mediate IKK activation by BCL10 and MALT1 in T lymphocytes. Mol Cell **14:**289-301.

338. **Sundrud, M. S., V. J. Torres, D. Unutmaz, and T. L. Cover.** 2004. Inhibition of primary human T cell proliferation by *Helicobacter pylori* vacuolating toxin (VacA) is independent of VacA effects on IL-2 secretion. Proc Natl Acad Sci U S A **101:**7727-32.

339. **Supajatura, V., H. Ushio, A. Wada, K. Yahiro, K. Okumura, H. Ogawa, T. Hirayama, and C. Ra.** 2002. Cutting edge: VacA, a vacuolating cytotoxin of *Helicobacter pylori*, directly activates mast cells for migration and production of proinflammatory cytokines. J Immunol **168:**2603-7.

340. **Sutton, P., C. Doidge, G. Pinczower, J. Wilson, S. Harbour, A. Swierczak, and A. Lee.** 2007. Effectiveness of vaccination with recombinant HpaA from *Helicobacter pylori* is influenced by host genetic background. FEMS Immunol Med Microbiol **50:**213-9.

341. **Suzuki, J., H. Ohnishi, A. Wada, T. Hirayama, H. Ohno, N. Ueda, H. Yasuda, T. Iiri, Y. Wada, M. Futai, and H. Mashima.** 2003. Involvement of syntaxin 7 in human gastric epithelial cell vacuolation induced by the *Helicobacter pylori*-produced cytotoxin VacA. J Biol Chem **278:**25585-90.

342. **Suzuki, J., H. Ohnsihi, H. Shibata, A. Wada, T. Hirayama, T. Iiri, N. Ueda, C. Kanamaru, T. Tsuchida, H. Mashima, H. Yasuda, and T. Fujita.** 2001. Dynamin is involved in human epithelial cell vacuolation caused by the *Helicobacter pylori*-produced cytotoxin VacA. J Clin Invest **107:**363-70.

343. **Suzuki, M., H. Mimuro, K. Kiga, M. Fukumatsu, N. Ishijima, H. Morikawa, S. Nagai, S. Koyasu, R. H. Gilman, D. Kersulyte, D. E. Berg, and C.**

Sasakawa. 2009. *Helicobacter pylori* CagA phosphorylation-independent function in epithelial proliferation and inflammation. Cell Host Microbe **5:**23-34.

344. **Suzuki, M., H. Mimuro, T. Suzuki, M. Park, T. Yamamoto, and C. Sasakawa.** 2005. Interaction of CagA with Crk plays an important role in *Helicobacter pylori*-induced loss of gastric epithelial cell adhesion. J Exp Med **202:**1235-47.

345. **Sykora, J., K. Siala, J. Varvarovska, P. Pazdiora, R. Pomahacova, and M. Huml.** 2009. Epidemiology of *Helicobacter pylori* infection in asymptomatic children: a prospective population-based study from the Czech Republic. Application of a monoclonal-based antigen-in-stool enzyme immunoassay. Helicobacter **14:**286-97.

346. **Szabo, I., S. Brutsche, F. Tombola, M. Moschioni, B. Satin, J. L. Telford, R. Rappuoli, C. Montecucco, E. Papini, and M. Zoratti.** 1999. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. EMBO J **18:**5517-27.

347. **Tabassam, F. H., D. Y. Graham, and Y. Yamaoka.** 2009. *Helicobacter pylori* activate epidermal growth factor receptor- and phosphatidylinositol 3-OH kinase-dependent Akt and glycogen synthase kinase 3beta phosphorylation. Cell Microbiol **11:**70-82.

348. **Takeshima, E., K. Tomimori, H. Kawakami, C. Ishikawa, S. Sawada, M. Tomita, M. Senba, F. Kinjo, H. Mimuro, C. Sasakawa, J. Fujita, and N. Mori.** 2009. NF-kappaB activation by *Helicobacter pylori* requires Akt-mediated phosphorylation of p65. BMC Microbiol **9:**36.

349. **Talley, N. J., A. R. Zinsmeister, A. Weaver, E. P. DiMagno, H. A. Carpenter, G. I. Perez-Perez, and M. J. Blaser.** 1991. Gastric adenocarcinoma and *Helicobacter pylori* infection. J Natl Cancer Inst **83:**1734-9.

350. **Tammer, I., S. Brandt, R. Hartig, W. Konig, and S. Backert.** 2007. Activation of Abl by *Helicobacter pylori*: a novel kinase for CagA and crucial mediator of host cell scattering. Gastroenterology **132:**1309-19.

351. **Tartaglia, M., C. M. Niemeyer, A. Fragale, X. Song, J. Buechner, A. Jung, K. Hahlen, H. Hasle, J. D. Licht, and B. D. Gelb.** 2003. Somatic mutations in *PTPN11* in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet **34:**148-50.

352. **Taylor, D. E., M. Eaton, N. Chang, and S. M. Salama.** 1992. Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. J Bacteriol **174:**6800-6.

353. **Tegtmeyer, N., D. Zabler, D. Schmidt, R. Hartig, S. Brandt, and S. Backert.** 2009. Importance of EGF receptor, HER2/Neu and Erk1/2 kinase signalling for host cell elongation and scattering induced by the *Helicobacter pylori* CagA protein: antagonistic effects of the vacuolating cytotoxin VacA. Cell Microbiol **11:**488-505.

354. **Telford, J. L., P. Ghiara, M. Dell'Orco, M. Comanducci, D. Burroni, M. Bugnoli, M. F. Tecce, S. Censini, A. Covacci, Z. Xiang, E. Papini, C. Montecucco, L. Parente, and R. Rappuoli.** 1994. Gene structure of the *Helicobacter pylori* cytotoxin and evidence of its key role in gastric disease. J Exp Med **179:**1653-58.

355. **Terebiznik, M. R., D. Raju, C. L. Vazquez, K. Torbricki, R. Kulkarni, S. R. Blanke, T. Yoshimori, M. I. Colombo, and N. L. Jones.** 2009. Effect of *Helicobacter pylori*'s vacuolating cytotoxin on the autophagy pathway in gastric epithelial cells. Autophagy **5:**370-9.

356. **Thomas, J. E., A. Dale, M. Harding, W. A. Coward, T. J. Cole, and L. T. Weaver.** 1999. *Helicobacter pylori* colonization in early life. Pediatr Res **45:**218-23.

357. **Tindberg, Y., C. Bengtsson, F. Granath, M. Blennow, O. Nyren, and M. Granstrom.** 2001. *Helicobacter pylori* infection in Swedish school children: lack

of evidence of child-to-child transmission outside the family. Gastroenterology **121:**310-6.

358. **Tokudome, S., R. Ando, R. Ghadimi, T. Tanaka, N. Hattori, Z. Yang, M. Marumoto, H. Agawa, K. Arakawa, Y. Osaka, H. Tanaka, A. Hosono, and M. A. Moore.** 2007. Are there any real *Helicobacter pylori* infection-negative gastric cancers in Asia? Asian Pac J Cancer Prev **8:**462-3.

359. **Tolwinski, N. S., and E. Wieschaus.** 2004. Rethinking WNT signaling. Trends Genet **20:**177-81.

360. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539-47.

361. **Tombola, F., L. Morbiato, G. Del Giudice, R. Rappuoli, M. Zoratti, and E. Papini.** 2001. The *Helicobacter pylori* VacA toxin is a urea permease that promotes urea diffusion across epithelia. J Clin Invest **108:**929-37.

362. **Torres, V. J., M. S. McClain, and T. L. Cover.** 2004. Interactions between p-33 and p-55 domains of the *Helicobacter pylori* vacuolating cytotoxin (VacA). J Biol Chem **279:**2324-31.

363. **Torres, V. J., S. E. VanCompernolle, M. S. Sundrud, D. Unutmaz, and T. L. Cover.** 2007. *Helicobacter pylori* vacuolating cytotoxin inhibits activation-induced proliferation of human T and B lymphocyte subsets. J Immunol **179:**5433-40.

364. **Tsugane, S., Y. Tei, T. Takahashi, S. Watanabe, and K. Sugano.** 1994. Salty food intake and risk of *Helicobacter pylori* infection. Jpn J Cancer Res **85:**474-8.

365. **Tsutsumi, R., H. Higashi, M. Higuchi, M. Okada, and M. Hatakeyama.** 2003. Attenuation of *Helicobacter pylor*i CagA-SHP-2 signaling by interaction between CagA and C-terminal Src kinase. J Biol Chem **278:**3664-70.

366. **Tsutsumi, R., A. Takahashi, T. Azuma, H. Higashi, and M. Hatakeyama.** 2006. Focal adhesion kinase is a substrate and downstream effector of SHP-2 complexed with *Helicobacter pylori* CagA. Mol Cell Biol **26:**261-76.

367. **Tummuru, M. K., T. L. Cover, and M. J. Blaser.** 1993. Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. Infect Immun **61:**1799-809.

368. **Tuo, B., P. Song, G. Wen, X. Sewald, B. Gebert-Vogl, R. Haas, M. Manns, and U. Seidler.** 2009. *Helicobacter pylori* vacuolating cytotoxin inhibits duodenal bicarbonate secretion by a histamine-dependent mechanism in mice. J Infect Dis **199:**505-12.

369. **Udhayakumar, G., V. Jayanthi, N. Devaraj, and H. Devaraj.** 2007. Interaction of MUC1 with beta-catenin modulates the Wnt target gene *cyclinD1* in *H. pylori*-induced gastric cancer. Mol Carcinog **46:**807-17.

370. **Umeda, M., N. Murata-Kamiya, Y. Saito, Y. Ohba, M. Takahashi, and M. Hatakeyama.** 2009. *Helicobacter pylori* CagA causes mitotic impairment and induces chromosomal instability. J Biol Chem **284:**22166-72.

371. **Umehara, S., H. Higashi, N. Ohnishi, M. Asaka, and M. Hatakeyama.** 2003. Effects of *Helicobacter pylori* CagA protein on the growth and survival of B lymphocytes, the origin of MALT lymphoma. Oncogene **22:**8337-42.

372. **Viala, J., C. Chaput, I. G. Boneca, A. Cardona, S. E. Girardin, A. P. Moran, R. Athman, S. Memet, M. R. Huerre, A. J. Coyle, P. S. DiStefano, P. J. Sansonetti, A. Labigne, J. Bertin, D. J. Philpott, and R. L. Ferrero.** 2004. Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori cag* pathogenicity island. Nat Immunol **5:**1166-74.

373. **Vincent, P.** 1995. Transmission and acquisition of *Helicobacter pylori* infection: evidences and hypothesis. Biomed Pharmacother **49:**11-8.

374. **Vinion-Dubiel, A. D., M. S. McClain, D. M. Czajkowsky, H. Iwamoto, D. Ye, P. Cao, W. Schraw, G. Szabo, S. R. Blanke, Z. Shao, and T. L. Cover.** 1999. A dominant negative mutant of *Helicobacter pylori* vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. J Biol Chem **274:**37736-42.

375. **Wang, F., P. Xia, F. Wu, D. Wang, W. Wang, T. Ward, Y. Liu, F. Aikhionbare, Z. Guo, M. Powell, B. Liu, F. Bi, A. Shaw, Z. Zhu, A. Elmoselhi, D. Fan, T. L. Cover, X. Ding, and X. Yao.** 2008. *Helicobacter pylori* VacA disrupts apical membrane-cytoskeletal interactions in gastric parietal cells. J Biol Chem **283:**26714-25.

376. **Wang, H. J., and W. C. Wang.** 2000. Expression and binding analysis of GST-VacA fusions reveals that the C-terminal approximately 100-residue segment of

exotoxin is crucial for binding in HeLa cells. Biochem Biophys Res Commun **278:**449-54.

377. **Wang, W. C., H. J. Wang, and C. H. Kuo.** 2001. Two distinctive cell binding patterns by vacuolating toxin fused with glutathione S-transferase: one high-affinity m1-specific binding and the other lower-affinity binding for variant m forms. Biochemistry **40:**11887-96.

378. **Watanabe, T., N. Asano, S. Fichtner-Feigl, P. L. Gorelick, Y. Tsuji, Y. Matsumoto, T. Chiba, I. J. Fuss, A. Kitani, and W. Strober.** 2010. NOD1 contributes to mouse host defense against *Helicobacter pylori* via induction of type I IFN and activation of the ISGF3 signaling pathway. J Clin Invest **120:**1645-62.

379. **Watanabe, T., M. Tada, H. Nagai, S. Sasaki, and M. Nakao.** 1998. *Helicobacter pylor*i infection induces gastric cancer in Mongolian gerbils. Gastroenterology **115:**642-8.

380. **Webb, P. M., T. Knight, D. G. Newell, J. B. Elder, and D. Forman.** 1996. *Helicobacter pylori* transmission: evidence from a comparison with hepatitis A virus. Eur J Gastroenterol Hepatol **8:**439-41.

381. **Weydig, C., A. Starzinski-Powitz, G. Carra, J. Lower, and S. Wessler.** 2007. CagA-independent disruption of adherence junction complexes involves E-cadherin shedding and implies multiple steps in *Helicobacter pylori* pathogenicity. Exp Cell Res **313:**3459-71.

382. **Whary, M. T., N. Sundina, L. E. Bravo, P. Correa, F. Quinones, F. Caro, and J. G. Fox.** 2005. Intestinal helminthiasis in Colombian children promotes a Th2 response to *Helicobacter pylori*: possible implications for gastric carcinogenesis. Cancer Epidemiol Biomarkers Prev **14:**1464-9.

383. **Willhite, D. C., and S. R. Blanke.** 2004. *Helicobacter pylori* vacuolating cytotoxin enters cells, localizes to the mitochondria, and induces mitochondrial membrane permeability changes correlated to toxin channel activity. Cell Microbiol **6:**143-54.

384. **Willhite, D. C., T. L. Cover, and S. R. Blanke.** 2003. Cellular vacuolation and mitochondrial cytochrome c release are independent outcomes of *Helicobacter pylori* vacuolating cytotoxin activity that are each dependent on membrane channel formation. J Biol Chem **278:**48204-9.

385. **Willhite, D. C., D. Ye, and S. R. Blanke.** 2002. Fluorescence resonance energy transfer microscopy of the *Helicobacter pylori* vacuolating cytotoxin within mammalian cells. Infect Immun **70:**3824-32.

386. **Wirth, H. P., M. H. Beins, M. Yang, K. T. Tham, and M. J. Blaser.** 1998. Experimental infection of Mongolian gerbils with wild-type and mutant *Helicobacter pylori* strains. Infect Immun **66:**4856-66.

387. **Wroblewski, L. E., R. M. Peek, Jr., and K. T. Wilson.** 2010. *Helicobacter pylori* and gastric cancer: factors that modulate disease risk. Clin Microbiol Rev **23:**713-39.

388. **Xiang, Z., S. Censini, P. F. Bayeli, J. L. Telford, N. Figura, R. Rappuoli, and A. Covacci.** 1995. Analysis of expression of CagA and VacA virulence factors in 43 strains of *Helicobacter pylori* reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. Infect Immun **63:**94-8.

389. **Yahiro, K., T. Niidome, M. Kimura, T. Hatakeyama, H. Aoyagi, H. Kurazono, K. Imagawa, A. Wada, J. Moss, and T. Hirayama.** 1999. Activation of *Helicobacter pylori* VacA toxin by alkaline or acid conditions increases its binding to a 250-kDa receptor protein-tyrosine phosphatase beta. J Biol Chem **274:**36693-9.

390. **Yahiro, K., A. Wada, M. Nakayama, T. Kimura, K. Ogushi, T. Niidome, H. Aoyagi, K. Yoshino, K. Yonezawa, J. Moss, and T. Hirayama.** 2003. Protein-

tyrosine phosphatase alpha, RPTP alpha, is a *Helicobacter pylori* VacA receptor. J Biol Chem **278:**19183-9.

391. **Yahiro, K., A. Wada, E. Yamasaki, M. Nakayama, Y. Nishi, J. Hisatsune, N. Morinaga, J. Sap, M. Noda, J. Moss, and T. Hirayama.** 2004. Essential domain of receptor tyrosine phosphatase beta (RPTPbeta) for interaction with *Helicobacter pylori* vacuolating cytotoxin. J Biol Chem **279:**51013-21.

392. **Yamamoto, S.** 2001. Stomach cancer incidence in the world. Jpn J Clin Oncol **31:**471.

393. **Yamaoka, Y.** 2009. *Helicobacter pylori* typing as a tool for tracking human migration. Clin Microbiol Infect **15:**829-34.

394. **Yamaoka, Y., H. M. El-Zimaity, O. Gutierrez, N. Figura, J. G. Kim, T. Kodama, K. Kashima, and D. Y. Graham.** 1999. Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. Gastroenterology **117:**342-9.

395. **Yamaoka, Y., S. Kikuchi, H. M. el-Zimaity, O. Gutierrez, M. S. Osato, and D. Y. Graham.** 2002. Importance of *Helicobacter pylori oipA* in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. Gastroenterology **123:**414-24.

396. **Yamaoka, Y., D. H. Kwon, and D. Y. Graham.** 2000. A M(r) 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. Proc Natl Acad Sci U S A **97:**7533-8.

397. **Yamasaki, E., A. Wada, A. Kumatori, I. Nakagawa, J. Funao, M. Nakayama, J. Hisatsune, M. Kimura, J. Moss, and T. Hirayama.** 2006. *Helicobacter pylori* vacuolating cytotoxin induces activation of the proapoptotic proteins Bax and Bak, leading to cytochrome c release and cell death, independent of vacuolation. J Biol Chem **281:**11250-9.

398. **Yamazaki, S., A. Yamakawa, Y. Ito, M. Ohtani, H. Higashi, M. Hatakeyama, and T. Azuma.** 2003. The CagA protein of *Helicobacter pylori* is translocated into epithelial cells and binds to SHP-2 in human gastric mucosa. J Infect Dis **187:**334-7.

399. **Yang, W. L., J. Wang, C. H. Chan, S. W. Lee, A. D. Campos, B. Lamothe, L. Hur, B. C. Grabiner, X. Lin, B. G. Darnay, and H. K. Lin.** 2009. The E3 ligase TRAF6 regulates Akt ubiquitination and activation. Science **325:**1134-8.

400. **Ye, D., and S. R. Blanke.** 2002. Functional complementation reveals the importance of intermolecular monomer interactions for *Helicobacter pylori* VacA vacuolating activity. Mol Microbiol **43:**1243-53.

401. **Ye, D., and S. R. Blanke.** 2000. Mutational analysis of the *Helicobacter pylori* vacuolating toxin amino terminus: identification of amino acids essential for cellular vacuolation. Infect Immun **68:**4354-7.

402. **Ye, D., D. C. Willhite, and S. R. Blanke.** 1999. Identification of the minimal intracellular vacuolating domain of the *Helicobacter pylori* vacuolating toxin. J Biol Chem **274:**9277-82.

403. **Yim, J. Y., N. Kim, S. H. Choi, Y. S. Kim, K. R. Cho, S. S. Kim, G. S. Seo, H. U. Kim, G. H. Baik, C. S. Sin, S. H. Cho, and B. H. Oh.** 2007. Seroprevalence of *Helicobacter pylori* in South Korea. Helicobacter **12:**333-40.

404. **Yokoyama, K., H. Higashi, S. Ishikawa, Y. Fujii, S. Kondo, H. Kato, T. Azuma, A. Wada, T. Hirayama, H. Aburatani, and M. Hatakeyama.** 2005. Functional antagonism between *Helicobacter pylori* CagA and vacuolating toxin VacA in control of the NFAT signaling pathway in gastric epithelial cells. Proc Natl Acad Sci U S A **102:**9661-6.

405. **Yun, M. S., S. E. Kim, S. H. Jeon, J. S. Lee, and K. Y. Choi.** 2005. Both ERK and Wnt/beta-catenin pathways are involved in Wnt3a-induced proliferation. J Cell Sci **118:**313-22.

406. **Zeaiter, Z., D. Cohen, A. Musch, F. Bagnoli, A. Covacci, and M. Stein.** 2008. Analysis of detergent-resistant membranes of *Helicobacter pylori* infected gastric adenocarcinoma cells reveals a role for MARK2/Par1b in CagA-mediated disruption of cellular polarity. Cell Microbiol **10:**781-94.

407. **Zhao, W., W. Wu, and X. Xu.** 2007. Oral vaccination with liposome-encapsulated recombinant fusion peptide of urease B epitope and cholera toxin B subunit affords prophylactic and therapeutic effects against *H. pylori* infection in BALB/c mice. Vaccine **25:**7664-73.

408. **Zheng, P. Y., and N. L. Jones.** 2003. *Helicobacter pylor*i strains expressing the vacuolating cytotoxin interrupt phagosome maturation in macrophages by recruiting and retaining TACO (coronin 1) protein. Cell Microbiol **5:**25-40.

409. **Zhu, Y., C. Wang, J. Huang, Z. Ge, Q. Dong, X. Zhong, Y. Su, and S. Zheng.** 2007. The *Helicobacter pylori* virulence factor CagA promotes Erk1/2-mediated Bad phosphorylation in lymphocytes: a mechanism of CagA-inhibited lymphocyte apoptosis. Cell Microbiol **9:**952-61.

Chapter Two

Polymorphism in the CagA EPIYA Motif Impacts Development of Gastric Cancer

Published as: **Jones, K.R.[*], Y.M. Joo, S. Jang[*], Y.J. Yoo, H.S. Lee, I.S. Chung, C.H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. *J Clin Microbiol* 47:959-68.

[*] These authors contributed equally to this work.

The work presented in this chapter is the sole work of K. R. Jones with the following exceptions: Y.M Joo and S. Jang performed the *cag* genotyping, Y.J. Yoo and J.H. Cha assisted with experimental design, H.S. Lee originally isolated *H. pylori* strains, and I.S. Chung performed the biopsies and supplied the diagnoses, C.H. Olsen assisted with the statistical analysis, and J.M. Whitmire assisted with microscopy, IL-8 ELISAs, and in generating the figures.

Abstract

Helicobacter pylori causes diseases ranging from gastritis to peptic ulcer disease to gastric cancer. Geographically, areas with high incidences of *H. pylori* infection often overlap with areas with high incidences of gastric cancer, which remains one of the leading causes of cancer-related death worldwide. Strains of *H. pylori* that carry the virulence factor cytotoxin-associated gene A (*cagA*) are much more likely to be associated with development of gastric cancer. Moreover, particular C-terminal

polymorphisms in CagA vary by geography, and have been suggested to influence disease development. We conducted a large-scale molecular epidemiologic analysis of South Korean strains and herein report a statistical link between the East Asian CagA, EPIYA-ABD genotype and development of gastric cancer. Characterization of a subset of the Korean isolates showed that all strains from cancer patients expressed and delivered phosphorylatable CagA to host cells, whereas the presence of the *cagA* gene did not strictly correlate to expression and delivery of CagA in all non-cancer strains.

Introduction

Helicobacter pylori is a medically important pathogen, and although infection rates vary geographically, this bacterium colonizes more than 50% of the world's population (1, 28). This spiral shaped, Gram-negative, microaerophilic bacterium chronically inhabits the unforgiving environment of the stomach, and causes subclinical gastritis in the majority of patients. However, in some individuals, *H. pylori* colonization results in peptic ulcer disease; 75% of gastric ulcers and 90% of duodenal ulcers are attributed to *H. pylori* infection (20). In its most severe sequelae, *H. pylori* infection can lead to the development of two forms of gastric cancer: adenocarcinoma and mucosa-associated lymphoid tissue (MALT) lymphoma (12, 33, 34, 43). The association of *H. pylori* with stomach cancer led the World Health Organization to classify it as a class I carcinogen in 1994 (26). It currently remains the only bacterium to obtain this perilous distinction.

Gastric cancer is the second most common cause of cancer death worldwide, and this fact could be reflective of the high incidence of *H. pylori* infection (19, 30, 32, 47).

Interestingly, geographic areas with the highest level of gastric cancer, which include most East Asian countries, also have the highest rate of *H. pylori* infection (2, 19, 47). Additionally, in East Asian countries, 90% of strains carry the cytotoxin-associated gene A (*cagA*; 27), which has emerged as a major contributor to disease severity. In fact, *cagA*-positive *H. pylori* strains are at least twice as likely to cause cancer than *H. pylori* strains without *cagA* (13, 22).

cagA is carried on the *cag* pathogenicity island (PAI), which carries genes that produce a type IV secretion apparatus that is used to directly inject CagA into host cells (16). Within the cells, CagA is phosphorylated by host cell kinases, forms a complex with the SHP-2 (Src homology region 2-containing phosphatase 2; 25), and alters multiple host signaling pathways (23-25, 29, 36, 46). The phosphorylation of CagA occurs in the carboxy terminus of conserved tyrosine residues that are part of a repeated five amino acid sequence (Glu-Pro-Ile-Tyr-Ala) referred to as the EPIYA motif (24, 25).

Initial studies showed that CagA proteins from various *H. pylori* isolates migrated differently on denaturing gels (18). It was subsequently shown that a number of *cagA* alleles exist, and that variation in the carboxy terminus of the protein is the major difference between the different alleles. Polymorphisms in the C terminus occur in the EPIYA region and typically involve changes in the amino acid sequences flanking the five amino acid repeat. The most common motifs have been designated as EPIYA-A, -B, -C, and -D (24), and are found in two distinct combinations by geographic location. Western CagA consists of a combination of EPIYA-A, -B, and -C motifs (up to five -C motifs have been identified), whereas East Asian CagA contains a combination of EPIYA-A, -B, and -D motifs (5, 18, 24, 25, 31, 42).

EPIYA -C and -D serve as the primary CagA phosphorylation sites and are required for binding to SHP-2 (24). Among Western isolates, molecular epidemiological studies have indicated a correlation between disease severity and increased number of EPIYA -C motifs (6, 24, 48). Indeed, in cases where Western strains are associated with cancer, most have multiple EPIYA -C motifs (9, 39). This increase may be due to elevated morphological transformation as a result of increased CagA phosphorylation and SHP-2 binding (24).

East Asian CagA containing the EPIYA -D motif demonstrates higher affinity for SHP-2 than Western CagA. This leads to greater morphological changes in infected cells (24), as well as greater levels of inflammation and atrophy (10). These findings, along with the fact that East Asian strains predominate in countries with the highest rates of gastric cancer, suggest that East Asian CagA may have the potential to induce more severe forms of gastric disease (2, 19, 47).

In order to assess the correlation between *cagA* genotype and *H. pylori*-induced disease severity, we examined a collection of isolates from South Korea, which has one of the highest rates of *H. pylori* colonization (44), and one of the highest rates of gastric cancer in the world (22, 41). Additionally, the majority of South Korean strains carry the East Asian *cagA* allele (17). Here we present molecular epidemiologic evidence that there is a significant association between the development of gastric cancer and infection with *H. pylori* strains carrying the EPIYA -ABD genotype.

Materials and Methods

Bacterial Strains

Korean bacterial strains along with G27-MA (3), and its isogenic derivatives G27-MA Δ*cagA* (4) and G27-MA ΔPAI [provided by Manuel Amieva and constructed as described in Galgani, *et al.* (21)], were cultured as previously described (15). Briefly, bacterial stocks preserved at -80°C were grown and expanded on antibiotic supplemented horse blood agar (HBA) plates. Overnight liquid cultures, brucella broth (BB; Acumedia, Lansing, MI) containing 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and 10 μg/ml vancomycin (Amresco, Solon, OH), were subcultured to an optical density at 600 nm of 0.05 in fresh media and were grown for 18 hours under microaerophilic conditions created by an Anoxomat evacuation/replacement system (Spiral Biotech, Norwood, MA).

Clinical Isolate Acquisition

Isolates were obtained from patients presenting with gastric symptoms to the Division of Gastroenterology in the Department of Internal Medicine at the College of Medicine of The Catholic University of Korea in Seoul, South Korea. Written informed consent was received from each patient, and the protocol was approved by the Institutional Review Board of Human Research at The Catholic University of Korea. Biopsies were collected at the site of visible mucosal disturbance and histology was performed to provide a diagnosis along with culture for the presence of Gram-negative, spiral shaped bacteria that produced a functional urease enzyme. Subsequently, a single colony isolate was selected from each biopsy for further characterization. An extensive breakdown of the epidemiological characteristics of the patients can be found in Table 1,

and a complete list of strains is available as Table S1 in the supplemental material. Strains are named such that the letters following the strain number indicate the disease state as follows: cancer (-CA), duodenal ulcer (-DU), gastritis (-G), gastric ulcer (-GU).

cagA *Genotyping*

All the primers used in this study are listed in Table 2. Genomic DNA was extracted using the Easy DNA kit (Invitrogen, Carlsbad, CA), and genotyping of the C terminus of *cagA* was performed by PCR using a modified version of that developed by Argent *et al.*(8) as schematically depicted in Fig. 4A, 4B, and Supplementary Fig. S1 in the supplemental material. Briefly, amplification with primers cagA28F or cag2 and cagA-P1C or cagA-pA-1 (R) identifies an EPIYA-A motif. Amplification with primers cagA28F or cag2 and a 1:1 mixture of primers cagA-P2TA and cagA-P2CG indicates an EPIYA-B motif. Amplification using primers cagA28F or cag2 and cagA-P3E identifies either the presence of an EPIYA-C or an EPIYA-D motif, and an additional amplicon with cagA28F or cag2 and the unique cagA-pD (R) primer categorizes the CagA as having an EPIYA-D motif. In some cases where the PCR amplification was inconclusive and to confirm results of the PCR genotyping, *cagA* was amplified with cag2 and either cagA seq (R) or grace2 (which lies within the downstream conserved *glutamate racemase* gene) and then sequenced using primers cag2 and cagA seq (R). Sanger dideoxy sequencing was performed by the Uniformed Services University of the Health Science Biomedical Instrumentation Center (Bethesda, MD) or at Cosmo Genetech Co, Ltd (Seoul, Korea). The resulting DNA sequences were analyzed using Vector NTI version 9.1 (Invitrogen) and Sequencher 4.5 (Gene Codes Corp., Ann Arbor, MI).

Table 1: Primer Sequences

Primer	**Sequence (5'-3')**	**Reference**
cag 2*	GGAACCCTAGTCGGTAATG	(37)
cagA28F	TTCTCAAAGGAGCAATTGGC	(8)
cagA-P1C	GTCCTGCTTTCTTTTTATTAACTTKAGC	(8)
cagA-pA1 (R)	CTTGTCCTGYTTTCTTTTTATTAAC	This Study
cagA-P2TA	TTTAGCAACTTGAGTATAAATGGG	(8)
cagA-P2CG	TTTAGCAACTTGAGCGTAAATGGG	(8)
cagA-P3E	ATCAATTGTAGCGTAAATGGG	(8)
cagA-pD (R)	TTGATTTGCCTCATCAAAATC	This Study
cagA seq (R)*	TGGTTGAATCCAATTTTATC	This Study
grace2	TCATGCGAGCGGCGATGT	This Study

*Also used for sequencing

CagA Protein Expression

After 24 hours, lawns of bacteria were harvested from horse blood agar plates, pelleted, resuspended in 1 X phosphate buffered saline (PBS), and mixed with 5 X Laemmli sample buffer. Bacterial lysates were separated by sodium dodecyl sulfate-polyacrylamide gele electrophoresis, using a 10% separating gel and a 4% stacking gel, and proteins were transferred to nitrocellulose membranes using a semidry transfer apparatus (OWL, Thermo Scientific, Rochester, NY). Membranes were probed with a 1:5,000 dilution of mouse immunoglobulin G1 (IgG1) anti-CagA monoclonal antibody (Austral Biologicals, San Ramon, CA) followed by a 1:20,000 dilution of horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG secondary antibody (Santa Cruz Biotechnology, Santa Cruz, CA). Alternatively, membranes were probed with a 1:5,000 dilution of rabbit IgG anti-CagA polyclonal antibody, B-300 (Santa Cruz Biotechnology) followed by a 1:20,000 dilution of HRP conjugated bovine anti-rabbit IgG secondary antibody (Santa Cruz Biotechnology). Proteins were detected using the Pierce ECL Western Blotting Substrate kit (Thermo Scientific/Pierce, Rockford, IL) and photographic film with a Series XXXV A Rapid Processor (S&W Imaging, Frederick, MD) or using the Super Signal West Pico Chemiluminescent Substrate kit (Thermo Scientific/Pierce) and a LAS-3000 Intelligent Dark Box with LAS-3000 Lite capture software (FujiFilm, Stamford, CT).

CagA Phosphorylation Assays

The phosphorylation assays were essentially conducted as previously described (14). Briefly, six-well tissue culture plates were seeded with 3.5 X 10^5 AGS cells per

well and allowed to grow for three days in normal cell culture media, Dulbecco's Modified Eagle's Media without L-Glutamine (Quality Biological, Inc., Gaithersburg, MD) supplemented with 10% fetal bovine serum, 10 μg/ml vancomycin, and 2 nM L-glutamine (Quality Biological, Inc.). Two hours prior to infection, AGS cells were washed with 1 X PBS, and 3 ml of fresh media were added to each well. Liquid cultures of *H. pylori* were resuspended in 1 ml of 1 X PBS, and used to infect the AGS cells at an MOI of 100. Infections were allowed to proceed for 5 hours at which point the media was removed, and the cells were washed with 1 X PBS and lysed with 5 X Laemmli sample buffer. Infected cell lysates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, using a 6% separating gel and a 4% stacking gel. Proteins were transferred to nitrocellulose membranes by semidry transfer. Membranes were then probed with a 1:5,000 dilution of an anti-phospho-tyrosine monoclonal antibody, pY100 (Cell Signaling Technology, Danvers, MA) followed by a 1:20,000 dilution of HRP conjugated goat anti-mouse IgG secondary antibody and detection as described above. Membranes were subsequently stripped (using a heated 10-mM dithiothreitol solution) and re-probed with polyclonal anti-CagA antibody, B-300 as described above. Densitometry was performed using MultiGauge software (FujiFilm, Stamford, CT).

A low level of cross-reactivity for CagA with the anti-phospho-tyrosine antibodies (pY100 and pY99; BD Biosciences, San Jose, CA) was observed. Hence bacterial lysates were run adjacent to their corresponding infected cell lysates so that any cross reactivity could be accounted for when comparing the ratio of phosphorylated CagA to total CagA of infected cell lysates (Table 3).

Cell Elongation Studies

The cell elongation studies were essentially conducted as previously described (14). Six-well tissue culture plates were seeded with 2.7 X 10^5 AGS cells per well and allowed to grow in normal cell culture media for 22 hours. After 22 hours and approximately two hours before infection, the media was removed, the cells were washed with 1 X PBS, and 1 ml of fresh cell culture media supplemented with 10% BB was added to each well. Eighteen-hour liquid cultures of *H. pylori* that were suspended in 1 ml of the BB supplemented cell culture media were then used to infect at an MOI of 100.

Infections were allowed to proceed for nine hours, at which point cells were fixed with 2% paraformaldehyde in 100 mM phosphate buffer (pH7.4), and stained with Giemsa (Sigma-Aldrich, Inc, St. Louis, MO) per manufacture's directions. Cells were analyzed using an Olympus BX60 (Olympus America Inc, Center Valley, PA) and were digitally photographed using a Spot RT color camera (Diagnostic Instruments, Sterling Heights, MI). One hundred cells were counted to assess the number of cells displaying the "hummingbird" phenotype, which is characterized by the presence of finger-like protrusions (40). In each case, the infections and analysis were replicated to verify the reproducibility of the results. Any strain that on average induced greater than 60% of the AGS cells to display the "hummingbird" phenotype in biologically independent experiments was considered positive for the presence of a functional CagA.

IL-8 Induction Assay

Six-well tissue culture plates were seeded with 4.2 X 10^5 AGS cells per well and allowed to grow in normal cell culture media for 24 hours. At this point the media was

removed and replaced with media lacking serum for a period of 24 hours before *H. pylori* infection. Approximately two hours prior to infection, the cells were washed with 1 X PBS, and 1 ml of fresh media without serum was added to each well. Eighteen-hour liquid cultures of *H. pylori* were pelleted and resuspended in 700 µl of cell culture media without serum and used to infect the semi-confluent AGS cells at an MOI of 100. After five hours, the cell culture supernatant was collected, samples were centrifuged at 16,100 relative centrifugal force for 10 minutes, and the supernatant was transferred to a new tube, and stored at -20°C until later use. Human interleukin-8 (IL-8) concentration was measured using an enzyme-linked immunosorbent assay kit (R&D Systems, Minneapolis, MN) following the manufacturer's direction. The change in IL-8 concentration was calculated in comparison to G27-MA ΔPAI. An independent biological repeat of each infection was conducted, and strains were considered to induce IL-8 if the average fold change was more than 10 fold.

Statistical Analysis

The Fisher exact test was used to analyze the association between disease state and EPIYA motif genotype. Log linear modeling was used to assess whether this association was consistent across the age and sex subgroups. We fit a saturated model using categorical variables representing genotype, disease state, sex, and age groups. A backwards selection algorithm identified two- and three-way association among these variables that were statistically significant at the 5% level. Data was analyzed using SPSS version 14 software (SPSS Inc., Chicago, IL).

Nucleotide Sequence Accession Numbers

The sequences of the C-terminal region of CagA from 47 strains have been deposited in GenBank, accession numbers FJ458117-FJ458163.

Results

*Sample Acquisition/*cagA *Genotyping*

A total of 260 *H. pylori* clinical isolates were obtained from patients presenting with gastric maladies (Table 2). Six of these were missing the epidemiological data of age and gender. Of the remaining 254, the mean patient age was 51 years, with an age range of 14 to 86 years. There were 126 females (49.6%), with a mean age of 52 years and an age range of 21 to 86 years, and 128 males (50.4%), with a mean age of 50 years and an age range of 14 to 82 years. Of the 254 samples, 45.3% were from patients with gastritis, 43% with ulcers (21.7% gastric ulcers and 21.3% duodenal ulcers), and 11.8% with cancer.

Four different PCR reactions were conducted for each strain in order to genotype *cagA* (Fig. 4A and 4B; see Fig. S1 in the supplemental material). As previously described, three of these PCR reactions identify different EPIYA motifs; one identifies the EPIYA-A motif, one EPIYA-D motif, we also designed and employed primer cagA-pD (R), which is well conserved among strains carrying the EPIYA-D motif and one of the first primers designed to specifically amplify the EPIYA-D motif. Using this technique, we were able to genotype 234 strains (Table 2). These strains displayed the same age range as the full collection and a mean age of 50 years. Again the proportion of females (112) to males (116) was virtually identical to the larger collection, 49.1% to

Table 2: Epidemiological Breakdown of the Korean Collection

			Disease State			
	Total	*Genotyped Samples*	*Gastritis*	*Gastric Ulcer*	*Duodenal Ulcer*	*Gastric Cancer*
Overall Total	260*	234* (90%)	108** (46%)	42 (18%)	54*** (23%)	30 (13%)
Age Range	14-86	14-86	19-82	34-84	14-72	37-86
Mean	51	50	49	55	45	58
Females	126 (50%)	112 (49%)	69 (64%)	10 (24%)	19 (39%)	14 (47%)
Age Range	21-86	21-86	21-82	46-84	31-72	37-86
Mean	52	52	49	57	51	61
Males	128 (50%)	116 (51%)	38 (36%)	32 (76%)	30 (61%)	16 (53%)
Age Range	14-82	14-82	19-78	34-82	14-70	38-70
Mean	50	49	48	54	41	55
-ABD *cagA*		200* (85%)	87** (81%)	38 (90%)	45*** (83%)	30 (100%)
Age Range		14-86	19-78	34-84	14-72	37-86
Mean		51	49	55	44	58
Females		90 (46%)	57 (66%)	8 (21%)	11 (28%)	14 (47%)
Age Range		21-86	21-75	46-84	36-72	37-86

Mean	52	49	58	52	61
Males	104 (54%)	29 (34%)	30 (79%)	29 (73%)	16 (53%)
Age Range	14-82	19-78	34-82	14-70	38-70
Mean	49	48	53	42	55
All Other genotypes	34 (15%)	21 (19%)	4 (10%)	9 (17%)	0 (0%)
Age Range	28-82	28-82	48-81	31-72	0
Mean	50	49	61	49	0
Females	22 (65%)	12 (57%)	2 (50%)	8 (89%)	0
Age Range	28-82	28-82	48-56	31-72	0
Mean	51	50	52	51	0
Males	12	9 (43%)	2 (50%)	1 (11%)	0
Age Range	33-81	36-61	58-81	33	0
Mean	49	46	70	N/A	0

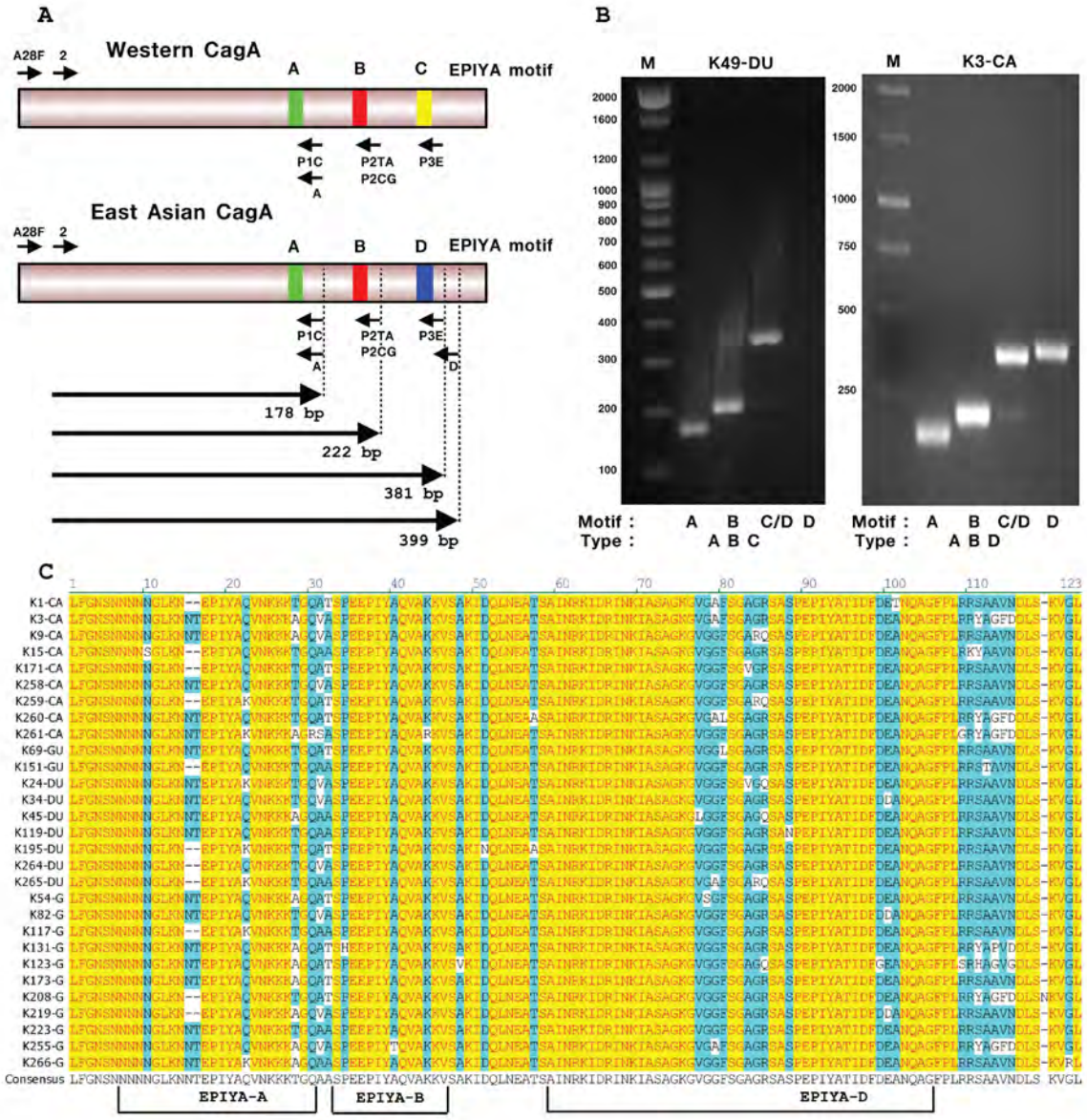
*6 w/o age or sex information
**One w/o age
***5 w/o age

50.9%, respectively. The remaining 26 strains, which all came from non-cancer patients, failed to yield PCR products or gave incorrectly sized bands, and thus were not further analyzed.

To confirm the *cagA* genotyping results, we sequenced the C-terminal region of the *cagA* gene from 47 of the 234 genotyped strains. These sequences verified that the PCR genotyping method was accurate. Alignments of the predicted amino acid sequences of those strains carrying the -ABD motif can be found in Fig. 4C.
Of the 234 genotyped strains, 208 isolates (88.9%) carried an EPIYA-D motif, therefore classifying them as East Asian, and 26 isolates (11.1%) were determined to carry Western CagA (see Table S1 in the supplemental material). Among the East Asian strains, eight carried an incomplete -ABD motif or contained additions of one or more motifs (EPIYA-AABD, -BD, -BBD, -ABAB*D, and -AB*D, with a mutation within the EPIYA-B motif as designated by the asterick), and thus, we subdivided the strains based on the presence of CagA containing a complete EPIYA-ABD motif versus all other EPIYA motifs. Given these characteristics, 34 individuals were determined to have "other genotypes," which includes alternative -ABD as well as Western motifs. We next analyzed the distribution of *cagA* genotype among disease states. There were 108 gastritis patients, 54 duodenal ulcer patients, 42 gastric ulcers patients, and 30 gastric cancer patients. EPIYA-ABD CagA composed 80.6% of gastritis patients, 83.3% of duodenal ulcer patients, 90.5% of gastric ulcer patients and 100% of gastric cancer patients. Stratification of the patients based on age, sex and disease categories can be found in Table 2, and a schematic depiction of the distribution of the *cagA* genotypes stratified by disease state within this Korean population can be found in Fig. 5.

Figure 4: Genotyping of the cagA *variable-EPIYA motif.* A. Schematic representation of the *cagA* variable region. Western CagA (EPIYA-ABC) is depicted on the top, and East Asian CagA (EPIYA-ABD) is depicted on the bottom. The annealing positions (small arrows), names of the primers used in this study, and the expected sizes (large arrows) of the amplified specific EPIYA motif products as based on strain K3-CA DNA sequence are shown. Primer names are abbreviated as follows: cagA28F is A28F, cag2 is 2, cagA-P1C is P1C, cagA-pA1 (R) is A, cagA-P2TA is P2TA, cagA-P2CG is P2CG, cagA-P3E is P3E, and cagA-pD (R) is D. B. PCR amplicons of K49-DU and K3-CA using the forward primer 2 and the reverse primer A, P2TA and P2CG (equimolar mixture), P3E, or D. "M" designates the size markers (in base pairs). Type indicates resulting EPIYA motif identified. C. Amino Acid alignment of the carboxy terminus of CagA from 29 Korean strains encoding the EPIYA-ABD motif. The EPIYA-A, -B and -D motifs are indicated below the consensus sequence.

Figure 4: Genotyping of the cagA *variable-EPIYA motif*

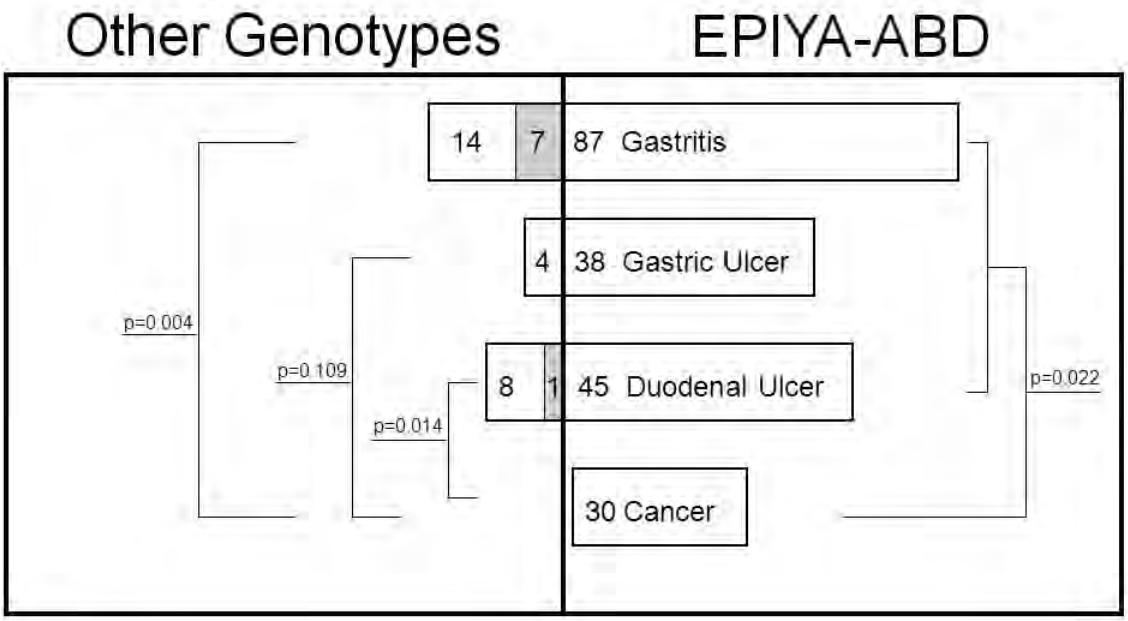


Even though this collection was evenly distributed for the factors of age and gender, several trends for each were observed based on statistical analysis. Not surprisingly, age is statistically linked to disease state (P<0.001). Moreover, as has been suggested in other studies (38), males were more likely to have ulcers (odds ratio, 3.89; confidence interval, .81 to 8.36), whereas females were more likely to have gastritis (odds ratio, 1.91; confidence interval, 1.91 to 5.67). Conversely, the cancer patients were evenly distributed by gender, with 46.7% being female and 53.3% being male. This differs from what is most often seen in literature, which shows that men are anywhere from 1.5 to 2.5 times more likely to be afflicted with gastric cancer than women (as reviewed in 35). A significant three-way association was also observed using log linear modeling between gender, disease, and *cagA* allele (P=0.009).

Given the fact that 100% of the gastric cancer patients were infected with *H. pylori* encoding CagA with the EPIYA-ABD motif, we conducted statistical analysis to assess the relationship between disease state and genotype. The Fisher's exact test showed that the proportion of patients with the EPIYA-ABD genotype varied significantly (P=0.022) by diagnosis (Fig. 5). In fact, the proportion of cancer patients with the EPIYA-ABD genotype (100%) was significantly higher than the proportion in gastritis patients (80.6%; P=0.004) or duodenal ulcer patient (83.3%; P= 0.014), but not in gastric ulcer patients (90.5%; P=0.109; Fig. 5). Taken together, these data suggest that there is a definitive link between infection with *H. pylori* strains carrying *cagA* which encodes the EPIYA-ABD motif and the development of gastric cancer.

Figure 5: Schematic depiction of the distribution of the cagA *genotypes stratified by disease state within this Korean population.* Distribution of *cagA* genotypes, EPIYA-ABD versus all other EPIYA motifs, within the four different disease states: gastritis, gastric ulcers, duodenal ulcers, and cancer. The shaded portion within the other genotypes subgrouping corresponds to the isolates that contain an alternative EPIYA-ABD motif. Calculated P values, using the Fisher's exact test are shown.

Figure 5: Schematic depiction of the distribution of the cagA *genotypes stratified by disease state within this Korean population*



CagA Protein Expression

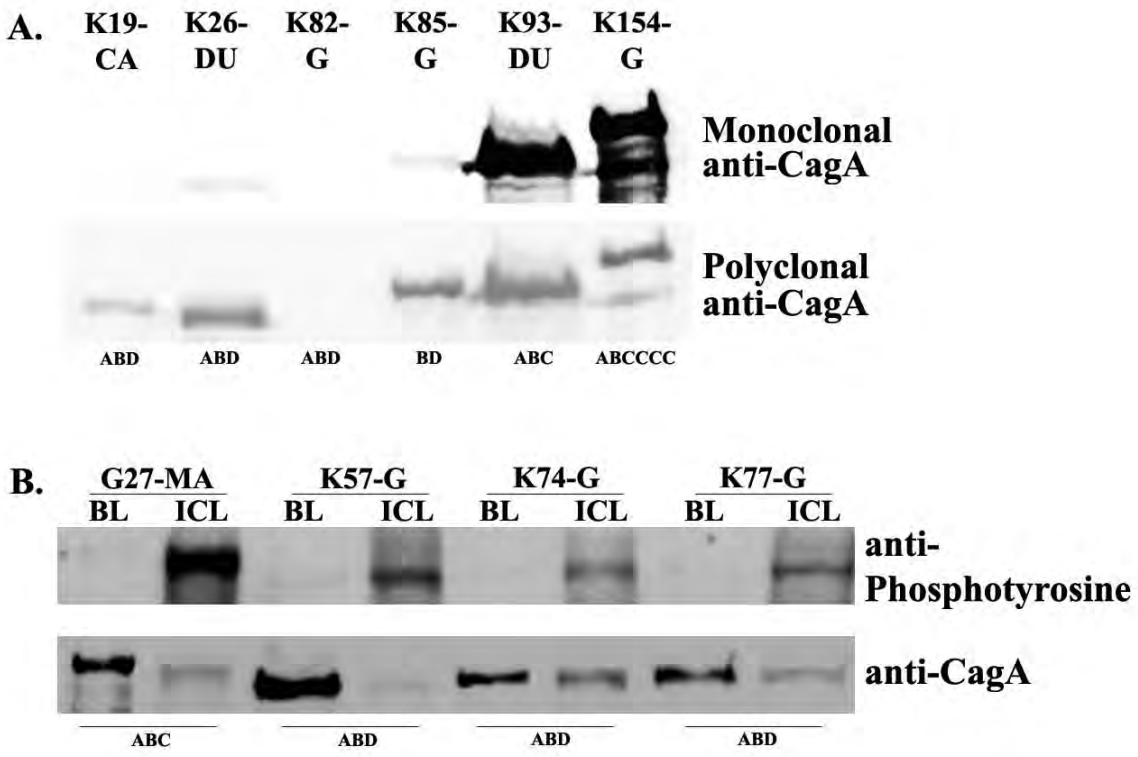
Given the fact that we saw a significant statistical link between the presence of the EPIYA-ABD genotype and gastric cancer but that some strains that carry *cagA* do not actually express the CagA protein (25), we next sought to determine if genotypically *cagA*$^{+}$ strains were phenotypically CagA positive. Bacterial lysates from a subset of 77 randomly chosen strains were assessed for expression of CagA. Of the 77 isolates examined using a monoclonal antibody, four samples (K19-CA, K82-G, K255-G, and K264-DU) showed no appreciable CagA expression (Fig.6A and data not shown). Given the fact that CagA shows heterogeneity in the carboxy terminus that may affect protein structure and monoclonal antibody recognition, we also utilized a polyclonal anti-CagA antibody to ensure that CagA was actually not expressed in these strains. As shown in Figure 6A, the polyclonal antibody was better able to detect CagA in the majority of strains. This included K19-CA for which CagA was not detected with the monoclonal antibody. Using this assay, three of the 77 strains (K82-G, K255-G, and K264-DU) expressed no detectable level of CagA (Table 3 and data not shown).

Delivery and Phosphorylation of CagA

Once CagA is expressed, it must be delivered to host cells via the type IV secretion apparatus and phosphorylated by host cell kinases to be biologically active (40). Therefore, we next conducted phosphorylation assays to determine if CagA could be delivered to and phosphorylated in host cells. Of the 77 strains tested, 59 of the isolates efficiently delivered CagA to the host cells as detected by the appearance of a strongly phosphorylated CagA band (Fig. 6B and Table 3). Of the remaining 18 isolates, an

Figure 6: Expression, Delivery, and Phosphorylation of CagA. A. Western blot analysis of bacterial lysates from the six indicated Korean strains was conducted using a monoclonal anti-CagA antibody (top), or the polyclonal anti-CagA antibody, B300 (bottom). B. Lysates from the bacterial cells alone (BL) and AGS cells infected with the same bacterial strain (ICL) were assessed for delivery and phosphorylation of CagA. Membranes were probed with an anti-phosphotyrosine antibody, pY100 (top), stripped, and subsequently reprobed with polyclonal anti-CagA antibody, B300 (bottom). Data are shown from the positive control, G27-MA, and three indicated Korean isolates.

Figure 6: Expression, Delivery, and Phosphorylation of CagA



intermediate level of phosphorylated CagA was detected for 13 strains, and no detectable phosphorylated CagA was found for five strains (K82-G, K111-DU, K123-G, K255-G, and K264-DU; Table 3). Importantly, the three strains shown to be negative for CagA expression were included among these five, and K19-CA, which was only detected with the polyclonal CagA antibody, was positive for phosphorylation.

Cell Elongation Assay

Upon injection of CagA into host cells, it becomes phosphorylated and causes striking host cell elongation, which is known as the "hummingbird" phenotype (40). Thus, to reassess the presence of functional CagA in the 18 isolates that produced either an intermediate phenotype or no detectable level of phosphorylated CagA, the ability to induce the "hummingbird" phenotype was assessed in cultured AGS cells. A wild-type strain, G27-MA, and its isogenic Δ*cagA* mutant were used as positive and negative controls, respectively. The percentage of cells displaying the "hummingbird" phenotype for the G27-MA Δ*cagA* infected cells was 23% and for the G27-MA-infected cells was 80% (Fig. 7A and B). The range for the 18 Korean isolates was between 31.5% (K82-G) and 79% (K25-DU; Fig. 7C and D; Table 3). Given the large range of changes, we conservatively required that a strain induce at least 60% of AGS cells to display the "hummingbird" phenotype to be considered positive for delivery of functional CagA. Four of the 18 samples tested did not meet this threshold (K82-G, K123-G, K255-G, and K264-DU). These samples also showed no detectable level of phosphorylated CagA via the phosphorylation assay.

Table 3: Analysis of CagA Expression and Function

Strain	EPIYA Motif	CagA Expression	Phosphorylation of CagA[a]	Induction of "Hummingbird" Phenotype[b]	IL-8 Induction[c]
K82-G	ABD	-	-	-	-
K255-G	ABD	-	-	-	-
K264-DU	ABD	-	-	-	+/-
K111-DU	ABD	+	-	+	
K123-G	ABD	+	-	-	-
K17-CA	ABD	+	+/-	+	+
K26-DU	ABD	+	+/-	+	+
K208-G	ABD	+	+/-	+	+
K21-CA	ABD	+	+/-	+	
K23-DU	ABD	+	+/-	+	
K25-DU	ABD	+	+/-	+	
K42-DU	ABD	+	+/-	+	
K104-CA	ABD	+	+/-	+	
K182-DU	ABD	+	+/-	+	
K193-G	ABD	+	+/-	+	
K238-DU	ABD	+	+/-	+	
K248-G	ABD	+	+/-	+	
K259-CA	ABD	+	+/-	+	
K3-CA	ABD	+	+		
K6-CA	ABD	+	+		
K10-CA	ABD	+	+		
K16-CA	ABD	+	+		

K19-CA	ABD	+	+
K28-DU	ABD	+	+
K34-DU	ABD	+	+
K35-DU	ABD	+	+
K36-DU	ABD	+	+
K37-DU	ABD	+	+
K41-DU	ABD	+	+
K43-DU	ABD	+	+
K44-DU	ABD	+	+
K45-DU	ABD	+	+
K46-DU	ABD	+	+
K47-DU	ABD	+	+
K48-DU	ABD	+	+
K57-G	ABD	+	+
K60-G	ABC	+	+
K64-G	ABCC	+	+
K74-G	ABD	+	+
K77-G	ABD	+	+
K78-G	AABD	+	+
K80-CA	ABD	+	+
K85-G	BD	+	+
K93-DU	ABC	+	+
K107-DU	ABD	+	+
K109-G	ABD	+	+
K112-G	ABD	+	+
K113-G	ABD	+	+

K115-G	ABC	+	+
K117-G	ABD	+	+
K131-G	ABD	+	+
K162-G	ABD	+	+
K172-G	ABCC	+	+
K175-G	ABD	+	+
K178-G	ABD	+	+
K183-G	ABD	+	+
K185-G	ABD	+	+
K196-G	ABD	+	+
K197-G	ABD	+	+
K209-G	ABD	+	+
K218-G	ABD	+	+
K220-DU	ABD	+	+
K223-G	ABD	+	+
K235-G	ABD	+	+
K241-G	ABD	+	+
K258-CA	ABD	+	+
K260-CA	ABD	+	+
K261-CA	ABD	+	+
K262-G	ABC	+	+
K263-G	ABABD*	+	+
K265-DU	ABD	+	+
K266-G	ABD	+	+
K24-DU	ABD	+	++
K27-DU	ABD	+	++

K146-G	ABD**	+	++
K154-G	ABCCCC	+	++
K165-G	ABD	+	++

[a]**Phosphorylation of CagA:** ++ near or above G27-MA (positive control), + slightly below the level of G27-MA, +/- below the level of G27-MA but above background, - not detectable.

[b]**Induction of Hummingbird Phenotype:** + greater than 60% of cells displayed the hummingbird phenotype, - less than 60% of cells displayed the hummingbird phenotype.

[c]**Induction of IL-8:** + induction greater than 10-fold increase over G27-MA ΔPAI, - induction less than 10-fold increase over G27-MA ΔPAI

* -ABABD second -B motif is replaced with leucine, ELIYA

** -B motif's proline is replaced with a serine, ESIYA

Induction of IL-8

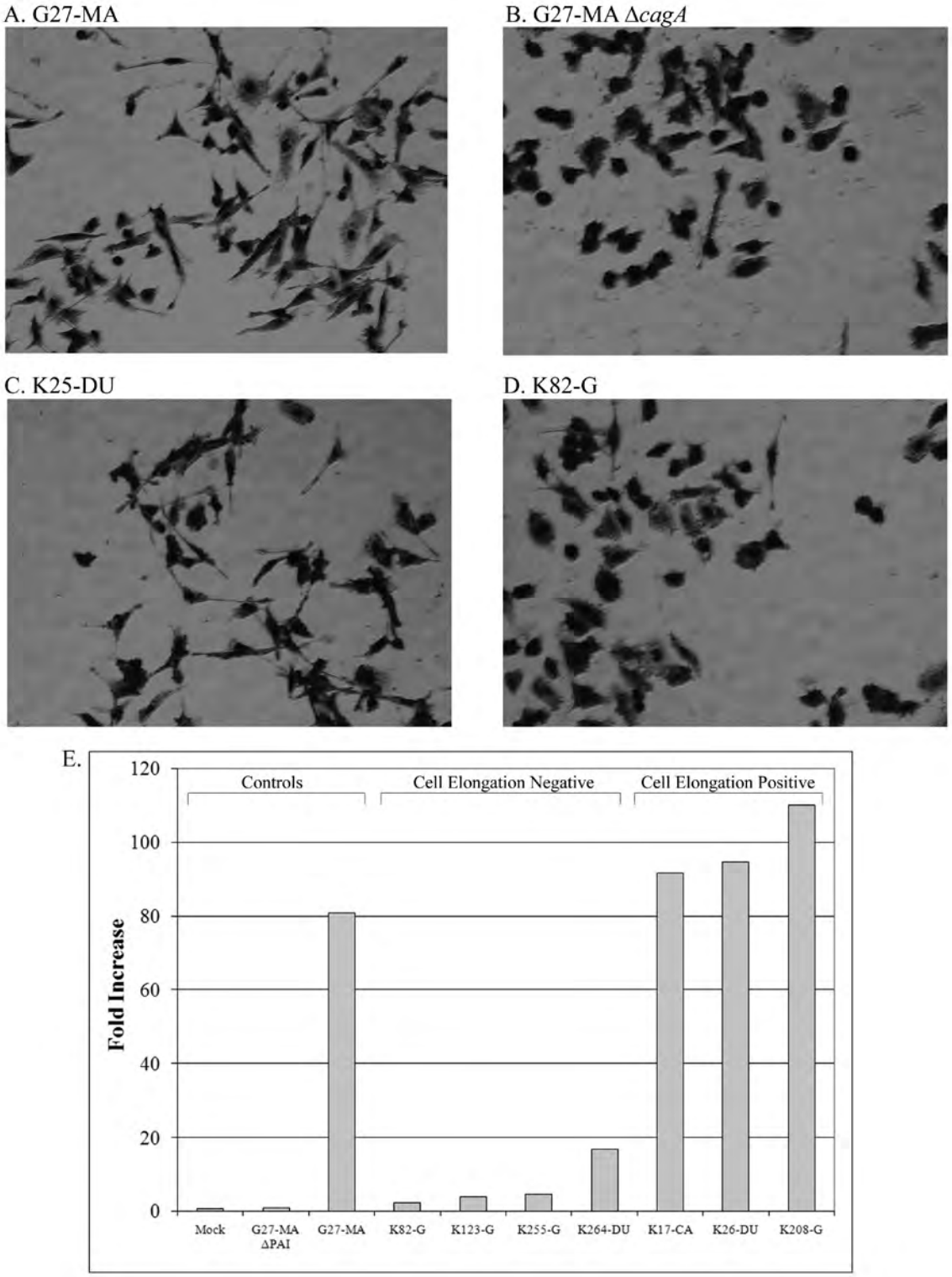
Since several isolates were identified that did not express any detectable level of functional CagA as measured by the phosphorylation assay and induction of the "hummingbird" phenotype, we finally assessed assembly of the type IV secretion system on the bacterial surface in this subset of strains. Proper assembly of the type IV secretion system has been shown to result in the induction of IL-8 in cultured AGS cells (11). Therefore we assessed IL-8 induction with the four strains that failed to produce phosphorylated CagA and failed to induce the "hummingbird" phenotype. Additionally, as a positive control, we analyzed several strains that did induce the "hummingbird" phenotype. One (K264-DU) out of the four "hummingbird" phenotype-negative samples induced IL-8 (at least 10 fold above the level induced by G27-MA ΔPAI) indicating that the failure to detect CagA and phosphorylated CagA or to induce the "hummingbird" phenotype was not impacted by a lack of a functional type IV secretion system (Fig. 7E and Table 3).

Discussion

Herein we show a significant statistical link between the presence of the CagA EPIYA-ABD motif and development of gastric cancer. In fact, 100% of gastric cancer patients analyzed in this South Korean population were infected with *H. pylori* strains encoding CagA containing the EPIYA-ABD motif. Statistical analysis with the Fisher exact test showed that the proportion of EPIYA-ABD genotype varied significantly by diagnosis (P=0.022), and that this distribution was statistically different than that of gastritis patients (P=0.004) or duodenal ulcer patients (P=0.014; Fig. 5).

Figure 7: Morphological Changes and Induction of IL-8. Induction of morphological changes in AGS cells when measured after 9 hours of infection with the following strains: G27-MA (positive control) (A), G27-MA Δ*cagA* (negative control) (B), K25-DU (C), and K82-G (D). E. Induction of IL-8 from the indicated strains expressed as increase above the induction elicited by G27-MA ΔPAI (negative control) after a five-hour infection.

Figure 7: Morphological Changes and Induction of IL-8



These data suggest that the distribution of alleles is not random and is important in the case of gastric cancer. While on a whole, the presence of the *cagA* gene did not strictly correlate to the expression and delivery of CagA to host cells, all of the analyzed cancer strains did express a functional CagA that could be delivered to and phosphorylated in host cells. This is the first time that a specific *cagA* allele has been statistically linked to gastric cancer. However, it should be noted that the EPIYA-ABD allele us not necessarily a predictor of cancer since there was a high percentage of peptic ulcer patients, both gastric (90%) and duodenal (83%), infected with isolates containing EPIYA-ABD CagA. Alternatively, these data could indicate that patients infected with H. *pylori* containing non EPIYA-ABD motifs are more likely not to develop cancer.

This association between the East Asian *cag*A genotype and gastric cancer may be due to higher affinity for SHP-2 (25). Binding of CagA with SHP-2 occurs via interaction of the phosphorylated EPIYA motifs and the SH2 domains from the host cell protein. This interaction changes the conformation of SHP-2 to its active form. Thus, the stronger affinity of East Asian CagA for SHP-2 results in longer periods of SHP-2 activity. This likely explains why East Asian strains cause greater morphological damage and greater level of induction of multiple cellular pathways, resulting in increased proliferation, morphogenesis, and cell motility than Western strains (23-25, 29, 36, 46).

While we did isolate eight EPIYA-AABD, -BD, -BBD, -ABAB*D, and –AB*D motifs, the vast majority of East Asian strains we examined showed strong conservation, and lack of duplication in the EPIYA-D region. This suggests that variation in the East Asian *cagA* is not as favorable as in Western isolates where the EPIYA-C motif is found

to vary widely among isolates (5, 7). Similar results have been seen by Argent et al., who showed by sequence analysis of 500 East Asian strains available in GenBank that the percentage of other East Asian alleles compared to those coding for EPIYA-ABD CagA was fairly small: 88.3% of East Asian strains contained an EPIYA-ABD CagA(5). Also, it is interesting that all of the cancer strains were specifically EPIYA-ABD, which suggests that among East Asian isolates this combination of EPIYA motifs is most favored for cancer development. It should be noted that in a study of Japanese cancer patients, all *H. pylori* isolates contained the EPIYA-D motif, and the majority of those isolates (84%) contained an EPIYA-ABD motif (10). Alternatively, the other 16% were made up of isolates carrying the EPIYA-AABD, -ABBD, - ABABD, and -ABDBD motifs (10).

The reason for the EPIYA-ABD conservation is unknown, but perhaps a single EPIYA-D motif allows for optimal SHP-2 binding. The presence of extra motifs may contort CagA's conformation and destabilize binding to SHP-2. Additionally, it is known that phosphorylated CagA at EPIYA-A and -B motifs binds to Csk and activates a negative feedback loop that inactivates the Src family kinases, and ultimately reduces the level of phosphorylated CagA in the cell (45). Thus, it is reasonable to suggest that the presence of additional EPIYA-A or -B motifs in association with EPIYA-D motif would more strongly activate this negative feedback loop (5). In support of the importance of conservation of the EPIYA-ABD motif in the disease state, seven out of the eight isolates containing a EPIYA-D motif but not a complete standard EPIYA-ABD motif only caused gastritis.

Multiple host, dietary, environmental, and bacterial virulence factors have been shown to play a role in *H. pylori*-induced disease. In this molecular epidemiologic study, we have shown a definitive statistical difference in the distribution of the *cagA* allele coding for the EPIYA-ABD motif in cancer versus other disease states; 100% of the cancer patients were infected with *H. pylori* strains carrying the EPIYA-ABD genotype. Currently the reason for this correlation is unclear, and further study is required to elucidate the molecular role that the EPIYA motif plays in cancer development.

Acknowledgements

We would like to thank Beth Carpenter for critical reading of the manuscript, M. Amieva for the G27-MA ΔPAI strain, and B.I. Kim for initial statistical analysis. This work was supported by the Korea Research Foundation Grant, which was funded by the Korean Government (MOEHRD; grant KRF-2006-311-E00083).

References

1. 1993. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. The EUROGAST Study Group. Gut **34:**1672-6.

2. **Ahn, Y. O., B. J. Park, K. Y. Yoo, N. K. Kim, D. S. Heo, J. K. Lee, H. S. Ahn, D. H. Kang, H. Kim, M. S. Lee, and et al.** 1991. Incidence estimation of stomach cancer among Koreans. J Korean Med Sci **6:**7-14.

3. **Amieva, M. R., N. R. Salama, L. S. Tompkins, and S. Falkow.** 2002. *Helicobacter pylori* enter and survive within multivesicular vacuoles of epithelial cells. Cell Microbiol **4:**677-90.

4. **Amieva, M. R., R. Vogelmann, A. Covacci, L. S. Tompkins, W. J. Nelson, and S. Falkow.** 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science **300:**1430-4.

5. **Argent, R. H., J. L. Hale, E. M. El-Omar, and J. C. Atherton.** 2008. Differences in *Helicobacter pylori* CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. J Med Microbiol **57:**1062-7.

6. **Argent, R. H., M. Kidd, R. J. Owen, R. J. Thomas, M. C. Limb, and J. C. Atherton.** 2004. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of *Helicobacter pylori*. Gastroenterology **127:**514-23.

7. **Argent, R. H., R. J. Thomas, F. Aviles-Jimenez, D. P. Letley, M. C. Limb, E. M. El-Omar, and J. C. Atherton.** 2008. Toxigenic *Helicobacter pylori* infection precedes gastric hypochlorhydria in cancer relatives, and *H. pylori* virulence evolves in these families. Clin Cancer Res **14:**2227-35.

8. **Argent, R. H., Y. Zhang, and J. C. Atherton.** 2005. Simple method for determination of the number of *Helicobacter pylori* CagA variable-region EPIYA tyrosine phosphorylation motifs by PCR. J Clin Microbiol **43:**791-5.

9. **Azuma, T., A. Yamakawa, S. Yamazaki, K. Fukuta, M. Ohtani, Y. Ito, M. Dojo, Y. Yamazaki, and M. Kuriyama.** 2002. Correlation between variation of the 3' region of the *cagA* gene in *Helicobacter pylori* and disease outcome in Japan. J Infect Dis **186:**1621-30.

10. **Azuma, T., S. Yamazaki, A. Yamakawa, M. Ohtani, A. Muramatsu, H. Suto, Y. Ito, M. Dojo, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Higashi, and M. Hatakeyama.** 2004. Association between diversity in the Src homology 2

domain--containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis **189:**820-7.

11. **Backhed, F., E. Torstensson, D. Seguin, A. Richter-Dahlfors, and B. Rokbi.** 2003. *Helicobacter pylori* infection induces interleukin-8 receptor expression in the human gastric epithelium. Infect Immun **71:**3357-60.

12. **Blaser, M. J.** 1998. *Helicobacter pylori* and gastric diseases. Bmj **316:**1507-10.

13. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with *Helicobacter pylori* strains possessing cagA is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

14. **Bourzac, K. M., L. A. Satkamp, and K. Guillemin.** 2006. The *Helicobacter pylori cag* pathogenicity island protein CagN is a bacterial membrane-associated protein that is processed at its C terminus. Infect Immun **74:**2537-43.

15. **Carpenter, B. M., T. K. McDaniel, J. M. Whitmire, H. Gancz, S. Guidotti, S. Censini, and D. S. Merrell.** 2007. Expanding the *Helicobacter pylori* genetic toolbox: modification of an endogenous plasmid for use as a transcriptional reporter and complementation vector. Appl Environ Microbiol **73:**7506-14.

16. **Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiara, M. Borodovsky, R. Rappuoli, and A. Covacci.** 1996. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc Natl Acad Sci U S A **93:**14648-53.

17. **Choi, K. D., N. Kim, D. H. Lee, J. M. Kim, J. S. Kim, H. C. Jung, and I. S. Song.** 2007. Analysis of the 3' variable region of the *cagA* gene of *Helicobacter pylori* isolated in Koreans. Dig Dis Sci **52:**960-6.

18. **Covacci, A., S. Censini, M. Bugnoli, R. Petracca, D. Burroni, G. Macchia, A. Massone, E. Papini, Z. Xiang, N. Figura, and et al.** 1993. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci U S A **90:**5791-5.

19. **Crew, K. D., and A. I. Neugut.** 2006. Epidemiology of gastric cancer. World J Gastroenterol **12:**354-62.

20. **Ernst, P. B., and B. D. Gold.** 2000. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu Rev Microbiol **54:**615-40.

21. **Galgani, M., I. Busiello, S. Censini, S. Zappacosta, L. Racioppi, and R. Zarrilli.** 2004. *Helicobacter pylori* induces apoptosis of human monocytes but not monocyte-derived dendritic cells: role of the cag pathogenicity island. Infect Immun **72:**4480-5.

22. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

23. **Higashi, H., A. Nakaya, R. Tsutsumi, K. Yokoyama, Y. Fujii, S. Ishikawa, M. Higuchi, A. Takahashi, Y. Kurashima, Y. Teishikata, S. Tanaka, T. Azuma, and M. Hatakeyama.** 2004. *Helicobacter pylori* CagA induces Ras-independent morphogenetic response through SHP-2 recruitment and activation. J Biol Chem **279:**17205-16.

24. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

25. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **295:**683-6.

26. **International Agency for Research on Cancer.** 1994. Infection with *Helicobacter pylori*, p. 177-240, Monographs on the Evaluation of Carcinogenic Risks to Humans, vol. 61, Lyon.

27. **Kim, J. M., J. S. Kim, H. C. Jung, I. S. Song, and C. Y. Kim.** 2000. Virulence factors of *Helicobacter pylori* in Korean isolates do not influence proinflammatory cytokine gene expression and apoptosis in human gastric epithelial cells, nor do these factors influence the clinical outcome. J Gastroenterol **35:**898-906.

28. **Matysiak-Budnik, T., and F. Megraud.** 1997. Epidemiology of *Helicobacter pylori* infection with special reference to professional risk. J Physiol Pharmacol **48 Suppl 4:**3-17.

29. **Neel, B. G., H. Gu, and L. Pao.** 2003. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends Biochem Sci **28:**284-93.

30. **Neugut, A. I., M. Hayek, and G. Howe.** 1996. Epidemiology of gastric cancer. Semin Oncol **23:**281-91.

31. **Nguyen, L. T., T. Uchida, K. Murakami, T. Fujioka, and M. Moriyama.** 2008. *Helicobacter pylori* virulence and the diversity of gastric cancer in Asia. J Med Microbiol **57:**1445-53.

32. **Parkin, D. M., F. Bray, J. Ferlay, and P. Pisani.** 2005. Global cancer statistics, 2002. CA Cancer J Clin **55:**74-108.

33. **Parsonnet, J., G. D. Friedman, D. P. Vandersteen, Y. Chang, J. H. Vogelman, N. Orentreich, and R. K. Sibley.** 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med **325:**1127-31.

34. **Parsonnet, J., S. Hansen, L. Rodriguez, A. B. Gelb, R. A. Warnke, E. Jellum, N. Orentreich, J. H. Vogelman, and G. D. Friedman.** 1994. *Helicobacter pylori* infection and gastric lymphoma. N Engl J Med **330:**1267-71.

35. **Roder, D. M.** 2002. The epidemiology of gastric cancer. Gastric Cancer **5 Suppl 1:**5-11.

36. **Roovers, K., and R. K. Assoian.** 2000. Integrating the MAP kinase signal into the G1 phase cell cycle machinery. Bioessays **22:**818-26.

37. **Rudi, J., C. Kolb, M. Maiwald, D. Kuck, A. Sieg, P. R. Galle, and W. Stremmel.** 1998. Diversity of *Helicobacter pylori vacA* and *cagA* genes and relationship to VacA and CagA protein expression, cytotoxin production, and associated diseases. J Clin Microbiol **36:**944-8.

38. **Salih, B. A., M. F. Abasiyanik, N. Bayyurt, and E. Sander.** 2007. *H pylori* infection and other risk factors associated with peptic ulcers in Turkish patients: a retrospective study. World J Gastroenterol **13:**3245-8.

39. **Satomi, S., A. Yamakawa, S. Matsunaga, R. Masaki, T. Inagaki, T. Okuda, H. Suto, Y. Ito, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Kutsumi, and T. Azuma.** 2006. Relationship between the diversity of the *cagA* gene of *Helicobacter pylori* and gastric cancer in Okinawa, Japan. J Gastroenterol **41:**668-73.

40. **Segal, E. D., J. Cha, J. Lo, S. Falkow, and L. S. Tompkins.** 1999. Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*. Proc Natl Acad Sci U S A **96:**14559-64.

41. **Shin, A., H. R. Shin, D. Kang, S. K. Park, C. S. Kim, and K. Y. Yoo.** 2005. A nested case-control study of the association of *Helicobacter pylori* infection with gastric adenocarcinoma in Korea. Br J Cancer **92:**1273-5.

42. **Stein, M., F. Bagnoli, R. Halenbeck, R. Rappuoli, W. J. Fantl, and A. Covacci.** 2002. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. Mol Microbiol **43:**971-80.

43. **Talley, N. J., A. R. Zinsmeister, A. Weaver, E. P. DiMagno, H. A. Carpenter, G. I. Perez-Perez, and M. J. Blaser.** 1991. Gastric adenocarcinoma and *Helicobacter pylori* infection. J Natl Cancer Inst **83:**1734-9.

44. **Tokudome, S., R. Ando, R. Ghadimi, T. Tanaka, N. Hattori, Z. Yang, M. Marumoto, H. Agawa, K. Arakawa, Y. Osaka, H. Tanaka, A. Hosono, and M. A. Moore.** 2007. Are there any real *Helicobacter pylor*i infection-negative gastric cancers in Asia? Asian Pac J Cancer Prev **8:**462-3.

45. **Tsutsumi, R., H. Higashi, M. Higuchi, M. Okada, and M. Hatakeyama.** 2003. Attenuation of *Helicobacter pylori* CagA-SHP-2 signaling by interaction between CagA and C-terminal Src kinase. J Biol Chem **278:**3664-70.

46. **Tsutsumi, R., A. Takahashi, T. Azuma, H. Higashi, and M. Hatakeyama.** 2006. Focal adhesion kinase is a substrate and downstream effector of SHP-2 complexed with *Helicobacter pylori* CagA. Mol Cell Biol **26:**261-76.

47. **Yamamoto, S.** 2001. Stomach cancer incidence in the world. Jpn J Clin Oncol **31:**471.

48. **Yamaoka, Y., H. M. El-Zimaity, O. Gutierrez, N. Figura, J. G. Kim, T. Kodama, K. Kashima, and D. Y. Graham.** 1999. Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. Gastroenterology **117:**342-9.

Table S1: Complete Korean Collection

Strain	**Disease**	**Sex**	**Age**	***cagA* EPIYA motif**	***cagA* accession number**
K1-CA	Cancer	F	68	ABD	FJ458117
K2-CA	Cancer	F	64	ABD	
K3-CA	Cancer	F	65	ABD	FJ458118
K4-CA	Cancer	F	37	ABD	
K5-CA	Cancer	M	70	ABD	
K6-CA	Cancer	F	45	ABD	
K7-CA	Cancer	M	56	ABD	
K8-CA	Cancer	M	56	ABD	
K9-CA	Cancer	M	58	ABD	FJ458119
K10-CA	Cancer	M	52	ABD	
K11-CA	Cancer	M	68	ABD	
K12-G	Gastritis	F	52	ABD	
K13-CA	Cancer	M	38	ABD	
K14-CA	Cancer	F	78	ABD	
K15C-CA	Cancer	F	66	ABD	FJ458120
K16-CA	Cancer	M	48	ABD	
K17-CA	Cancer	F	56	ABD	
K18-CA	Cancer	M	64	ABD	
K19-CA	Cancer	F	86	ABD	
K20-CA	Cancer	M	48	ABD	
K21-CA	Cancer	M	44	ABD	
K22-GU	Gastric Ulcer	M	42	ABD	

K23-DU	Duodenal Ulcer	M	47	ABD	
K24-DU	Duodenal Ulcer	M	38	ABD	FJ458121
K25-DU	Duodenal Ulcer	M	44	ABD	
K26-DU	Duodenal Ulcer	M	20	ABD	
K27-DU	Duodenal Ulcer	M	47	ABD	
K28-DU	Duodenal Ulcer	M	28	ABD	
K29-DU	Duodenal Ulcer	F	57	ABD	
K30-DU	Duodenal Ulcer	F	61	ABCCC	FJ458122
K31-DU	Duodenal Ulcer	M	33	ABD	
K32-DU	Duodenal Ulcer	F	41	ABD	
K33-DU	Duodenal Ulcer	F	31	ABC	FJ458123
K34-DU	Duodenal Ulcer	M	43	ABD	FJ458124
K35-DU	Duodenal Ulcer	F	56	ABD	
K36-DU	Duodenal Ulcer	M	46	ABD	
K37-DU	Duodenal Ulcer	M	61	ABD	
K38-DU	Duodenal Ulcer	F	39	ABC	
K39-DU	Duodenal Ulcer	M	59	ABD	
K40-DU	Duodenal Ulcer	M	53	ABD	
K41-DU	Duodenal Ulcer	M	55	ABD	
K42-DU	Duodenal Ulcer	F	48	ABD	
K43-DU	Duodenal Ulcer	M	70	ABD	
K44-DU	Duodenal Ulcer	M	42	ABD	
K45-DU	Duodenal Ulcer	M	22	ABD	FJ458125
K46-DU	Duodenal Ulcer	F	61	ABD	

K47-DU	Duodenal Ulcer	F	72	ABD	
K48-DU	Duodenal Ulcer	F	41	ABD	
K49-DU	Duodenal Ulcer	M	33	ABC	
K50-DU	Duodenal Ulcer	M	35	ABD	
K51-GU	Gastric Ulcer	M	54	ABD	
K52-GU	Gastric Ulcer	M	46	ABD	
K53-G	Gastritis	M	60	ABD	
K54-G	Gastritis	F	58	ABD	FJ458126
K55-GU	Gastric Ulcer	F	57	ABD	
K56-G	Gastritis	F	48	ABD	
K57-G	Gastritis	F	63	ABD	
K58-G	Gastritis	F	61	ABD	
K59-G	Gastritis	M	48	BBD	FJ458127
K60-G	Gastritis	M	53	ABC	
K61-GU	Gastric Ulcer	F	57	*	
K62-GU	Gastric Ulcer	F	65	*	
K63-GU	Gastric Ulcer	M	59	*	
K64-G	Gastritis	F	61	ABCC	FJ458128
K65-G	Gastritis	M	49	ABD	
K66-G	Gastritis	M	43	ABC	
K67-DU	Duodenal Ulcer	F	57	ABD	
K68-GU	Gastric Ulcer	F	46	ABD	
K69-GU	Gastric Ulcer	M	63	ABD	FJ458129
K70-G	Gastritis	F	68	ABC	FJ458130

K71-G	Gastritis	M	54	ABD	
K72-GU	Gastric Ulcer	M	34	ABD	
K73-GU	Gastric Ulcer	M	72	ABD	
K74-G	Gastritis	F	52	ABD	
K75-G	Gastritis	F	24	ABD	
K76-G	Gastritis	F	55	ABD	
K77-G	Gastritis	F	37	ABD	
K78-G	Gastritis	M	36	AABD	FJ458131
K79-GU	Gastric Ulcer	F	84	ABD	
K80-CA	Cancer	F	61	ABD	
K81-GU	Gastric Ulcer	M	47	ABD	
K82-G	Gastritis	M	39	ABD	FJ458132
K83-G	Gastritis	F	75	ABD	
K84-G	Gastritis	M	48	ABD	
K85-G	Gastritis	F	28	BD	
K86-G	Gastritis	M	37	ABCC	
K87-G	Gastritis	F	52	ABD	
K88-G	Gastritis	F	69	ABD	
K89-GU	Gastric Ulcer	M	38	ABD	
K90-GU	Gastric Ulcer	M	51	ABD	
K91-GU	Gastric Ulcer	M	82	ABD	
K92-GU	Gastric Ulcer	M	41	ABD	
K93-DU	Duodenal Ulcer	F	37	ABC	FJ458133
K94-GU	Gastric Ulcer	M	65	ABD	

K95-CA	Cancer	F	41	ABD	
K96-G	Gastritis	F	47	ABD	
K97-GU	Gastric Ulcer	M	51	ABD	
K98-DU	Duodenal Ulcer	M	23	ABD	
K99-G	Gastritis	F	54	ABD	
K100-GU	Gastric Ulcer	M	46	ABD	
K101-GU	Gastric Ulcer	F	61	ABD	
K102-DU	Duodenal Ulcer	M	38	ABD	
K103-G	Gastritis	M	32	ABD	
K104-CA	Cancer	M	46	ABD	
K105-GU	Gastric Ulcer	M	71	ABD	
K106-DU	Duodenal Ulcer	M	14	ABD	
K107-DU	Duodenal Ulcer	M	26	ABD	
K108-GU	Gastric Ulcer	M	62	ABD	
K109-G	Gastritis	M	40	ABD	
K110-GU	Gastric Ulcer	M	81	ABCC	FJ458134
K111-DU	Duodenal Ulcer	F	36	ABD	
K112-G	Gastritis	M	57	ABD	
K113-G	Gastritis	M	29	ABD	
K114-DU	Duodenal Ulcer	F	47	ABC	
K115-G	Gastritis	F	82	ABC	FJ458135
K116-G	Gastritis	M	59	ABD	
K117-G	Gastritis	F	21	ABD	FJ458136
K118-CA	Cancer	F	67	ABD	

K119-DU	Duodenal Ulcer	M	31	ABD	FJ458137
K120-G	Gastritis	F	41	ABC	
K121-GU	Gastric Ulcer	M	42	ABD	
K122-DU	Duodenal Ulcer	M	60	ABD	
K123-G	Gastritis	M	76	ABD	FJ458138
K125-G	Gastritis	M	59	ABD	
K126-GU	Gastric Ulcer	M	69	ABD	
K127-GU	Gastric Ulcer	M	71	ABD	
K128-GU	Gastric Ulcer	M	58	ABC	
K131-G	Gastritis	F	61	ABD	FJ458139
K136-G	Gastritis	F	52	*	
K137-G	Gastritis	F	62	*	
K138-GU	Gastric Ulcer	F	21	*	
K139-GU	Gastric Ulcer	F	49	*	
K140-G	Gastritis	F	49	*	
K141-G	Gastritis	M	57	*	
K142-GU	Gastric Ulcer	M	65	*	
K143-GU	Gastric Ulcer	F	71	*	
K144-GU	Gastric Ulcer	F	53	*	
K145-GU	Gastric Ulcer	M	62	*	
K146-G	Gastritis	M	40	ABD**	FJ458140
K147-GU	Gastric Ulcer	M	62	*	
K148-GU	Gastric Ulcer	M	37	*	
K149-GU	Gastric Ulcer	M	71	*	

K150-G	Gastritis	F	26	ABD	
K151-GU	Gastric Ulcer	M	65	ABD	FJ458141
K152-G	Gastritis	F	62	ABD	
K153-GU	Gastric Ulcer	M	42	ABD	
K154-G	Gastritis	F	55	ABCCCC	FJ458142
K155-DU	Duodenal Ulcer	N/A	N/A	ABD	
K156-DU	Duodenal Ulcer	F	47	ABD	
K157-G	Gastritis	M	43	ABD	
K158-G	Gastritis	M	60	ABD	
K159-G	Gastritis	F	35	ABD	
K160-DU	Duodenal Ulcer	M	30	ABD	
K161-G	Gastritis	F	65	ABD	
K162-G	Gastritis	F	63	ABD	
K163-G	Gastritis	F	66	ABD	
K164-G	Gastritis	M	43	ABD	
K165-G	Gastritis	M	28	ABD	
K166-G	Gastritis	F	38	ABC	
K167-G	Gastritis	F	27	ABD	
K169-G	Gastritis	F	47	ABD	
K170-G	Gastritis	F	41	ABD**	FJ458143
K171-CA	Cancer	F	72	ABD	FJ458144
K172-G	Gastritis	F	31	ABCC	FJ458145
K173-G	Gastritis	F	45	ABD	FJ458146
K174-G	Gastritis	N/A	N/A	ABD	

K175-G	Gastritis	F	41	ABD	
K176-DU	Duodenal Ulcer	N/A	N/A	ABD	
K177-G	Gastritis	F	39	ABD	
K178-G	Gastritis	F	40	ABD	
K179-G	Gastritis	F	38	ABCCC	FJ458147
K180-G	Gastritis (polyps)	F	50	ABD	
K181-DU	Duodenal Ulcer	F	57	ABD	
K182-DU	Duodenal Ulcer	N/A	N/A	ABD	
K183-G	Gastritis	M	40	ABD	
K184-G	Gastritis	F	55	ABD	
K185-G	Gastritis	F	52	ABD	
K186-G	Gastritis	M	41	ABD	
K188-G	Gastritis (IM)	F	43	ABD	
K190-G	Gastritis	M	61	ABC	
K192-DU	Duodenal Ulcer	F	61	AABD	FJ458148
K193-G	Gastritis	F	50	ABD	
K194-GU	Gastric Ulcer	M	48	*	
K195-GU	Gastric Ulcer	F	48	ABD	FJ458149
K196-G	Gastritis	F	50	ABD	
K197-G	Gastritis	F	45	ABD	
K198-GU	Gastric Ulcer	F	56	ABC	FJ458150
K199-GU	Gastric Ulcer	M	50	ABD	
K200-GU	Gastric Ulcer	M	63	ABD	
K201-GU	Gastric Ulcer	M	55	ABD	

K202-GU	Gastric Ulcer	M	41	ABD	
K203-G	Gastritis	F	55	ABD	
K204-GU	Gastric Ulcer	F	63	ABD	
K205-GU	Gastric Ulcer	F	57	ABD	
K206-GU	Gastric Ulcer	F	51	ABD	
K207-G	Gastritis	M	39	ABD	
K208-G	Gastritis	F	56	ABD	FJ458151
K209-G	Gastritis	F	24	ABD	
K210-G	Gastritis	F	61	ABD	
K211-G	Gastritis	F	54	ABD	
K212-G	Gastritis	F	45	ABD	
K216-G	Gastritis	F	67	ABD	
K217-G	Gastritis	M	77	ABD	
K218-G	Gastritis	F	62	ABD	
K219-G	Gastritis	M	40	ABD	FJ458152
K220-DU	Duodenal Ulcer	N/A	N/A	ABD	
K221-G	Gastritis	M	37	BD	
K222-GU	Gastric Ulcer	M	41	ABD	
K223-G	Gastritis	F	25	ABD	FJ458153
K224-G	Gastritis	F	35	ABD	
K225-DU	Duodenal Ulcer	F	60	ABC	
K226-GU	Gastric Ulcer	M	42	ABD	
K227-G	Gastritis	F	31	ABD	
K228-GU	Gastric Ulcer	M	54	ABD	

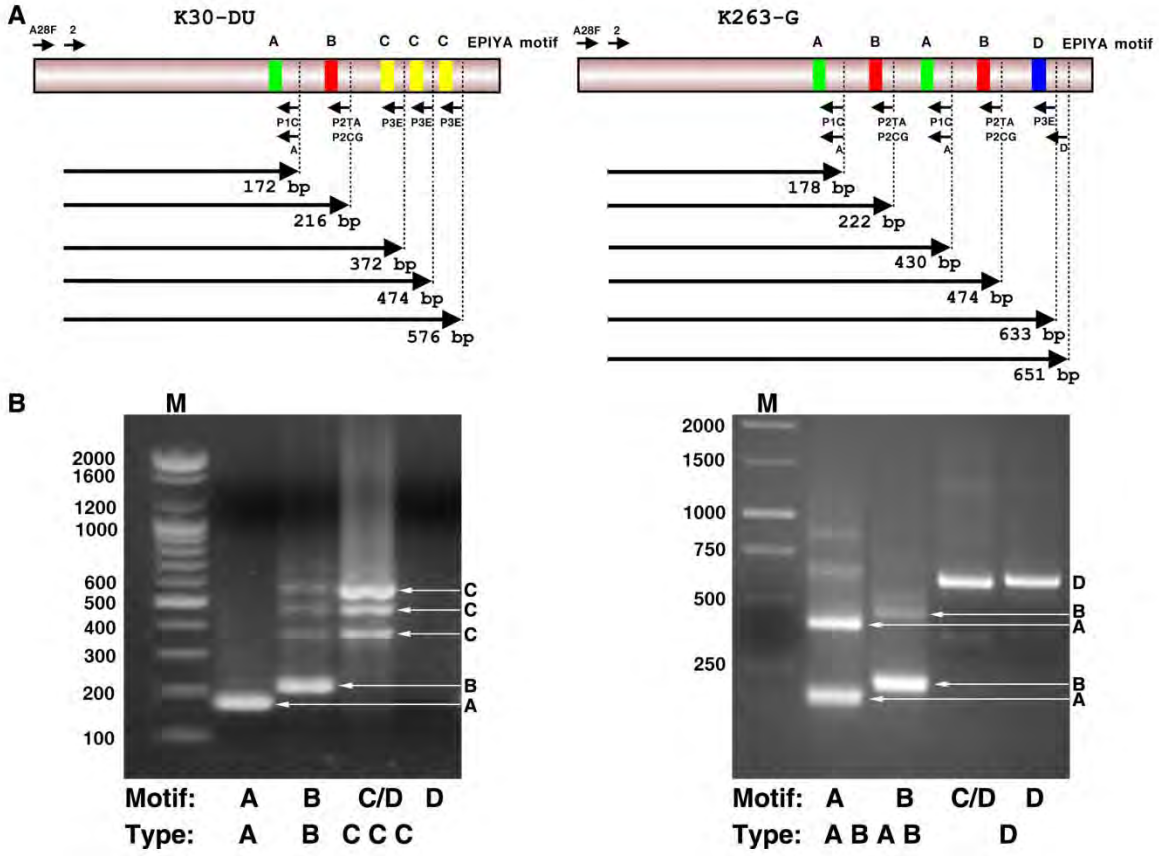
K229-GU	Gastric Ulcer	M	62	ABD	
K230-G	Gastritis (IM)	M	56	ABD	
K231-GU	Gastric Ulcer	M	42	ABD	
K232-G	Gastritis	F	56	ABD	
K233-G	Gastritis	M	38	ABD	
K234-DU	Duodenal Ulcer	M	41	ABD	
K235-G	Gastritis	F	50	ABD	
K236-G	Gastritis	F	64	ABD	
K237-G	Gastritis	F	48	ABD	
K238-DU	Duodenal Ulcer	M	55	ABD	
K239-G	Gastritis	M	46	ABD	
K240-G	Gastritis	F	41	ABD	
K241-G	Gastritis	M	41	ABD	
K242-G	Gastritis	M	78	ABD	
K243-G	Gastritis	F	60	BC	
K244-DU	Duodenal Ulcer	N/A	N/A	ABD	
K245-G	Gastritis	M	19	ABD	
K246-G	Gastritis	F	40	ABD	
K247-G	Gastritis	F	56	ABD	
K248-G	Gastritis	M	58	ABD	
K249-GU	Gastric Ulcer	F	48	ABC	
K250-G	Gastritis	F	53	ABD	
K251-DU	Duodenal Ulcer	M	70	ABD	
K253-DU	Duodenal Ulcer	F	61	ABC	

K254-G	Gastritis	M	54	ABD	
K255-G	Gastritis	F	74	ABD	FJ458154
K256-G	Gastritis	F	51	ABD	
K257-CA	Cancer	M	64	ABD	
K258-CA	Cancer	M	68	ABD	FJ458155
K259-CA	Cancer	M	44	ABD	FJ458156
K260-CA	Cancer	M	58	ABD	FJ458157
K261-CA	Cancer	F	48	ABD	FJ458158
K262-G	Gastritis	F	56	ABC	FJ458159
K263-G	Gastritis	M	59	ABABD***	FJ458160
K264-DU	Duodenal Ulcer	M	32	ABD	FJ458161
K265-DU	Duodenal Ulcer	M	42	ABD	FJ458162
K266-G	Gastritis	F	34	ABD	FJ458163

*Indeterminate in genotyping assay
**-B motif's proline is replaced with a serine, ESIYA, therefore classified as other
***-ABABD second -B motif's proline is replaced with leucine, ELIYA, therefore classified as other

Figure S1: Genotyping of the complex cagA *variable-EPIYA motifs.* A. Schematic representation of the variable region of K30-DU CagA and K263-G CagA is depicted on the left and right, respectively. The annealing positions (small arrows) and names of the primers used in this study and the expected sizes of the amplified specific EPIYA motif products according to the DNA sequence are shown. B. PCR amplicons of K30-DU and K263-G are shown using the forward primer 2 and the reverse primers A, P2TA and P2CG (equimolar mixture), P3E, or D. M designates the size markers (in base pairs). Type indicates resulting EPIYA motif identified.

Figure S1: Genotyping of the complex cagA *variable EPIYA motifs*



Chapter Three

An Epidemiological Link Between Gastric Disease and Polymorphisms in VacA and CagA

Published as: **Jang, S.[*], K.R. Jones[*], C.H. Olsen, Y.M. Joo, , Y.J. Yoo, I.S. Chung, J.H. Cha, and D. S. Merrell.** 2010. Epidemiological link between gastric disease and polymorphism in VacA and CagA. *J Clin Microbiol* 48:959-68.

[*] These authors contributed equally to this work.

The work presented in this chapter is the sole work of K. R. Jones with the following exceptions: Y.M Joo and S. Jang performed the *vacA* genotyping, Y.J. Yoo and J.H. Cha assisted with experimental design, H.S. Lee originally isolated *H. pylori* strains. I.S. Chung performed the biopsies and supplied the diagnoses, and C.H. Olsen assisted with the statistical analysis.

Abstract

Gastritis, peptic ulcer disease and gastric cancer are a few of the diverse disease manifestations that have been shown to be associated with infection by *Helicobacter pylori.* Why some individuals develop more severe forms of disease remains largely unknown. In this study, 225 South Korean strains were genotyped for *vacA* and then analyzed to determine if particular genotypes varied across disease state, sex, or *cagA* allele. Of these strains, 206 strains carried an s1/i1/m1 allele, 11 strains carried an s1/i1/m2 allele, and 8 strains carried an s1/i2/m2 allele. By using Fisher's exact test, a

statistical association between variations in the *cagA* and *vacA* alleles was identified (P= 0.0007), and by using log linear modeling, this variation was shown to affect the severity of disease outcome (P= 0.027). Additionally, we present evidence that variation within the middle region of VacA contributes significantly to the distribution of *vacA* alleles across gender (P=0.008) as well as the association with disease outcome (P= 0.011). In this South Korean population, the majority of *H. pylori* strains carry the *vacA* s1/i1/m1 allele and the CagA EPIYA-ABD allele. These facts may contribute to the high incidence of gastric maladies, including gastric cancer.

Introduction

Helicobacter pylori is a Gram-negative bacterium (27) that chronically infects the gastric mucosa of over half of the world's population (28, 50) and is associated with the development of chronic gastritis, gastric and duodenal ulcers, and gastric adenocarcinoma and mucosa associated lymphoid tissue (MALT) lymphoma (6, 10, 14, 40). Given *H. pylori*'s high prevalence, chronic persistence, and its link to gastric cancer, it is no wonder that gastric cancer is the second most common cause of cancer associated death (35) with the mortality rate being especially high in East Asian countries such as China, Japan, and Korea (18).

H. pylori strains express various toxins that enable the bacteria to cause host cell damage and interact with the host immune response. Included among these toxins are the cytotoxin associated gene A (CagA) and the vacuolating cytotoxin (VacA; 33). CagA has emerged as a major contributor to disease severity, and there is a direct link between presence of CagA and increased cancer risk (7, 16). CagA induces various pathologic

changes by modulating host cell signaling pathways, primarily after tyrosine phosphorylation at the EPIYA motif (17, 19-21, 34, 44, 54). Interestingly, CagA is polymorphic, and the distribution of EPIYA motif combinations differs geographically.

VacA is another important toxin that is produced and secreted by all *H. pylori* strains (4, 11), and was previously shown to have various modes of action (12, 15, 26, 33, 38, 48, 49, 52, 56). Like CagA, VacA has been shown to contain a number of polymorphisms. Currently, three polymorphic regions of *vacA* have been identified: signal (s), intermediate (i), and middle (m) regions. Each of these polymorphic regions has two main types that divide them further into type 1 and type 2 (3, 42). The s region encodes the N-terminal signal sequence (29, 41), and polymorphisms in the s region affect anion channel-forming efficiency of the toxin (29); the s1 type has an increased ability to form membrane channels (29). Polymorphisms in the m region affect the cell tropism of the toxin (22); the m1 type of VacA shows toxicity to a broader range of cells than the m2 type (1, 37). The i region, which is located between the s and m regions and also displays two main polymorphisms (42). The i1 type of VacA has stronger vacuolating activity than the i2 type (42). Individually, the s1, i1, and m1 types have been shown to be associated with more severe forms of *H. pylori* induced disease (5, 42).

Recently, we presented molecular epidemiologic evidence that there is a significant association between the development of gastric cancer and infection with *H. pylori* strains carrying the EPIYA-ABD *cagA* genotype in South Korea, which has one of the highest rates of *H. pylori* colonization (51) and one of the highest rates of gastric cancer in the world (16, 46). Given the mounting body of evidence that indicates that *cagA* and *vacA* interact, herein we assess *vacA* polymorphisms across various *cagA*

alleles and in relation to disease development and we show a significant 3-way association between *vacA*, *cagA*, and disease.

Material and Methods

Bacterial Strains and Culture Conditions

The South Korean *H. pylori* clinical isolates used in this study were previously described in Jones *et al.* (23), and included 115 gastritis isolates, 55 gastric ulcer isolates, 54 duodenal ulcer isolates, and 30 gastric cancer isolates with epidemiological data on age and gender. *H. pylori* stocks preserved at -80°C were grown and expanded on antibiotic-supplemented horse blood agar plates under microaerophilic conditions created by an Anoxomat evacuation/replacement system (Spiral Biotech, Norwood, MA), as previously described (9, 23).

vacA *Genotyping*

The primers used for *vacA* genotyping and sequencing of the i region are listed in Table 4. Chromosomal DNA of all 254 *H. pylori* strains was isolated using the Easy-DNA kit (Invitrogen, Carlsbad, CA). Four individual PCRs were performed to identify the *vacA* genotype of each strain (Fig. 8). The s region was identified by amplifying with primers VA1-F and VA1-R. The s1 region produced a 259 base pair (bp) amplicon where as the s2 region produced a 286 bp amplicon (3). The m1 and m2 regions were determined by amplifying with primers VAG-F and VAG-R yielding a 567 bp and 642 bp product, respectively (4). The i region was genotyped using two independent PCR reactions using a universal forward primer (VacF1) and different i region type-specific

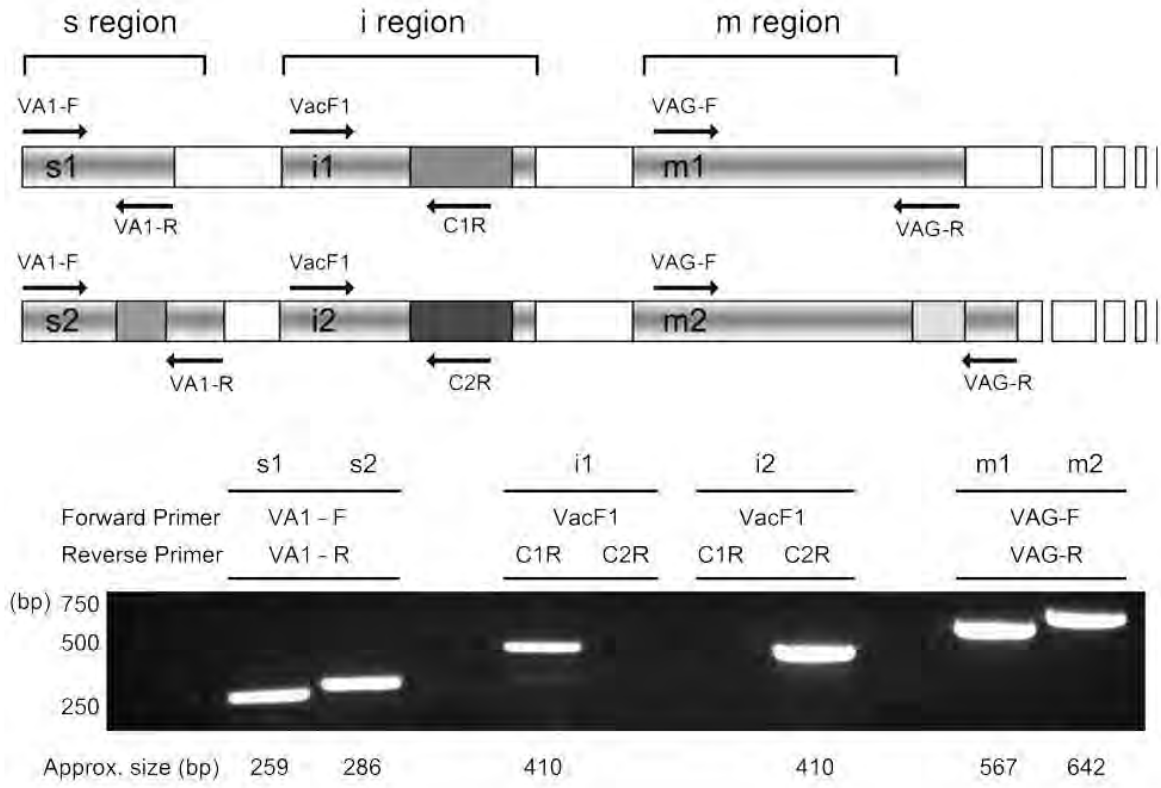
Table 4: Primer sequences

Primer	Sequence (5'-3')	Reference
VA1-F[a]	ATGGAAATACAACAAACACAC	(3)
VA1-R[a]	CTGCTTGAATGCGCCAAAC	(3)
VAG-F[b]	CAATCTGTCCAATCAAGCGAG	(4)
VAG-R[b]	GCGTCAAAATAATTCCAAGG	(4)
C1R[c]	TTAATTTAACGCTGTTTGAAG	(42)
C2R[c]	GATCAACGCTCTGATTTGA	(42)
VacF1[c,d]	GTTGGGATTGGGGGAATGCCG	(42)
VacR9[d]	TGTTTATCGTGCTGTATGAAGG	(42)

[a] This primer pair was used to genotype the s region.
[b] This primer pair was used to genotype the m region.
[c] These primers were used to genotype the i region.
[d] These primers were used for the i region sequencing.

Figure 8: Genotyping of the vacA *polymorphic regions.* (Top) Schematic representation of the *vacA* alleles, where an s1/i1/m1 allele is shown on the top and an s2/i2/m2 allele is depicted on the bottom. The annealing positions (arrows) and names of the primers used in this study are shown. (Bottom) PCR amplicons of each polymorphic region using the primers listed above the gel are depicted. The approximate size of the amplicon is listed below each band.

Figure 8: Genotyping of the vacA *polymorphic regions*



reverse primers (C1R and C2R) as described by Rhead *et al.* (42). C1R and C2R specifically anneal with the i1 and i2 *vacA* allele, respectively (42).

Sequencing of 60 isolates from patients suffering from gastritis (24 isolates), duodenal ulcers (10 isolates), gastric ulcers (8 isolates), and cancer (18 isolates) was conducted to identify specific amino acid changes in the i1 allele. Sanger dideoxy sequencing was performed at the Uniformed Services University of the Health Science Biomedical Instrumentation Center (Bethesda, MD). The resulting DNA sequences were analyzed using Vector NTI version 9.1 (Invitrogen, Carlsbad, CA) and Sequencher 4.5 (Gene Codes Corp., Ann Arbor, MI).

Statistical Analysis

The Fisher's exact test was used to analyze the association between the *vacA* allele and disease state or *cagA* allele (based on the EPIYA motif). Log linear modeling was used to assess higher order associations. We fit a saturated model using categorical variables representing *vacA* genotype, *cagA* genotype, disease state, gender, and age categories. A backward selection algorithm identified higher-order associations among these variables, which were statistically significant at the 5% level. Data were analyzed using SPSS version 14 or 16 software (SPSS Inc., Chicago, IL) or SAS version 9.1 software (SAS Institute Inc., Cary, NC).

Nucleotide Sequence Accession Numbers

The sequences for the i1 region of *vacA* from 60 strains have been deposited in GenBank under accession numbers GQ338184 to GQ338243 (see Table S2 in the supplemental material).

Results

*Sample Acquisition/*vacA *Genotyping*

The strains used for this study were previously used for characterization of the distribution of *cagA* alleles (23), and represent 260 strains obtained from patients presenting with gastric maladies; 254 of those strains have contain complete epidemiological data (see Table S2 in the supplemental material). The mean patient age was 51 years, with an age range of 14 to 86 years (Table 5). Within this population 49.6% (126 patients) were female, with an age range of 21 to 86 years and a mean age of 52 years, and 50.4% (128 patients) were male, with a mean age of 50 years and an age range of 14 to 82 years. Of these samples, 11.8% were from patients with cancer, 42.9% were from patients with peptic ulcer disease (21.7% gastric ulcers and 21.3% duodenal ulcers), and 45.3% were from patients with gastritis (23).

Four different PCRs were conducted for each strain in order to genotype *vacA* (Fig. 8 and see Table S2 in the supplemental material). The s region was identified and differentiated by amplifying with the VA1-F and VA1-R primers (3) and the m region was determined by amplifying with the primers VAG-F and VAG-R (4). The i region was genotyped using two independent PCR reactions using a universal forward primer

Table 5: vacA *genotyped isolates and disease state*

	Total	*vacA* genotyped Isolates	Disease state of *vacA* genotyped isolates			
			Gastritis	Gastric Ulcer	Duodenal Ulcer	Gastric Cancer
Overall Total	254	225	103	43	49	30
Age Range	14-86	14-86	19-82	34-84	14-72	37-86
Mean	51	50	49	55	45	58
Females	126	111	68	10	19	14
Age Range	21-86	21-86	21-82	46-84	31-72	37-86
Mean	52	52	49	57	51	61
Males	128	114	35	33	30	16
Age Range	14-82	14-82	19-78	34-82	14-70	38-70
Mean	50	49	47	54	41	55
s1/i1/m1		206	95	41	42	28
Age Range		14-86	19-78	34-84	14-70	37-86
Mean		50	48	55	43	58
Females		96	61	8	14	13
Age Range		21-86	21-75	46-84	31-61	37-86
Mean		51	49	57	48	60
Males		110	34	33	28	15
Age Range		14-82	19-78	34-82	14-70	38-70
Mean		49	47	54	41	56

s1/i1/m2	11	4	1	4	2
Age Range	38-82	38-82	63	41-57	46-68
Mean	54	54	N/A	51	56
Females	9	4	1	3	1
Age Range	38-82	38-82	63	48-57	68
Mean	56	54	N/A	54	NA
Males	2	0	0	1	1
Age Range	41-44	0	0	41	46
Mean	43	0	0	N/A	N/A
s1/i2/m2	8	4	1	3	0
Age Range	38-72	38-68	56	61-72	0
Mean	57	51	N/A	65	0
Females	6	3	1	2	0
Age Range	38-72	38-68	56	61-72	0
Mean	58	53	N/A	67	0
Males	2	1	0	1	0
Age Range	43-61	43	0	61	0
Mean	52	N/A	0	N/A	0

*Total number of samples only includes the 254 that contained complete epidemiological data for age and gender.

•There was a statistical association between the m allele and gender (P=0.0233).

(VacF1) and one of two i region type-specific reverse primers (C1R for the i1 type and C2R for the i2 type) as described by Rhead *et al.* (42).

The distribution of *vacA* polymorphisms is shown in Table 5. Of the 254 strains containing complete epidemiological data, 225 strains were successfully genotyped for the *vacA* allele (Table 5 and see Table S2 in the supplemental material). The strains that were not successfully genotyped for *vacA* failed to yield PCR products or gave incorrectly sized bands and thus were not further analyzed. The genotyped strains were obtained from patients with a mean age of 50 years and an age range of 14 to 86 years. These patients included 111 females (49.3%) with a mean age of 52 years and an age range of 21 to 86 years, and 114 males (50.7%) with a mean age of 49 years and an age range of 14 to 82 years. Of these 225 strains, 206 strains (91.6%) carried an s1/i1/m1 *vacA* allele (mean patient age of 50 years and an age range of 14 to 86 years). Of the strains carrrying the s1/i1/m1 allele, 96 (46.6%) were from female patients (mean age of 51 years and an age range of 21 to 86 years), and 110 (53.4%) were from male patients (mean age of 49 years and an age range of 14 to 82 years). Eleven strains (4.9%) carried for an s1/i1/m2 *vacA* genotype and were obtained from patients with a mean age of 54 years and an age range of 38 to 82 years. Of these strains, 9 (81.8%) were obtained from female patients (mean age of 56 years and an age range of 38 to 82 years), and 2 (18.2%) were obtained from male patients (mean age of 43 years and an age range of 41 to 44 years). Eight strains (3.9%) carried the s1/i2/m2 *vacA* allele and were obtained from patients with a mean age of 57 years and an age range of 38 to 72 years. Six of these strains (75.0%) were obtained from female patients (mean age of 58 years with an age range of 38 to 72 years), and 2 (25.0%) were obtained from male patients (mean age of

52 years and an age range of 43 to 61 years). Of note, neither s2 nor s1/i2/m1 *vacA* alleles were found.

Distribution of vacA *Allele and Gender*

Statistical analysis of allele distributions showed a significant association between the *vacA* allele and gender (P=0.0233). Patients that carried non s1/i1/m1 strains were 4.3 times more likely to be female than male. This difference is likely driven by the m region, since the distribution of any combination that contained the m region when compared to gender was statistically significant (P=0.005 for s/m and P=0.0233 for m/i), whereas the combination lacking m (P=0.1672 for s/i) was not significant. Moreover, when the distribution of polymorphisms within each region was analyzed alone versus gender, only the distribution of the m polymorphisms was significant (P=0.008). In fact, if a patient contained an m2 allele they were 3.75 times more likely to be female than male. This fact combined with the fact that the m1 allele appears to affect toxicity to a larger variety of cells (22), may contribute to the finding that males are 1.5 to 2.5 times more likely to develop gastric cancer than females (reviewed in 43).

Associations among vacA*,* cagA*, and Disease*

Given the diversity of the identified roles of the VacA toxin, we assessed whether the distribution of the *vacA* alleles had a direct impact on disease state. First, the individual regions were assessed for their impact on disease development. The distribution of polymorphisms in the m region and i region among disease state was not statistically significant (P=0.5397 and P=0.7399, respectively), and the distribution of

polymorphisms in the s region in relationship to disease state could not be determined because only the s1 allele was found within this population. Statistical analysis for two way associations using SASS software showed no statistical association between the distribution of the *vacA* alleles and disease state (P= 0.7499). However, log linear modeling taking into account age and gender, did reveal a two way association between *vacA* allele and disease only in the East Asian (EPIYA-ABD) strains (P=0.030). The majority of East Asian EPIYA-ABD CagA strains carrying non s1/i1/m1 *vacA* alleles were associated with duodenal ulcers. Conversely, strains carrying non s1/i1/m1 *vacA* alleles and any other genotype of CagA were associated with gastritis. A complete breakdown of the *vacA* allele and disease state is provided in Table 5.

VacA was previously suggested to interact synergistically with the *H. pylori* virulence factor, CagA (2, 36). Thus, we next analyzed whether there was any association between the distribution of *vacA* alleles, the distribution of *cagA* alleles, and disease state. Of the 225 strains that were genotyped for the *vacA* allele, 224 of these strains had previously been genotyped for the *cagA* allele. Of these strains, 199 isolates (88.8%) can be classified as East Asian (encoding for an EPIYA-D motif), and 25 isolates (11.2%) were determined to be classified as Western strains (encoding for at least one EPIYA-C motif; see Table S2 in the supplemental material). Eight East Asian strains carried an EPIYA motif other than a defined –ABD motif, either incomplete or containing the addition of one or more motifs, including -AABD, -BD, -BBD, -ABAB*D, as well as -AB*D, where a mutation within the EPIYA-B motif is designated by the asterick. Based on these distributions, the strains were subdivided based on the

Figure 9: Schematic depiction of the distribution of the vacA *genotypes stratified by disease state and* cagA *allele within this South Korean population.* Shown is the distribution of *vacA* genotypes within the four different disease states: gastritis, gastric ulcers, duodenal ulcers, and cancer. The shaded portions within the disease state and *vacA* genotype subgroupings correspond to the isolates that carry a *cagA* EPIYA-ABD motif.

Figure 9: Schematic depiction of the distribution of the vacA *genotypes stratified by disease state and* cagA *allele within this South Korean population*

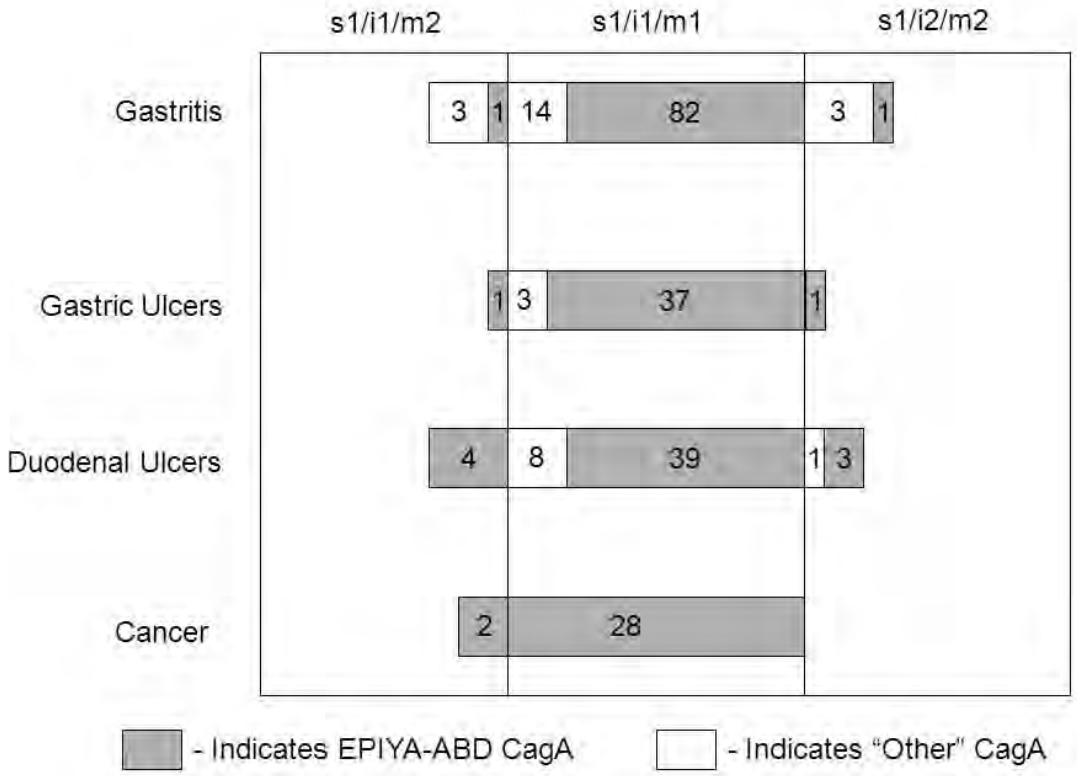


Table 6: vacA *genotype and* cagA *allele*.

	***vacA* genotyped Isolates**	***vacA* and *cagA* genotyped isolates**	***cagA* Allele**	
			EPIYA -ABD	**Other alleles**
Overall Total	225	224	191	33
Age Range	14-86	14-86	14-86	28-82
Mean	50	50	50	49
Females	111	111	90	21
Age Range	21-86	21-86	21-86	28-82
Mean	52	52	52	50
Males	114	113	101	12
Age Range	14-82	14-82	14-82	33-81
Mean	49	49	49	49
s1/i1/m1	206	205	180	25
Age Range	14-86	14-86	14-86	28-81
Mean	50	50	50	47
Females	96	96	82	14
Age Range	21-86	21-86	21-86	28-61
Mean	51	51	52	46
Males	110	109	98	11
Age Range	14-82	14-82	14-82	33-81
Mean	49	49	49	49

s1/i1/m2	11	11	8	3
Age Range	38-82	38-82	41-68	38-82
Mean	54	54	52	58
Females	9	9	6	3
Age Range	38-82	38-82	41-68	38-82
Mean	56	56	56	58
Males	2	2	2	0
Age Range	41-44	41-44	41-44	N/A
Mean	43	43	43	N/A
s1/i2/m2	8	8	3	5
Age Range	38-72	38-72	54-72	38-68
Mean	57	57	62	53
Females	6	6	2	4
Age Range	38-72	38-72	54-72	38-68
Mean	58	58	63	56
Males	2	2	1	1
Age Range	43-61	43-61	61	43
Mean	52	52	N/A	N/A

* indicates any other genotype besides EPIYA-ABD, including Western strains and EPIYA-AABD, -BD, -BBD, -ABAB**D, as well as -AB**D where a mutation within the EPIYA-B motif is designated by the **

[r]There is a statistical two way association between the distribution of *cagA* alleles and the distribution of *vacA* alleles (P=0.0007).

presence of a complete CagA EPIYA-ABD motif versus all other EPIYA motifs, yielding 33 isolates that were determined to have "other genotypes." When the distribution of the *vacA* alleles (s1/i1/m1, s1/i1/m2, si/i2/m2) was assessed among the distribution of *cagA* alleles (EPIYA-ABD versus all other genotypes), a strong two way association was identified (P=0.0007; Fig. 9 and Table 6). People infected with strains carrying non-EPIYA-ABD *cagA* alleles were associated with two-times-higher probability of carrying the s1/i1/m2 *vacA* allele and 10-times-higher probability of carrying the s1/i2/m2 *vacA* allele than people infected with the EPIYA-ADB *cagA* allele (Table 6. When combinations of the regions were compared, the distribution of polymorphisms among the *cagA* alleles was statistically significant for every combination, s/m (P=0.0006), i/m (P=0.0007), and s/i (P=0.0019), and the distribution of individual regions of *vacA*, m (P=0.0019) and i (P=0.0019) versus the *cagA* allele were also statistically significant. Once again, the distribution of polymorphisms in the s region among *cagA* alleles could not be determined because only the s1 allele was found within this population. This indicates that each of these regions is important for the distribution of the *vacA* allele with *cagA* allele.

Given the strong correlation between *cagA* allele and disease state we previously observed (23), we next wondered if the *vacA* allele affected this distribution. Indeed, log linear modeling revealed a significant three-way association among *vacA* allele, *cagA* allele, and disease state (P=0.027). As with the case of gender, the distribution of any combination that contained the m region when assessed via the distribution of *cagA* allele and disease state was statistically significant (P=0.004 for s/m and P=0.025 for i/m), whereas the non m combination (P=0.586 for s/i) was not significant. Moreover, when

Table 7: vacA *allele,* cagA *allele and disease state*

	***vacA* and *cagA* genotyped isolates**	**Disease state of *vacA* genotyped isolates**			
		Gastritis	**Gastric Ulcer**	**Duodenal Ulcer**	**Gastric Cancer**
Overall Total	224	103	42	49	30
EPIYA -ABD	192	83	38	40	30
Other genotypes*	32	20	4	9	0
Females	111	68	10	19	14
EPIYA -ABD	90	57	8	11	14
Other genotypes*	21	11	2	8	0
Males	113	35	32	30	16
EPIYA -ABD	101	26	30	29	16
Other genotypes*	12	9	2	1	0
s1/i1/m1	205	95	40	42	28
EPIYA -ABD	180	81	37	34	28
Other genotypes*	25	14	3	8	0
Females	96	61	8	14	13
EPIYA -ABD	82	55	7	7	13
Other genotypes*	14	6	1	7	0
Males	109	34	32	28	15
EPIYA -ABD	98	26	30	27	15
Other genotypes*	11	8	2	1	0

s1/i1/m2	11	4	1	4	2
EPIYA -ABD	8	1	1	4	2
Other genotypes*	3	3	0	0	0
Females	9	4	1	3	1
EPIYA -ABD	6	1	1	3	1
Other genotypes*	3	3	0	0	0
Males	2	0	0	1	1
EPIYA -ABD	2	0	0	1	1
Other genotypes*	0	0	0	0	0
s1/i2/m2	8	4	1	3	0
EPIYA -ABD	3	1	0	2	0
Other genotypes*	5	3	1	1	0
Females	6	3	1	2	0
EPIYA -ABD	2	1	0	1	0
Other genotypes*	4	2	1	1	0
Males	2	1	0	1	0
EPIYA -ABD	1	0	0	1	0
Other genotypes*	1	1	0	0	0

* indicates any other genotype besides EPIYA-ABD, including Western strains and EPIYA-AABD, -BD, -BBD, -ABAB**D, as well as -AB**D where a mutation within the EPIYA-B motif is designated by the **

·There is a statistical three way association between the distribution of *cagA* alleles. the distribution of *vacA* alleles, and disease state (P=0.027).

the distribution of polymorphisms within each region was analyzed individually, only the distribution of the m polymorphism (P=0.011) was significant. A complete breakdown of the strains based on disease state, *vacA* allele, and *cagA* allele is provided in Table 7.

Sequence Analysis of the Intermediate 1 Region

The i region was previously suggested to be the best indicator of pathology caused by VacA, and the i1 type is more virulent than i2 (2, 42). Two specific amino acid sequences, a phenylalanine at position 178 in cluster A and a methionine at position 254 in cluster C (Fig. 10), have been identified as markers within the Taiwanese population and a particular amino acid substitution at position 231 in cluster B has been suggested to affect disease severity (45). Given these reasons, we determined the i1 type amino acid sequence for 60 strains (18 from cancer patients, 8 from gastric ulcer patients, 10 from duodenal ulcer patients, and 24 from gastritis patients; Fig. 10).

Sequence analysis of the i1 type in our strains revealed that a phenylalanine at position 178 in cluster A was present in 91.7% of the strains in the South Korean population. Additionally, the substitution of a methionine at position 254 in cluster C was well conserved in the SouthKorean population; 98.3% of strains contained this substitution.

Substitution of a glycine for a serine at position 231 in cluster B was previously suggested to be statistically linked to disease development within the Taiwanese population (45). However, at this site within strains among the South Korean population, neither the distribution of the amino acids with regard to disease (P=0.8082) nor the distribution of glycine as compared to any other amino acid with regards to disease

Figure 10: Amino acid alignment of the i1 type of VacA. This amino acid alignment is from 60 Korean strains of various disease states: 24 from gastritis (G) patients, 10 from duodenal ulcer (DU) patients, 8 from gastric ulcer (GU) patients, and 18 from gastric cancer (CA) patients. The abbreviations listed after the strain, correspond to the disease state of the strain. The three defined clusters of the i1 region, cluster A, B, and C, are indicated.

Figure 10: Amino acid alignment of the i1 type of VacA

```
                150       160       170       180       190       200       210       220       230       240       250
K1-CA   KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K2-CA   KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K3-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K4-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K5-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K6-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K7-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K8-CA   KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K9-CA   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K16-CA  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K17-CA  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K21-CA  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K80-CA  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K104-CA KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K258-CA KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K259-CA KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K260-CA KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVNFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEVSLYDGATLNLASNSVKLMGN
K261-CA KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K51-GU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K52-GU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K55-GU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K69-GU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSTDRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K72-GU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K81-GU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K128-GU KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K249-GU KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K23-DU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSYKDAANRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEKITSRENAEISLYDGATLNLASNSVKLMGN
K25-DU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K26-DU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K34-DU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K38-DU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K49-DU  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K93-DU  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K111-DU KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K182-DU KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K238-DU KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K57-G   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K58-G   KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K74-G   KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K77-G   KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K109-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K112-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSYKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K117-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K120-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K123-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K131-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K162-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K178-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDGADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K183-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K185-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K190-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K193-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K196-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K197-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K207-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K208-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K209-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSDKNAEISLYDGATLNLASNSVKLMGN
K248-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNISIDNFLEINNRVGSGAGRKARSTVLTLQASEGITSGKNAEISLYDGATLNLASNSVKLMGN
K255-G  KLSGLRNYTGGDLDVNMQKATLRLGQFNGNSFTSFKDSADRTTRVDFNAKNILIDNFLEINNRVGSGAGRKASSTVLTLQASEGITSNKNAEISLYDGATLNLASNSVKLMGN
K262-G  KLSGLRNFTGGDLDVNMQKATLRLGQFNGNSFTSYKDSADRTTRVNFNAKNISIENFVEINNRVGSGAGRKASSTVLTLQASEGITSSKNAEISLYDGATLNLASNSVKLNGN
                                          Cluster A                                   Cluster B                   Cluster C
```

(P=0.5214) was statistically significant. Additionally, there was no statistical significance when the distribution of glycine was determined with regards to individual disease states; gastritis (P=0.7871), gastric ulcer (P=0.4329), duodenal ulcer (P=0.7287) or cancer (P=0.2414). However, there was a statistical association between the distribution of the glycine amino acid and the distribution of *cagA* alleles (P=0.0318).

Sequence analysis of the i1 type revealed two additional amino acid polymorphisms across the South Korean population. At position 151, the majority of the strains carried a phenylalanine (37 strains) or a tyrosine (23 strains). This distribution of amino acids at this position was not statistically linked to disease (P=0.3886) or the distribution of *cagA* allele (P=0.6983). Additionally, polymorphism at position 196 leads to either a serine or a leucine at this position. The distribution of amino acids at this position had no statistical association with the distribution of *cagA* alleles (P=1.0000) or disease state (P=0.0669).

Discussion

The majority (92%) of the South Korean isolates analyzed in this study carried the s1/i1/m1 *vacA* allele, which was previously suggested to be the most virulent form of the toxin (3, 24, 30-32, 42). The fact that the majority of *H. pylori* strains in this population carry the most toxic form of both VacA and CagA, may explain the high rate of severe gastric disease among the South Korean population. When age and gender were taken into account, a two way association between the distribution of *vacA* alleles and disease state was found within the strains carrying EPIYA-ABD CagA. Non

s1/i1/m1 *vacA* alleles were associated with duodenal ulcers within the population carrying the East Asian EPIYA-ABD CagA, and gastritis within the population carrying any other genotype of CagA.

The distribution of the m alleles varied significantly across gender and the *cagA* allele and had a significant impact on the three-way association between the *cagA* allele and disease state. This suggests that the polymorphism within the m region is the major contributor to the association of the *vacA* allele, the *cagA* allele and disease state within this population. This is in concordance with a previously reported meta-analysis that found that the m1 region increased the risk for gastric cancer in Latin American (odds ratio [OR]=3.59) and African (OR=10.18) populations (47). The increased tropism of the m1 *vacA* allele (37) combined with the finding that the patients infected with *H. pylori* strains encoding the m2 allele are more likely to be female may explain why males are overall more likely to develop gastric cancer (reviewed in 43). To our knowledge, this is the first time that the m allele distribution has been linked to gender. To determine the role that the m allele has the association between gender and disease state, populations where the m2 allele is more prevalent, such as in regions of China (39) and Poland (25), should be analyzed. However, it should be noted that this region alone is not a good predictor of disease state for the South Korean population, since two of the strains that carried the m2 allele were from cancer patients (Table 5).

Previous work with Western strains suggested that the i region of *vacA* is the major determinant of vacuolating activity, and is the most important region for disease development (2, 13, 42). However, we found that within this population of predominantly East Asian strains, the i region was not a major determinant of disease

state. This may indicate that the i region is more important within the context of strains that express the Western *cagA* allele or that there are other factors that mask the importance of this region in East Asian isolates.

Three clusters in the i1 region where sequence difference occur within i1 have been reported: cluster A, B, and C (42). Amino acid substitutions (tyrosine to phenylalanine in cluster A and an asparagine to methionine in cluster C) are conserved and predicted to serve as a marker for Taiwanese VacA (45). Also, the substitution of a glycine for a serine at the ninth amino acid in cluster B was statistically linked to disease development in the Taiwanese population (45). The amino acid substitutions within cluster A and cluster C were also conserved in the South Korean isolates, indicating that these substitutions are likely a marker of East Asian VacA. No correlation between the ninth amino acid substitutions in cluster B and disease severity was identified for our South Korean population, which suggests this amino acid does not play a role in disease progression or is important in combination with another virulence factor.

Two additional positions within the i1 region that showed polymorphism were identified: positions 151 and 196. Neither the phenylalanine nor the tyrosine found at position 151 was linked to disease state or the distribution of *cagA* allele. While the distribution of the amino acids at position 196 had no statistical association with the *cagA* allele or disease, there was a trend toward significance: ~78% of strains from cancer patients and 100% of strains from gastric ulcers patients carry a serine at this position. This suggests that additional populations should be assessed to determine if a serine at position 196 has an impact on disease development and severity.

Previously reported studies have identified an association between *vacA* and *cagA* that appears to affect *H. pylori* toxicity and disease severity (55, 57). *Basso et al.* found that increasing numbers of CagA EPIYA-C motifs impacted cancer risk, and that i region polymorphisms of VacA were a major indicator for the development of peptic ulcers (5). Additionally, infection with strains carrying CagA and s1/m1 VacA results in highly active corpus gastritis (32), which is linked to development of gastric cancer (30-32). In this study, log linear modeling, taking into consideration age and gender, identified a two way association between the *vacA* allele and disease state for East Asian (CagA EPIYA-ABD) strains. Not surprisingly, since the majority of these strains carried both CagA EPIYA-ABD and *vacA* s1/i1/m1, the majority of cancer strains (28 out of 30) carry this combination. This suggests that the role of the *vacA* allele could differentially affect disease progression based on other virulence factors. This could be due to the finding that VacA acts as an immune modulator (15, 52) and perhaps changes the immune response to the immunogenic CagA. Like CagA, VacA was found to disorganize the cytoskeleton of gastric epithelial cells, leading to increased cell spreading and growth (38). Thus, this phenotype may help compensate for the presence of a less virulent *cagA* allele or synergistically contribute to severe gastric maladies in conjunction with East-Asian CagA. Evidence suggests that the combination of CagA and VacA may dampen the effect of each protein alone, possibly leading to increased survival of infected host cells (2). This, perhaps, occurs through CagA preventing VacA-induced apoptosis (36, 56) or by inhibiting the autophagy pathway induced by VacA (49).

CagA and VacA are the two best studied virulence factors of *H. pylori.* Interestingly, both toxins exhibit a high degree of polymorphism, and it is becoming

increasingly evident that these polymorphisms, alone and in concert, affect *H. pylori*-induced disease. Indeed, the finding that the majority of South Korean *H. pylori* strains carry the most toxic form of CagA and VacA may explain the reason for the high prevalence of gastric disease and mortality of patients with gastric cancer in South Korea. However, the reason why only a portion of the population develops gastric cancer still remains unclear. Other bacterial virulence factors as well as multiple host, dietary, and environmental factors have been indicated as being participants in *H. pylori*-induced disease (reviewed in 8 and 53). Further study is required to determine which factors are involved and what role they have in the development of *H. pylori*-induced gastric cancer.

Acknowledgements

We thank Jang Gi Cho for help with graphics, as well as Jeannette Whitmire and Beth Carpenter for critical reading of this manuscript. This work was supported by National Research Foundation of Korea grant funded by the South Korean Government (grant 2009-0073957) (to J.-H.C.), and by grant AI065529 from the NIAID (to DSM).

Contents of this manuscript are the sole responsibility of the authors and do not necessarily represent the official views of the NIH or the DOD. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. **Amieva, M. R., and E. M. El-Omar.** 2008. Host-bacterial interactions in *Helicobacter pylori* infection. Gastroenterology **134:**306-23.

2. **Argent, R. H., R. J. Thomas, D. P. Letley, M. G. Rittig, K. R. Hardie, and J. C. Atherton.** 2008. Functional association between the *Helicobacter pylori* virulence factors VacA and CagA. J Med Microbiol **57:**145-50.

3. **Atherton, J. C., P. Cao, R. M. Peek, Jr., M. K. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori.* Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem **270:**17771-7.

4. **Atherton, J. C., T. L. Cover, R. J. Twells, M. R. Morales, C. J. Hawkey, and M. J. Blaser.** 1999. Simple and accurate PCR-based system for typing vacuolating cytotoxin alleles of *Helicobacter pylori*. J Clin Microbiol **37:**2979-82.

5. **Basso, D., C. F. Zambon, D. P. Letley, A. Stranges, A. Marchet, J. L. Rhead, S. Schiavon, G. Guariso, M. Ceroti, D. Nitti, M. Rugge, M. Plebani, and J. C. Atherton.** 2008. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology **135:**91-9.

6. **Blaser, M. J.** 1998. *Helicobacter pylori* and gastric diseases. BMJ **316:**1507-10.

7. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

8. **Brenner, H., D. Rothenbacher, and V. Arndt.** 2009. Epidemiology of stomach cancer. Methods Mol Biol **472:**467-77.

9. **Carpenter, B. M., T. K. McDaniel, J. M. Whitmire, H. Gancz, S. Guidotti, S. Censini, and D. S. Merrell.** 2007. Expanding the *Helicobacter pylori* genetic toolbox: modification of an endogenous plasmid for use as a transcriptional reporter and complementation vector. Appl Environ Microbiol **73:**7506-14.

10. **Covacci, A., J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328-33.

11. **Cover, T. L., and S. R. Blanke.** 2005. *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality. Nat Rev Microbiol **3:**320-32.

12. **Cover, T. L., S. A. Halter, and M. J. Blaser.** 1992. Characterization of HeLa cell vacuoles induced by *Helicobacter pylori* broth culture supernatant. Hum Pathol **23:**1004-10.

13. **Douraghi, M., Y. Talebkhan, H. Zeraati, F. Ebrahimzadeh, A. Nahvijoo, A. Morakabati, M. Ghafarpour, M. Esmaili, M. Bababeik, A. Oghalaie, N. Rakhshani, M. E. Hosseini, M. A. Mohagheghi, and M. Mohammadi.** 2009. Multiple Gene Status in *Helicobacter pylori* Strains and Risk of Gastric Cancer Development. Digestion **80:**200-207.

14. **Dunn, B. E., H. Cohen, and M. J. Blaser.** 1997. *Helicobacter pylori*. Clin Microbiol Rev **10:**720-41.

15. **Gebert, B., W. Fischer, E. Weiss, R. Hoffmann, and R. Haas.** 2003. *Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation. Science **301:**1099-102.

16. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

17. **Hatakeyama, M.** 2008. Linking epithelial polarity and carcinogenesis by multitasking *Helicobacter pylori* virulence factor CagA. Oncogene **27:**7047-54.

18. **Hatakeyama, M.** 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat Rev Cancer **4:**688-94.

19. **Higashi, H., A. Nakaya, R. Tsutsumi, K. Yokoyama, Y. Fujii, S. Ishikawa, M. Higuchi, A. Takahashi, Y. Kurashima, Y. Teishikata, S. Tanaka, T. Azuma, and M. Hatakeyama.** 2004. *Helicobacter pylori* CagA induces Ras-independent morphogenetic response through SHP-2 recruitment and activation. J Biol Chem **279:**17205-16.

20. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

21. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **295:**683-6.

22. **Ji, X., T. Fernandez, D. Burroni, C. Pagliaccia, J. C. Atherton, J. M. Reyrat, R. Rappuoli, and J. L. Telford.** 2000. Cell specificity of *Helicobacter pylori* cytotoxin is determined by a short region in the polymorphic midregion. Infect Immun **68:**3754-7.

23. **Jones, K. R., Y. M. Joo, S. Jang, Y. J. Yoo, H. S. Lee, I. S. Chung, C. H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol **47:**959-68.

24. **Letley, D. P., and J. C. Atherton.** 2000. Natural diversity in the N terminus of the mature vacuolating cytotoxin of *Helicobacter pylori* determines cytotoxin activity. J Bacteriol **182:**3278-80.

25. **Maciorkowska, E., I. Roszko, O. Kowalczuk, M. Kaczmarski, L. Chyczewski, and A. Kemona.** 2007. The evaluation of *vacA* gene alleles frequency in *Helicobacter pylori* strains in children and adults in Podlaskie region. Folia Histochem Cytobiol **45:**215-9.

26. **Manente, L., A. Perna, E. Buommino, L. Altucci, A. Lucariello, G. Citro, A. Baldi, G. Iaquinto, M. A. Tufano, and A. De Luca.** 2008. The *Helicobacter pylori's* protein VacA has direct effects on the regulation of cell cycle and apoptosis in gastric epithelial cells. J Cell Physiol **214:**582-7.

27. **Marshall, B. J., and J. R. Warren.** 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet **1:**1311-5.

28. **Matysiak-Budnik, T., and F. Megraud.** 1997. Epidemiology of *Helicobacter pylori* infection with special reference to professional risk. J Physiol Pharmacol **48 Suppl 4:**3-17.

29. **McClain, M. S., P. Cao, H. Iwamoto, A. D. Vinion-Dubiel, G. Szabo, Z. Shao, and T. L. Cover.** 2001. A 12-amino-acid segment, present in type s2 but not type s1 *Helicobacter pylori* VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol **183:**6499-508.

30. **Meining, A., M. Stolte, R. Hatz, N. Lehn, S. Miehlke, A. Morgner, and E. Bayerdorffer.** 1997. Differing degree and distribution of gastritis in *Helicobacter pylori*-associated diseases. Virchows Arch **431:**11-5.

31. **Miehlke, S., A. Hackelsberger, A. Meining, R. Hatz, N. Lehn, P. Malfertheiner, M. Stolte, and E. Bayerdorffer.** 1998. Severe expression of corpus gastritis is characteristic in gastric cancer patients infected with *Helicobacter pylori*. Br J Cancer **78:**263-6.

32. **Miehlke, S., C. Kirsch, K. Agha-Amiri, T. Gunther, N. Lehn, P. Malfertheiner, M. Stolte, G. Ehninger, and E. Bayerdorffer.** 2000. The *Helicobacter pylori vacA* s1, m1 genotype and *cagA* is associated with gastric carcinoma in Germany. Int J Cancer **87:**322-7.

33. **Montecucco, C., and R. Rappuoli.** 2001. Living dangerously: how *Helicobacter pylori* survives in the human stomach. Nat Rev Mol Cell Biol **2:**457-66.

34. **Neel, B. G., H. Gu, and L. Pao.** 2003. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends Biochem Sci **28:**284-93.

35. **Neugut, A. I., M. Hayek, and G. Howe.** 1996. Epidemiology of gastric cancer. Semin Oncol **23:**281-91.

36. **Oldani, A., M. Cormont, V. Hofman, V. Chiozzi, O. Oregioni, A. Canonici, A. Sciullo, P. Sommi, A. Fabbri, V. Ricci, and P. Boquet.** 2009. *Helicobacter pylori* counteracts the apoptotic action of its VacA toxin by injecting the CagA protein into gastric epithelial cells. PLoS Pathog **5:**e1000603.

37. **Pagliaccia, C., M. de Bernard, P. Lupetti, X. Ji, D. Burroni, T. L. Cover, E. Papini, R. Rappuoli, J. L. Telford, and J. M. Reyrat.** 1998. The m2 form of the *Helicobacter pylori* cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A **95:**10212-7.

38. **Pai, R., T. L. Cover, and A. S. Tarnawski.** 1999. *Helicobacter pylori* vacuolating cytotoxin (VacA) disorganizes the cytoskeletal architecture of gastric epithelial cells. Biochem Biophys Res Commun **262:**245-50.

39. **Pan, Z. J., D. E. Berg, R. W. van der Hulst, W. W. Su, A. Raudonikiene, S. D. Xiao, J. Dankert, G. N. Tytgat, and A. van der Ende.** 1998. Prevalence of vacuolating cytotoxin production and distribution of distinct *vacA* alleles in *Helicobacter pylori* from China. J Infect Dis **178:**220-6.

40. **Parsonnet, J., G. D. Friedman, D. P. Vandersteen, Y. Chang, J. H. Vogelman, N. Orentreich, and R. K. Sibley.** 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med **325:**1127-31.

41. **Pugsley, A. P.** 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiol Rev **57:**50-108.

42. **Rhead, J. L., D. P. Letley, M. Mohammadi, N. Hussein, M. A. Mohagheghi, M. Eshagh Hosseini, and J. C. Atherton.** 2007. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology **133:**926-36.

43. **Roder, D. M.** 2002. The epidemiology of gastric cancer. Gastric Cancer **5 Suppl 1:**5-11.

44. **Roovers, K., and R. K. Assoian.** 2000. Integrating the MAP kinase signal into the G1 phase cell cycle machinery. Bioessays **22:**818-26.

45. **Sheu, S. M., K. H. Hung, B. S. Sheu, H. B. Yang, and J. J. Wu.** 2009. Association of nonsynonymous substitutions in the intermediate region of the *vacA* gene of *Helicobacter pylori* with gastric diseases in Taiwan. J Clin Microbiol **47:**249-51.

46. **Shin, A., H. R. Shin, D. Kang, S. K. Park, C. S. Kim, and K. Y. Yoo.** 2005. A nested case-control study of the association of *Helicobacter pylori* infection with gastric adenocarcinoma in Korea. Br J Cancer **92:**1273-5.

47. **Sugimoto, M., and Y. Yamaoka.** 2009. The association of *vacA* genotype and *Helicobacter pylori*-related disease in Latin American and African populations. Clin Microbiol Infect.

48. **Szabo, I., S. Brutsche, F. Tombola, M. Moschioni, B. Satin, J. L. Telford, R. Rappuoli, C. Montecucco, E. Papini, and M. Zoratti.** 1999. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. EMBO J **18:**5517-27.

49. **Terebiznik, M. R., D. Raju, C. L. Vazquez, K. Torbricki, R. Kulkarni, S. R. Blanke, T. Yoshimori, M. I. Colombo, and N. L. Jones.** 2009. Effect of *Helicobacter pylori*'s vacuolating cytotoxin on the autophagy pathway in gastric epithelial cells. Autophagy **5:**370-9.

50. **The EUROGAST Study Group.** 1993. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. The EUROGAST Study Group. Gut **34:**1672-6.

51. **Tokudome, S., R. Ando, R. Ghadimi, T. Tanaka, N. Hattori, Z. Yang, M. Marumoto, H. Agawa, K. Arakawa, Y. Osaka, H. Tanaka, A. Hosono, and M. A. Moore.** 2007. Are there any real *Helicobacter pylori* infection-negative gastric cancers in Asia? Asian Pac J Cancer Prev **8:**462-3.

52. **Torres, V. J., S. E. VanCompernolle, M. S. Sundrud, D. Unutmaz, and T. L. Cover.** 2007. *Helicobacter pylori* vacuolating cytotoxin inhibits activation-induced proliferation of human T and B lymphocyte subsets. J Immunol **179:**5433-40.

53. **Tsugane, S., and S. Sasazuki.** 2007. Diet and the risk of gastric cancer: review of epidemiological evidence. Gastric Cancer **10:**75-83.

54. **Tsutsumi, R., A. Takahashi, T. Azuma, H. Higashi, and M. Hatakeyama.** 2006. Focal adhesion kinase is a substrate and downstream effector of SHP-2 complexed with *Helicobacter pylori* CagA. Mol Cell Biol **26:**261-76.

55. **Van Doorn, L. J., C. Figueiredo, F. Megraud, S. Pena, P. Midolo, D. M. Queiroz, F. Carneiro, B. Vanderborght, M. D. Pegado, R. Sanna, W. De Boer, P. M. Schneeberger, P. Correa, E. K. Ng, J. Atherton, M. J. Blaser, and W. G. Quint.** 1999. Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. Gastroenterology **116:**823-30.

56. **Willhite, D. C., and S. R. Blanke.** 2004. *Helicobacter pylori* vacuolating cytotoxin enters cells, localizes to the mitochondria, and induces mitochondrial membrane permeability changes correlated to toxin channel activity. Cell Microbiol **6:**143-54.

57. **Yamazaki, S., A. Yamakawa, T. Okuda, M. Ohtani, H. Suto, Y. Ito, Y. Yamazaki, Y. Keida, H. Higashi, M. Hatakeyama, and T. Azuma.** 2005. Distinct diversity of *vacA*, *cagA*, and *cagE* genes of *Helicobacter pylori* associated with peptic ulcer in Japan. J Clin Microbiol **43:**3906-16

Table S2: Complete Korean Collection

Strain	**Disease**	**Sex**	**Age**	***cagA* EPIYA motif**	***cagA* accession number**	***vacA allele***	***vacA* accession number**
K1-CA	Cancer	F	68	ABD	FJ458117	s1/i1/m2	GQ338184
K2-CA	Cancer	F	64	ABD		s1/i1/m1	GQ338205
K3-CA	Cancer	F	65	ABD	FJ458118	s1/i1/m1	GQ338222
K4-CA	Cancer	F	37	ABD		s1/i1/m1	GQ338225
K5-CA	Cancer	M	70	ABD		s1/i1/m1	GQ338227
K6-CA	Cancer	F	45	ABD		s1/i1/m1	GQ338233
K7-CA	Cancer	M	56	ABD		s1/i1/m1	GQ338235
K8-CA	Cancer	M	56	ABD		s1/i1/m1	GQ338239
K9-CA	Cancer	M	58	ABD	FJ458119	s1/i1/m1	GQ338242
K10-CA	Cancer	M	52	ABD		s1/i1/m1	
K11-CA	Cancer	M	68	ABD		s1/i1/m1	
K12-G	Gastritis	F	52	ABD		s1/i1/m1	
K13-CA	Cancer	M	38	ABD		s1/i1/m1	
K14-CA	Cancer	F	78	ABD		s1/i1/m1	
K15C-CA	Cancer	F	66	ABD	FJ458120	s1/i1/m1	
K16-CA	Cancer	M	48	ABD		s1/i1/m1	GQ338194
K17-CA	Cancer	F	56	ABD		s1/i1/m1	GQ338196
K18-CA	Cancer	M	64	ABD		s1/i1/m1	
K19-CA	Cancer	F	86	ABD		s1/i1/m1	
K20-CA	Cancer	M	48	ABD		s1/i1/m1	

K21-CA	Cancer	M	44	ABD		s1/i1/m1	GQ338209
K22-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	
K23-DU	Duodenal Ulcer	M	47	ABD		s1/i1/m1	GQ338210
K24-DU	Duodenal Ulcer	M	38	ABD	FJ458121	s1/i1/m1	
K25-DU	Duodenal Ulcer	M	44	ABD		s1/i1/m1	GQ338214
K26-DU	Duodenal Ulcer	M	20	ABD		s1/i1/m1	GQ338218
K27-DU	Duodenal Ulcer	M	47	ABD		s1/i1/m1	
K28-DU	Duodenal Ulcer	M	28	ABD		s1/i1/m1	
K29-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m1	
K30-DU	Duodenal Ulcer	F	61	ABCCC	FJ458122	s1/i2/m2	
K31-DU	Duodenal Ulcer	M	33	ABD		s1/i1/m1	
K32-DU	Duodenal Ulcer	F	41	ABD		s1/i1/m1	
K33-DU	Duodenal Ulcer	F	31	ABC	FJ458123	s1/i1/m1	
K34-DU	Duodenal Ulcer	M	43	ABD	FJ458124	s1/i1/m1	GQ338223
K35-DU	Duodenal Ulcer	F	56	ABD		s1/i1/m2	
K36-DU	Duodenal Ulcer	M	46	ABD		s1/i1/m1	
K37-DU	Duodenal Ulcer	M	61	ABD		s1/i2/m2	
K38-DU	Duodenal Ulcer	F	39	ABC		s1/i1/m1	GQ338224
K39-DU	Duodenal Ulcer	M	59	ABD		s1/i1/m1	
K40-DU	Duodenal Ulcer	M	53	ABD		s1/i1/m1	
K41-DU	Duodenal Ulcer	M	55	ABD		s1/i1/m1	
K42-DU	Duodenal Ulcer	F	48	ABD		s1/i1/m2	
K43-DU	Duodenal Ulcer	M	70	ABD		s1/i1/m1	
K44-DU	Duodenal Ulcer	M	42	ABD		s1/i1/m1	

K45-DU	Duodenal Ulcer	M	22	ABD	FJ458125	s1/i1/m1	
K46-DU	Duodenal Ulcer	F	61	ABD		s1/i1/m1	
K47-DU	Duodenal Ulcer	F	72	ABD		s1/i2/m2	
K48-DU	Duodenal Ulcer	F	41	ABD		s1/i1/m1	
K49-DU	Duodenal Ulcer	M	33	ABC		s1/i1/m1	GQ338226
K50-DU	Duodenal Ulcer	M	35	ABD		s1/i1/m1	
K51-GU	Gastric Ulcer	M	54	ABD		s1/i1/m1	GQ338228
K52-GU	Gastric Ulcer	M	46	ABD		s1/i1/m1	GQ338229
K53-G	Gastritis	M	60	ABD		s1/i1/m1	
K54-G	Gastritis	F	58	ABD	FJ458126	s1/i1/m1	
K55-GU	Gastric Ulcer	F	57	ABD		s1/i1/m1	GQ338230
K56-G	Gastritis	F	48	ABD		s1/i1/m1	
K57-G	Gastritis	F	63	ABD		s1/i1/m1	GQ338231
K58-G	Gastritis	F	61	ABD		s1/i1/m1	GQ338232
K59-G	Gastritis	M	48	BBD	FJ458127	s1/i1/m1	
K60-G	Gastritis	M	53	ABC		s1/i1/m1	
K61-GU	Gastric Ulcer	F	57	*		*	
K62-GU	Gastric Ulcer	F	65	*		*	
K63-GU	Gastric Ulcer	M	59	*		*	
K64-G	Gastritis	F	61	ABCC	FJ458128	s1/i1/m1	
K65-G	Gastritis	M	49	ABD		s1/i1/m1	
K66-G	Gastritis	M	43	ABC		s1/i2/m2	
K67-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m1	
K68-GU	Gastric Ulcer	F	46	ABD		s1/i1/m1	

K69-GU	Gastric Ulcer	M	63	ABD	FJ458129	s1/i1/m1	GQ338234
K70-G	Gastritis	F	68	ABC	FJ458130	s1/i2/m2	
K71-G	Gastritis	M	54	ABD		s1/i1/m1	
K72-GU	Gastric Ulcer	M	34	ABD		s1/i1/m1	GQ338236
K73-GU	Gastric Ulcer	M	72	ABD		s1/i1/m1	
K74-G	Gastritis	F	52	ABD		s1/i1/m1	GQ338237
K75-G	Gastritis	F	24	ABD		s1/i1/m1	
K76-G	Gastritis	F	55	ABD		s1/i1/m1	
K77-G	Gastritis	F	37	ABD		s1/i1/m1	GQ338238
K78-G	Gastritis	M	36	AABD	FJ458131	s1/i1/m1	
K79-GU	Gastric Ulcer	F	84	ABD		s1/i1/m1	
K80-CA	Cancer	F	61	ABD		s1/i1/m1	GQ338240
K81-GU	Gastric Ulcer	M	47	ABD		s1/i1/m1	GQ338241
K82-G	Gastritis	M	39	ABD	FJ458132	s1/i1/m1	
K83-G	Gastritis	F	75	ABD		s1/i1/m1	
K84-G	Gastritis	M	48	ABD		s1/i1/m1	
K85-G	Gastritis	F	28	BD		s1/i1/m1	
K86-G	Gastritis	M	37	ABCC		s1/i1/m1	
K87-G	Gastritis	F	52	ABD		s1/i1/m1	
K88-G	Gastritis	F	69	ABD		s1/i1/m1	
K89-GU	Gastric Ulcer	M	38	ABD		s1/i1/m1	
K90-GU	Gastric Ulcer	M	51	ABD		s1/i1/m1	
K91-GU	Gastric Ulcer	M	82	ABD		s1/i1/m1	
K92-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	

K93-DU	Duodenal Ulcer	F	37	ABC	FJ458133	s1/i1/m1	GQ338243
K94-GU	Gastric Ulcer	M	65	ABD		s1/i1/m1	
K95-CA	Cancer	F	41	ABD		s1/i1/m1	
K96-G	Gastritis	F	47	ABD		s1/i1/m1	
K97-GU	Gastric Ulcer	M	51	ABD		s1/i1/m1	
K98-DU	Duodenal Ulcer	M	23	ABD		s1/i1/m1	
K99-G	Gastritis	F	54	ABD		s1/i2/m2	
K100-GU	Gastric Ulcer	M	46	ABD		s1/i1/m1	
K101-GU	Gastric Ulcer	F	61	ABD		s1/i1/m1	
K102-DU	Duodenal Ulcer	M	38	ABD		s1/i1/m1	
K103-G	Gastritis	M	32	ABD		s1/i1/m1	
K104-CA	Cancer	M	46	ABD		s1/i1/m1	GQ338185
K105-GU	Gastric Ulcer	M	71	ABD		s1/i1/m1	
K106-DU	Duodenal Ulcer	M	14	ABD		s1/i1/m1	
K107-DU	Duodenal Ulcer	M	26	ABD		s1/i1/m1	
K108-GU	Gastric Ulcer	M	62	ABD		s1/i1/m1	
K109-G	Gastritis	M	40	ABD		s1/i1/m1	GQ338186
K110-GU	Gastric Ulcer	M	81	ABCC	FJ458134	s1/i1/m1	
K111-DU	Duodenal Ulcer	F	36	ABD		s1/i1/m1	GQ338187
K112-G	Gastritis	M	57	ABD		s1/i1/m1	GQ338188
K113-G	Gastritis	M	29	ABD		s1/i1/m1	
K114-DU	Duodenal Ulcer	F	47	ABC		s1/i1/m1	
K115-G	Gastritis	F	82	ABC	FJ458135	s1/i1/m2	
K116-G	Gastritis	M	59	ABD		s1/i1/m1	

K117-G	Gastritis	F	21	ABD	FJ458136	s1/i1/m1	GQ338189
K118-CA	Cancer	F	67	ABD		s1/i1/m1	
K119-DU	Duodenal Ulcer	M	31	ABD	FJ458137	s1/i1/m1	
K120-G	Gastritis	F	41	ABC		s1/i1/m1	GQ338190
K121-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	
K122-DU	Duodenal Ulcer	M	60	ABD		s1/i1/m1	
K123-G	Gastritis	M	76	ABD	FJ458138	s1/i1/m1	GQ338191
K125-G	Gastritis	M	59	ABD		s1/i1/m1	
K126-GU	Gastric Ulcer	M	69	ABD		s1/i1/m1	
K127-GU	Gastric Ulcer	M	71	ABD		s1/i1/m1	
K128-GU	Gastric Ulcer	M	58	ABC		s1/i1/m1	GQ338192
K129-GU	Gastric Ulcer	M	36	*		*	
K130-G	Gastritis	F	64	*		*	
K131-G	Gastritis	F	61	ABD	FJ458139	s1/i1/m1	GQ338193
K132-GU	Gastric Ulcer	M	23	*		*	
K133-GU	Gastric Ulcer	M	63	*		*	
K134-GU	Gastric Ulcer	M	46	*		*	
K135-DU	Duodenal Ulcer	F	62	*		*	
K136-G	Gastritis	F	52	*		*	
K137-G	Gastritis	F	62	*		*	
K138-GU	Gastric Ulcer	F	21	*		*	
K139-GU	Gastric Ulcer	F	49	*		*	
K140-G	Gastritis	F	49	*		*	
K141-G	Gastritis	M	57	*		*	

K142-GU	Gastric Ulcer	M	65	*		*	
K143-GU	Gastric Ulcer	F	71	*		*	
K144-GU	Gastric Ulcer	F	53	*		*	
K145-GU	Gastric Ulcer	M	62	*		*	
K146-G	Gastritis	M	40	ABD**	FJ458140	s1/i1/m1	
K147-GU	Gastric Ulcer	M	62	*		*	
K148-GU	Gastric Ulcer	M	37	*		*	
K149-GU	Gastric Ulcer	M	71	*		*	
K150-G	Gastritis	F	26	ABD		s1/i1/m1	
K151-GU	Gastric Ulcer	M	65	ABD	FJ458141	s1/i1/m1	
K152-G	Gastritis	F	62	ABD		s1/i1/m1	
K153-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	
K154-G	Gastritis	F	55	ABCCCC	FJ458142	s1/i1/m2	
K155-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	
K156-DU	Duodenal Ulcer	F	47	ABD		s1/i1/m1	
K157-G	Gastritis	M	43	ABD		s1/i1/m1	
K158-G	Gastritis	M	60	ABD		s1/i1/m1	
K159-G	Gastritis	F	35	ABD		s1/i1/m1	
K160-DU	Duodenal Ulcer	M	30	ABD		s1/i1/m1	
K161-G	Gastritis	F	65	ABD		s1/i1/m1	
K162-G	Gastritis	F	63	ABD		s1/i1/m1	GQ338195
K163-G	Gastritis	F	66	ABD		s1/i1/m1	
K164-G	Gastritis	M	43	ABD		s1/i1/m1	
K165-G	Gastritis	M	28	ABD		s1/i1/m1	

K166-G	Gastritis	F	38	ABC		s1/i2/m2	
K167-G	Gastritis	F	27	ABD		s1/i1/m1	
K169-G	Gastritis	F	47	ABD		s1/i1/m1	
K170-G	Gastritis	F	41	ABD**	FJ458143	s1/i1/m1	
K171-CA	Cancer	F	72	ABD	FJ458144	s1/i1/m1	
K172-G	Gastritis	F	31	ABCC	FJ458145	s1/i1/m1	
K173-G	Gastritis	F	45	ABD	FJ458146	s1/i1/m1	
K174-G	Gastritis	N/A	N/A	ABD		s1/i1/m1	
K175-G	Gastritis	F	41	ABD		s1/i1/m2	
K176-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	
K177-G	Gastritis	F	39	ABD		s1/i1/m1	
K178-G	Gastritis	F	40	ABD		s1/i1/m1	GQ338197
K179-G	Gastritis	F	38	ABCCC	FJ458147	s1/i1/m2	
K180-G	Gastritis (polyps)	F	50	ABD		s1/i1/m1	
K181-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m2	
K182-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	GQ338198
K183-G	Gastritis	M	40	ABD		s1/i1/m1	GQ338199
K184-G	Gastritis	F	55	ABD		s1/i1/m1	
K185-G	Gastritis	F	52	ABD		s1/i1/m1	GQ338200
K186-G	Gastritis	M	41	ABD		s1/i1/m1	
K188-G	Gastritis (IM)	F	43	ABD		s1/i1/m1	
K190-G	Gastritis	M	61	ABC		s1/i1/m1	GQ338201
K192-DU	Duodenal Ulcer	F	61	AABD	FJ458148	s1/i1/m1	

K193-G	Gastritis	F	50	ABD		s1/i1/m1	GQ338202
K194-GU	Gastric Ulcer	M	48	*		s1/i1/m1	
K195-GU	Gastric Ulcer	F	48	ABD	FJ458149	s1/i1/m1	
K196-G	Gastritis	F	50	ABD		s1/i1/m1	GQ338203
K197-G	Gastritis	F	45	ABD		s1/i1/m1	GQ338204
K198-GU	Gastric Ulcer	F	56	ABC	FJ458150	s1/i2/m2	
K199-GU	Gastric Ulcer	M	50	ABD		s1/i1/m1	
K200-GU	Gastric Ulcer	M	63	ABD		s1/i1/m1	
K201-GU	Gastric Ulcer	M	55	ABD		s1/i1/m1	
K202-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	
K203-G	Gastritis	F	55	ABD		s1/i1/m1	
K204-GU	Gastric Ulcer	F	63	ABD		s1/i1/m2	
K205-GU	Gastric Ulcer	F	57	ABD		s1/i1/m1	
K206-GU	Gastric Ulcer	F	51	ABD		s1/i1/m1	
K207-G	Gastritis	M	39	ABD		*	GQ338206
K208-G	Gastritis	F	56	ABD	FJ458151	s1/i1/m1	GQ338207
K209-G	Gastritis	F	24	ABD		s1/i1/m1	GQ338208
K210-G	Gastritis	F	61	ABD		s1/i1/m1	
K211-G	Gastritis	F	54	ABD		s1/i1/m1	
K212-G	Gastritis	F	45	ABD		s1/i1/m1	
K213-G	Gastritis	F	52	*		*	
K214-G	Gastritis	F	53	*		*	
K215-GU	Gastric Ulcer	F	73	*		*	
K216-G	Gastritis	F	67	ABD		s1/i1/m1	

K217-G	Gastritis	M	77	ABD		*	
K218-G	Gastritis	F	62	ABD		s1/i1/m1	
K219-G	Gastritis	M	40	ABD	FJ458152	s1/i1/m1	
K220-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	
K221-G	Gastritis	M	37	BD		s1/i1/m1	
K222-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	
K223-G	Gastritis	F	25	ABD	FJ458153	s1/i1/m1	
K224-G	Gastritis	F	35	ABD		s1/i1/m1	
K225-DU	Duodenal Ulcer	F	60	ABC		s1/i1/m1	
K226-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	
K227-G	Gastritis	F	31	ABD		s1/i1/m1	
K228-GU	Gastric Ulcer	M	54	ABD		s1/i1/m1	
K229-GU	Gastric Ulcer	M	62	ABD		s1/i1/m1	
K230-G	Gastritis (IM)	M	56	ABD		s1/i1/m1	
K231-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	
K232-G	Gastritis	F	56	ABD		s1/i1/m1	
K233-G	Gastritis	M	38	ABD		s1/i1/m1	
K234-DU	Duodenal Ulcer	M	41	ABD		s1/i1/m2	
K235-G	Gastritis	F	50	ABD		s1/i1/m1	
K236-G	Gastritis	F	64	ABD		s1/i1/m1	
K237-G	Gastritis	F	48	ABD		s1/i1/m1	
K238-DU	Duodenal Ulcer	M	55	ABD		s1/i1/m1	GQ338211
K239-G	Gastritis	M	46	ABD		*	
K240-G	Gastritis	F	41	ABD		s1/i1/m1	

K241-G	Gastritis	M	41	ABD		s1/i1/m1	
K242-G	Gastritis	M	78	ABD		s1/i1/m1	
K243-G	Gastritis	F	60	BC		*	
K244-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	
K245-G	Gastritis	M	19	ABD		s1/i1/m1	
K246-G	Gastritis	F	40	ABD		s1/i1/m1	
K247-G	Gastritis	F	56	ABD		s1/i1/m1	
K248-G	Gastritis	M	58	ABD		s1/i1/m1	GQ338212
K249-GU	Gastric Ulcer	F	48	ABC		s1/i1/m1	GQ338213
K250-G	Gastritis	F	53	ABD		s1/i1/m1	
K251-DU	Duodenal Ulcer	M	70	ABD		s1/i1/m1	
K253-DU	Duodenal Ulcer	F	61	ABC		s1/i1/m1	
K254-G	Gastritis	M	54	ABD		s1/i1/m1	
K255-G	Gastritis	F	74	ABD	FJ458154	s1/i1/m1	GQ338215
K256-G	Gastritis	F	51	ABD		s1/i1/m1	
K257-CA	Cancer	M	64	ABD		s1/i1/m1	
K258-CA	Cancer	M	68	ABD	FJ458155	s1/i1/m1	GQ338216
K259-CA	Cancer	M	44	ABD	FJ458156	s1/i1/m2	GQ338217
K260-CA	Cancer	M	58	ABD	FJ458157	s1/i1/m1	GQ338219
K261-CA	Cancer	F	48	ABD	FJ458158	s1/i1/m1	GQ338220
K262-G	Gastritis	F	56	ABC	FJ458159	s1/i1/m1	GQ338221
K263-G	Gastritis	M	59	ABABD***	FJ458160	s1/i1/m1	
K264-DU	Duodenal Ulcer	M	32	ABD	FJ458161	s1/i1/m1	
K265-DU	Duodenal Ulcer	M	42	ABD	FJ458162	s1/i1/m1	

K266-G	Gastritis	F	34	ABD	FJ458163	s1/i1/m1	

*Indeterminate in genotyping assay
**-B motif's proline is replaced with a serine, ESIYA, therefore classified as other
***-ABABD second -B motif's proline is replaced with leucine, ELIYA, therefore classified as other

Chapter Four

Polymorphisms in the Intermediate Region of VacA Impact Helicobacter pylori – Induced Disease Development

Published as: **Jones, K.R.[*], S.Jang[*], J.Y. Chang, J. Kim, I.S. Chung, C.H. Olsen, D.S. Merrell and J.H. Cha and D. S. Merrell.** 2010. Polymorphisms in the Intermediate Region of VacA Impact *Helicobacter pylori* – Induced Disease Development. *J Clin Microbiol* 49:101-10.

[*] These authors contributed equally to this work.

The work presented in this chapter is the sole work of K. R. Jones with the following exceptions: S. Jang and J.Y. Chang performed the *vacA* sequencing, J.H. Cha assisted with experimental design, J. Kim assisted in Genebank registration, I.S. Chung performed the biopsies and supplied the diagnoses, and C.H. Olsen assisted with the statistical analysis.

Abstract

Helicobacter pylori is the etiological agent of diseases such as gastritis, gastric and duodenal ulcers, and two types of gastric cancers. While some insight has been gained into the etiology of these diverse manifestations, by and large, the reason that some individuals develop more severe disease remains elusive. Recent studies have focused on the role of the *H. pylori* toxins, CagA and VacA, on the disease process and

have suggested that both toxins are intimately involved. Moreover, CagA and VacA are polymorphic within different *H. pylori* strains and particular polymorphisms seem to show a correlation to development of particular disease states. Among VacA polymorphisms, the intermediate region has recently been proposed to play a major role in disease outcome. Herein we describe a detailed sequence analysis of the polymorphic intermediate region of *vacA* from strains obtained from a large South Korean population. We show that polymorphisms found at amino acid position 196 are associated with more severe disease manifestations. Additionally, polymorphisms found at amino acid position 231 are linked to disease in strains that carry the non EPIYA-ABD allele of CagA. Collectively, these data help explain the impact of the VacA intermediate region on disease and lead to the hypothesis that there are allele-driven interactions between VacA and CagA.

Introduction

The medically important microbe, *Helicobacter pylori* colonizes the inhospitable niche of the gastric mucosa of over fifty percent of the world's population (34, 54). *H. pylori* is a spiral-shaped, microaerophilic, Gram-negative bacterium (33) that is the etiological agent of a multitude of diseases, including gastritis, peptic ulcers (both duodenal and gastric ulcers), as well as adenocarcinoma and MALT lymphoma (6, 12, 17, 18, 45). This class I carcinogen contributes to gastric cancer mortality, which is still one of the most common causes of mortality due to cancer (18, 42). This is especially true in East Asian countries such as China, Korea, and Japan (23).

Two *H. pylori* toxins that facilitate host cellular damage and directly interplay with the host immune system are the cytotoxin associated gene A (CagA) and the vacuolating cytotoxin (VacA). A growing number of studies have begun to suggest that VacA and CagA interact in such a way as to affect disease severity (2, 27, 59, 62). VacA was identified within a few years of the discovery of *H. pylori*, and appears to be produced and secreted by all strains (4, 13). This toxin was initially identified and named due to its ability to cause large cytoplasmic vacuoles in intoxicated host cells (14). However, VacA has subsequently been shown to have a multitude of functions. For instance, when inserted into the plasma membrane, VacA can act as an anion-selective channel (52), which may aid bacterial survival through leakage of host cytosolic anions that can be utilized by the bacterium (40). VacA also has the ability to induce apoptosis through permeabilization of the mitochondrial membrane, thereby causing cytochrome *c* release (32, 60). Furthermore, VacA can induce the autophagy pathway (53), disorganize the host cell cytoskeleton to cause spreading of host cells, inhibit T cell activation, and block T cell and B cell proliferation (20, 56).

CagA is directly injected into host cells through a type IV secretion apparatus (10), and is phosphorylated within host cells by host-cell kinases. This phosphorylation event subsequently makes CagA competent for interaction with the Src homology 2 domain-containing protein tyrosine phosphatase (SHP-2; 23, 26). The downstream affects of this interaction include alterations in numerous host signaling pathways (22, 24-26, 41, 48, 58), which are believed to be responsible for the increased cancer risk associated with infection by strains that express CagA (7, 21). Phosphorylation of CagA occurs in the carboxy-terminus of the protein at conserved tyrosine residues that exist as

part of a repeated five amino acid sequence (Glu-Pro-Ile-Tyr-Ala) referred to as the EPIYA repeat (25, 26). The numbers of these EPIYA repeats and the flanking amino acid regions surrounding these repeats vary dramatically across strains. Based on flanking sequences, four distinct EPIYA motifs have been identified (-A, -B, -C, -D; 25, 61). Strains that carry various combinations of these motifs can be divided into two main geographical distributions, which are hallmarked by differences in the primary phosphorylation sites, EPIYA -C or -D (25); East Asian strains contain EPIYA-ABD whereas Western strains contain EPIYA-ABC, where the EPIYA-C motif may be repeated up to five times (2, 25, 26, 43). These different EPIYA combinations have been suggested to impact disease progression (29, 61).

Similar to the polymorphic nature of CagA, VacA contains three distinct segments that exhibit variation within the amino-terminus. These areas of variation are broadly defined as the signal (s), intermediate (i), and middle (m) regions and two or more primary variants have been described for each region: signal, s1 or s2, intermediate, i1, i2, or i3, and middle, m1 or m2 (3, 11, 47). Various combinations of each s, i, and m region are then combined within each *H. pylori* strain to yield a particular *vacA* allele. The s region of VacA appears to influence the efficiency of anion channel formation based on the hydrophobicity of amino acid residues that are found near a proteolytic cleavage site found in this region (35, 46); the s1 form contains a hydrophobic region adjacent to the proteolytic cleavage site that increases membrane insertion and formation of membrane channels (30, 35). The m region affects host cell tropism (28); VacA toxins encoding the m1 region are toxic to a broader range of host cells (1, 44). The i region is positioned between the s and m regions and is the most recent region to be

described. The i1 variants of VacA have been shown to have stronger vacuolating activity than toxins containing the i2 regions (47). Due to the increased anion channel formation capability, broader cell tropism and enhanced vacuolating activity, individual associations between the s1, m1, and i1 types and more severe forms of *H. pylori* induced disease have been identified (5, 47). Furthermore, several studies have linked strains carrying the s1/m1 allele of the toxin to more severe disease outcomes, since these strains show the strongest vacuolating activity to the broadest range of cells (31). However, recently the i region has been suggested to be a better predictor of disease severity than either the s or m region, though i appears to co-vary with the s and m regions (47). This means that the more toxic i1 region is often associated with s1/m1 (47).

Within the i region, three specific clusters have been identified as the main areas to contain polymorphisms: clusters A, B, and C. Among these, cluster B and C have been shown to impact the vacuolating activity of the toxin (47). Because of this link to toxin activity, researchers have sought to determine the role of natural individual amino acid changes within these clusters in ultimate disease outcome. For example, variation in the ninth amino acid in cluster B (amino acid 231 of the protoxin) has been linked to disease development in the Taiwanese population (50). However, this amino acid appeared to have no impact within the South Korean population (27). Recently, we identified two additional positions in the VacA intermediate region that contained polymorphisms: position 151 and 196 (27). While neither of these amino acids showed a statistically significant link to disease severity in the small population of samples that were examined, the distribution of amino acids found at position 196 displayed a trend toward significance (27). Given this, we sequenced the *vacA* intermediate region from 231

South Korean isolates and then analyzed the distribution of polymorphisms across the entire region. Furthermore, we compared these polymorphic *vacA* distributions to the various *cagA* alleles carried by each strain as well as to the ultimate disease development. Herein, we present an expanded i1 and i2 consensus sequence, show that amino acid 196 impacts disease development, and show that amino acid 231 is important for disease development, but only within strains that carry a non-EPIYA-ABD *cagA* allele.

Material and Methods

Bacterial Strains and Culture Conditions

This South Korean population of 260 strains has been previously described and includes 115 gastritis isolates, 60 gastric ulcer isolates, 55 duodenal ulcer isolates, and 30 gastric cancer isolates (27, 29). Isolates were preserved as stocks at -80°C and then grown and expanded on antibiotic-supplemented horse blood agar plates under microaerophilic conditions created by an Anoxomat evacuation/replacement system (Spiral Biotech, Norwood, MA) exactly as previously described (9, 29).

vacA *i Region Sequencing*

Chromosomal DNA of each of the 260 *H. pylori* strains was isolated using the Easy-DNA kit (Invitrogen, Carlsbad, CA). The *vacA* intermediate region was amplified and then Sanger dideoxy sequenced using the primers previously described by Rhead, *et al.*: VacF1 3'-GTTGGGATTGGGGGAATGCCG-5' and VacR9 3'-TGTTTATCGTGCTGTATGAAGG-5' (47). Sanger dideoxy sequencing was performed at both the Uniformed Services University of the Health Science Biomedical

Instrumentation Center (Bethesda, MD) and Cosmo Genetech Co., Ltd (Seoul, South Korea). Resulting DNA sequences were analyzed using Vector NTI version 9.1 (Invitrogen, Carlsbad, CA) and Sequencher 4.5 (Gene Codes Corp., Ann Arbor, MI). The amino acid numbering system used in this study is based off the VacA sequence of strain G27; numbering begins at the translational start such that amino acid 1 is the first methionine of the translated protein.

Statistical Analysis

The Fisher's exact test was used to analyze the association between the *vacA* allele, disease state, *cagA* allele, and specific amino acids within the intermediate region. Log linear modeling was used to assess higher order associations that were significant at the 5% level. We fit a saturated model using categorical variables representing *vacA* genotype, *cagA* genotype, disease state, gender, and amino acids within the i region using a backward selection algorithm, which eliminates the least significant association at each step and then reforms the model to look for associations. Data were analyzed using SPSS version 16 software (SPSS Inc., Chicago, IL) or SAS version 9.1 software (SAS Institute Inc., Cary, NC).

Nucleotide Sequence Accession Numbers

The sequences for the i region of *vacA* from the 60 original strains analyzed (27) were previously deposited in Genbank under accession numbers GQ338184 to GQ338243 and the i region sequences of the additional 171 strains have been deposited

in GenBank under accession numbers HM047564 to HM047592 and HM047594 to HM047735 (see Table S3 in the supplemental material).

Results

*Sample Acquisition/*vacA *i Region Sequencing*

The strains used in this study have been previously characterized for distribution of both the *cagA* allele (27, 29) and the *vacA* allele (27). However, the previous study characterizing the *vacA* allele relied primarily on PCR-based typing methods and only analyzed the *vacA* intermediate region sequence from a subset of strains carrying the i1 allele; whereas the current study represents a detailed analysis of *vacA* sequences from the complete collection of South Korean strains. The complete collection of 260 strains contained 254 strains for which we had complete epidemiological data. These 254 strains were obtained from patients with a mean age of 51 years and an age range of 14 to 86 years. These strains were evenly distributed by gender; 126 strains were from female patients with a mean age of 53 years and an age range of 21 to 86 years, and 128 strains were from male patients with a range of 14 to 82 years and a mean age of 50 years. These strains were distributed across various *H. pylori*-induced disease states: 45% from patients diagnosed with gastritis, 22% with gastric ulcers, 21% with duodenal ulcers, and 12% with gastric cancer (29).

Of the 260 strains in this entire collection, 231 strains were successfully sequenced for the *vacA* i region and 222 of these strains contained complete epidemiological data as well as CagA and VacA genotypes (Fig. 11, Table 8, and Table S3 in the supplemental material). These sequenced regions were from strains from

Table 8: Distribution of cagA *and* vacA *allele across the different disease states.*

			Gastritis	Duodenal Ulcers	Gastric Ulcers	Gastric Cancer
Overall						
Total		222	101	49	42	30
	Age Range	14-86	19-82	14-72	34-84	37-86
	Mean Age	50.3	48.3	45.1	55.1	56.6
Male		112	34	30	32	16
	Age Range	14-82	19-76	14-70	34-82	38-70
	Mean Age	48.9	46.6	41.3	54.5	55.1
Female		110	67	19	10	14
	Age Range	21-86	21-82	31-72	46-84	37-86
	Mean Age	51.7	49.2	51.1	57.1	61
CagA						
EPIYA-ABD		189	81	40	38	30
	Age Range	14-86	19-78	14-72	34-84	37-86

	Mean Age	50.5	48.9	44.5	54.5	57.9
Male		100	25	29	30	16
	Age Range	14-82	19-78	14-70	34-82	38-70
	Mean Age	48.9	48.1	41.6	53.5	55.1
Female		89	56	11	8	14
	Age Range	21-86	21-75	36-72	46-84	37-86
	Mean Age	52.2	49.2	52.1	58.4	61
Other*		33	20	9	4	0
	Age Range	28-82	28-82	31-61	48-81	N/A
	Mean Age	49.3	47.7	47.8	60.8	N/A
Male		12	9	1	2	0
	Age Range	33-81	36-61	33	58-81	N/A
	Mean Age	48.8	46	N/A	69.5	N/A
Female		21	11	8	2	0
	Age Range	28-	28-82	31-61	48-56	N/A

		82				
	Mean Age	49.5	49	49.6	52	N/A
VacA						
s1i1m1		200	92	41	40	27
	Age Range	14-84	19-78	14-70	34-84	37-78
	Mean Age	49.8	48.5	43	54.9	57
Male		104	32	27	32	15
	Age Range	14-82	19-78	14-70	34-82	38-70
	Mean Age	50.1	48.2	40.3	54.5	55.9
Female		94	60	14	8	12
	Age Range	21-84	21-75	31-61	46-84	37-78
	Mean Age	50.5	48.7	48.3	56.5	58.3
s1i1m2		11	4	4	1	2
	Age Range	38-82	38-82	41-57	63	44-68

	Mean Age	53.9	54	50.5	N/A	56
Male		2	0	1	0	1
	Age Range	41-44	N/A	41	N/A	44
	Mean Age	42.5	N/A	N/A	N/A	N/A
Female		9	4	3	1	1
	Age Range	38-82	38-82	48-57	63	68
	Mean Age	56.4	54	53.7	N/A	N/A
s1i2m2		8	4	3	1	0
	Age Range	38-72	38-68	61-72	56	N/A
	Mean Age	56.6	50.8	64.7	N/A	N/A
Male		2	1	1	0	0
	Age Range	43-61	43	61	N/A	N/A
	Mean Age	52	N/A	N/A	N/A	N/A
Female		6	3	2	1	0

	Age Range	38-72	38-68	61-72	56	N/A
	Mean Age	58.2	53.3	66.5	N/A	N/A
s1i3m1		3	1	1	0	1
	Age Range	32-86	32	47	N/A	86
	Mean Age	55	N/A	N/A	N/A	N/A
Male		2	1	1	0	0
	Age Range	32-47	32	47	N/A	N/A
	Mean Age	39.5	N/A	N/A	N/A	N/A
Female		1	0	0	0	1
	Age Range	86	N/A	N/A	N/A	86
	Mean Age	N/A	N/A	N/A	N/A	N/A

*indicates any other genotype besides EPIYA-ABD, including Western strains and EPIYA-AABD, -BD, -BBD, -ABAB**D, as well as -AB**D where a mutation within the EPIYA-B motif is designated by the **

N/A stands for Not Applicable

Figure 11: Weblogo showing the VacA intermediate region's major polymorphic domains within the South Korean population. The WebLogo was created by inputting the amino acid sequence from each of the three different intermediate alleles (221-i1 sequences, 8-i2 sequences, and 3-i3 sequences into http://weblogo.threelusone.com/ (15, 49). The three primary regions of polymorphism (clusters A, B, and C) as well as the two amino acids shown to impact disease development (196 and 231) are indicated. The logo represents the alignment at each position by a stack of letters, where the height of each letter is proportional to the observed frequency of the corresponding amino acid and the overall height of each stack is proportional to the sequence conservation, measured in bits, at that position (15).

Figure 11: Weblogo showing the VacA intermediate region's major polymorphic domains within the South Korean population

patient with an age range of 14 to 86 years and a mean age of 50.3 years. Of the 222 isolates, 112 were from male patients aged 14 to 82 years with a mean age of 48.9 years, and 110 were from female patients aged 21 to 85 years with a mean age of 51.7 years. The sequenced strains distribution across disease states was similar to that of the overall population; 45.5% were from gastritis patients, 41% were from patients suffering from ulcers (22.1% duodenal ulcers and 18.9% gastric ulcers), and 13.5% were from cancer patients.

Vac i Region, i1, i2, and i3

The fact that the *vacA* i region shows polymorphism that may affect toxin activity was only recently described. At that time, consensus sequences were identified for the i1 and i2 regions based on sequences from strains 60190 and Tx30a, respectively (Fig. 11 and Fig. 12; 47). Subsequently, the consensus sequences for the i1 and i2 regions were verified by analysis of 123 strains from four distinct populations by Chung, *et al.*, who also described an i3 region that appears to be a hybrid of the i1 and i2 sequences (11).

Based on this new (i1, i2, and i3) nomenclature, our South Korean population contained four different *vacA* alleles, s1/i1/m1 (200 isolates), s1/i1/m2 (11 isolates), s1/i2/m2 (eight isolates), and s1/i3/m1 (three isolates; Table 8). All three of the isolates defined as having an i3 region contained an i2 cluster B consensus sequence and an i1 cluster C consensus sequence (Fig. 11 and Fig. 12). Of the s1/i1/m1 strains, 92 were from gastritis patients, 41 were from duodenal ulcer patients, 40 were from gastric ulcer patients, and 27 were from gastric cancer patients. Four of the s1/i1/m2 strains were

from gastritis patients, four were from duodenal ulcer patients, one was from a gastric ulcer patient, and two were from cancer patients. Of the eight s1/i2/m2 strains, four were from gastritis patients, three were from duodenal ulcer patients, and one from a gastric ulcer patient. The three s1/i3/m1 strains were from one gastritis patient, one duodenal ulcer patient, and one cancer patient (Table 8). There was no statistical association between the distribution of the *vacA* alleles across the different disease states (P=0.6865).

Given the fact that we were able to determine the i region sequence from 231 strains, we next examined the i sequences to identify amino acids that predominated in East Asian strains, as well as to determine the amino acid differences between the i1 and i2 *vacA* alleles that might differ from the defined consensus sequence (47). For this analysis, we defined the consensus sequence in our population as the amino acids encoded for by at least 85% of all strains. Comparison of our consensus sequence to that defined by Rhead *et al*. revealed the following: in cluster A we found 1) substitution of a phenylalanine instead of tyrosine at the first amino acid for the i1 allele, 2) reversal of the twelfth and fourteenth amino acids of the i2 allele (asparagine, phenylalanine, aspartic acid versus aspartic acid, phenylalanine, asparagine in our population), and 3) that the main difference between the i1 and i2 *vacA* alleles in our population were the fourth, and sixth amino acids (i1: SAD and i2: GAN; Fig. 12; 47). In cluster B we found 1) our population displayed variability at the ninth amino acid of the i1 allele as compared to the previously described serine, 2) the first two amino acids of the i2 alleles in our population were similar to the i1 sequence and contained a glutamine followed by alanine rather than the lysine serine combination found in other populations, and 3) the major differences in cluster B between i1 and i2 alleles in our population were the fifth, ninth, and tenth amino

Figure 12: Consensus Sequences for the vacA *intermediate region.* The consensus sequences of the i1 and i2 regions were previously defined by Read, *et al.* and Chung, *et al.* (11, 47) and are shown in comparison to sequences obtained in this study. Dark gray shading indicates difference between the different i1 and i2 consensus sequences or differences among the i3 sequences. Light gray shading indicates amino acids that are different between i1 and i2 strains. Positions, where there is no consensus for a particular amino acid are designated by a period. Dashes represent the points of insertion of additional amino acids that are only found in Cluster C of the i2 allele.

Figure 12: Consensus Sequences for the vacA *intermediate region*

i1 Consensus Sequence

Cluster A	Cluster B	Cluster C	Citation
MKDSADRTTRVDFN	QASEGITSSK	ASNS---VKLNGN	Rhead, *et al.* (47)
	QASE.ITS	AS.---SV	Chung, *et al.*(11)
FKDSADRTTRVDFN	QASEGITS.K	ASNS---VKLMGN	This publication

i2 Consensus Sequence

Cluster A	Cluster B	Cluster C	Citation
FKDGANRTTRVNFD	KSSEKITSRE	VSSSNQSVDLYGK	Rhead, *et al.* (47)
	KSSEKITSRE	VSSSN.SVDLYEK	Chung, *et al.* (11)
FKDGANRTTRVDFN	QASEKITSRE	VSSSNHSVDLYGK	This publication

i3 Sequences

Strains	Cluster A	Cluster B	Cluster C
i1 Consensus	FKDSADRTTRVDFN	QASEGITS.K	ASNS---VKLMGN
i2 Consensus	FKDGANRTTRVDFN	QASEKITSRE	VSSSNHSVDLYGK
K19-CA	FKDGANRTTRVDFN	QASEKITSRE	ASNS---VKLMGN
K23-DU	FKDAANRTTRVDFN	QASEKITSRE	ASNS---VKLMGN
K103-G	FKDSADRTTRVDFN	QASEKITSRE	ASNS---VKLMGN

acids (Fig. 12; 11, 47). In cluster C we found 1) our population had substituted a methionine for an asparagine at the eighth amino acid of the i1 allele, 2) the replacement of a histidine for a glutamine at the sixth amino acid of the i2 allele, 3) the major differences in cluster C between the i1 and i2 alleles were the insertion of three additional amino acids in the i2 allele, an asparagine, histidine, and serine, after the fourth amino acid and then subsequent differences in the first, third, sixth, eighth, and tenth amino acids of the i1 allele (correlating with the first, third, ninth, eleventh and thirteenth amino acids of the i2 allele; Fig. 12; 11, 47).

The newest allele of the intermediate region, the i3 region has been defined as one that contains a cluster B from either i1 or i2 and a cluster C from the other allele (11). This South Korean population contained three i3 strains. All of these strains contained a cluster B from an i2 allele and a cluster C from an i1 allele. The difference in cluster A in this population between an i1 and an i2 are the fourth and sixth amino acids. Of note, the three i3 sequences showed no consensus between these amino acids, one was identical to the i1, another identical to the i2, and one contained a different combination all together (Fig. 12).

Amino Acid 151

In depth sequence analysis of the entire i region revealed that, as previously described (27), amino acid 151 showed amino acid polymorphism. Strains contained either a tyrosine (91 isolates, 41%) or phenylalanine (131 isolates, 59%) at this position. The distribution of amino acids at this position was not associated with gender (P=0.1749); the distribution between males and females was fairly even (Table 9). In

Table 9: Distribution of amino acids at polymorphic positions within the intermediate region.

			Gastritis	**Duodenal Ulcers**	**Gastric Ulcers**	**Gastric Cancer**
Amino Acid 151						
Tyrosine		91	41	15	20	15
	Age Range	14-84	25-74	14-70	34-84	38-78
	Mean Age	49.8	46.7	45.5	54.3	56.4
Male		51	17	10	17	7
	Age Range	32-81	28-74	14-70	34-81	38-64
	Mean Age	45.9	45.1	42.4	52.7	51.1
Female		40	24	5	3	8
	Age Range	37-84	25-74	41-57	48-84	41-78
	Mean Age	49	47.9	51.6	63	61

Phenylalanine		131	60	34	22	15
	Age Range	19-86	19-82	20-72	41-82	37-86
	Mean Age	50.7	49.9	44.9	55.9	59.3
Male		61	17	20	15	9
	Age Range	19-82	19-78	20-70	41-80	44-70
	Mean Age	49.8	50	40.7	56.5	58.2
Female		70	43	14	7	6
	Age Range	21-86	21-82	31-72	46-63	37-86
	Mean Age	51.5	49.9	50.9	54.6	61
Amino Acid 231						
Glycine		153	67	34	30	22
	Age Range	14-84	19-82	14-70	34-82	37-78
	Mean Age	50.3	50.2	42.6	54.6	57

Male		83	25	23	24	11
	Age Range	14-84	19-78	14-70	34-82	38-68
	Mean Age	48.6	49.3	40.7	53	54.5
Female		70	42	11	6	11
	Age Range	25-84	25-82	31-61	48-84	37-78
	Mean Age	52.3	50.7	46.8	60.7	59.5
Other AA*		69	34	15	12	8
	Age Range	21-86	27-74	31-72	41-81	44-86
	Mean Age	50.4	45.5	50.5	57.3	60.3
Male		29	9	7	8	5
	Age Range	29-81	29-61	31-61	41-81	44-70
	Mean Age	50	42.6	43.3	60.1	47.6

Female		40	25	8	4	3
	Age Range	21-86	21-74	37-72	46-57	48-86
	Mean Age	50.7	46.6	56.9	51.8	66.7
Amino Acid 196						
Serine		148	64	28	31	25
	Age Range	19-84	19-82	23-70	34-84	37-78
	Mean Age	50.9	48.8	43.6	56.7	57.4
Male		76	22	18	24	12
	Age Range	19-82	19-76	23-70	34-82	38-68
	Mean Age	49.6	46.6	40.7	56.1	55.5

Female		72	42	10	7	13
	Age Range	21-84	21-82	31-61	48-84	37-78
	Mean Age	52.3	48.4	48.7	58.9	59.1
Leucine		74	37	21	11	5
	Age Range	14-86	24-78	14-72	38-69	44-86
	Mean Age	49.1	48.4	47	50.5	60.4
Male		36	12	12	8	4
	Age Range	14-78	28-78	14-70	38-69	44-70
	Mean Age	48.8	49.2	42.1	49.6	54
Female		38	25	9	3	1
	Age Range	24-86	24-74	39-72	46-57	86
	Mean Age	50.7	48	53.7	53	N/A

* Other AA indicates strains containing any amino acid other than a glycine at this position

N/A stands for Not Applicable

agreement with the smaller subset of isolates we previously analyzed (27), there was no association with variation at this residue and the *cagA* allele (P=0.4433). Variation was also not associated with the overall *vacA* allele (P=0.2177), nor with either the i (P=0.1692) or m (P=0.8097) sub-regions of the *vacA* allele. Polymorphisms at this amino acid also did not impact disease state, even when individual disease states were compared directly to each other (Table 10).

Amino Acid 231

Amino acid 231, which is the ninth amino acid in cluster B, was previously shown to contain amino acid polymorphisms that are important for disease development in a Taiwanese population (50). In our population, we identified five different amino acids at this position: glycine, serine, aspartic acid, asparagine, and arginine. To determine if distribution of residues at this position was important, statistical associations were analyzed across the distribution of all the different amino acids as well as glycine versus all other amino acids combined; a previous study identified the presence of the glycine residue as important for the progression to more severe disease (50). Since we identified no difference between which associations were statistically significant and since the previous literature assessed the glycine residue versus any other amino acid (50), the numbers we present compare the distribution of glycine versus all other amino acids found at position 231. The distribution of amino acids at position 231 was not associated with gender (P=0.1109; Table 9). The strong association we previously identified using a subset of South Korean isolates (27) between the distribution of amino acid polymorphisms at this position and the distribution of the *cagA* allele was maintained in

this larger population (P=0.0081). More specifically, among EPIYA-ABD strains, glycine was more prevalent at this position (72.5%) than among strains carrying a non EPIYA-ABD *cagA* allele (48.5%). Polymorphisms at this site were also associated with the overall *vacA* allele (P<0.0001), every sub-region of the *vacA* allele (i region P<0.0001 or m region P=0.0168), and every combination of sub-regions (s and i P<0.0001, m and i P<0.0001 and s and m P=0.0168). However, even though glycine at this position was previously identified as important for the progression to more severe disease (50), we found that polymorphisms at this position did not impact disease state. Additionally there was no significance when comparing across any individual disease, regardless if individual disease states were compared alone or if a combination of disease states was analyzed (Table 10). These data suggest that this amino acid is not important for disease progress or that other factors mask the contribution of this amino acid to disease state in the South Korean population.

Amino Acid 196

As we previously described in a subset of South Korean strains (27), amino acid 196 was identified as a position that contained amino acid polymorphisms. In this larger population, either a serine (148 isolates, 67%) or a leucine (74 isolates, 33%) was found at this position. The distribution of amino acids at this position was not associated with gender (P=0.7762) and was fairly evenly distributed between males and females (Table 9). Additionally, there was still no association with the *cagA* allele (P=0.6928). However, polymorphisms at this residue were associated with the *vacA* allele (P=0.0003). This association was present for the i region of the *vacA* allele (P=0.0001), or any

combination that contained the i region (s and i P=0.0001 and m and i regions P=0.0003). However, there was no association when the m region was assessed alone (P=0.7600) or when the combination of regions did not include the i region (s and m P=0.0760). In fact, all of the i2 strains contain a leucine at this position (Fig. 11).

Polymorphisms at this amino acid did not impact disease state (P=0.0624) as a whole (Table 10). Additionally, this position was not significant when gastritis was compared to any non-cancer disease states: gastritis versus all other disease states (duodenal ulcers, gastric ulcers, and gastric cancers; P=0.3916), gastritis versus peptic ulcers (both duodenal and gastric; P=0.8808), gastritis versus duodenal ulcers (P=0.4796), or gastritis versus gastric ulcers (P=0.2500). It was also not significant when peptic ulcers were compared to non-cancer disease states: duodenal ulcers versus all other disease states (P=0.1237), duodenal ulcers versus gastric ulcers (P=0.1250), or gastric ulcers versus all other disease states (P=0.3636). There was also no association with the amino acid at position 196 with gastric cancers versus peptic ulcers (P=0.0689). However, we did find an association between the amino acid at this position and the development of gastric cancer when compared to development of all other disease states (P=0.0389), versus duodenal ulcers alone (P=0.0254), and versus gastritis alone (P=0.0460). Additionally, there was a statistical association between polymorphisms at position 196 and more severe disease manifestations (gastric ulcers and gastric cancer) versus less severe disease manifestations (gastritis and duodenal ulcers; P=0.0155; Table 10). While the presence of a serine at this position was more prevalent across all patients, patients suffering from gastric cancer were five times more likely than other patients to carry a serine at this location. These data suggest that position 196 does impact disease

Table 10: P values of the distribution of amino acids at polymorphic positions across various disease states.

	Distribution of Amino Acids At Position		
	151	**231**	**196**
Across all Disease States	0.2639	0.8914	0.0624
Gastritis versus all other disease states (duodenal ulcers, gastric ulcers, and gastric cancers)	1.0000	0.4694	0.3916
Gastritis versus peptic ulcers (both duodenal and gastric)	0.7700	0.6418	0.8808
Gastritis versus duodenal ulcers	0.2818	0.8529	0.4796
Gastritis versus gastric ulcers	0.4627	0.6946	0.2500
Gastritis versus gastric cancers	0.4044	0.5132	**0.0460**
Duodenal ulcers versus all other disease states	0.1025	1.0000	0.1237

Duodenal ulcers versus gastric ulcers	0.1306	1.0000	0.1250
Duodenal ulcers versus gastric cancer	0.0994	0.8015	**0.0254**
Gastric ulcers versus all other disease states	0.3847	0.8533	0.3636
Gastric ulcers versus gastric cancer	1.0000	1.0000	0.3994
Gastric cancer versus all other gastric diseases	0.3207	0.6732	**0.0389**
Gastric cancer versus peptic ulcers	0.2910	0.8201	0.0689

Associations that were statistically significant are in bold script with the corresponding P value shaded in gray.

development, but that the overall impact is masked by the progression of this disease through gastric ulcers.

Higher Order Associations

Log linear modeling using a combination of available data revealed two direct three way associations: disease state, *cagA* allele, and variation at amino acid 231 (P=0.012), and disease state and variation at amino acids 196 and 231 (P=0.029). As mentioned above, we found no direct association with variation of amino acid 231 and disease state in our population (P=0.8914). However, we noted that the two direct three way associations described above involved disease state and amino acid 231, or disease state, this amino acid, and the *cagA* allele. We therefore asked if the presence of a Western or East Asian *cagA* allele affected the ability of variation at residue 231 to be associated with disease progression. We found that a two way association did exist between disease state and amino acid 231 but only within the non EPIYA-ABD population (P=0.0367). This again suggests that the effect of different virulence factors or polymorphisms within these virulence factors may be masked by which *cagA* allele is present.

Discussion

Polymorphisms within *vacA* have been studied for several years, but have primarily focused on the s and m regions. However, the newest identified polymorphic region of *vacA,* the i region, has been suggested to be a determinant of vacuolating activity, as well as the best indicator of disease pathology, at least in Western strains (2,

16, 47). Within this region, three clusters of polymorphisms have been reported: cluster A, B, and C (47). In fact, two different amino acid substitutions within i1 clusters have been identified as potential marker for Taiwanese VacA (50). However, as previously suggested by our group (27) both of these amino acid substitutions are conserved within the South Korean population: we found a phenylalanine instead of a tyrosine for the first amino acid in cluster A (94.4%) and a methionine for the second asparagine within cluster C (99.5%). This suggests that these amino acids substitution are not limited to the Taiwanese population, but could serve as a general marker for East Asian VacA.

Further analysis of the amino acid consensus sequence of clusters A, B, and C, revealed several differences from the consensus sequence previously reported by Rhead *et al.* (47) and Chung *et al*. (11). This South Korean population provides the largest number of VacA i sequences analyzed to date, allowing us to better identify amino acids that differ between the i1 and i2 alleles (Fig. 12). Cluster A was well conserved across strains and the i1 and i2 consensus sequences were very similar to one another. Interestingly, each of the three i3 strains that we identified showed the presence of differences in cluster A; one of the strains contained the exact i1 consensus sequence, one an exact i2 consensus sequence, and that one was completely different (Fig. 12), indicating that these strains may be in the process of evolving from one allele into the other allele.

Within cluster B, the fifth, ninth and the tenth amino acids represent the main difference between the i1 and i2 alleles. The ninth amino acid (residue 231) was previously shown to have a role in disease (50). We found this residue was not variable in the small subset of i2 strains within our population, though it was variable within the i1

strains. Moreover, this variability at amino acid 231 does impact disease development, but only within strains containing a non EPIYA-ABD CagA. This finding suggests that within strains carrying the most virulent *cagA* allele, this residue is less important.

The differences between the i1 and i2 alleles were most pronounced in cluster C. This is primarily due to the addition of three polar amino acids, asparagine, histidine, and serine, but there were an additional five amino acids that differ between the i1 and i2 alleles (Fig. 12). Since both clusters B and C have been suggested to affect toxin activity, studies directed towards examination of the specific role of these amino acids in vacuolating activity would be of interest. Creation and examination of the activity of isogenic toxin derivatives varying only in the residues of interest would aide in this effort. Furthermore, given that the i3 strains appear to be hybrids of the i1 and i2 alleles at cluster B and C, it would be of significant interest to determine if there is a functional difference between vacuolating activity of i1, i2, and i3 toxins. To further investigate the role of these clusters, toxin activity could be assessed among our identified i3 strains, as well as those that contain an i1 cluster B and i2 cluster C. Finally, even though it has not been suggested to be important for activity, it would also be interesting to assess the sequences from cluster A to analyze the variance of amino acids four and six, potentially providing insight into VacA evolution.

Interestingly, five possible different amino acids were found to occur at amino acid 231 (the ninth amino acid of cluster B): glycine, serine, arginine, asparagine, and aspartic acid. Of these, glycine is the only non polar amino acid within this group, suggesting that perhaps this residue is important for the conformation or folding of VacA. Indeed, this could explain the statistical association between the distribution of amino

acids at this position with every region of VacA. Previous work showed that a glycine found at amino acid 231 was linked to disease development within the Taiwanese population (50). However, among our South Korean population, the distribution of amino acids found at this position had no overall impact on disease development. Conversely, there was a strong association with variation at this residue and the *cagA* allele, as well as two direct three-way associations between disease state, the distribution of amino acids at this position, and either *cagA* allele or amino acid 196, both of which affect the development of cancer in this population (Table 10; 27, 29). Given the association with *cagA*, the impact of this amino acid across East Asian and Western *cagA* alleles was examined, and a direct two way association between amino acid 231 and disease state was found among the strains that carry non EPIYA-ABD *cagA* alleles. This emphasizes the importance of the i region when carried within the context of Western strains, and suggests that there may be other factors that are more important in disease development or that mask the importance of the i region in severe disease development among East Asian strains. Indeed, even among Western strains, the *cagA* allele was the most important virulence factor for development of gastric cancer, whereas the VacA i region was the best indicator for development of peptic ulcer disease (5). Further exploration of differences between Western and East Asian strains may help to explain the exact mechanism of interaction of CagA and VacA.

Variation at amino acid 196 was statistically linked only to the *vacA* i region, indicating that it may be a true indicator of i region associated impacts. While overall this amino acid was not linked to disease state nor was there an association between the less severe disease states and distribution of amino acids at position 196, there was a

statistical difference in the distribution of amino acids at this position when gastric cancer was compared against any other disease state, gastritis, or duodenal ulcers. However, there was no statistical difference in the distribution of amino acids at this position between gastric cancer and gastric ulcers. This may be due to the fact that gastric ulcers can be a precursor to gastric cancer (19, 36). Given this, more severe disease states (gastric cancer and ulcers) were next compared to less severe disease manifestations (duodenal ulcers and gastritis), and a statistical association was identified. This suggests that variation at residue 196 is important for progression to severe disease, and that the overall significance of the amino acids found at this position is probably masked by the lack of association between gastric cancer and gastric ulcers.

While a serine at this position was more prevalent across the overall population, patients suffering from gastric cancer were five times more likely to carry a serine at this location. However, all of the i2 alleles carried a leucine at this position, and if this trend holds true for a larger population of i2 strains, then its contribution to disease development may become more evident. Since, the distribution of amino acids at this position was not only linked to the overall Vac allele, but specifically to the i region, the different prevalence of serine found at this position may explain why a previous study concluded that the i region was the best predictor of disease (47).

Since both CagA and VacA polymorphisms influence disease development, perhaps it is not surprising that within a population of isolates from a country with one of the highest rates of *H. pylori* colonization and gastric cancer (21, 51, 55), the majority of strains encode for the most toxic form of both CagA and VacA. While polymorphisms important for disease severity have been identified within both CagA and VacA

individually, strains that are CagA positive and carry the VacA s1/m1 allele have been shown to induce highly active corpus gastritis, which has been associated with the progression to gastric cancer (37-39). Additionally, we recently presented evidence that within this high risk population for gastric cancer development, there is a significant interaction between the *vacA* allele, *cagA* allele, and disease state (27). However, the reason only a small percentage of the population develops cancer is still unclear. Though it is evident that host, environmental, and bacterial factors play a role in *H. pylori*-induced disease (reviewed in 8, 57), additional studies are required to determine the contribution of all of these factors, both individually and in conjunction with each other, to the development of *H. pylori*-induced gastric cancer.

Acknowledgements

We thank Jeannette Whitmire and Beth Carpenter for critical reading of this manuscript. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (R13-2003-013-04002-0) (to JHC). Contents of this manuscript are the sole responsibility of the authors and do not necessarily represent the official views of the DOD.

References

1. **Amieva, M. R., and E. M. El-Omar.** 2008. Host-bacterial interactions in *Helicobacter pylori* infection. Gastroenterology **134:**306-23.

2. **Argent, R. H., R. J. Thomas, D. P. Letley, M. G. Rittig, K. R. Hardie, and J. C. Atherton.** 2008. Functional association between the *Helicobacter pylori* virulence factors VacA and CagA. J Med Microbiol **57:**145-50.

3. **Atherton, J. C., P. Cao, R. M. Peek, Jr., M. K. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori.* Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem **270:**17771-7.

4. **Atherton, J. C., T. L. Cover, R. J. Twells, M. R. Morales, C. J. Hawkey, and M. J. Blaser.** 1999. Simple and accurate PCR-based system for typing vacuolating cytotoxin alleles of *Helicobacter pylori*. J Clin Microbiol **37:**2979-82.

5. **Basso, D., C. F. Zambon, D. P. Letley, A. Stranges, A. Marchet, J. L. Rhead, S. Schiavon, G. Guariso, M. Ceroti, D. Nitti, M. Rugge, M. Plebani, and J. C. Atherton.** 2008. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology **135:**91-9.

6. **Blaser, M. J.** 1998. *Helicobacter pylori* and gastric diseases. BMJ **316:**1507-10.

7. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

8. **Brenner, H., D. Rothenbacher, and V. Arndt.** 2009. Epidemiology of stomach cancer. Methods Mol Biol **472:**467-77.

9. **Carpenter, B. M., T. K. McDaniel, J. M. Whitmire, H. Gancz, S. Guidotti, S. Censini, and D. S. Merrell.** 2007. Expanding the *Helicobacter pylori* genetic toolbox: modification of an endogenous plasmid for use as a transcriptional reporter and complementation vector. Appl Environ Microbiol **73:**7506-14.

10. **Censini, S., C. Lange, Z. Xiang, J. E. Crabtree, P. Ghiara, M. Borodovsky, R. Rappuoli, and A. Covacci.** 1996. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc Natl Acad Sci U S A **93:**14648-53.

11. **Chung, C., A. Olivares, E. Torres, O. Yilmaz, H. Cohen, and G. Perez-Perez.** 2010. Diversity of VacA intermediate region among *Helicobacter pylori* strains from several regions of the world. J Clin Microbiol **48:**690-6.

12. **Covacci, A., J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328-33.

13. **Cover, T. L., and S. R. Blanke.** 2005. *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality. Nat Rev Microbiol **3:**320-32.

14. **Cover, T. L., S. A. Halter, and M. J. Blaser.** 1992. Characterization of HeLa cell vacuoles induced by *Helicobacter pylori* broth culture supernatant. Hum Pathol **23:**1004-10.

15. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res **14:**1188-90.

16. **Douraghi, M., Y. Talebkhan, H. Zeraati, F. Ebrahimzadeh, A. Nahvijoo, A. Morakabati, M. Ghafarpour, M. Esmaili, M. Bababeik, A. Oghalaie, N. Rakhshani, M. E. Hosseini, M. A. Mohagheghi, and M. Mohammadi.** 2009. Multiple Gene Status in *Helicobacter pylori* Strains and Risk of Gastric Cancer Development. Digestion **80:**200-207.

17. **Dunn, B. E., H. Cohen, and M. J. Blaser.** 1997. *Helicobacter pylori*. Clin Microbiol Rev **10:**720-41.

18. **Ernst, P. B., and B. D. Gold.** 2000. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu Rev Microbiol **54:**615-40.

19. **Furuta, T., S. Baba, M. Takashima, H. Futami, H. Arai, M. Kajimura, H. Hanai, and E. Kaneko.** 1998. Effect of *Helicobacter pylori* infection on gastric juice pH. Scand J Gastroenterol **33:**357-63.

20. **Gebert, B., W. Fischer, E. Weiss, R. Hoffmann, and R. Haas.** 2003. *Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation. Science **301:**1099-102.

21. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

22. **Hatakeyama, M.** 2008. Linking epithelial polarity and carcinogenesis by multitasking *Helicobacter pylori* virulence factor CagA. Oncogene **27:**7047-54.

23. **Hatakeyama, M.** 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat Rev Cancer **4:**688-94.

24. **Higashi, H., A. Nakaya, R. Tsutsumi, K. Yokoyama, Y. Fujii, S. Ishikawa, M. Higuchi, A. Takahashi, Y. Kurashima, Y. Teishikata, S. Tanaka, T. Azuma, and M. Hatakeyama.** 2004. *Helicobacter pylori* CagA induces Ras-independent morphogenetic response through SHP-2 recruitment and activation. J Biol Chem **279:**17205-16.

25. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

26. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein. Science **295:**683-6.

27. **Jang, S., K. R. Jones, C. H. Olsen, Y. M. Joo, Y. J. Yoo, I. S. Chung, J. H. Cha, and D. S. Merrell.** 2010. Epidemiological link between gastric disease and polymorphisms in VacA and CagA. J Clin Microbiol **48:**559-67.

28. **Ji, X., T. Fernandez, D. Burroni, C. Pagliaccia, J. C. Atherton, J. M. Reyrat, R. Rappuoli, and J. L. Telford.** 2000. Cell specificity of *Helicobacter pylori* cytotoxin is determined by a short region in the polymorphic midregion. Infect Immun **68:**3754-7.

29. **Jones, K. R., Y. M. Joo, S. Jang, Y. J. Yoo, H. S. Lee, I. S. Chung, C. H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol **47:**959-68.

30. **Letley, D. P., and J. C. Atherton.** 2000. Natural diversity in the N terminus of the mature vacuolating cytotoxin of *Helicobacter pylori* determines cytotoxin activity. J Bacteriol **182:**3278-80.

31. **Letley, D. P., J. L. Rhead, R. J. Twells, B. Dove, and J. C. Atherton.** 2003. Determinants of non-toxicity in the gastric pathogen *Helicobacter pylori*. J Biol Chem **278:**26734-41.

32. **Manente, L., A. Perna, E. Buommino, L. Altucci, A. Lucariello, G. Citro, A. Baldi, G. Iaquinto, M. A. Tufano, and A. De Luca.** 2008. The *Helicobacter pylori's* protein VacA has direct effects on the regulation of cell cycle and apoptosis in gastric epithelial cells. J Cell Physiol **214:**582-7.

33. **Marshall, B. J., and J. R. Warren.** 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet **1:**1311-5.

34. **Matysiak-Budnik, T., and F. Megraud.** 1997. Epidemiology of *Helicobacter pylori* infection with special reference to professional risk. J Physiol Pharmacol **48 Suppl 4:**3-17.

35. **McClain, M. S., P. Cao, H. Iwamoto, A. D. Vinion-Dubiel, G. Szabo, Z. Shao, and T. L. Cover.** 2001. A 12-amino-acid segment, present in type s2 but not type s1 *Helicobacter pylori* VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol **183:**6499-508.

36. **McColl, K. E., and E. el-Omar.** 1996. *Helicobacter pylori* and disturbance of gastric function associated with duodenal ulcer disease and gastric cancer. Scand J Gastroenterol Suppl **215:**32-7.

37. **Meining, A., M. Stolte, R. Hatz, N. Lehn, S. Miehlke, A. Morgner, and E. Bayerdorffer.** 1997. Differing degree and distribution of gastritis in *Helicobacter pylori*-associated diseases. Virchows Arch **431:**11-5.

38. **Miehlke, S., A. Hackelsberger, A. Meining, R. Hatz, N. Lehn, P. Malfertheiner, M. Stolte, and E. Bayerdorffer.** 1998. Severe expression of corpus gastritis is characteristic in gastric cancer patients infected with *Helicobacter pylori*. Br J Cancer **78:**263-6.

39. **Miehlke, S., C. Kirsch, K. Agha-Amiri, T. Gunther, N. Lehn, P. Malfertheiner, M. Stolte, G. Ehninger, and E. Bayerdorffer.** 2000. The *Helicobacter pylori vacA* s1, m1 genotype and *cagA* is associated with gastric carcinoma in Germany. Int J Cancer **87:**322-7.

40. **Montecucco, C., and R. Rappuoli.** 2001. Living dangerously: how *Helicobacter pylori* survives in the human stomach. Nat Rev Mol Cell Biol **2:**457-66.

41. **Neel, B. G., H. Gu, and L. Pao.** 2003. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends Biochem Sci **28:**284-93.

42. **Neugut, A. I., M. Hayek, and G. Howe.** 1996. Epidemiology of gastric cancer. Semin Oncol **23:**281-91.

43. **Nguyen, L. T., T. Uchida, K. Murakami, T. Fujioka, and M. Moriyama.** 2008. *Helicobacter pylori* virulence and the diversity of gastric cancer in Asia. J Med Microbiol **57:**1445-53.

44. **Pagliaccia, C., M. de Bernard, P. Lupetti, X. Ji, D. Burroni, T. L. Cover, E. Papini, R. Rappuoli, J. L. Telford, and J. M. Reyrat.** 1998. The m2 form of the *Helicobacter pylori* cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A **95:**10212-7.

45. **Parsonnet, J., G. D. Friedman, D. P. Vandersteen, Y. Chang, J. H. Vogelman, N. Orentreich, and R. K. Sibley.** 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med **325:**1127-31.

46. **Pugsley, A. P.** 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiol Rev **57:**50-108.

47. **Rhead, J. L., D. P. Letley, M. Mohammadi, N. Hussein, M. A. Mohagheghi, M. Eshagh Hosseini, and J. C. Atherton.** 2007. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology **133:**926-36.

48. **Roovers, K., and R. K. Assoian.** 2000. Integrating the MAP kinase signal into the G1 phase cell cycle machinery. Bioessays **22:**818-26.

49. **Schneider, T. D., and R. M. Stephens.** 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res **18:**6097-100.

50. **Sheu, S. M., K. H. Hung, B. S. Sheu, H. B. Yang, and J. J. Wu.** 2009. Association of nonsynonymous substitutions in the intermediate region of the *vacA* gene of *Helicobacter pylori* with gastric diseases in Taiwan. J Clin Microbiol **47:**249-51.

51. **Shin, A., H. R. Shin, D. Kang, S. K. Park, C. S. Kim, and K. Y. Yoo.** 2005. A nested case-control study of the association of *Helicobacter pylori* infection with gastric adenocarcinoma in Korea. Br J Cancer **92:**1273-5.

52. **Szabo, I., S. Brutsche, F. Tombola, M. Moschioni, B. Satin, J. L. Telford, R. Rappuoli, C. Montecucco, E. Papini, and M. Zoratti.** 1999. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. EMBO J **18:**5517-27.

53. **Terebiznik, M. R., D. Raju, C. L. Vazquez, K. Torbricki, R. Kulkarni, S. R. Blanke, T. Yoshimori, M. I. Colombo, and N. L. Jones.** 2009. Effect of *Helicobacter pylori'*s vacuolating cytotoxin on the autophagy pathway in gastric epithelial cells. Autophagy **5:**370-9.

54. **The EUROGAST Study Group.** 1993. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. The EUROGAST Study Group. Gut **34:**1672-6.

55. **Tokudome, S., R. Ando, R. Ghadimi, T. Tanaka, N. Hattori, Z. Yang, M. Marumoto, H. Agawa, K. Arakawa, Y. Osaka, H. Tanaka, A. Hosono, and M. A. Moore.** 2007. Are there any real *Helicobacter pylori* infection-negative gastric cancers in Asia? Asian Pac J Cancer Prev **8:**462-3.

56. **Torres, V. J., S. E. VanCompernolle, M. S. Sundrud, D. Unutmaz, and T. L. Cover.** 2007. *Helicobacter pylori* vacuolating cytotoxin inhibits activation-induced proliferation of human T and B lymphocyte subsets. J Immunol **179:**5433-40.

57. **Tsugane, S., and S. Sasazuki.** 2007. Diet and the risk of gastric cancer: review of epidemiological evidence. Gastric Cancer **10:**75-83.

58. **Tsutsumi, R., A. Takahashi, T. Azuma, H. Higashi, and M. Hatakeyama.** 2006. Focal adhesion kinase is a substrate and downstream effector of SHP-2 complexed with *Helicobacter pylori* CagA. Mol Cell Biol **26:**261-76.

59. **Van Doorn, L. J., C. Figueiredo, F. Megraud, S. Pena, P. Midolo, D. M. Queiroz, F. Carneiro, B. Vanderborght, M. D. Pegado, R. Sanna, W. De Boer, P. M. Schneeberger, P. Correa, E. K. Ng, J. Atherton, M. J. Blaser, and W. G. Quint.** 1999. Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. Gastroenterology **116:**823-30.

60. **Willhite, D. C., and S. R. Blanke.** 2004. *Helicobacter pylori* vacuolating cytotoxin enters cells, localizes to the mitochondria, and induces mitochondrial membrane permeability changes correlated to toxin channel activity. Cell Microbiol **6:**143-54.

61. **Yamaoka, Y., T. Kodama, K. Kashima, D. Y. Graham, and A. R. Sepulveda.** 1998. Variants of the 3' region of the *cagA* gene in *Helicobacter pylori* isolates from patients with different *H. pylori*-associated diseases. J Clin Microbiol **36:**2258-63.

62. **Yamazaki, S., A. Yamakawa, T. Okuda, M. Ohtani, H. Suto, Y. Ito, Y. Yamazaki, Y. Keida, H. Higashi, M. Hatakeyama, and T. Azuma.** 2005. Distinct diversity of *vacA*, *cagA*, and *cagE* genes of *Helicobacter pylori* associated with peptic ulcer in Japan. J Clin Microbiol **43:**3906-16.

Table S3: Complete Korean Collection

Strain	**Disease**	**Sex**	**Age**	**CagA EPIYA motif**	***cagA* accession number**	***vacA allele***	***vacA* accession number**
K1-CA	Cancer	F	68	ABD	FJ458117	s1/i1/m2	GQ338184
K2-CA	Cancer	F	64	ABD		s1/i1/m1	GQ338205
K3-CA	Cancer	F	65	ABD	FJ458118	s1/i1/m1	GQ338222
K4-CA	Cancer	F	37	ABD		s1/i1/m1	GQ338225
K5-CA	Cancer	M	70	ABD		s1/i1/m1	GQ338227
K6-CA	Cancer	F	45	ABD		s1/i1/m1	GQ338233
K7-CA	Cancer	M	56	ABD		s1/i1/m1	GQ338235
K8-CA	Cancer	M	56	ABD		s1/i1/m1	GQ338239
K9-CA	Cancer	M	58	ABD	FJ458119	s1/i1/m1	GQ338242
K10-CA	Cancer	M	52	ABD		s1/i1/m1	HM047617
K11-CA	Cancer	M	68	ABD		s1/i1/m1	HM047618
K12-G	Gastritis	F	52	ABD		s1/i1/m1	HM047619
K13-CA	Cancer	M	38	ABD		s1/i1/m1	HM047620
K14-CA	Cancer	F	78	ABD		s1/i1/m1	HM047621
K15-CA	Cancer	F	66	ABD	FJ458120	s1/i1/m1	HM047622
K16-CA	Cancer	M	48	ABD		s1/i1/m1	GQ338194
K17-CA	Cancer	F	56	ABD		s1/i1/m1	GQ338196
K18-CA	Cancer	M	64	ABD		s1/i1/m1	HM047623
K19-CA	Cancer	F	86	ABD		s1/i3/m1	HM047624
K20-CA	Cancer	M	48	ABD		s1/i1/m1	HM047625
K21-CA	Cancer	M	44	ABD		s1/i1/m1	GQ338209

K22-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	HM047626
K23-DU	Duodenal Ulcer	M	47	ABD		s1/i3/m1	GQ338210
K24-DU	Duodenal Ulcer	M	38	ABD	FJ458121	s1/i1/m1	HM047627
K25-DU	Duodenal Ulcer	M	44	ABD		s1/i1/m1	GQ338214
K26-DU	Duodenal Ulcer	M	20	ABD		s1/i1/m1	GQ338218
K27-DU	Duodenal Ulcer	M	47	ABD		s1/i1/m1	HM047628
K28-DU	Duodenal Ulcer	M	28	ABD		s1/i1/m1	HM047629
K29-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m1	HM047630
K30-DU	Duodenal Ulcer	F	61	ABCCC	FJ458122	s1/i2/m2	HM047631
K31-DU	Duodenal Ulcer	M	33	ABD		s1/i1/m1	HM047632
K32-DU	Duodenal Ulcer	F	41	ABD		s1/i1/m1	HM047633
K33-DU	Duodenal Ulcer	F	31	ABC	FJ458123	s1/i1/m1	HM047634
K34-DU	Duodenal Ulcer	M	43	ABD	FJ458124	s1/i1/m1	GQ338223
K35-DU	Duodenal Ulcer	F	56	ABD		s1/i1/m2	HM047635
K36-DU	Duodenal Ulcer	M	46	ABD		s1/i1/m1	HM047636
K37-DU	Duodenal Ulcer	M	61	ABD		s1/i2/m2	HM047637
K38-DU	Duodenal Ulcer	F	39	ABC		s1/i1/m1	GQ338224
K39-DU	Duodenal Ulcer	M	59	ABD		s1/i1/m1	HM047638
K40-DU	Duodenal Ulcer	M	53	ABD		s1/i1/m1	HM047639
K41-DU	Duodenal Ulcer	M	55	ABD		s1/i1/m1	HM047640
K42-DU	Duodenal Ulcer	F	48	ABD		s1/i1/m2	HM047641
K43-DU	Duodenal Ulcer	M	70	ABD		s1/i1/m1	HM047642
K44-DU	Duodenal Ulcer	M	42	ABD		s1/i1/m1	HM047643
K45-DU	Duodenal Ulcer	M	22	ABD	FJ458125	s1/i1/m1	HM047644

K46-DU	Duodenal Ulcer	F	61	ABD		s1/i1/m1	HM047645
K47-DU	Duodenal Ulcer	F	72	ABD		s1/i2/m2	HM047646
K48-DU	Duodenal Ulcer	F	41	ABD		s1/i1/m1	HM047647
K49-DU	Duodenal Ulcer	M	33	ABC		s1/i1/m1	GQ338226
K50-DU	Duodenal Ulcer	M	35	ABD		s1/i1/m1	HM047648
K51-GU	Gastric Ulcer	M	54	ABD		s1/i1/m1	GQ338228
K52-GU	Gastric Ulcer	M	46	ABD		s1/i1/m1	GQ338229
K53-G	Gastritis	M	60	ABD		s1/i1/m1	HM047649
K54-G	Gastritis	F	58	ABD	FJ458126	s1/i1/m1	HM047650
K55-GU	Gastric Ulcer	F	57	ABD		s1/i1/m1	GQ338230
K56-G	Gastritis	F	48	ABD		s1/i1/m1	HM047651
K57-G	Gastritis	F	63	ABD		s1/i1/m1	GQ338231
K58-G	Gastritis	F	61	ABD		s1/i1/m1	GQ338232
K59-G	Gastritis	M	48	BBD	FJ458127	s1/i1/m1	HM047652
K60-G	Gastritis	M	53	ABC		s1/i1/m1	HM047653
K61-GU	Gastric Ulcer	F	57	*		*	
K62-GU	Gastric Ulcer	F	65	*		*	
K63-GU	Gastric Ulcer	M	59	*		*	
K64-G	Gastritis	F	61	ABCC	FJ458128	s1/i1/m1	HM047654
K65-G	Gastritis	M	49	ABD		s1/i1/m1	HM047655
K66-G	Gastritis	M	43	ABC		s1/i2/m2	HM047656
K67-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m1	HM047657
K68-GU	Gastric Ulcer	F	46	ABD		s1/i1/m1	HM047658
K69-GU	Gastric Ulcer	M	63	ABD	FJ458129	s1/i1/m1	GQ338234

K70-G	Gastritis	F	68	ABC	FJ458130	s1/i2/m2	HM047659
K71-G	Gastritis	M	54	ABD		s1/i1/m1	HM047660
K72-GU	Gastric Ulcer	M	34	ABD		s1/i1/m1	GQ338236
K73-GU	Gastric Ulcer	M	72	ABD		s1/i1/m1	HM047661
K74-G	Gastritis	F	52	ABD		s1/i1/m1	GQ338237
K75-G	Gastritis	F	24	ABD		s1/i1/m1	HM047662
K76-G	Gastritis	F	55	ABD		s1/i1/m1	HM047663
K77-G	Gastritis	F	37	ABD		s1/i1/m1	GQ338238
K78-G	Gastritis	M	36	AABD	FJ458131	s1/i1/m1	HM047664
K79-GU	Gastric Ulcer	F	84	ABD		s1/i1/m1	HM047665
K80-CA	Cancer	F	61	ABD		s1/i1/m1	GQ338240
K81-GU	Gastric Ulcer	M	47	ABD		s1/i1/m1	GQ338241
K82-G	Gastritis	M	39	ABD	FJ458132	s1/i1/m1^{+}	
K83-G	Gastritis	F	75	ABD		s1/i1/m1	HM047666
K84-G	Gastritis	M	48	ABD		s1/i1/m1	HM047667
K85-G	Gastritis	F	28	BD		s1/i1/m1	HM047668
K86-G	Gastritis	M	37	ABCC		s1/i1/m1	HM047669
K87-G	Gastritis	F	52	ABD		s1/i1/m1	HM047670
K88-G	Gastritis	F	69	ABD		s1/i1/m1	HM047671
K89-GU	Gastric Ulcer	M	38	ABD		s1/i1/m1	HM047672
K90-GU	Gastric Ulcer	M	51	ABD		s1/i1/m1	HM047673
K91-GU	Gastric Ulcer	M	82	ABD		s1/i1/m1	HM047674
K92-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	HM047675
K93-DU	Duodenal Ulcer	F	37	ABC	FJ458133	s1/i1/m1	GQ338243

K94-GU	Gastric Ulcer	M	65	ABD		s1/i1/m1	HM047676
K95-CA	Cancer	F	41	ABD		s1/i1/m1	HM047677
K96-G	Gastritis	F	47	ABD		s1/i1/m1	HM047678
K97-GU	Gastric Ulcer	M	51	ABD		s1/i1/m1	HM047679
K98-DU	Duodenal Ulcer	M	23	ABD		s1/i1/m1	HM047680
K99-G	Gastritis	F	54	ABD		s1/i2/m2	HM047681
K100-GU	Gastric Ulcer	M	46	ABD		s1/i1/m1	HM047682
K101-GU	Gastric Ulcer	F	61	ABD		s1/i1/m1	HM047683
K102-DU	Duodenal Ulcer	M	38	ABD		s1/i1/m1	HM047684
K103-G	Gastritis	M	32	ABD		s1/i3/m1	HM047685
K104-CA	Cancer	M	46	ABD		s1/i1/m1	GQ338185
K105-GU	Gastric Ulcer	M	71	ABD		s1/i1/m1	HM047686
K106-DU	Duodenal Ulcer	M	14	ABD		s1/i1/m1	HM047687
K107-DU	Duodenal Ulcer	M	26	ABD		s1/i1/m1	HM047688
K108-GU	Gastric Ulcer	M	62	ABD		s1/i1/m1	HM047689
K109-G	Gastritis	M	40	ABD		s1/i1/m1	GQ338186
K110-GU	Gastric Ulcer	M	81	ABCC	FJ458134	s1/i1/m1	HM047690
K111-DU	Duodenal Ulcer	F	36	ABD		s1/i1/m1	GQ338187
K112-G	Gastritis	M	57	ABD		s1/i1/m1	GQ338188
K113-G	Gastritis	M	29	ABD		s1/i1/m1	HM047691
K114-DU	Duodenal Ulcer	F	47	ABC		s1/i1/m1	HM047692
K115-G	Gastritis	F	82	ABC	FJ458135	s1/i1/m2	HM047693
K116-G	Gastritis	M	59	ABD		s1/i1/m1	HM047694
K117-G	Gastritis	F	21	ABD	FJ458136	s1/i1/m1	GQ338189

K118-CA	Cancer	F	67	ABD		s1/i1/m1	HM047695
K119-DU	Duodenal Ulcer	M	31	ABD	FJ458137	s1/i1/m1	HM047696
K120-G	Gastritis	F	41	ABC		s1/i1/m1	GQ338190
K121-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	HM047697
K122-DU	Duodenal Ulcer	M	60	ABD		s1/i1/m1	HM047698
K123-G	Gastritis	M	76	ABD	FJ458138	s1/i1/m1	GQ338191
K125-G	Gastritis	M	59	ABD		s1/i1/m1	HM047699
K126-GU	Gastric Ulcer	M	69	ABD		s1/i1/m1	HM047700
K127-GU	Gastric Ulcer	M	71	ABD		s1/i1/m1	HM047701
K128-GU	Gastric Ulcer	M	58	ABC		s1/i1/m1	GQ338192
K129-GU	Gastric Ulcer	M	36	*		*	
K130-G	Gastritis	F	64	*		*	
K131-G	Gastritis	F	61	ABD	FJ458139	s1/i1/m1	GQ338193
K132-GU	Gastric Ulcer	M	23	*		*	
K133-GU	Gastric Ulcer	M	63	*		*	
K134-GU	Gastric Ulcer	M	46	*		*	
K135-DU	Duodenal Ulcer	F	62	*		*	
K136-G	Gastritis	F	52	*		*	
K137-G	Gastritis	F	62	*		*	
K138-GU	Gastric Ulcer	F	21	*		*	
K139-GU	Gastric Ulcer	F	49	*		*	
K140-G	Gastritis	F	49	*		*	
K141-G	Gastritis	M	57	*		*	
K142-GU	Gastric Ulcer	M	65	*		*	

K143-GU	Gastric Ulcer	F	71	*		*	
K144-GU	Gastric Ulcer	F	53	*		*	
K145-GU	Gastric Ulcer	M	62	*		*	
K146-G	Gastritis	M	40	ABD**	FJ458140	s1/i1/m1	HM047702
K147-GU	Gastric Ulcer	M	62	*		*	
K148-GU	Gastric Ulcer	M	37	*		*	
K149-GU	Gastric Ulcer	M	71	*		*	
K150-G	Gastritis	F	26	ABD		s1/i1/m1	HM047703
K151-GU	Gastric Ulcer	M	65	ABD	FJ458141	s1/i1/m1	HM047704
K152-G	Gastritis	F	62	ABD		s1/i1/m1	HM047705
K153-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	HM047706
K154-G	Gastritis	F	55	ABCCCC	FJ458142	s1/i1/m2	HM047707
K155-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	HM047708
K156-DU	Duodenal Ulcer	F	47	ABD		s1/i1/m1	HM047709
K157-G	Gastritis	M	43	ABD		s1/i1/m1	HM047710
K158-G	Gastritis	M	60	ABD		s1/i1/m1	HM047711
K159-G	Gastritis	F	35	ABD		s1/i1/m1	HM047712
K160-DU	Duodenal Ulcer	M	30	ABD		s1/i1/m1	HM047713
K161-G	Gastritis	F	65	ABD		s1/i1/m1	HM047714
K162-G	Gastritis	F	63	ABD		s1/i1/m1	GQ338195
K163-G	Gastritis	F	66	ABD		s1/i1/m1	HM047715
K164-G	Gastritis	M	43	ABD		s1/i1/m1	HM047716
K165-G	Gastritis	M	28	ABD		s1/i1/m1	HM047717
K166-G	Gastritis	F	38	ABC		s1/i2/m2	HM047718
K167-G	Gastritis	F	27	ABD		s1/i1/m1	HM047719
K169-G	Gastritis	F	47	ABD		s1/i1/m1	HM047720

K170-G	Gastritis	F	41	ABD**	FJ458143	s1/i1/m1	HM047721
K171-CA	Cancer	F	72	ABD	FJ458144	s1/i1/m1	HM047722
K172-G	Gastritis	F	31	ABCC	FJ458145	s1/i1/m1	HM047723
K173-G	Gastritis	F	45	ABD	FJ458146	s1/i1/m1	HM047724
K174-G	Gastritis	N/A	N/A	ABD		s1/i1/m1	HM047725
K175-G	Gastritis	F	41	ABD		s1/i1/m2	HM047726
K176-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	HM047727
K177-G	Gastritis	F	39	ABD		s1/i1/m1	HM047728
K178-G	Gastritis	F	40	ABD		s1/i1/m1	GQ338197
K179-G	Gastritis	F	38	ABCCC	FJ458147	s1/i1/m2	HM047729
K180-G	Gastritis (polyps)	F	50	ABD		s1/i1/m1	HM047730
K181-DU	Duodenal Ulcer	F	57	ABD		s1/i1/m2	HM047731
K182-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	GQ338198
K183-G	Gastritis	M	40	ABD		s1/i1/m1	GQ338199
K184-G	Gastritis	F	55	ABD		s1/i1/m1	HM047732
K185-G	Gastritis	F	52	ABD		s1/i1/m1	GQ338200
K186-G	Gastritis	M	41	ABD		s1/i1/m1	HM047733
K188-G	Gastritis (IM)	F	43	ABD		s1/i1/m1	HM047734
K190-G	Gastritis	M	61	ABC		s1/i1/m1	GQ338201
K192-DU	Duodenal Ulcer	F	61	AABD	FJ458148	s1/i1/m1	HM047735
K193-G	Gastritis	F	50	ABD		s1/i1/m1	GQ338202
K194-GU	Gastric Ulcer	M	48	*		s1/i1/m1	HM047564
K195-GU	Gastric Ulcer	F	48	ABD	FJ458149	s1/i1/m1	HM047565
K196-G	Gastritis	F	50	ABD		s1/i1/m1	GQ338203
K197-G	Gastritis	F	45	ABD		s1/i1/m1	GQ338204

K198-GU	Gastric Ulcer	F	56	ABC	FJ458150	s1/i2/m2	HM047566
K199-GU	Gastric Ulcer	M	50	ABD		s1/i1/m1	HM047567
K200-GU	Gastric Ulcer	M	63	ABD		s1/i1/m1	HM047568
K201-GU	Gastric Ulcer	M	55	ABD		s1/i1/m1	HM047569
K202-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	HM047570
K203-G	Gastritis	F	55	ABD		s1/i1/m1	HM047571
K204-GU	Gastric Ulcer	F	63	ABD		s1/i1/m2	HM047572
K205-GU	Gastric Ulcer	F	57	ABD		s1/i1/m1	HM047573
K206-GU	Gastric Ulcer	F	51	ABD		s1/i1/m1	HM047574
K207-G	Gastritis	M	39	ABD		*	GQ338206
K208-G	Gastritis	F	56	ABD	FJ458151	s1/i1/m1	GQ338207
K209-G	Gastritis	F	24	ABD		s1/i1/m1	GQ338208
K210-G	Gastritis	F	61	ABD		s1/i1/m1	HM047575
K211-G	Gastritis	F	54	ABD		s1/i1/m1	HM047576
K212-G	Gastritis	F	45	ABD		s1/i1/m1	HM047577
K213-G	Gastritis	F	52	*		*	
K214-G	Gastritis	F	53	*		*	
K215-GU	Gastric Ulcer	F	73	*		*	
K216-G	Gastritis	F	67	ABD		s1/i1/m1	HM047578
K217-G	Gastritis	M	77	ABD		*	
K218-G	Gastritis	F	62	ABD		s1/i1/m1	HM047579
K219-G	Gastritis	M	40	ABD	FJ458152	s1/i1/m1	HM047580
K220-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	HM047581
K221-G	Gastritis	M	37	BD		s1/i1/m1	HM047582
K222-GU	Gastric Ulcer	M	41	ABD		s1/i1/m1	HM047583
K223-G	Gastritis	F	25	ABD	FJ458153	s1/i1/m1	HM047584

K224-G	Gastritis	F	35	ABD		s1/i1/m1	HM047585
K225-DU	Duodenal Ulcer	F	60	ABC		s1/i1/m1	HM047586
K226-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	HM047587
K227-G	Gastritis	F	31	ABD		s1/i1/m1	HM047588
K228-GU	Gastric Ulcer	M	54	ABD		s1/i1/m1	HM047589
K229-GU	Gastric Ulcer	M	62	ABD		s1/i1/m1	HM047590
K230-G	Gastritis (IM)	M	56	ABD		s1/i1/m1	HM047591
K231-GU	Gastric Ulcer	M	42	ABD		s1/i1/m1	HM047592
K232-G	Gastritis	F	56	ABD		*	
K233-G	Gastritis	M	38	ABD		s1/i1/m1	HM047594
K234-DU	Duodenal Ulcer	M	41	ABD		s1/i1/m2	HM047595
K235-G	Gastritis	F	50	ABD		s1/i1/m1	HM047596
K236-G	Gastritis	F	64	ABD		s1/i1/m1	HM047597
K237-G	Gastritis	F	48	ABD		s1/i1/m1	HM047598
K238-DU	Duodenal Ulcer	M	55	ABD		s1/i1/m1	GQ338211
K239-G	Gastritis	M	46	ABD		*	HM047599
K240-G	Gastritis	F	41	ABD		s1/i1/m1	HM047600
K241-G	Gastritis	M	41	ABD		s1/i1/m1	HM047601
K242-G	Gastritis	M	78	ABD		s1/i1/m1	HM047602
K243-G	Gastritis	F	60	BC		*	
K244-DU	Duodenal Ulcer	N/A	N/A	ABD		s1/i1/m1	HM047603
K245-G	Gastritis	M	19	ABD		s1/i1/m1	HM047604
K246-G	Gastritis	F	40	ABD		s1/i1/m1	HM047605
K247-G	Gastritis	F	56	ABD		s1/i1/m1	HM047606
K248-G	Gastritis	M	58	ABD		s1/i1/m1	GQ338212
K249-GU	Gastric Ulcer	F	48	ABC		s1/i1/m1	GQ338213

K250-G	Gastritis	F	53	ABD		s1/i1/m1	HM047607
K251-DU	Duodenal Ulcer	M	70	ABD		s1/i1/m1	HM047608
K253-DU	Duodenal Ulcer	F	61	ABC		s1/i1/m1	HM047609
K254-G	Gastritis	M	54	ABD		s1/i1/m1	HM047610
K255-G	Gastritis	F	74	ABD	FJ458154	s1/i1/m1	GQ338215
K256-G	Gastritis	F	51	ABD		s1/i1/m1	HM047611
K257-CA	Cancer	M	64	ABD		s1/i1/m1	HM047612
K258-CA	Cancer	M	68	ABD	FJ458155	s1/i1/m1	GQ338216
K259-CA	Cancer	M	44	ABD	FJ458156	s1/i1/m2	GQ338217
K260-CA	Cancer	M	58	ABD	FJ458157	s1/i1/m1	GQ338219
K261-CA	Cancer	F	48	ABD	FJ458158	s1/i1/m1	GQ338220
K262-G	Gastritis	F	56	ABC	FJ458159	s1/i1/m1	GQ338221
K263-G	Gastritis	M	59	ABABD***	FJ458160	s1/i1/m1	HM047613
K264-DU	Duodenal Ulcer	M	32	ABD	FJ458161	s1/i1/m1	HM047614
K265-DU	Duodenal Ulcer	M	42	ABD	FJ458162	s1/i1/m1	HM047615
K266-G	Gastritis	F	34	ABD	FJ458163	s1/i1/m1	HM047616

*Indeterminate in genotyping assay

**-B motif's proline is replaced with a serine, ESIYA, therefore classified as other

***-ABABD second -B motif's proline is replaced with leucine, ELIYA, therefore classified as other

a + indicates frameshift in the i region

Chapter Five

Construction and Analysis of Isogenic Strains of Helicobacter pylori *That Differ Only in the CagA EPIYA Motif*

Kathleen R. Jones, Jeannette M. Whitmire, Shana Miles, Sungil Jang, Jeong-Heon Cha, and D. Scott Merrell

The work presented in this chapter is the sole work of K. R. Jones with the following exceptions: J.H. Cha assisted with creation of the isogenic strains, and S. Jang assisted with the creation of the new restorant strain. S. Miles assisted with the animal work, and J.M. Whitmire assisted with immunofluorescent staining, microscopy, morphological assays, and figure generation.

Introduction

The process of *H. pylori*-induced pathogenesis, including development of gastric cancer, is not well understood. This is despite the fact that progress has been made in elucidating some key virulence factors that impact disease progression. In fact, a couple of virulence factors have been shown to have profound effects via the deregulation of host cell signaling pathways. The most well characterized of these virulence factors is the cytotoxin-associated gene A, *cagA*, which has emerged as a major contributor to the development of gastric cancer (6, 18). A number of *cagA* alleles exist, and these alleles differ in the carboxy-terminus of the encoded protein. Variation specifically occurs in the

EPIYA region, and typically involves changes in the amino acid sequences flanking the five-amino-acid repeat (19, 21, 22).

As mentioned in the introduction to this thesis, the sequence surrounding the EPIYA motif shows divergence across strains, which has led to the classification of four distinct EPIYA motifs: EPIYA-A, -B, -C, and -D (21). Western CagA contains a combination of EPIYA-A, -B, and -C motifs (isolates with up to five EPIYA-C motifs have been identified) (3), while East Asian CagA contains a combination of EPIYA-A, -B, and -D motifs (11, 21, 22, 28). These EPIYA repeats are not only important for CagA phosphorylation dependent effects, but a multimerization domain within the EPIYA repeat region also exists that affects downstream signaling pathways in a phosphorylation independent manner (26).

Numerous lines of evidence indicate that CagA polymorphisms may dramatically impact *H. pylori*-induced disease etiology. In fact, *in vitro* assays have demonstrated a dose-dependent response in the levels of CagA tyrosine phosphorylation, SHP-2 binding, and host cell morphological changes that occur as a result of increasing numbers of EPIYA-C motifs (22). There are also differences in induced inflammation depending on which CagA form is present in the infecting strain. Epidemiological studies have shown that among patients infected with *H. pylori* strains containing Western CagA, increased inflammation and increased disease severity correlate with an increasing number of EPIYA-C motifs (30). *In vitro* studies comparing East Asian CagA to Western CagA have also shown that East-Asian CagA binds SHP-2 with greater affinity and induces more significant morphological changes than Western CagA containing up to three EPIYA-C motifs (20, 21). Furthermore, epidemiological studies have shown that there is

significantly more inflammation and atrophy in patients infected with *H. pylori* strains carrying the East Asian CagA (5). *En masse* these data indicate that differences within the EPIYA motifs of CagA impact disease development (3-5, 14, 15, 30).

While this information has been useful and strongly suggests that differences in CagA impact disease development, all of these previous studies employed epidemiology data, used non-isogenic clinical isolates, or relied on transfection models. While the results indicate interesting trends, since the genetic variability of *H. pylori* strains is between 5-7%, there is concern that work with nonisogenic strans may not be a true indicator of the exact role of the EPIYA motif in disease development (1, 17, 24, 29). Furthermore, transfection assays likely do not recapitulate biological delivery of CagA by an infecting bacterium.

The goal of this study was to elucidate the role of the various EPIYA motifs in disease development. To this end we used splicing by overlapping extension (SOE) PCR and transformation to create isogenic strains of *H. pylori* that varied only in the EPIYA region of CagA. These strains were then assessed for their growth kinetics, ability to express CagA, localization differences on host cells, adherence to and internalization into host cells, ability to translocate and phosphorylate CagA, and ability to deregulate host cell pathways as assessed by changes in host cell morphology. Moreover, since a recent Mongolian gerbil model was shown to reproducibly develop CagA-dependent gastric cancer within a period of 6-12 weeks after infection with *H. pylori* strain 7.13 (14, 15), we used this model to investigate the role of the different EPIYA motifs in disease progression. Through the course of all these studies, it became apparent that there were secondary mutations within the isogenic strains that would prevent our ability to

appropriately compare the strains. However, the methods optimized during this study should be useful to the lab in the future. To this end, this chapter details the optimized methods for each of the assays mentioned above and briefly mentions the results that indicated the presence of secondary mutations in our strains.

Materials and Methods and Results

Bacterial Strains and Growth Conditions

Bacterial strains and plasmids are listed in Table 11, and primers are listed in Table 12. *E. coli* strains were maintained as frozen (-80ºC) stocks in LB broth (MoBio, Carlsbad, CA) supplemented with 40% glycerol (EMD Chemicals, Inc., Gibbstown, NJ), and expanded on LB agar (Mo Bio, Carlsbad, CA) plates or in LB broth liquid cultures. All cultures of *E. coli* were grown at 37°C and liquid cultures were maintained with shaking at 200 rpm. *H. pylori* strains were maintained as frozen (-80ºC) stocks in brain heart infusion broth (BD, Sparks, MD) supplemented with 20% glycerol and 10% fetal bovine serum (FBS; Gibco, Carlsbad, CA). All *H. pylori* strains were grown at 37°C and expanded on antibiotic-supplemented horse blood agar plates consisting of 4% Columbia agar base (Neogen Corporation, Lansing, MI), 5% defibrinated horse blood (HemoStat Laboratories, Dixon, CA), 0.2% β-cyclodextrin (Sigma, St. Louis, MO), 8 µg/ml amphotericin B (Amresco, Solon, OH), 2.5 U/ml polymyxin B (Sigma, St. Louis, MO), 5 µg/ml cefsulodin (Sigma, St. Louis, MO), 5 µg/ml trimethoprim (Sigma, St. Louis, MO), and 10 µg/ml vancomycin (Amresco, Solon, OH). *H. pylori* liquid cultures consisted of brucella broth (BB) (Neogen Corporation, Lansing, MI) containing 10% fetal bovine

Table 11: Bacterial strains and plasmids

Strain Name	Organism	EPIYA motif	Antibiotic resistance	Citation
7.13	*H. pylori*	EPIYA-ABtC		(14, 15)
DSM3	*E. coli* DH5α		Kan	(10)
DSM598	*E. coli* Top10		Amp	This study
DSM599	*E. coli* Top10		Amp & Kan	This study
DSM600	*H. pylori*	Δ*cagA*	Kan	This study
DSM530	*E. coli* Top10		Amp	This study
DSM531	*E. coli* Top10		Amp & Kan	This study
DSM577	*H. pylori*	ΔEPIYA	Kan	This study
DSM926	*H. pylori*	ΔEPIYA	Kan	This study
DSM591	*H. pylori*	EPIYA-ABtCCCC		(23)
DSM532	*E. coli* Top10	EPIYA-ABt	Amp	This study
DSM601	*H. pylori*	EPIYA-ABt		This study
DSM570	*E. coli* Top10	EPIYA-ABtC	Amp	This study
DSM602	*H. pylori*	EPIYA-ABtC		This study
DSM571	*E. coli* Top10	EPIYA-ABtCC	Amp	This study
DSM605	*H. pylori*	EPIYA-ABtCC		This study

DSM572	*E. coli* Top10	EPIYA-ABtCCC	Amp	This study
DSM606	*H. pylori*	EPIYA-ABtCCC		This study
DSM573	*E. coli* Top10	EPIYA-ABtCCCC	Amp	This study
DSM609	*H. pylori*	EPIYA-ABtCCCC		This study
DSM590	*H. pylori*	EPIYA-ABD		(23)
DSM533	*E. coli* Top10	EPIYA-ABD	Amp	This study
DSM547	*E. coli* Top10	EPIYA-ABtD	Amp	This study
DSM616	*H. pylori*	EPIYA-ABtD		This study
DSM641	*E. coli* Top10	EPIYA-ABtC	Amp	This study
DSM613	*H. pylori*	EPIYA-ABtC (WT)		This study
DSM927	*H. pylori*	EPIYA-ABtC (WT)		This study

Table 12: Primer Sequences

Primer Name	Primer Sequence	Citation
7.13del CagA-fp	CGTCTTTAACACAAGCAACACG	This study
7.13del CagA M-rp	GATTTTTGGCCCGGGAGGCTCGAGCATTGTTTCTCCTTACTATACC	This study
7.13 del CagA M-fp	GAAACAATGCTCGAGCCTCCCGGGCCAAAAATCTTAAAGGATTAAGG	This study
7.13 del CagA-rp	GTTTATGCTCTCTTTATAACCCC	This study
SacBSCN-F2	CGAATCGAATTCAGGAAC	(9)
Grace 1	GGTTGCACGCATTTTCCC	This study
7.13 del EPIYA-5-fp	GTCTGATAAGTTTGAAAACATC	This study
7.13 del EPIYA M-rp	GTCTATCCCCGGGAGGCTCGAGCCCATTACCGACTAGGGTTCC	This study
7.13 del EPIYA M-fp	GTAATGGGCTCGAGCCTCCCGGGGATAGACAAGCTCAAAGATTC	This study
7.13 del EPIYA-3-rp	CCTTGTTTTTAGCAAGGGGTGG	This study
K154-7IS-M1rp	GGCTTCTGCTTGAGATAACCCATTACCGACTAGGGTTCC	This study
K154-7IS-M2fp	GATAGACAAGCTCAAAGATTCTAC	This study
K154-7IS-M1fp	GGAACCCTAGTCGGTAATGGGTTATCTCAAGCAGAAGCC	This study
K154-7IS-M12rp	GTTTCAATTCTTGCTCCCTTGAAAGCCCTACCTTACTGAG	This study
K154-7IS-M12fp	GGGCTTTCAAGGGAGCAAGAATTGAAAC	This study
K154-7IS-M2rp	GTAGAATCTTTGAGCTTGTCTATC	This study
7.13cagA-2742-fp	GGAAGCAAAAGCTCAAGCTAACAGC	This study
7.13cagA-4561-rp	TACAGGTCTCACACATCATATCTCC	This study
K3-7IS-M1rp	CGTTGTGGCTTCTGTTTTAGATAACCCATTACCGACTAGGGTTCC	This study
K3-7IS-M2fp	GGTCATTTTGGCAAACTAGAACAAAAGATAGACAAGCTCAAAGATTC	This study
K3-7IS-M1fp	GGAACCCTAGTCGGTAATGGGTTATCTAAAACAGAAGCCACAACG	This study
K3-7IS- M2rp	GAATCTTTGAGCTTGTCTATCTTTTGTTCTAGTTTGCCAAAATGACC	This study
K3(A/T)-rp	CACCTTTTTAGCAACTTGAGTGTAAATGGGCTCTTCAGGGC	This study
K3(A/T)-fp	GCCCTGAAGAGCCCATTTACACTCAAGTTGCTAAAAAGGTG	This study

serum and 10 µg/ml vancomycin. For experiments, overnight liquid cultures of *H. pylori* were subcultured in fresh medium to an optical density at 600 nm of 0.05 and were grown for either 12 or 18 hours shaking at 100 rpm under microaerobic conditions (10% CO2, 5% O2, and 85% N2) created by an Anoxomat evacuation/replacement system (Spiral Biotech, Norwood, MA).

When appropriate, bacterial cultures were supplemented with the following antibiotics, as noted in Table 11: ampicillin (Amp) (USB Corporation, Santa Clara, CA) at 100 µg/ml and/or kanamycin (Kan) (Gibco, Carlsbad, CA) at 25 µg/ml. Additionally, 5% sucrose (Sigma, St. Louis, MO) was added to horse blood agar plates when needed.

Creation of Isogenic Strains

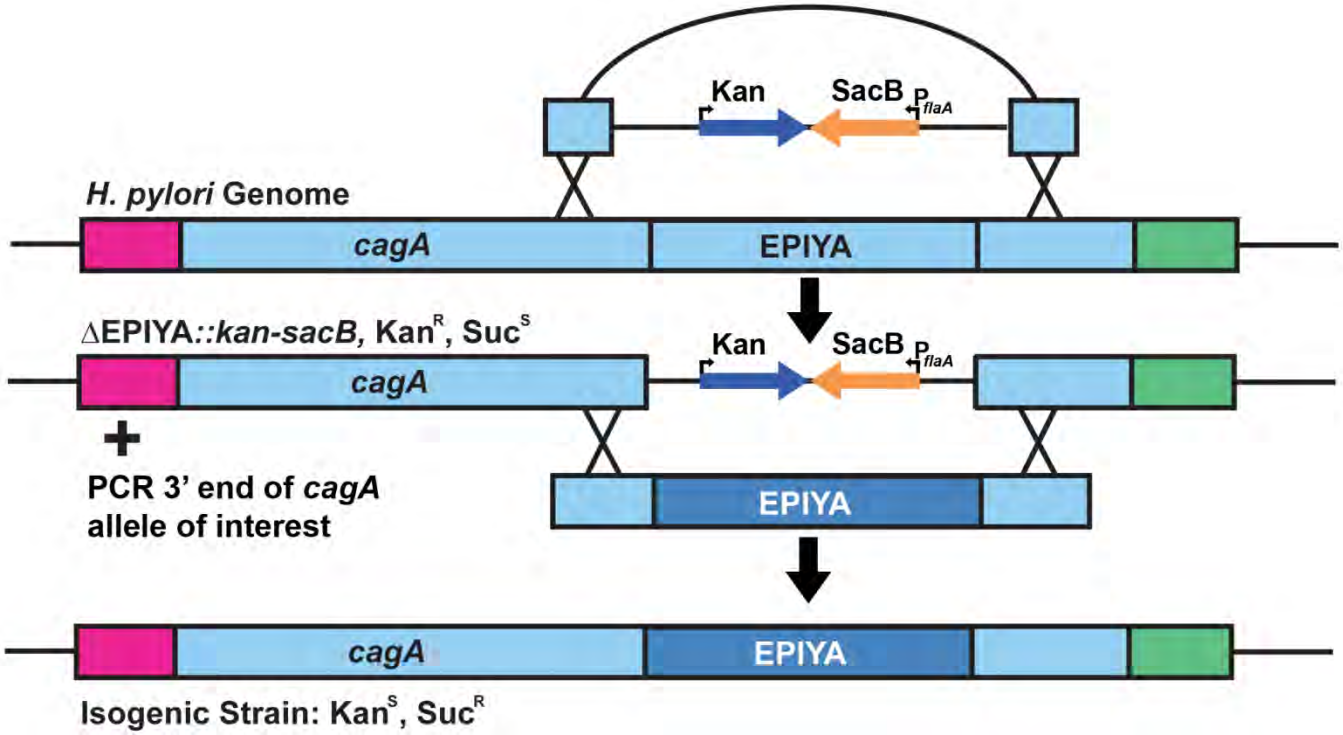
All strains were created in the wild type strain 7.13 background (14, 15). The Δ*cagA*, ΔEPIYA, and all of the isogenic strains were created using splicing by overlap extension (SOE) PCR, using the primers listed in Table 12 and Fig. 14. The basic strategy used is depicted in Fig. 13. Briefly, the EPIYA region of strain 7.13 was replaced with the counter-selectable kan-*sacB* cassette, and then different EPIYA motifs of interest were used to replace the kan-*sacB* cassette; thus, yielding the different isogenic strains.

ΔEPIYA Strain

The first strain that was created was the ΔEPIYA strain, which then served as the parental strain for construction of all of the isogenic strains. To create this strain, a ΔEPIYA construct was created that contained a fusion product consisting of the region

Figure 13: Creation of isogenic strains. A schematic showing the general strategy used to create the isogenic strains is shown. The wild type (7.13) EPIYA region was replaced with the counter selectable kan-*sacB* cassette, yielding DSM577, which is Kan^R and $Sucrose^S$. DSM577 (Δ EPIYA) was then used as the parental strain background for all isogenic strain construction. The different EPIYA regions were engineered to be flanked by the wild type upstream and downstream *cagA* regions, and to replace the kan-*sacB* cassette via double homologous recombination.

Figure 13: Creation of isogenic strains



upstream (5' to the region to be replaced) of the EPIYA motifs and the region downstream (3' of the region to be replaced) of the EPIYA motifs (Fig. 14A). Next, this construct was digested and ligated with the kan-*sacB* cassette to yield a construct that contained the upstream region of 7.13 *cagA* - the kan-*sacB* cassette - the downstream region of 7.13 *cagA*. The resulting construct was then transformed into *H. pylori*, and recombinants were selected as described below.

The first step in construction of the ΔEPIYA construct was accomplished by amplifying the upstream region of 7.13 *cagA* with the primers 7.13 del EPIYA-5-fp and 7.13 del EPIYA M-rp, the later of which was engineered to contain an XhoI restriction site (Fig. 14A). The downstream region was amplified using primer 7.13 del EPIYA M-fp, which was engineered to contain a SmaI restriction site, and primer 7.13 del EPIYA-3-rp (Fig. 14A). In order to amplify a single fused product, these upstream and downstream PCR products were gel purified using the QIAquick gel extraction kit (Qiagen, Germantown, MD), and the purified products were combined in a SOE reaction using the 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp primers (Fig. 14A). The amplified fusion product was cloned into pGEM-T Easy (Promega, Madison, WI), and the resulting strain was named DSM530. Orientation of the construct was confirmed by EcoRI (New England BioLabs, Inc., Ipswich, MA), XhoI (Invitrogen, Carlsbad, CA), and SmaI (New England BioLabs, Inc., Ipswich, MA) restriction digestion and the fusion was also sequenced using the T7 and SP6 primers. Next, a construct containing the kan-*sacB* cassette was created. The kan-*sacB* cassette was purified from pDSM3 by digestion with XhoI and SmaI (10). pDSM530 was then similarly digested with XhoI and SmaI and

Figure 14: Creation of the various EPIYA motifs. A schematic showing the strategy used to create the different EPIYA regions is shown. A.) The basic strategy used to create the homologous regions for the ΔEPIYA region replacement is shown with primers listed. The upstream region of homology, which is 5' of the region to be replaced is indicated by an 5', and the downstream region of homology, which is 3' of the region to be replaced is indicated by a 3'. This construct was digested and ligated to the kan-*sacB* cassette to create the ΔEPIYA strain, which has the EPIYA region replaced by the kan-*sacB* cassette (Fig. 1). For Western strains (EPIYA-AB^t, -AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC), the upstream region (5') of homology was amplified with primers 7.13 del EPIYA-5-fp and K154-7IS-M1rp, and the downstream region (3') of homology was amplified with primers K154-7IS-M2fp and 7.13 del EPIYA-3-rp. For East Asian strain (EPIYA-AB^{t}D), the upstream *cagA* region (5') of homology was amplified with primers 7.13 del EPIYA-5-fp and K3-7IS-M1rp, and the downstream *cagA* region (3') of homology was amplified with primers K3-7IS-M2fp and 7.13 del EPIYA-3-rp. The regions of homology from 7.13 were used in the creation of all the isogenic strains, except for the Δ*cagA* and restorant strains. B.) A schematic of how the EPIYA-AB^t strain was created with the different PCR primers and SOE products is shown. The regions of homology are from 7.13 and are depicted by the light blue boxes (5' and 3'). C.) A schematic of how the EPIYA-AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC strains were created with the different PCR primers and SOE products is shown. The regions of homology are from 7.13 and are depicted by the light blue boxes (5' and 3'). Figure legend is continued on page 307.

Figure 14: Creation of the various EPIYA motifs

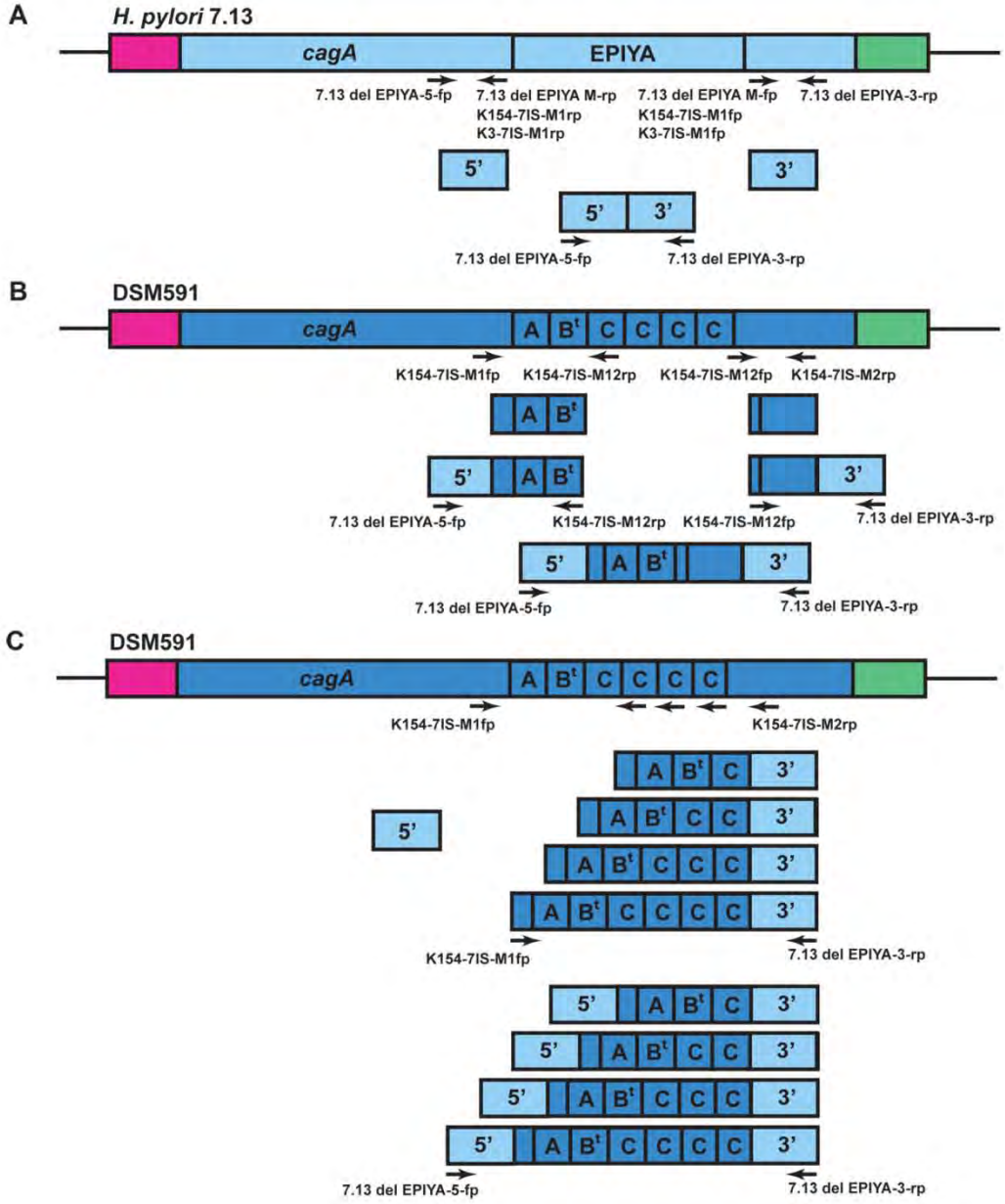
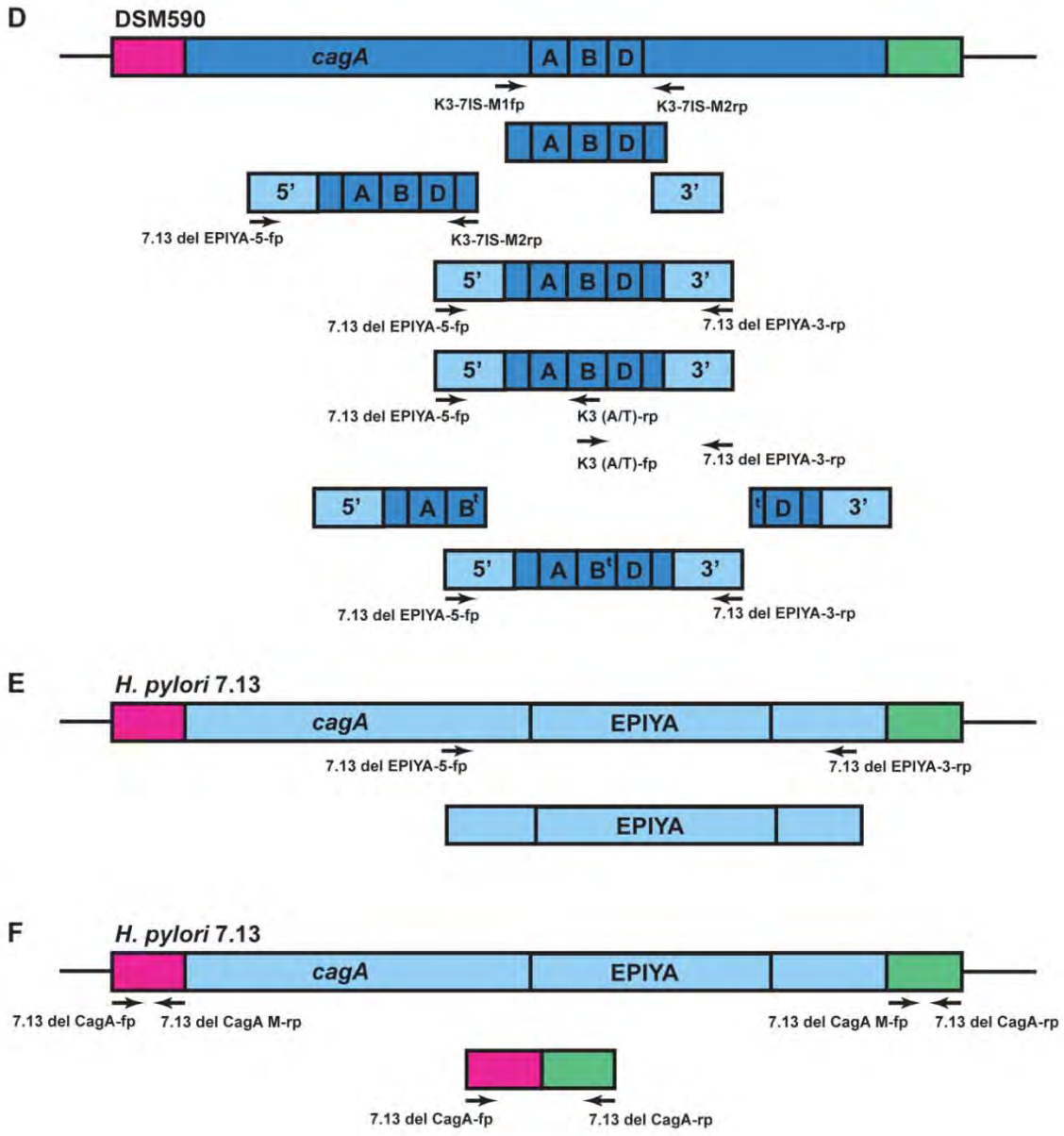


Figure 14b: Creation of the various EPIYA motifs. D.) A schematic of how the EPIYA-AB^{t}D strain was created with the different PCR primers and SOE products is shown. Site directed mutagenesis primers are designated. Again, the regions of homology are from 7.13 and are depicted by the light blue boxes (5' and 3'). E.) A schematic of how the restorant strain was created and the PCR primers used are indicated. F.) The basic strategy used to create the homologous regions for the replacement of the *cagA* gene by the kan-*sacB* cassette is depicted. PCR primers and the SOE product are shown. This construct was digested and ligated to the kan-*sacB* cassette and used to create the Δ*cagA* strain, which has the *cagA* gene replaced by the kan-*sacB* cassette.

Figure 14b: Creation of the various EPIYA motifs



ligated to the purified kan-*sacB* cassette; thus, yielding a plasmid that contained the 7.13 *cagA* upstream region - the kan-*sacB* cassette - the 7.13 *cagA* downstream region construct (Fig. 13). The proper size of this construct was verified by EcoRI and SmaI digestion. The strain bearing this construct was named DSM531. pDSM531 was then transformed into *H. pylori* 7.13, where it integrated into the chromosome via double homologous recombination. Transformants were selected based on kanamycin resistance. Due to the difference in size between the wild type EPIYA motif region and the EPIYA region replaced with the kan-*sacB* cassette, integration was confirmed via PCR using the 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp primers. Proper integration was also confirmed by successful amplification using a primer that lays within the kan-*sacB* cassette (SacBSCN-F2), and a primer in the glutamine racemase gene immediately downstream of *cagA*. The resulting ΔEPIYA strain was named DSM577. After the preliminary *in vitro* and *in vivo* characterization of the isogenic strains, which indicated that there were second site mutations within these strains, a new ΔEPIYA strain (DSM926) was created using identical methods.

Western Strains

The Western strains (EPIYA-AB^t, -AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC) were created by amplification of the EPIYA repeat region from the Korean clinical isolate DSM591. This strain was isolated from a 55 year old, female patient with gastritis and contained an EPIYA-AB^{t}CCCC motif (23). The ^{t} designates a natural change of the alanine in the EPIYA-B repeat to a threonine (EPIYT). The complete EPIYA sequence can be found in Gen Bank under the accession number FJ458142 (23).

EPIYA-AB^t Strain

The EPIYA-AB^t strain was created through a series of PCR reactions. The basic construct was designed to contain the *cagA* upstream region from 7.13 (for homologous recombination) - the EPIYA-AB^t region from DSM591 - the extreme C-terminus of *cagA* from DSM591 - the *cagA* downstream region from 7.13 (for homologous recombination; Fig. 14). The first PCR reaction amplified the 7.13 upstream *cagA* region using *H. pylori* strain 7.13 as a template and primers 7.13 del EPIYA-5-fp and K154-7IS-M1rp (Fig. 14A). The second PCR reaction amplified the EPIYA-AB^t region from DSM591 using the K154-7IS-M1fp and K154-7IS-M12rp primers (Fig. 14B). The third PCR reaction used DSM591 as the template and primers K154-7IS-M12fp and K154-7IS-M2rp to amplify the sequence from the end of the last EPIYA-C motif to the end of the replaced EPIYA region (Fig. 14B). The fourth PCR reaction amplified the downstream *cagA* region from 7.13 using primers K154-7IS-M2fp and 7.13 del EPIYA-3-rp (Fig. 14B). These purified amplicons were then used in three different SOE reactions. The first SOE reaction fused the 7.13 upstream region and the EPIYA-AB^t region from DSM591 (products from the first and second PCR reactions) and used primers 7.13 del EPIYA-5-fp and K154-7IS-M12rp (Fig. 14B). The second SOE reaction fused the extreme C-terminus of CagA to the 7.13 downstream *cagA* region (products from the third and fourth PCR reactions) and utilized primers K154-7IS-M12fp and 7.13 del EPIYA-3-rp (Fig. 14B). The final SOE reaction fused the products from the first two SOE reactions and amplified the full construct: 7.13 upstream *cagA* region - the EPIYA-AB^t region from DSM591 - the extreme *cagA* C-Terminus - the 7.13 downstream *cagA* region. This was accomplished with primers 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp (Fig.

14B). This full length product was next cloned into pGEM-T Easy, and proper orientation was verified by EcoRI and SmaI digestion. The construct was also sequenced with T7 (Promega, Madison, WI), SP6 (Promega, Madison, WI), K154-7IS-M1fp and K154-7IS-M2rp primers, and the resulting strain was named DSM532. DSM577 (7.13 ΔEPIYA) was next transformed with pDSM532, and the resulting double crossover event resulted in replacement of the kan-*sacB* cassette with the EPIYA-AB^{t} construct, making the strain now Suc^{R} and Kan^{S}. Transformants were selected for on 5% sucrose HBA plates. Since sucrose resistance can easily arise via spontaneous mutation, the transformants were further screened for sensitivity to kanamycin, and screened using the K154-7IS-M1fp and K154-7IS-M2rp primers. Next, an internal portion of *cagA* encompassing the EPIYA region was amplified using primers 7.13cagA-2742-fp and 7.13cagA-4561-rp, and sequenced with primers 7.13 del EPIYA-5-fp, 7.13 del EPIYA-3-rp, K154-7IS-M1fp, and K154-7IS-M2rp to ensure accurate sequencing of the replaced EPIYA region. The resulting strain was named DSM601.

EPIYA-AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC Strains

The EPIYA-AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC strains were all created using the same strategy, through a series of five PCR reactions (Fig. 14). The overall construct was the upstream *cagA* region from 7.13 (for homologous recombination) - the different EPIYA region combinations from DSM591 - the downstream *cagA* region from 7.13 (for homologous recombination). The first PCR reaction used *H. pylori* strain 7.13 as the template and primers 7.13 del EPIYA-5-fp and K154-7IS-M1rp (Fig. 14A) to produce the upstream region from 7.13. The second PCR reaction amplified the EPIYA

motifs of interest from DSM591 with primers K154-7IS-M1fp and K154-7IS-M2rp (Fig. 14C). Due to the fact that the EPIYA region of DSM591 contains the EPIYA-AB^{t}CCCC motif and the fact that the EPIYA-C motifs are repeats, amplification of this region produced a ladder of PCR products corresponding to the different number of EPIYA-C repeats, from one to four repeats (Fig. 14C). The third PCR reaction used *H. pylori* strain 7.13 as the template and primers K154-7IS-M2fp and 7.13 del EPIYA-3-rp (Fig. 14A) to produce the downstream *cagA* region from 7.13. These individual products were purified and used in two SOE reactions. The first SOE reaction created a single product that contained a single EPIYA motif region (-AB^{t}C, -AB^{t}CC, -AB^{t}CCC, and -AB^{t}CCCC) from DSM591 and the 7.13 downstream *cagA* region (products from the second and third PCR reactions) with primers K154-7IS-M1fp and 7.13 del EPIYA-3-rp (Fig. 14C). This step was repeated with each of the four different EPIYA motifs indicated above. The second SOE reaction used primers 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp and created the fused 7.13 upstream *cagA* region - EPIYA motif region from DSM591 - 7.13 downstream *cagA* region constructs (Fig. 14C).

Each of these constructs was purified and subcloned into pGEM-T Easy. Proper orientation was verified by EcoRI and SmaI digestion, and the constructs were sequenced with T7, SP6, K154-7IS-M1fp and K154-7IS-M2rp primers to ensure accurate sequencing of the entire EPIYA region. This yielded the following strains: DSM570 (EPIYA-AB^{t}C), DSM571 (EPIYA-AB^{t}CC), DSM572 (EPIYA-AB^{t}CCC), and DSM573 (EPIYA-AB^{t}CCCC).

These plasmids were individually transformed into DSM577 (7.13 ΔEPIYA strain), and the resulting double crossover event resulted in replacement of the kan-*sacB*

cassette with the different Western EPIYA constructs making the strains now Suc^R and Kan^S. Transformants were selected on 5% sucrose HBA plates, but since sucrose resistance can arise via spontaneous mutation, the transformants were further screened for sensitivity to kanamycin, and screened using the K154-7IS-M1fp and K154-7IS-M2rp primers. Next, an internal portion of *cagA* that encompasses the EPIYA region was amplified using primers 7.13cagA-2742-fp and 7.13cagA-4561-rp, and sequenced with primers 7.13 del EPIYA-5-fp, 7.13 del EPIYA-3-rp, K154-7IS-M1fp, and K154-7IS-M2rp to ensure accurate sequencing of the replaced EPIYA region. The resulting *H. pylori* strains were named the following: DSM602 (EPIYA-AB^tC), DSM605 (EPIYA-AB^tCC), DSM606 (EPIYA-AB^tCCC), and DSM609 (EPIYA-AB^tCCCC).

East Asian Strains

The East Asian strain was created by modification of the EPIYA repeat region of the Korean isolate DSM590. This strain was isolated from a 65 year old, male gastric cancer patient and contained an EPIYA-ABD motif (23). The complete EPIYA sequence can be found in Gen Bank under the accession number FJ458118 (23). Of note, the EPIYA-B motif contained the normal alanine (as compared to the created Western strains), so site directed mutagenesis was used to substitute the alanine for a threonine so as to provide consistency between the Western and East Asian strains.

EPIYA-ABD Strain

Construction of an EPIYA-ABD construct was made through a series of five PCR reactions (Fig. 14D). This construct contained the 7.13 upstream *cagA* region - the

EPIYA-ABD region from strain DSM590- the 7.13 downstream *cagA* region. The first PCR reaction amplified the upstream region of 7.13 with primers 7.13delEPIYA-5'-fp and K3-7IS-M1rp (Fig. 14A). The second PCR reaction amplified the EPIYA-ABD region from DSM590 using the K3-7IS-M1fp and K3-7IS- M2rp primers (Fig. 14D). The third PCR reaction used 7.13 as a template and the K3-7IS-M2fp and 7.13delEPIYA-3'-rp primers in order to amplify the downstream *cagA* region of 7.13 (Fig. 14A). The purified amplicons were then used in three different SOE reactions. The first SOE reaction fused the 7.13 upstream *cagA* region and the EPIYA-ABD region from strain DSM590 (the products from the first and second PCR reaction) with primers 7.13delEPIYA-5'-fp and K3-7IS-M2rp (Fig. 14D). These products were then used in the final SOE PCR reaction to create the final construct (7.13 upstream *cagA* region - DSM590 EPIYA-ABD region - 7.13 downstream *cagA* region) using primers 7.13delEPIYA-5'-fp and 7.13delEPIYA-3'-rp (Fig. 14D). This construct was then cloned into pGEM-T Easy, and the insertion was verified by EcoRI and SmaI digestion. Furthermore, it was also sequenced using the T7 and SP6 primers, and the resulting strain was named DSM533.

EPIYA-AB^{t}D Strain

In order to create the EPIYA-AB^{t}D strain, site-directed mutagenesis was accomplished using pDSM533 as the template. The substitution of the threonine for the alanine in the EPIYA-B motif was completed in three PCR reactions (Fig. 14D). Two different PCR reactions were performed using pDSM533 as the template and using primers 7.13 del EPIYA-5-fp and K3(A/T)-rp and K3(A/T)-fp and 7.13 del EPIYA-3-rp,

respectively (Fig. 14D). The K3(A/T)-rp and K3(A/T)-fp primers were designed to overlap and replace the alanine with the wanted threonine. The products of these PCR reactions were fused to create the EPIYA-AB^{t}D region using primers 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp (Fig. 14D). This construct was then cloned into pGEM-T Easy and verified via sequencing with T7 and SP6 primers. The resulting strain was named DSM547. DSM577 (7.13 ΔEPIYA) was next transformed with pDSM547, and the resulting double crossover event resulted in replacement of the kan-*sacB* cassette with the EPIYA-AB^{t}D construct making the strain now Suc^{R} and Kan^{S}. Transformants were selected on 5% sucrose HBA plates. Again, since sucrose resistance can arise via spontaneous mutation, the transformants were further screened for sensitivity to kanamycin, and screened using K3-7IS-M1fp and K3-7IS-M2rp primers. Next, an internal portion of *cagA* encompassing the EPIYA region was amplified using primers 7.13cagA-2742-fp and 7.13cagA-4561-rp, and sequenced using primers 7.13 del EPIYA-5-fp, 7.13 del EPIYA-3-rp, K3-7IS-M1fp, and K3-7IS-M2rp to ensure accurate sequencing of the replaced EPIYA region. The resulting strain was named DSM616.

Restorant Strain

As an important control for possible secondary mutations that arose during genetic manipulation, a restorant strain was also created. This strain contains the exact genomic sequence as 7.13, which is an EPIYA-AB^{t}C strain. The restorant was created by amplification of the entire EPIYA region of 7.13 with primers 7.13 del CagA-fp and 7.13 del CagA-rp (Fig. 14E). This product was purified and cloned into pGEM-T Easy. Proper orientation was verified by EcoRI and SmaI digestion, and the construct was

sequenced with the T7 and SP6 primers. The resulting strain was named DSM641. DSM577 (7.13 ΔEPIYA) was next transformed with pDSM641, and the resulting double crossover event resulted in replacement of the kan-*sacB* cassette with the wild type EPIYA construct making the strain now Suc^R and Kan^S. Transformants were selected on 5% sucrose HBA plates. The transformants were further screened for sensitivity to kanamycin, and screened using 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp primers. Again, due to a size difference between the wild type EPIYA region and the EPIYA region containing the kan-*sacB* cassette, amplification using these primers could assess integration of the new EPIYA motif and loss of the kan-*sacB* cassette. Next, an internal portion of *cagA* encompassing the EPIYA region was amplified using primers 7.13cagA-2742-fp and 7.13cagA-4561-rp, and sequenced using primers 7.13 del EPIYA-5-fp and 7.13 del EPIYA-3-rp to ensure accurate sequencing of the replaced EPIYA region. The resulting strain was named DSM613.

After the preliminary *in vitro* and *in vivo* characterization of the isogenic strains, which indicated that there were second site mutations within these strains, a new restorant strain was created using the same process as described above. However, this restorant strain was created by transforming DSM926 (7.13 ΔEPIYA) with pDSM641. The process to select and screen transformants was identical to the process used to create DSM613. The resulting strain was named DSM927.

ΔcagA *Strain*

Another important control to prove the role of *cagA* in any observed effects, was a Δ*cagA* strain. Again, like the ΔEPIYA strain, the creation of this strain was a multistep

process: (1) a construct consisting of the fused upstream (5') region and downstream (3') region of *cagA* was created, (2) a construct containing the region upstream of *cagA* - the kan-*sacB* cassette - the region downstream of *cagA* was created, and (3) this construct was integrated into *H. pylori* strain 7.13 via double homologous recombination. The region upstream of *cagA* was amplified with primers 7.13del CagA-fp and 7.13del CagA M-rp, the later of which was engineered with a XhoI restriction site. The downstream region of *cagA* was amplified using the 7.13 del CagA M-fp primer, which was engineered with a SmaI restriction site, and the 7.13 del CagA-rp primer (Fig. 14F). Purified PCR products from these reactions were then combined in a SOE reaction and amplified using the 7.13 del CagA-fp and 7.13 del CagA-rp primers to yield a fused product of the regions upstream and downstream of *cagA* (Fig. 14F). The resulting PCR product was cloned into pGEM-T Easy and size was verified by EcoRI, Xho, and SmaI digestion. This construct was also sequenced with the T7 and SP6 primers, and the resulting strain was named DSM598. Next, pDSM598 was digested with XhoI and SmaI and ligated to the purified kan-*sacB* cassette obtained by similar digestion of pDSM3. The resulting construct contained the region upstream of *cagA*- the kan-*sacB* cassette-and the region downstream of *cagA*. This construct was cloned into pGEM-T Easy, and proper orientation was confirmed by EcoRI and SmaI digestion. The resulting strain was named DSM599. pDSM599 was then transformed into *H. pylori* 7.13, where it integrated into the chromosome via double homologous recombination. Transformants were selected based on kanamycin resistance. Due to the difference in size between the wild type *cagA* and the *cagA* replaced with the kan-*sacB* cassette, integration was confirmed via PCR using the 7.13 del CagA-fp and 7.13 del CagA-rp primers. Proper

integration was also confirmed by successful amplification using a primer that lays within the kan-*sacB* cassette (SacBSCN-F2), and a primer in the glutamine racemase gene immediately downstream of *cagA*. The resulting ΔEPIYA strain was named DSM600.

The Δ*cagA*, ΔEPIYA, EPIYA-AB^t, -AB^{t}C, -AB^{t}CC, -AB^{t}CCC, -AB^{t}CCCC, -AB^{t}D, and restorant strains were successfully created and verified by sequencing.

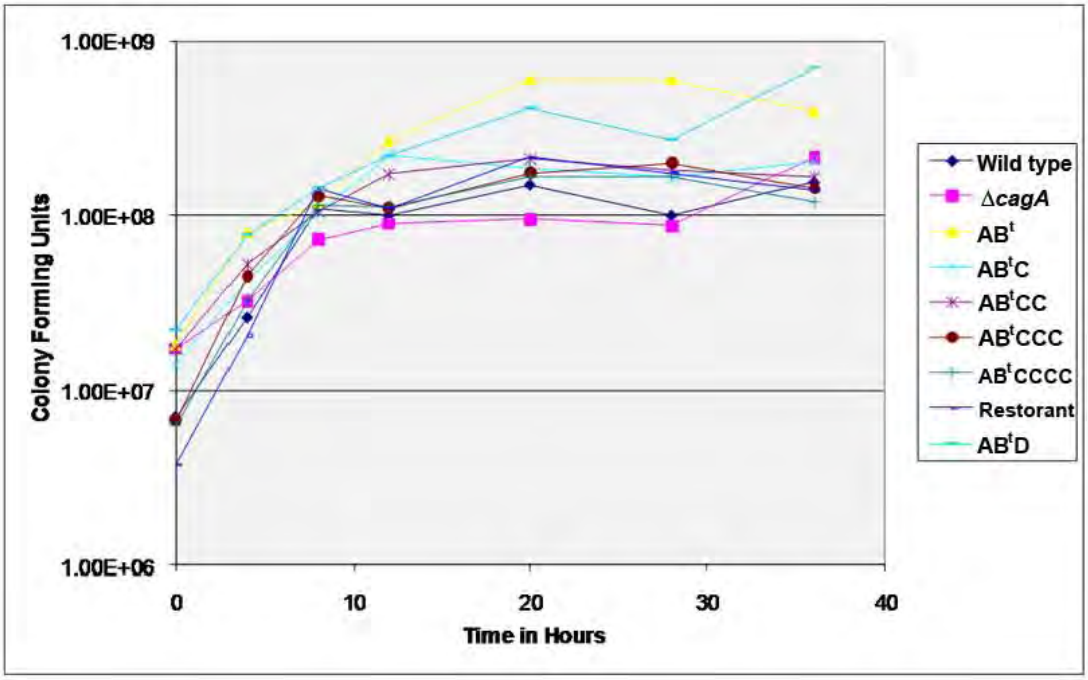
Growth Dynamics

To confirm that each of the isogenic strains behaved phenotypically like the wild type 7.13 strain, growth kinetics were monitored for each strain. Bacterial strains were grown and expanded on HBA plates for approximately 20 hours and used to inoculate 25 mL liquid cultures, which were subsequently grown for approximately 18 hours. These starter cultures were then used to inoculate a 75 mL liquid culture of each strain at a starting optical density at 600 nm of 0.05. The cultures were grown microaeorobically with shaking at 100 rpm at 37ºC. For each strain, a 3 mL aliquot was taken at time=0, 4, 8, 12, 20, 28, and 36 hours to assess optical density as well as colony forming units (CFU) per ml of culture, which was determined by serial dilution and plating on HBA plates. Two biologically independent replicates of this experiment were performed.

As shown in Fig. 15, each of the strains showed a pattern of growth that virtually mirrored that of the wild type *H. pylori* 7.13 strain. These data suggested that genetic manipulation of the strains did not result in any overt secondary mutations that slowed the growth rate of any of the isogenic strains.

Figure 15: Growth dynamics of the isogenic strains. Samples were taken from OD controlled liquid cultures across various time points, and serial dilutions were plated to obtain single colonies. This figure is representative of two biological repeats.

Figure 15: Growth dynamics of the isogenic strains



Expression of CagA

After growth dynamics of each strain were assessed, CagA expression was verified for each of the isogenic strains. Expression was assessed through Western blot analysis using bacteria grown on plates and in liquid culture. First, 18-20 hour lawns of each strain were harvested from HBA plates, pelleted, resuspended in 1 X phosphate-buffered saline (PBS; EMD Chemicals, Inc., Gibbstown, NJ), and then mixed with 5 X Laemmli sample buffer for qualitative Western blot analysis. For quantitative CagA protein analysis, an additional aliquot of the 18-20 hour lawn grown bacterial lysates was lysed with 300 µL of lysis buffer [150 mM NaCl, 50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 1% sodium dodecyl sulfate (SDS), and 10% glycerol, containing one Complete Mini Protease Inhibitor Cocktail Tablet (Roche Diagnostic, Indianapolis, IN) and 100 µL 10 mM sodium orthovanadate (added at the time of use) per 10 mL of lysis buffer]. Lysates were sonicated and then centrifuged to remove unlysed cellular debris. The amount of protein in each lysate was subsequently quantified using the BCA Protein Assay Kit (Thermo Scientific/Pierce, Rockford, IL).

Two different experiments were performed to assess CagA expression in liquid culture across a time course. The first experiment used aliquots taken from one large liquid culture at various time points, while the second experiment involved inoculation of one large liquid culture, which was then divided into smaller cultures. These smaller cultures were analyzed due to growth differences observed with various sizes of liquid culture. For both of these experiments, bacteria were grown and expanded on HBA pates for 20-24 hours. The bacteria were then used to inoculate starter liquid cultures, which were grown for approximately 18 hours. These starter cultures were used to inoculate an

OD controlled liquid culture of each strain at a starting optical density at 600 nm of 0.05. In the first experiment, aliquots from a 120 mL OD controlled liquid culture were taken at time=0, 9, 18, 27, and 40 hours for Western blot analysis. These bacterial lysates were lysed with 300 µL of the lysis buffer mentioned above, sonicated and then centrifuged to remove unlysed cellular debris. In experiment two, starter cultures were used to inoculate a 40 mL liquid culture of each strain at a starting optical density at 600 nm of 0.05. This OD controlled liquid culture was then divided into five 7 mL cultures. One culture from each strain was collected at time=0, 9, 18, 27, and 40 hours for Western blot analysis. These bacterial lysates were lysed with 300 µL of the lysis buffer mentioned above, sonicated and then centrifuged to remove unlysed cellular debris. Subsequent protein quantification was performed using the BCA Protein Assay Kit.

For quantitative Western blot analysis, equal amounts of protein plus 6 µL of 5 X Laemmli sample buffer were added to each well. Bacterial lysates were separated by sodium dodecyl sulfate- polyacrylamide gel electrophoresis using a 6% separating gel and a 4% stacking gel. Proteins were then transferred to nitrocellulose membranes (Thermo Scientific, Rockford, IL) using a semidry transfer apparatus (Owl; Thermo Scientific, Rochester, NY) at 300 mAmp for 45 minutes.

Membranes were probed with a 1:5,000 dilution of rabbit IgG anti-CagA polyclonal antibody b-300 (Santa Cruz Biotechnology, Santa Cruz, CA), followed by a 1:20,000 dilution of HRP-conjugated bovine anti-rabbit IgG secondary antibody (Santa Cruz Biotechnology, Santa Cruz, CA). Proteins were detected using the SuperSignal West Pico chemiluminescent substrate kit (Thermo Scientific/Pierce, Rockford, IL) and a LAS-3000 Intelligent Dark Box with LAS-3000 Lite capture software (Fujifilm,

Stamford, CT). Densitometry was performed using MultiGauge software (Fujifilm, Stamford, CT).

For this series of experiments, the qualitative Western blots of plate grown bacteria demonstrated that each of the isogenic strains (except the Δ*cagA* and ΔEPIYA strains) expressed CagA. Moreover, the size of the resulting protein increased with an increasing number of EPIYA motifs (Fig. 16A). Four biological repeats of quantitative Western blots on the plate grown bacteria suggested that each of the isogenic strains expressed less CagA than the wild type strain. The biggest difference was seen with the EPIYA-AB^{t}CCCC strain, which expressed exactly half of the amount of CagA as the wild type strain (Fig. 16B). Of note, the restorant strain, which should be genetically identical to the wild type strain, also expressed slightly lower levels of CagA.

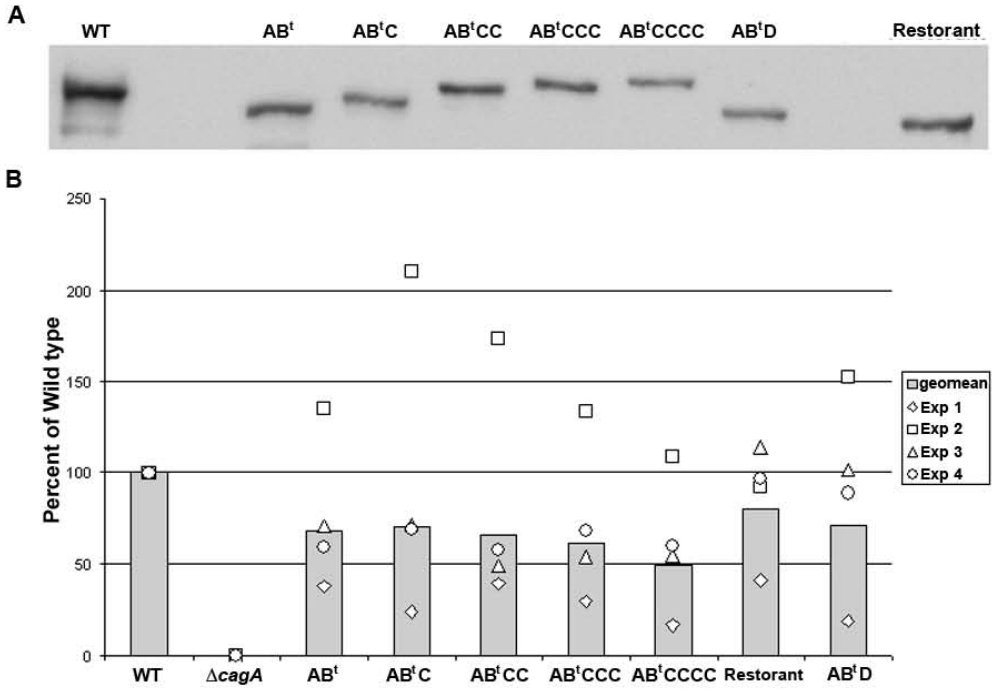
Analysis of the bacterial lysates harvested from liquid grown cultures was more problematic. Samples from the larger liquid cultures showed variable amounts of full length CagA as well as the appearance of numberous degradation products. Indeed, technical repeats showed that the majority of CagA was found in the degraded CagA bands. More full-length CagA was apparent in the smaller individually grown liquid cultures, however there was still more degraded CagA in these lysates than in the lawn grown bacteria. The reason for this difference remains unclear.

Localization of Bacteria

A previous study suggested that CagA affects bacteria localization on the surface of host cells (2). Therefore, we next wanted to examine if the EPIYA region impacted where the bacteria would localize on the host cells. Cover slips were placed in six well

Figure 16: CagA Expression. A. The qualitative Western blot on plate grown bacteria shows expression of CagA for each isogenic strain. The blot illustrates that the strains expected to express CagA do so at the expected size, which varies based on the number of EPIYA repeats. B. This histogram shows quantitative CagA expression. Total CagA expression was determined by densitometric analysis of quantitative Western blots of four biological repeats of lawn grown bacteria. Numbers are presented as percent of wild type and the column represents the geometric mean.

Figure 16: CagA Expression



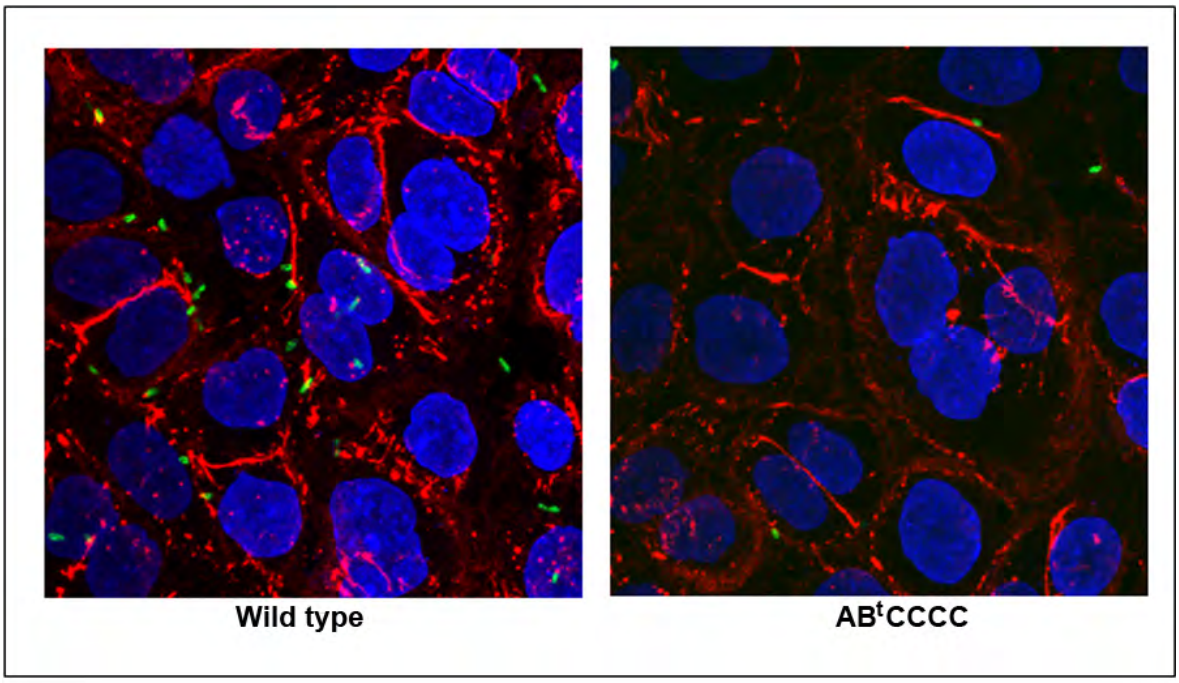
plates and left under a UV light for 20 minutes. One mL of collagen was then placed in each well and the cover slips were allowed to incubate at room temperature for one hour. Collagen was then aspirated and re-frozen for future use. Wells were washed three times with three mL of 1 X PBS and then allowed to air dry. The plates were then placed under the UV light for an additional 20 minutes and stored at 4ºC for less than one week. The resulting collagen coated cover slips were seeded with 4 x 10^5 AGS cells and allowed to grow to confluency for approximately four days. One hour before the infection, cells were washed with 1 X PBS and 1 mL of fresh warm cell culture medium was added: Dulbecco's modified Eagle's medium without L-glutamine (Quality Biological, Inc., Gaithersburg, MD), supplemented with 10% fetal bovine serum, 10 µg/ml vancomycin, and 2 nM L-glutamine (Quality Biological, Inc.). Of note, warm medium is needed to prevent cell stress and subsequent detachment. Twelve hour OD controlled liquid cultures were then used to infect cells at a multiplicity of infection (MOI) of 100. Infections were allowed to proceed for 5 minutes and then the cells were washed three times with warm cell culture medium and allowed to incubate for 10 minutes before being washed with 1 X PBS and fixed with 2% paraformaldehyde in 100 mM phosphate buffer (pH 7.4). The cells on the cover slips were then permeabilized and blocked in a 1 X PBS solution containing 3% BSA, 1% saponin (ACROS Organics/Thermo Scientific, New Jersey), and 0.05% sodium azide (EM Science, Gibbstown, NJ). The cover slips were subsequently simultaneously incubated with the primary antibodies, mouse monoclonal IgG1 anti-ZO-1 (Invitrogen) and rabbit polyclonal IgG anti-*H. pylori* (Thermo Scientific), which were diluted to a final concentration of 1 µg/mL and 1.25

μg/mL, respectively, in a 1 X PBS solution containing 3% BSA and 0.05% sodium azide. The cover slips were next washed three times in a 1 X PBS solution containing 3% BSA and 0.05% sodium azide for 10 minutes each time. The secondary antibodies, Alexa Fluor 555-conjugated goat anti-mouse IgG1 (Invitrogen), Alexa Fluor 488-conjugated goat anti-rabbit IgG (Invitrogen), and DAPI were diluted to final concentrations of 2 μg/mL, 2.5 μg/mL, and 10 μg/mL, respectively, in a 1 X PBS solution containing 3% BSA and 0.05% sodium azide. The cover slips were incubated with the secondary antibodies and DAPI for 45 minutes, and were then washed three times in a 1 X PBS solution containing 3% BSA and 0.05% sodium azide for 10 minutes each time. The cover slips were next rinsed twice with 1 X PBS and once with sterile double distilled water, removed, and placed cell side down on a drop of VectaShield (Vector Laboratories, Inc., Burlingame, CA) on a pre-cleaned slide and were sealed on alternating corners with nail polish. The nucleus (blue - DAPI), cellular junctions (red – ZO-1), and *H. pylori* (green) were visualized by collecting z-stacks with a Zeiss LSM 710 confocal microscope and projecting the stacks with ZEN2009 software.

Representative images from the wild type and the EPIYA-AB^{t}CCCCstrains are shown in Fig. 17, and represent infected cells stained for the nucleus (blue - Dapi), cellular junctions (red – ZO1), and *H. pylori* (green). No major differences in localization were seen among the various strains, including the Δ*cagA* strain, which has been shown previously to localize throughout the cell surface (2). In fact, all strains seemed to localize to the cellular tight junctions, regardless of *cagA* status.

Figure 17: Localization of bacteria on host cells. Confocal images of infected AGS cells stained for the nucleus (blue - Dapi), cellular junctions (red – ZO1), and *H. pylori* (green) are shown. The image on the left shows AGS cells infected with wild type bacteria, while the image on the right reveals AGS cells infected with the EPIYA-AB^{t}CCCC strain.

Figure 17: Localization of bacteria on host cells



Adherence and Internalization Assays

While *cagA* status has been suggested to affect localization, no difference in total adherence has been observed between wild type and Δ*cagA* strains of *H. pylori in vitro* (2). However, if there were significant differences in bacterial adherence or internalization, this could affect how much CagA could be translocated into host cells. Thus, the adherence and internalization assays were adapted from a previous study (12). Both the adherence and internalization assays were conducted on the same day with the same cultures in order to determine percent internalization. In short, adherence assays were completed by seeding 24 well plates with 2.2 x 10^5 AGS cells for 21 hours. One hour before the infection, cells were washed with 1 X PBS and 1 mL of fresh warm cell culture medium as described in the localization section was put on the cells. Again, warm medium was needed to prevent cell stress and subsequent detachment. Next, 12 hour OD controlled liquid cultures were used to infect three wells per strain at an MOI of 10. Infections were allowed to progress for 30 minutes and wells were then washed three times with warm cell culture medium. Cells were next lysed with 1 mL of 1% saponin in 1 X PBS. These lysates were then serial diluted and plated in triplicate per well for single colonies. Numbers between replicate plates were averaged, and then the number of colonies between the three replicate wells were averaged to yield a total number of adherent bacteria. This number was then divided by the inoculum to determine the percent of adherent bacteria. Four biological repeats were performed.

For internalization assays, 24 well plates were seeded with 2.2 x 10^5 AGS cells for 21 hours. One hour before the infection, cells were washed with 1 X PBS and 1 mL of fresh warm cell culture medium was put on the cells. The same 12 hour OD controlled

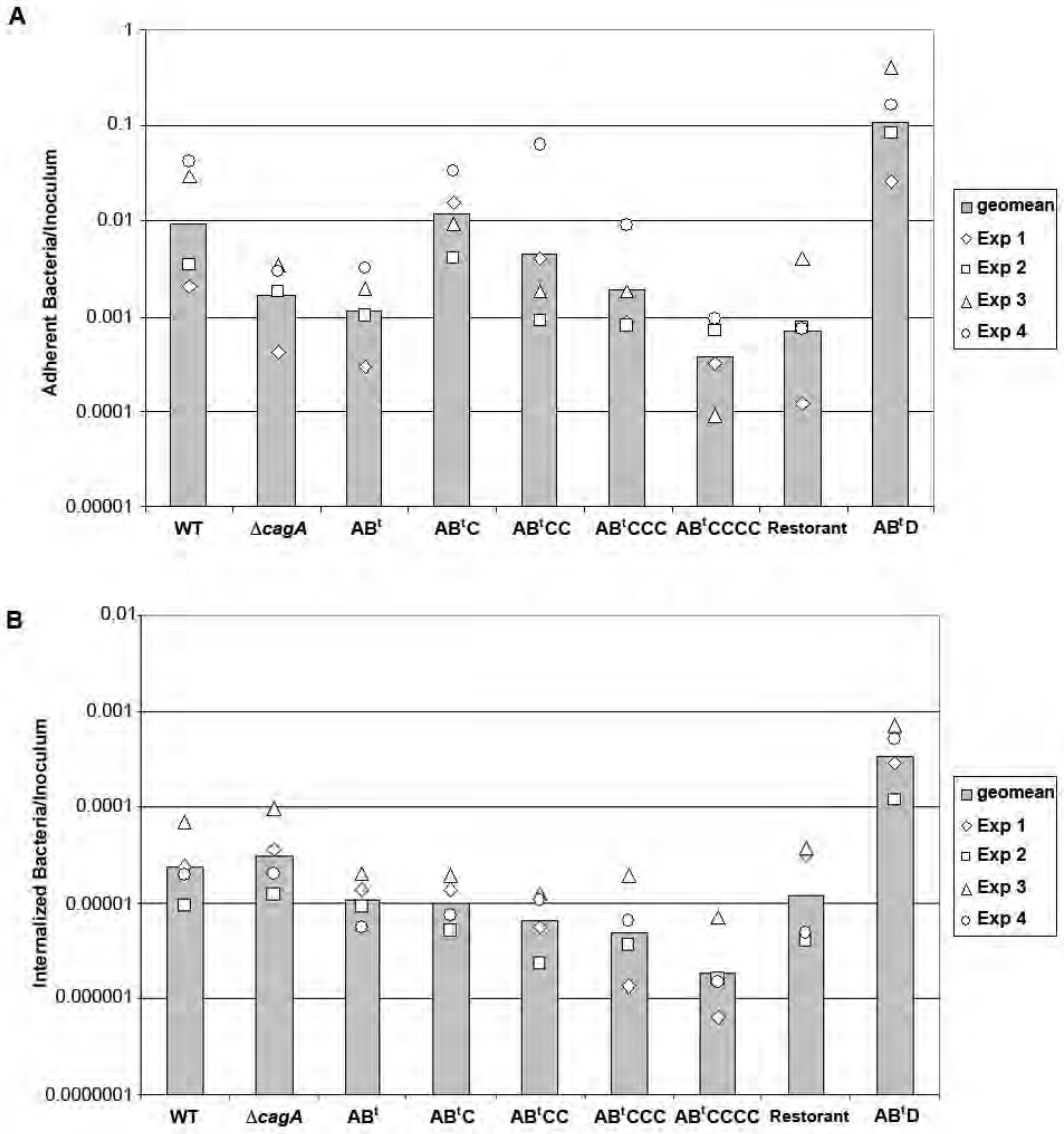
liquid cultures used in the adherence assay were used to infect three wells per strain at an MOI of 10. Infections were allowed to progress for 30 minutes and wells were then washed three times with warm cell culture medium. At that point, 1 mL of fresh warm cell culture medium supplemented with 200 µg/mL gentamicin (Gibco, Carlsbad, CA) was added to each well in order to kill all external bacteria. Infected cells were incubated for 2 hours and cells were washed again five times with warm cell culture medium. Cells were then lysed with 500 µL of 1% saponin. 100 µL of these lysates were then plated on five individual plates to obtain actual number of internalized bacteria per well. Numbers between the replicate wells per strain were then averaged and divided by the inoculum to obtain the percent of internalized bacteria. Four biological repeats were performed.

We were unable to demonstrate the high levels of adherence previously observed in other studies; however, this was likely because our bacterial cultures were not co-cultured with AGS cells (12). Adherence data revealed approximately a log difference in attachment between the wild type and the Δ*cagA* strains (Fig. 18A). The most adherent strain was the EPIYA-AB^{t}D strain, which showed a log increase in the amount of adherent bacteria as compared to the wild type strain. While the wild type and the EPIYA-AB^{t}C strains showed similar levels of adherent bacteria, there was an inverse trend between the number of EPIYA-C motifs and number of adherent bacteria. The least adherent strain was the EPIYA-AB^{t}CCCC strain (Figure 17 and 18A). Of note, the restorant strain, which should behave similar to the wild type strain, showed more than a log decrease in adherence as compared to the wild type and the EPIYA-AB^{t}C strains.

A fraction of the adherent bacteria actually are internalized into host cells (13). In our hands the percentage of adherent bacteria internalized ranged from 0.1% to 1.4%.

Figure 18: Adherence and Internalization. A. A histogram of the number of adherent bacteria per the innoculum for each of the isogenic strains is shown. Four biological repeats were performed, and the geometric mean is designated by the column. B. A histogram of the number of internalized bacteria per the innoculum for each isogenic strain is shown. Four biological repeats were performed, and the geometric mean is designated by the column.

Figure 18: Adherence and Internalization



Again, we observed little difference in the amount of wild type and Δ*cagA* bacteria able to be internalized into host cells (Fig. 18B). There was a minimal difference in internalization of the wild type, EPIYA-AB^{t}C and the restorant strains. Once again we observed an inverse trend with increasing number of EPIYA-C motifs and the number of internalized bacteria. The strain with the lowest percent of internalized bacteria was the EPIYA-AB^{t}CCCC strain. Conversely, the EPIYA-AB^{t}D strain showed the most internalized bacteria. Surprisingly, the strain with the highest percentage of internalization was the restorant strain, while the strain with the lowest percentage of internalization was the EPIYA-AB^{t}C strain. Taken *en masse* the fact that the restorant, EPIYA-AB^{t}C, and wild type strains did not behave the same in the adherence and internalization assays suggested that there may be unexpected differences in the strains.

CagA Phosphorylation Assays

In order for CagA to be biologically active, it must be expressed and translocated into host cells where it is phosphorylated by host cell kinases (27). Therefore, we next assessed the ability of the isogenic strains to translocate CagA into the host cells as evidenced by the appearance of the phosphorylated protein. The CagA phosphorylation assays were conducted essentially as previously described (7, 23). Briefly, six-well tissue culture plates were seeded with 3.5 x 10^5 AGS cells, and allowed to grow in normal cell culture medium for 3 days. Two hours prior to infection, AGS cells were washed with 1 X PBS, and 3 mL of fresh cell culture medium was added to each well. 18 hour OD controlled liquid cultures of each *H. pylori* strain were resuspended in 1 mL of 1 X PBS, and were used to infect AGS cells at a MOI of 100.

Infected cell lysates were collected from replicate wells every hour for the first 8 hours and again 16 and 24 hours post infection. Infected cells were lysed with 250 µL of the lysis buffer described in the CagA expression section. Lysates were sonicated and then centrifuged to remove unlysed cellular debris. Subsequent protein quantification was performed using the BCA Protein Assay Kit. Equal amounts of protein and 6 µL of 5 X Laemmli sample buffer were then added to each well. Infected cell lysates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, using a bilayer separating gel (6% and then 12%) and a 4% stacking gel, and proteins were transferred to nitrocellulose membranes using a semidry transfer apparatus at 300 mAmp for 90 minutes.

Membranes were probed with a 1:5,000 dilution of an anti-phospho-tyrosine monoclonal antibody, pY100 (Cell Signaling Technology, Danvers, MA), followed by a 1:20,000 dilution of HRP-conjugated goat anti-mouse IgG (Santa Cruz Biotechnology, Santa Cruz, CA) secondary antibody and detection was performed as described above. Membranes were subsequently stripped with periodic agitation at 55°C in a pre-warmed 10-mM dithiothreitol solution for 45 minutes. Resulting blots were then well rinsed with running deionized water to remove residual DTT. Membranes were probed with a 1:5,000 dilution of rabbit IgG anti-CagA polyclonal antibody, b-300 followed by a 1:20,000 dilution of HRP-conjugated bovine anti-rabbit IgG secondary antibody. Membranes were then subsequently re-stripped and reprobed with 1:1,000 dilution of goat anti-GAPDH IgG (Santa Cruz Biotechnology, Santa Cruz, CA) plus 1% BSA (EMD Chemicals, Inc., Gibbstown, NJ), followed by a 1:20,000 dilution of HRP-conjugated donkey anti-goat IgG (Santa Cruz Biotechnology, Santa Cruz, CA) secondary antibody

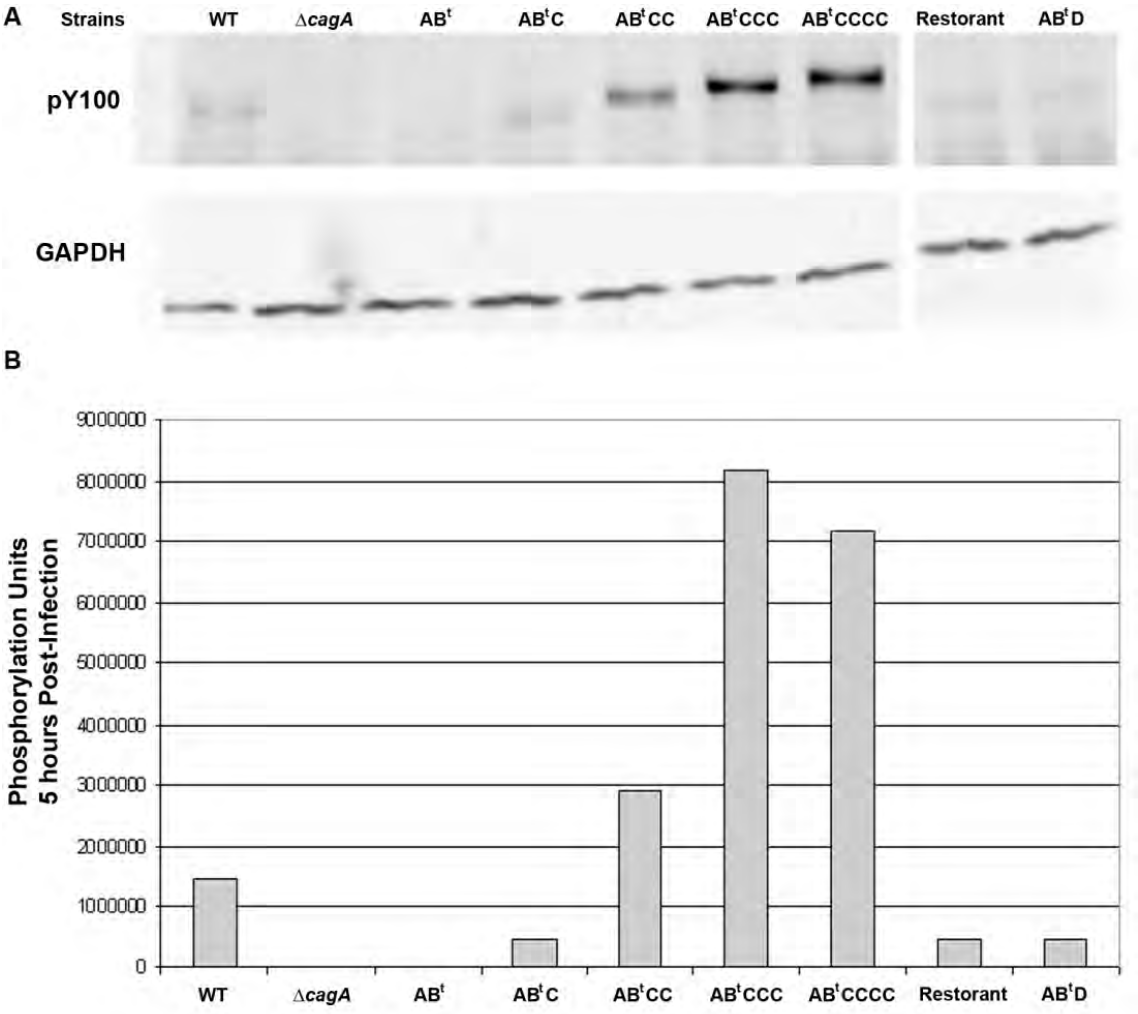
plus 1% BSA. Detection was conducted at each step as described above. Densitometry was performed using MultiGauge software, and the level of CagA phosphorylation was normalized to the level of GAPDH.

Since the primary phosphorylation sites are the tyrosine residues found within the EPIYA-C or -D motif, it was not surprising that the EPIYA-AB^t strain did not show any considerable accumulation of phosphorylated CagA at any time point. All other strains that expressed CagA showed a detectable amount of phosphorylated CagA by 2 hours post infection. The phosphorylation increased from 2 to 4-6 hours until peak phosphorylation was observed between 4 to 8 hours. Phosphorylation then decreased between 8 and 16 hours, and by 24 hours the amount of phosphorylated CagA was negligible.

When a single time point of 5 hours was assessed across the different strains we observed some expected trends (Fig. 19). There was no detectable phosphorylated CagA in the Δ*cagA*, or the EPIYA-AB^t strains. The three isogenic strains that contained only one EPIYA motif (EPIYA-AB^{t}C, EPIYA-AB^{t}D, and the restorant) showed similar levels of phosporylation. However, this amount of phosphorylated protein was almost 2.5 times less then the level of phosphorylated CagA seen in the wild type strain. Increasing amounts of phosphorylated CagA were detected with increasing numbers of EPIYA-C motifs (from one to three), and the greatest amount of phosphorylated CagA was seen with the EPIYA-AB^{t}CCC and EPIYA-AB^{t}CCCC strains. While the trends for the isogenic strains were what we expected based on the number of EPIYA motifs, the fact that the restorant, EPIYA-AB^{t}C, and EPIYA-AB^{t}D strains showed reduced levels of

Figure 19: Phosphorylation of CagA. A. A Western blot of *H. pylori* infected cell lysates is shown. Lysates were collected five hours post infection and subjected to Western blot analysis. The membrane was first probed with an anti-phosphotyrosine antibody (top panel) and then stripped and reprobed using an anti-GAPDH antibody (bottom panel) in order to normalize for equal loading. B. The histogram represents densitometric analysis of the Western blot shown above where the data were normalized for equal loading using the GAPDH blot.

Figure 19: Phosphorylation of CagA



phosphorylated CagA as compared to the wild type strain indicated once again that the strains may not be isogenic or may have secondary mutations.

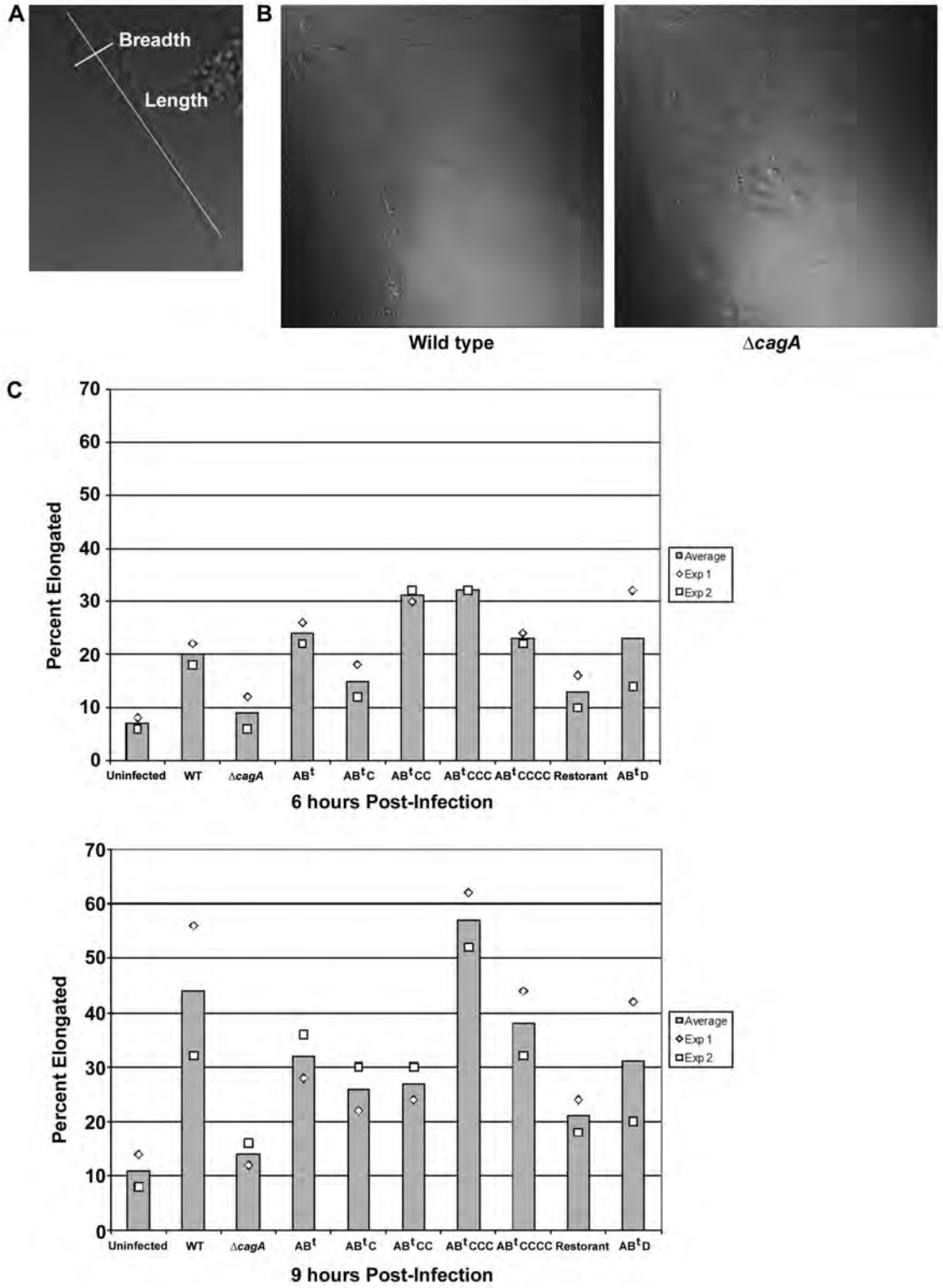
Morphological Changes

One of the easiest ways to grossly assess CagA-modulated host cell signalling pathways is to look at morphological changes induced by *H. pylori* in infected AGS cells. In these cells, CagA delivery results in the stereotypical "hummingbird" phenotype, which is characterized by long fingerlike protrusions from the cell (27). We therefore wanted to assess EPIYA-dependent differences in modulation of host signaling pathways. Morphology assays were adapted from a method described previously (7). Briefly, cover slips were placed in six well plates and collagen coated for one hour as described above. AGS cells were seeded at a density of 2.5 x 10^5 cells, and were allowed to grow in normal cell culture medium for approximately 24 hours. 18 hour OD controlled liquid cultures were used to infect the AGS cells at a MOI of 100. Infections were allowed to proceed for 3, 6, or 9 hours, and then samples were fixed with 2% paraformaldehyde in 100nM phosphate buffer (pH 7.4). Wells were given random numbers to blind the overall results. Images were captured on a Zeiss LSM Pascal confocal microscope, and 50 cells from each well were counted and the measurements for length and breadth of the cells were obtained (7). Two biological repeats were performed.

At six hours post infection, which was the time frame in which we saw maximal CagA phosphorylation, the number of elongated cells in the uninfected group was similar to cells infected with the Δ*cagA* strain. Surprisingly, the restorant and EPIYA-AB^{t}C strains showed only a slight increase in the number of elongated cells as compared to the

Figure 20: Morphology of infected host cells. A. An image of an elongated AGS cell with annotations for the breadth and length of the cells is shown. These depictions of breadth and length are representative of the areas used for measurements during the morphology experiments. The measurement methodology is adapted from Bouracz, *et al.* (7). B. The left panel is a confocal image of AGS cells infected with the wild type strain for nine hours, and the right panel is a confocal image of AGS cells infected with the Δ*cagA* strain for nine hours. C. The top histogram depicts the percent of elongated cells observed six hours post infection with the different isogenic strains. The bottom histogram depicts the percent of elongated cells recorded nine hours post infection with the different isogenic strains.

Figure 20: Morphology of infected host cells.



Δ*cagA* strain. Conversely, the wild type, EPIYA-AB^t, -AB^{t}CCCC, and -AB^{t}D strains all induced similar levels of elongated cells, and the EPIYA-AB^{t}CC and -AB^{t}CCC strains induced an increased number of elongated cells.

Some strains showed maximum phosphorylation at eight hours post infection. Thus, in order to assess for changes in those strains, we also looked at the morphological changes at nine hours post infection. Some trends continued at 9 hours post infection; cells treated with brucella broth (uninfected cells) or the Δ*cagA* strain showed a mimimal amount of cell elongation. While the number of elongated cells induced by the restorant increased, this strain still induced the least amount of elongated cells of all the strains that could deliver CagA. At this time point, the EPIYA-AB^t, -AB^{t}C, -AB^{t}CC, and -AB^{t}D strains all induced similar levels of elongated cells, and the wildtype and EPIYA-AB^{t}CCCC strains induced similar and increased levels of elongated cells. The strain which induced the largest percentage of elongated cells (~57%) was the EPIYA- AB^{t}CCC strain (Fig. 20). Once again, the fact that the wild type, EPIYA-AB^{t}C, and restorant strains all behaved so differently suggested that there is a second site mutation within at least some of the strains.

In vivo *analysis of isogenic strains*

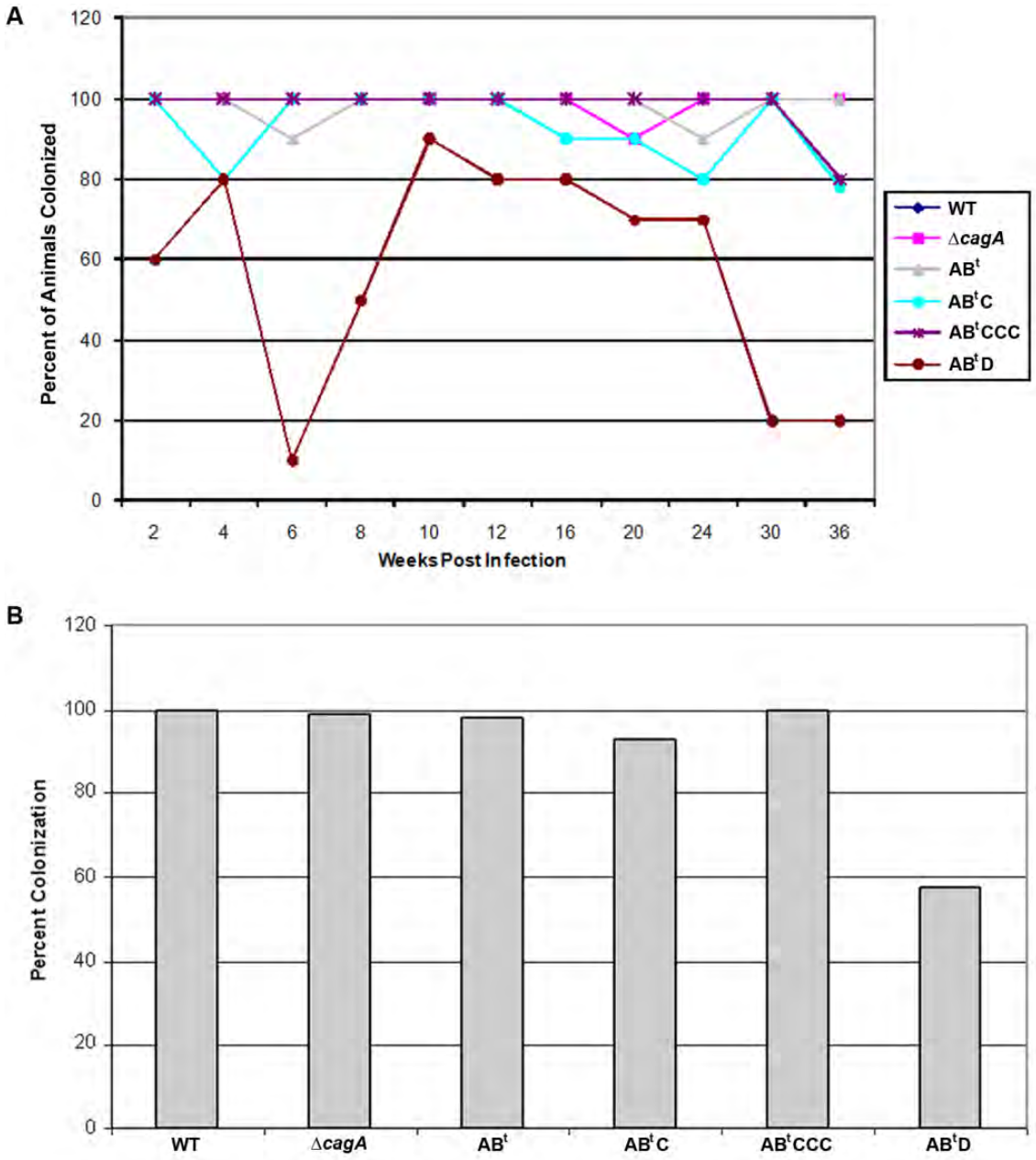
A large scale animal study was conducted to assess differences between the isogenic strains. Infection protocols were similar to those previously published (16, 25). Due to the large number of animals (N=715) to be infected, Mongolian gerbils were obtained from three different age groups: 28-35 days, 35-42 days, and 42-49 days. The different age groups were evenly divided across the different infection groups and the

animals were age matched at each time point. Male Mongolian gerbils (Charles River Laboratories International, Inc. Wilmington, MA) were fasted 12 hours prior to infection, and infected orogastrically with approximately 10^8 *H. pylori* cells for each of the different isogenic strains. Once again, due to the large number of animals (N=715) to be infected, the infection groups were divided into two groups. Group 1 consisted of animals that were mock infected (n=55), or were infected with the EPIYA-AB^{t}C strain (n=110), EPIYA-AB^{t}CCC strain (n=110), or the EPIYA-AB^{t}D strain (n=110). Group 2 consisted of animals infected with the Δ*cagA* strain (n=110), wild type 7.13 strain (n=110), or the EPIYA-AB^t strain (n=110). Animals were age matched and assigned random numbers for blinding purposes. A variety of time points were assessed (2, 4, 6, 8, 10, 12, 16, 20, 24, 30, and 36 weeks) and 10 animals per time point per strain were sacrificed. For the mock infected animals, 5 animals per time point were sacrificed. At the indicated times, the stomachs were harvested, and the glandular portion of the stomach was bisected. Half of the tissue was paraffin embedded, sectioned and stained with hemotoxylin and eosin. The other half of the stomach was weighed and homogenized in brucella broth with a mechanical homogenizer (Tissue, Tearer; Biospec Products Inc., Bartlesville, OK), and the number of viable CFU was determined by plating on HBA plates supplemented with 50 µg/mL vancomycin, 10 µg/mL nalidixic acid (Sigma, St. Louis, MO) and 100 µg/mL bacitracin (USB Corporation, Santa Clara, CA).

The most striking result from the *in vivo* study was the fact the EPIYA-AB^{t}D strain showed a dramatic decrease in colonization as compared to the other isogenic strains used to infect the Mongolian gerbils at every time point thoroughout the experiment (Fig. 21A). For convenience, colonization was defined by detection of a

Figure 21: In vivo *colonization*. A. A graph of the percent of animals colonized with the various isogenic strains for each time point is shown. Colonization was defined as our ability to detect at least one bacterium. The level of detection was decreased after the six week post infection time point. B. A histogram of the overall percentage of animals colonized per the various isogenic strains is shown.

Figure 21: In vivo *colonization*

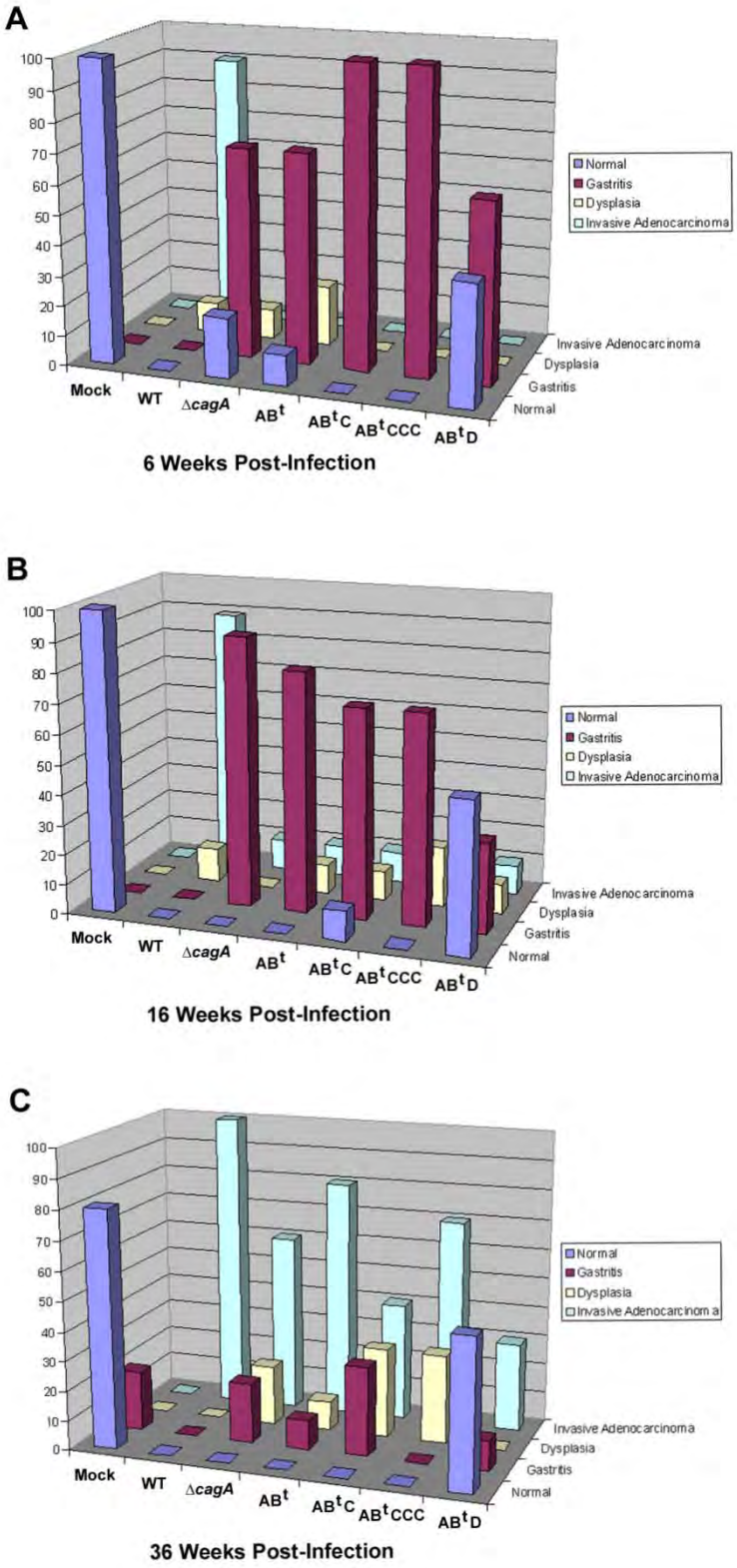


single bacterium. The overall colonization rate was 93% or above for animals infected with every strain (*ΔcagA* 99%, wild type 100%, EPIYA- AB^t 98%, EPIYA- AB^tC 93%, and EPIYA- AB^tCCC 100%), except for the EPIYA- AB^tD strain which was 57% (Fig 21B). However, within the animals infected with the EPIYA- AB^tD strain there were dramatic differences in colonization loads. Of the animals that were colonized with the EPIYA- AB^tD, only strain 67% showed colonization levels similar to wild type.

We also wanted to assess induced pathological differences, so a pathologist, blinded to the study, analyzed the sections for diagnosis. Animals infected with wild type bacteria progressed to cancer within six weeks, which is similar to what has been previously published (14, 15). However, at the six week time point, animals infected with the isogenic strains only developed gastritis. Of note these included a number of animals infected with the Δ*cagA* strain (Fig. 22A). This trend extended throughout most of the experiment (Fig. 22B). However, by 36 weeks, most animals had progressed to gastric cancer, again including those infected with the Δ*cagA* strain. Additionally, there was a high percentage of animals infected with the AB^tD strain that displayed normal gastric histology, even at 36 weeks post-infection. However, these results are likely confounded by the fact that these animals may not have been colonized. In fact, some animals that had no detectable levels of *H. pylori* infection developed gastric disease. This fact suggests that there may have been external factors influencing disease within these animals or that the strains contained secondary mutations that affected colonization and disease development.

Figure 22: Isogenic strain-induced disease states. Histograms of the percent of animals afflicted with each disease state at six weeks (A), 16 weeks (B), and 36 weeks (C) are shown. Animals with normal gastric mucosa are depicted by the blue bars, while animals that suffered from gastritis are depicted by maroon colored bars. Animals with signs of dysplasia are indicated by the yellow bars, and animals that were diagnosed with gastric cancer are represented by turquoise bars.

Figure 22: Isogenic strain-induced disease states



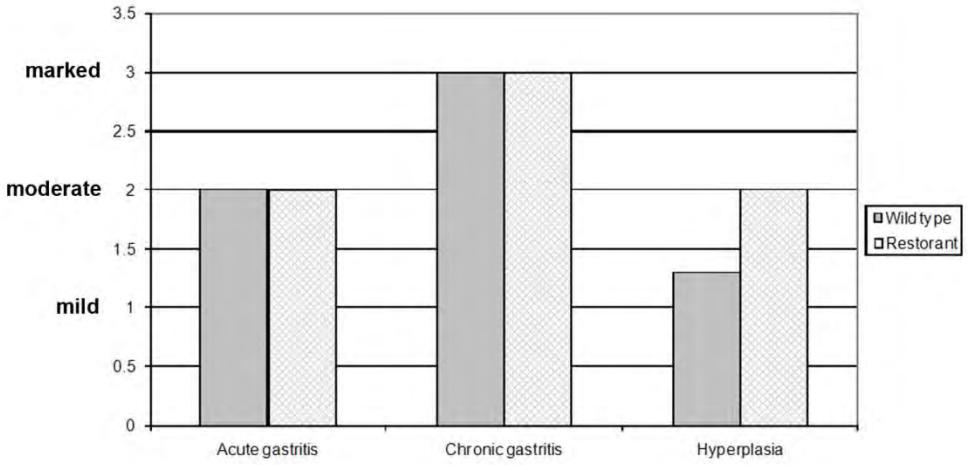
Due to the potential second site mutations, a small animal study was performed to assess pathologic differences between the wild type and a newly created restorant strain (DSM927) at the six week time point. Histological sections were graded for acute gastritis, chronic gastritis, and hyperplasia, using a scale from 1 (mild) to 3 (marked) as previously described (8). Animals infected with both wild type and restorant strains showed infiltration of polymorphonuclear neutrophils in the stroma (a grade of two; Fig. 23). Animals infected with both strains also showed formation of lymphoid follicles in the mucosa and submucosa (a grade of three; Fig. 11). Most animals infected with the wild type strain showed the presence of hyperplastic glands in the mucosa (score of one) with a few animals showing heterotopic proliferative glands adjacent to the submucosa (score of two). Similarly, infection with the restorant strain produced hyperplastic glands in the mucosa and heterotopic proliferative glands adjacent to the submucosa (score of two; Fig. 23).

Conclusions

Herein we present the first attempt to address the role of the CagA EPIYA motifs in *H. pylori*-induced disease progression by the creation and characterization of isogenic strains. Since it was shown to cause cancer in the Mongolian gerbil, we used the *H. pylori* 7.13 strain as the background strain and constructed Δ*cagA*, ΔEPIYA, EPIYA-AB^t, -AB^tC, -AB^tCC, -AB^tCCC, -AB^tCCCC, -AB^tD, and restorant isogenic strains. To this end, we were able to successfully optimize *in vitro* assays to assess localization of the isogenic strains on AGS cells, adherence to AGS cells, internalization into AGS cells, translocation and phosphorylation of CagA, as well as morphological changes in AGS cells, which is a surrogate for modulation of host signaling pathways. Additionally, we

Figure 23: Histopathology of the new restorant. A histogram depicting similarities in the level of disease between animals infected with the wild type strain and the newly created restorant strain, DSM927. Histological sections were graded for acute gastritis, chronic gastritis, and hyperplasia, using a scale from 1 (mild) to 3 (marked; 8). Acute gastritis is graded based on the infiltration of polymorphonuclear neutrophils, and chronic gastritis is graded based on the number and location of lymphoid follicles. Hyperplasia is measured by the presence and location of hyperplastic glands and heterotopic proliferative glands.

Figure 23: Histopathology of the new restorant



were also able to carry out a long term large scale animal study. While the growth dynamics of all of these strains mimicked the wild type strain, changes in the expression of CagA were evident across these strains. Since theoretically, all of the isogenic strains, with the exceptions of the Δ*cagA* and ΔEPIYA strains, should express equal levels of CagA, these differences in CagA expression were the first indication that the strains might contain secondary mutations. Increasing evidence suggested that there were secondary mutations within the strains, since the wild type, the EPIYA-AB^{t}C, and the restorant strains did not act similarly in the *in vitro* assays. Additionally, there were drastic differences in the pathology induced between the wild type strain and the EPIYA-AB^{t}C strain *in vivo*.

Since beginning these studies, we have learned that strain 7.13 loses its *in vivo* virulence after lab passage. Also, other work in our lab has shown drastic differences *in vivo* phenotype for strains that were identically manipulated to delete the same gene. Thus, 7.13 is likely not a good strain to complete our studies. However, to date each of the isogenic strains has been reconstructed in the genetically stable strain background of strain G27. Thus, the *in vitro* characterization of the EPIYA-motifs may still be accomplished using the techniques outlined in this chapter. Also, a new strain has been shown to cause gastric cancer in a mouse model. Thus, there are additional options to explore the role of the EPIYA motifs in *H. pylori*-induced disease progression.

References

1. **Alm, R. A., and T. J. Trust.** 1999. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. J Mol Med **77:**834-46.

2. **Amieva, M. R., R. Vogelmann, A. Covacci, L. S. Tompkins, W. J. Nelson, and S. Falkow.** 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science **300:**1430-4.

3. **Argent, R. H., J. L. Hale, E. M. El-Omar, and J. C. Atherton.** 2008. Differences in *Helicobacter pylori* CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. J Med Microbiol **57:**1062-7.

4. **Argent, R. H., M. Kidd, R. J. Owen, R. J. Thomas, M. C. Limb, and J. C. Atherton.** 2004. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of *Helicobacter pylori*. Gastroenterology **127:**514-23.

5. **Azuma, T., S. Yamazaki, A. Yamakawa, M. Ohtani, A. Muramatsu, H. Suto, Y. Ito, M. Dojo, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Higashi, and M. Hatakeyama.** 2004. Association between diversity in the Src homology 2 domain--containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis **189:**820-7.

6. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

7. **Bourzac, K. M., C. M. Botham, and K. Guillemin.** 2007. *Helicobacter pylori* CagA induces AGS cell elongation through a cell retraction defect that is independent of Cdc42, Rac1, and Arp2/3. Infect Immun **75:**1203-13.

8. **Cao, X., T. Tsukamoto, T. Seki, H. Tanaka, S. Morimura, L. Cao, T. Mizoshita, H. Ban, T. Toyoda, H. Maeda, and M. Tatematsu.** 2008. 4-Vinyl-2,6-dimethoxyphenol (canolol) suppresses oxidative stress and gastric carcinogenesis in *Helicobacter pylori*-infected carcinogen-treated Mongolian gerbils. Int J Cancer **122:**1445-54.

9. **Carpenter, B. M., H. Gancz, R. P. Gonzalez-Nieves, A. L. West, J. M. Whitmire, S. L. Michel, and D. S. Merrell.** 2009. A single nucleotide change affects *fur*-dependent regulation of *sodB* in *H. pylori*. PLoS One **4:**e5369.

10. **Copass, M., G. Grandi, and R. Rappuoli.** 1997. Introduction of unmarked mutations in the *Helicobacter pylori vacA* gene with a sucrose sensitivity marker. Infect Immun **65:**1949-52.

11. **Covacci, A., S. Censini, M. Bugnoli, R. Petracca, D. Burroni, G. Macchia, A. Massone, E. Papini, Z. Xiang, N. Figura, and R. Rappuoli.** 1993. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci U S A **90:**5791-5.

12. **El-Etr, S. H., A. Mueller, L. S. Tompkins, S. Falkow, and D. S. Merrell.** 2004. Phosphorylation-independent effects of CagA during interaction between *Helicobacter pylori* and T84 polarized monolayers. J Infect Dis **190:**1516-23.

13. **Evans, D. G., D. J. Evans, Jr., and D. Y. Graham.** 1992. Adherence and internalization of *Helicobacter pylori* by HEp-2 cells. Gastroenterology **102:**1557-67.

14. **Franco, A. T., D. A. Israel, M. K. Washington, U. Krishna, J. G. Fox, A. B. Rogers, A. S. Neish, L. Collier-Hyams, G. I. Perez-Perez, M. Hatakeyama, R. Whitehead, K. Gaus, D. P. O'Brien, J. Romero-Gallo, and R. M. Peek, Jr.** 2005. Activation of beta-catenin by carcinogenic *Helicobacter pylori*. Proc Natl Acad Sci U S A **102:**10646-51.

15. **Franco, A. T., E. Johnston, U. Krishna, Y. Yamaoka, D. A. Israel, T. A. Nagy, L. E. Wroblewski, M. B. Piazuelo, P. Correa, and R. M. Peek, Jr.** 2008.

Regulation of gastric carcinogenesis by *Helicobacter pylori* virulence factors. Cancer Res **68:**379-87.

16. **Gancz, H., S. Censini, and D. S. Merrell.** 2006. Iron and pH homeostasis intersect at the level of Fur regulation in the gastric pathogen *Helicobacter pylori*. Infect Immun **74:**602-14.

17. **Go, M. F., V. Kapur, D. Y. Graham, and J. M. Musser.** 1996. Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. J Bacteriol **178:**3934-8.

18. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

19. **Hatakeyama, M.** 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat Rev Cancer **4:**688-94.

20. **Hatakeyama, M., and H. Higashi.** 2005. *Helicobacter pylori* CagA: a new paradigm for bacterial carcinogenesis. Cancer Sci **96:**835-43.

21. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

22. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **295:**683-6.

23. **Jones, K. R., Y. M. Joo, S. Jang, Y. J. Yoo, H. S. Lee, I. S. Chung, C. H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol **47:**959-68.

24. **Majewski, S. I., and C. S. Goodwin.** 1988. Restriction endonuclease analysis of the genome of *Campylobacter pylori* with a rapid extraction method: evidence for considerable genomic variation. J Infect Dis **157:**465-71.

25. **Miles, S., M. B. Piazuelo, C. Semino-Mora, M. K. Washington, A. Dubois, R. M. Peek, Jr., P. Correa, and D. S. Merrell.** 2010. Detailed *in vivo* analysis of the role of *Helicobacter pylori* Fur in colonization and disease. Infect Immun **78:**3073-82.

26. **Ren, S., H. Higashi, H. Lu, T. Azuma, and M. Hatakeyama.** 2006. Structural basis and functional consequence of *Helicobacter pylori* CagA multimerization in cells. J Biol Chem **281:**32344-52.

27. **Segal, E. D., J. Cha, J. Lo, S. Falkow, and L. S. Tompkins.** 1999. Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*. Proc Natl Acad Sci U S A **96:**14559-64.

28. **Stein, M., F. Bagnoli, R. Halenbeck, R. Rappuoli, W. J. Fantl, and A. Covacci.** 2002. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. Mol Microbiol **43:**971-80.

29. **Taylor, D. E., M. Eaton, N. Chang, and S. M. Salama.** 1992. Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. J Bacteriol **174:**6800-6.

30. **Yamaoka, Y., H. M. El-Zimaity, O. Gutierrez, N. Figura, J. G. Kim, T. Kodama, K. Kashima, and D. Y. Graham.** 1999. Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. Gastroenterology **117:**342-9.

Chapter Six

Discussion

Although much has been learned about specific *H. pylori* virulence factors, little is currently understood about why some *H. pylori*-infected individuals progress to develop gastric cancer while others remain asymptomatic. The goal of this thesis was to better understand the association between different polymorphisms in CagA and VacA and disease outcome. Specifically, we showed that East Asian CagA (EPIYA-ABD) was linked to progression to gastric cancer in a South Korean population (39). In fact, all *H. pylori* strains from cancer patients expressed and delivered phosphorylatable CagA to host cells. In contrast, the presence of the *cagA* gene did not strictly correlate to expression and delivery of CagA with non-cancer strains (39). Our study was the first to statistically link a specific *cagA* allele to gastric cancer development in a human population. We next examined the role of VacA polymorphisms within that population, and found that while the distribution of *vacA* alleles was not directly associated with disease state, it was associated with the distribution of *cagA* alleles. Furthermore, the *vacA* allele was associated with the *cagA* allele and disease state. Next, we were able to analyze the contribution of the newly described i region of VacA to disease development. To this end, we identified an amino acid (196) that was associated with for development of gastric cancer. We were also able to identify some associations that were CagA-dependent, such as the association of VacA and disease state in the EPIYA-ABD population as well as amino acid distribution at position 231 and disease state in the non

EPIYA-ABD population. In addition, in the process of completing this thesis, we were able to optimize techniques that will ultimately be used to characterize CagA isogenic strains. Those future studies will help to elucidate the role of the EPIYA motifs in *H. pylori*-induced host cell damage both *in vitro* and *in vivo*. *En masse*, the data presented herein add to what we know about the complexity of *H. pylori*-induced pathogenesis. Overall, it is becoming increasingly more evident that polymorphisms within CagA and VacA, alone and in concert, affect *H. pylori*-induced disease. However, the reason why only a portion of the population develops gastric cancer still remains unclear. Other bacterial virulence factors, as well as multiple host, dietary, and environmental factors have been indicated as participants in *H. pylori*-induced disease. Clearly, further study is required to determine which factors are involved and what role they have in the development of *H. pylori*-induced gastric cancer.

Unanswered Questions Stemming from the Epidemiological Studies

CagA

A key question that should be addressed is, why is there a difference in the degree of CagA variation between Western and East Asian strains. Western isolates vary widely in the number of EPIYA-C motifs that are present (7, 8), whereas there is a distinct lack of variation in East Asian strains. In fact, one study examining Gen Bank sequences of 500 East Asian strains, found that 441 (88.2%) contained a canonical EPIYA-ABD motif (7). Indeed, additional studies confirmed this conservation among East Asian strains; greater than 84% of the examined strains across all three studies contained an EPIYA-

ABD motif (7, 12, 39). Moreover, in our molecular epidemiologic CagA study, the majority of strains (87.5%) contained an EPIYA-D motif. Interestingly, those isolates that contained a nonstandard EPIYA-ABD motif were associated with development of gastritis (39). This finding suggests that variation in East Asian *cagA* is not as favorable as in Western isolates and that variation may affect disease progression.

The reasons for strict conservation of the EPIYA-ABD motif are unknown, but may be explained by several possible theories. One theory is a difference in the degree of selective pressure for variation imposed on Western and East Asian strains. Western CagA shows a lower affinity for SHP-2 and is associated with less severe inflammation, host cell morphological changes, and disease as compared to East Asian strains (30, 31). Moreover, there is a dose response in the number of EPIYA-C motifs and the levels of tyrosine phosphorylation, SHP-2 binding, host cell morphological changes, and inflammation induced by Western CagA (32, 39, 87). Perhaps, increased inflammation enhances colonization, and therefore may act as a positive selective pressure to increase the number of EPIYA-C motifs. This pressure would not be experienced by East Asian strains since the EPIYA-D motif is already so biologically active. However, if increased inflammation is important for colonization of *H. pylori*, then there may be a selective pressure to keep a canonical EPIYA-ABD motif. For instance, perhaps extra EPIYA-A or -B motifs in association with an EPIYA-D motif more strongly activate the negative feedback loop that results from EPIYA-A or -B binding to Csk, thereby decreasing inflammation (7, 82). Additionally, a single EPIYA-D motif may be optimal for SHP-2 binding, and extra EPIYA-D motifs may contort CagA's conformation, thereby destabilizing this interaction, again resulting in a decrease in inflammation. The true

reason for this conservation among East Asian strains could help elucidate the impact of these motifs.

Work on the EPIYA motif region of CagA has primarily focused on differences in phosphorylation and subsequent modulation of phosphorylation-dependent host signaling pathways (reviewed in (40). Recently, however, a CagA multimerization domain was described that is located within the EPIYA region, and therefore varies as the EPIYA motifs vary. Some studies suggest that this domain is responsible for the differential modulation of some phosphorylation-independent host signaling pathways (Fig. 2; (59, 74, 77, 83). However, since the existence of this domain is a very recent discovery, more work is needed to clearly define the role of the multimerization domain on *H. pylori*-induced changes in host signaling pathways. Additionally, changes in this domain as a result of changes in the EPIYA motif will need to be investigated.

While we have gained much knowledge about the role of the C-terminus of CagA, not much is currently known about the N-terminus of the protein. Moreover, of the studies that have been completed, conflicting data have arisen. Specifically, it was demonstrated that the N-terminus of CagA is responsible for directing CagA to the plasma membrane (14), but other data demonstrated the EPIYA-motifs located in the C-terminus were responsible for proper localization (32, 33). Pelz *et al.* recently demonstrated that two independent domains, one in the N-terminus and one in the C-terminus, are responsible for directing CagA to the plasma membrane (69). In fact, these authors showed that the first 200 amino acids of CagA actually act as an inhibitory domain that dampens the host response to the C-terminus of CagA. This domain reduces activation of the oncogenic transcription factor β-catenin, reduces the length of the

"hummingbird" protrusions, and increases the speed and strength of new cell to cell contact (69). Thus, an interesting question would be to ask if the activity of the inhibitory domain varies in conjunction with varying motifs in the C-terminus. In other words, is the effect of this inhibitory domain different based on the variations of the different *cagA* alleles? This question could be addressed by removing the inhibitory domain, by deleting the first 200 amino acids of CagA as described previously (69), within the context of isogenic strains that differ only in the EPIYA region of *cagA*. Differences in induced host cell morphological changes could then be assessed when these strains were used to infect AGS cells, and fold changes in the number of elongated cells could be calculated and compared between the strains and their isogenic strains lacking the inhibitory domain. If the difference in fold change was similar across the different EPIYA isogenic strains, this would suggest that this inhibitory domain acts independently of the EPIYA motif. Additionally, levels of activated β-catenin could be measured in the infected host cells. If this inhibitory domain is influenced by the EPIYA motif that is present, it could also be interesting to determine whether there is also an interaction with the multermization domain. This hypothesis could be addressed by making phosphorylation-resistant mutants within the above mentioned strains, thereby abrogating the phosphorylation-dependent pathways; the tyrosine could be changed to an alanine or serine through site specific mutagenesis (5, 21, 32). The multermization domain is important for the activation of PARP-1. Determing the levels of PARP-1 activation between phosphorylation-resistant isolates containing the different multermization domains and the isogenic strains missing the inhibitory domain will identify the impact, if any, of this inhibitory domain on the activation of CagA phosphorylation-independent

pathways. While we know quite a lot about the contribution of CagA and the EPIYA motifs to *H. pylori*-induced disease, there are clearly more questions on the molecular level that need to be answered.

VacA

Many questions still remain to be addressed regarding the different *vacA* alleles. Included among these are questions concerning the signal (s) region of the protein. Geographic differences between the s1 region have been reported, and three subtypes have been identified: s1a, s1b, and s1c (11). Are these differences important for how VacA acts on host cells and/or its interplay with other virulence factors? Although this nomenclature is now seldom used, it would be interesting to determine if the different subtypes display any functional differences in activity, since the s region is responsible for most of the toxicity of VacA (52-54, 72). This objective could be accomplished by creating isogenic strains that vary only in the s1 subtype or more simply by intoxicating eukaryotic cells with identical concentrations of purified VacA containing one of the three different subtypes. One could assess the ability of the various VacA toxins to cause vacuoles within host cells as well as induce apoptosis, which could be measured through activated TUNEL assays or by measuring the amount of activated caspase 3.

Another question that has arisen from our VacA epidemiology work concerns the overall importance of the middle (m) region of the protein. The first aspect of the m region that should be addressed is the association between the m region and gender. Our work was the first study to observe such an association (37), and begs the question of why this association exists. Females appear more likely to be infected with strains

encoding the m2 *vacA* allele (37). Does this association exist in populations where the m2 allele is more prevalent, such as in regions of China (67) and Poland (49)? If it does exist, then there may be something physiologically different between the gastric environment of males and females. The nature of these differences may be unknown, there are many possibilities. For instance, is there a difference in the pH of the stomach acid that may consequently influence disease state? Is there a difference between the actual gastric epithelium between males and females? Minute differences in the thickness or composition of the mucus layer, which could in turn impact contact of *H. pylori* with the gastric epithelium, could affect the amount of toxin delivered to host cells. Are there differences in the amount or type of adherence receptors expressed in the gastric epithelium of males versus females? Furthermore, what affect does the endocrine system, more specifically changes in hormone levels, have on this process? Recently, Dr. Claire Fraser-Liggett has identified a gastric microbiome and is currently identifying differences in this bacterial population between individuals (2009 Bullard Lecture). It is possible that the differences in eating habits between the sexes influence this microbiome and thereby the distribution of the m *vacA* allele. Answers to these questions may help to explain the differences in distribution of the m1 *vacA* allele between men and women. Overall, the increased cellular tropism of the m1 *vacA* allele (66), the finding that the presence of the m1 region increases the risk for gastric cancer (79), and the fact that patients infected with *H. pylori* strains encoding for the m2 allele are more likely to be female (37) may explain why males are overall more likely to develop gastric cancer (reviewed in 76).

The second aspect of the m allele that should be examined is its contribution to the association between VacA and CagA, as well as between VacA, CagA, and disease state. Does the association between the m region and the *cagA* allele occur simply because the *vacA* regions that are responsible for more severe disease also co-vary among themselves? This possiblity does not seem likely in this population, since we only found the s1 allele and were still able to detect this association. Furthermore, is the three-way association between the m region, CagA, and disease state due to the fact that the m2 allele has a narrower cell tropism, thereby affecting the types of cells VacA can intoxicate? If this were the case, one would expect to see a direct correlation between the m region and disease state. However, this correlation is not seen within this South Korean population.

Several questions also still remain regarding the intermediate (i) region of VacA. The major difference between the i1 and i2 alleles is the addition of three polar amino acids (asparagine, histidine, and serine) in Cluster C (37, 38). Since both clusters B and C have been suggested to affect toxin activity, studies directed towards understanding the specific role of these amino acids in vacuolating activity would be interesting. In order to better examine this, an i1 allele could be genetically engineered through the addition of only these three amino acids. Perhaps, it would not be surprising if the addition of these amino acids decreased toxin activity, since it has been demonstrated that additional amino acids near the cleavage site in the s region decrease the ability of the toxin to integrate into the cellular membrane, which in turn decreases toxin activity (11, 53). Such a result might explain why the i region has been suggested to be a better predictor of disease.

Another aspect of the i region worth examining involves toxin evolution and the presence of the i3 allele. Is the presence of the i3 region really a snapshot of the evolution of the i region from one allele to another (i1 or i2)? It would be interesting to infect animals with *H. pylori* strains containing the i3 allele, and to sequence several recovered isolates after infection to see if the i3 allele has been replaced with either an i1 or i2 allele. Since evolution is not a rapid event, this experiment would consist of long term animal infections - 36 weeks. Isolates from these animals could be sequenced, but would likely be used to subsequently re-infect new animals, as multiple passages *in vivo* would likely be needed to evaluate any evolution of this allele. This experiment should be performed with an i3 strain where cluster B is from an i1 strain and cluster C is from an i2 strain, as well as an i3 strain where cluster B is from an i2 strain and cluster C is from an i1 strain. Results from these experiments could indicate if there is an overall selective pressure to evolve *in vivo*, and if the original cluster sequences influences that evolution. In this same vein, it would be interesting to determine if there is a functional difference between the i3 allele and the i1 and i2 alleles. If so, is there a functional difference between i3 strains that contain cluster B from an i1 strain and cluster C from an i2 strain as compared to an i3 strain that contains cluster B from an i2 strain and cluster C from an i1 strain? Alternatively, is toxicity dependent on the sequence of an individual cluster?

Finally, it would be interesting to look at a population containing an increased percentage of i2 alleles in order to assess the distribution of amino acids at position 196 on the *vacA* allele as well as on disease state. Our work demonstrated that the distribution of amino acids found at this position was linked to more severe disease,

specifically gastric cancer (38). In the South Korean population analyzed, all of the i2 alleles encoded for a leucine at amino acid 196 (38). Thus, perhaps it would be interesting to analyze a population where the i2 allele is more prevalent to see if this trend persists. If so, the amino acid found at this position may be partially responsible for the results of a previous study in which it was concluded that the i region was the best predictor of disease (75). Further studies could also investigate variation in the major amino acid differences seen between the i1 and i2 alleles. This line of research might better indicate which amino acids are critical for toxin activity.

Cag/Vac Interaction

Recent studies have identified an association between the *cagA* allele and the v*acA* allele that appears to affect *H. pylori* toxicity and disease severity (37, 84, 89). Infection with *H. pylori* strains that encode for CagA and s1/m1 VacA result in highly active corpus gastritis (57), which is linked to the development of gastric cancer (55-57). Our study also found an association between the *cagA* allele and *vacA* allele, as well as a three way association between the distribution of the *cagA* and *vacA* alleles and disease state (37). Indeed, in our South Korean population, the majority of *H. pylori* strains carry the most toxic form of both VacA and CagA, and this may explain the high rate of severe gastric disease among the South Korean population (37).

Conventionally, one might think that both toxins concomitantly exert drastic effects on the same host cell. However, recent data suggest that the converse is true; the presence of both CagA and VacA may dampen the effect of each protein alone, possibly leading to increased survival of infected host cells (9). In fact, when both toxins are

present, there is less VacA-induced apoptosis then when cells are intoxicated with VacA alone (9). Additionally, eukaryotic cells intoxicated with both toxins demonstrate less CagA-induced morphological changes as compared to cells intoxicated with CagA alone (9).

Since these results are still fairly new, many questions remain. For instance, is there a direct interaction between CagA and VacA, or more likely, is this effect the result of activation of competing pathways by the two toxins (Fig. 2 and 3)? If there is a direct interaction between these two proteins, this could possibly be detected by performing pull-down assays. Does VacA somehow amplify the function of the newly identified inhibitory domain in the N-terminus of CagA (69)? Also, in thinking about the chronology of *H. pylori* infection, does the bacterium utilize the two toxins to increase the life span of the host cell and thus, to prolong infection? Indeed, it seems plausible that the most severe forms of gastric disease would result from long term infection of cells, and therefore, long term *H. pylori*-induced inflammation. In terms of CagA and VacA "interaction," does an order of events exist that is important for the resulting effects? For example, since VacA is secreted while CagA is injected by *H. pylori*, do cells first need to be intoxicated by one or the other toxin to see the protective correlate? This question could be directly assessed through *in vitro* assays. First, eukaryotic cells will be tranfected with *cagA* under control of an inducible promoter. CagA can then be induced within these cells and then intoxicated with VacA at various time points, followed by analysis of these cells for induction of apoptosis and morphological changes. Additionally, is the damping effect of the toxins achieved by reaching a threshold of both toxins, or is the mere presence of any amount of the two toxins sufficient? This question

could be assessed by determining if there is a dose response to the toxins. For instance, the transfected eukaryotic cells described above could be intoxicated with increasing amounts of VacA and a set amount of induced CagA, or with a set amount of VacA and increasing amounts of induced CagA. Eukaryotic cells would then be analyzed for levels of apoptosis as well as morphological changes. Additionally, the order of toxin addition could be inversed, based on the findings from the above study assessing the importance of the order of intoxication.

Finally, considering CagA and VacA "interaction" in the context of our epidemiological data, what are the physiological consequences of intoxication with the different CagA or VacA variants, or combination of these different toxins? For instance, if a strain carries EPIYA-ABD CagA and s2/i2/m2 VacA, which typically shows no toxic activity, are the effects of CagA similar to a strain that carries an EPIYA-ABD CagA but no VacA? Alternatively, is the combination of Western CagA and s1/i1/m1 VacA more or less lethal to cells than East Asian CagA and s2/i2/m2 VacA? These and other allele based questions could be addressed by making VacA isogenic strains within the context of the CagA EPIYA isogenic strain background as discussed later in this chapter. Since evidence is increasing that the association between the *cagA* allele and the *vacA* allele impact disease development, I strongly believe that the future of pathogenesis studies in *H. pylori* will have to focus on the effect of combinations of these virulence factors.

A Hierarchy of Virulence Factors

Another emerging theory is that CagA may be the "master" virulence factor, and that other virulence factors or polymorphisms are important only in the context of which

cagA allele is present (13, 37, 38). In fact, studies have found that in terms of gastric cancer, the *cagA* allele carried is the most important bacterial risk factor (15). Conversely, the i region of VacA is the best predictor for duodenal ulcers (15). Indeed, in our epidemiological studies, we found that different associations existed in populations carrying particular *cagA* alleles and that these associations were not found in populations encoding for a different *cagA* allele. For instance, when age and gender were taken into account, a two way association between the distribution of the *vacA* allele and disease state was found only within the EPIYA-ABD CagA population. Furthermore, non-s1/i1/m1 *vacA* alleles were associated with duodenal ulcers within the population carrying the East Asian EPIYA-ABD CagA, but with gastritis in the population carrying any other genotype of CagA (37). We also found that an association existed between disease state and amino acid 231 of the VacA i region, but only within the non EPIYA-ABD population (38). These findings again suggest that the effect of different virulence factors or polymorphisms within these virulence factors may be masked by which *cagA* allele is present. Indeed, this fact may help explain the vast amount of conflicting literature concerning the importance of these different virulence factors. Employing the statistical technique of meta-analysis to survey the epidemiological data in different geographic regions might help to shed light on some of these reported differences.

Lessons Learned from the Current Project

Unfortunately, the major molecular biology study in my thesis project encountered problems since it was ultimately discovered that the isogenic *cagA* strains that were constructed actually contained secondary mutations. While we do not

understand exactly how these mutations arose, some clues may come from thinking about the inherent genetic variability of the bacterium. In fact, *H. pylori* also has an increased rate of spontaneous mutations as compared to *E. coli*, with initial studies demonstrating that the spontaneous mutation rate in *H. pylori* is 10^{-7}-10^{-8} (28, 35, 85). Again, this rate varies among strains and a rate as high as 3 x 10^{-4} has been observed (46). In fact, genetic polymorphisms seem to be normal between strains. In a study that examined the genetic sequence of a house-keeping gene (*glmM*) it was found that the sequence was unique in all the strains examined (44). Furthermore, this microdiversity was observed in a number of other genes (1, 3, 4, 34, 65, 78), as well as within strains taken from the same patient (36). Additionally, *H. pylori*'s genome contains multiple genes that are phase variable. Indeed, when a single reference strain was sequenced, up to 27 genes were identified that contained nucleotide repeats that could facilitate phase variation (51, 81); two examples of these genes that have been examined in more detail include *fliP* (41) and *oipA* (88).

Animal passage of strains has been shown to induce formation of large numbers of fragmented genes and repeated regions (22). Also, Mongolian gerbils are naturally infected with *H. bilis* and *H. pylori* strains are naturally competent. In fact, the *H. pylori* has the ability to take up new DNA *in vivo*, which creates a constant chance for genetic exchange since a host can be infected with multiple *H. pylori* strains or related bacteria (24, 25, 36, 43). This phenomenon has been well documented with one of the *H. pylori* reference strains (J99). When an archived isolate of J99 was compared to isolates from the same patient taken 6 years apart, there was a high level of genetic diversity (36). Collectively, the new isolates had lost up to 2.3% of the open reading frames compared to

the archived J99 strain. Additionally these strains had gained DNA that was not found in the original J99 strain (36). Overall, natural competence has been proposed to contribute to the vast allelic diversity of the organism, and to help account for the considerable genetic variability (6-7%) that is seen between strains (4, 26, 50, 80).

In order to reduce the potential for genetic variation that could affect our experiments, the Merrell group has adapted certain lab protocols. For instance, when we create a mutant strain, we select a single colony of the mutant strain and then never utilize single colonies again. Bacteria are expanded as patches of cells from the freezer stock (-80°C) on antibiotic-supplemented horse blood agar plates for 36-42 hours, which is the minimal amount of time for growth. Bacteria are then expanded as lawns from these patches for about 20 hours on plates. The lawn is then used to inoculate 18 hour liquid starter cultures that are ultimately used to inoculate OD controlled experimental liquid cultures. All of these protocols are performed in an attempt to minimize the number of lab passages of the strain, and to make sure that if genetic variability occurs, it does so in the context of a population of cells. Furthermore, when feasible, we also create an independent biological isolate of all mutant strains.

In devising our isogenic strain study, we first created a ΔEPIYA strain, which was used as the strain background for all subsequent strains. Moreover, we followed the aforementioned lab protocols for expanding bacteria from freezer stocks, transformation, selection, and growth, which were designed to minimize the possibility of variation. Therefore, there was no reason to believe that these strains would contain secondary mutations. However, upon reflection on the project, there were several different points throughout the process when the data suggested that there might be a secondary mutation

that would complicate our study. Sequence analysis as well as growth dynamics suggested that the strains were in fact isogenic. However, the first indication that the strains were not behaving as expected came with the quantification of CagA expression. Theoretically, all of the isogenic strains, with the exceptions of the Δ*cagA* and ΔEPIYA strains, should have expressed CagA at similar levels. However, up to a two-fold difference could be seen between the EPIYA-AB^{t}CCCC and wild type strains (Fig. 16). Conversely, there were no major differences when CagA expression was compared to the restorant strain, which I believed to be a reasonable comparison; the restorant strain should be genetically identical to the 7.13 wild type strain, and had undergone the same genetic manipulation as the other isogenic strains. Minor differences in the restorant suggested that perhaps the genetic manipulation of the strain slightly altered the overall expression level of CagA. At the time, we considered this as no surprise; however, in hindsight, this result may have been the first indication that there were problems with the strains.

While CagA has not been shown to affect the adherence of *H. pylori* (5), any difference in adherence of strains to host cells would potentially alter the amount of CagA that could be translocated and phosphorylated, thereby changing the deregulation of host cell signaling pathways and potentially affecting development of gastric disease. In our studies, the adherence assay was the first assay that showed marked differences between the isogenic strains. Indeed, the restorant, EPIYA-AB^{t}C, and wild type strains, which should have all adhered at similar levels, did not behave as expected; the restorant strain was 10-fold less adherent than the wild type or EPIYA-AB^{t}C strains. When these preliminary results showed the lack of consistency between the restorant, EPIYA-AB^{t}C,

and wild type strains, we sought to rule out the potential confounding effect of slight differences in the growth phase of the different isogenic strains (Fig. 15). Re-examination of the growth curves suggested that at 18 hours some of the isogenic strains may have entered stationary phase, which could adversely affect the adherence of the bacteria to the AGS cells. We therefore repeated the assay using 12 hour liquid cultures to infect the AGS cells. We found that while the number of adherent bacteria was increased, the trends stayed the same, suggesting that there were secondary mutations within the strains (Fig. 18).

We next assessed the ability of the strains to deliver CagA to host cells, where it is phosphorylated by host cell kinases, thereby causing morphological changes within the cells. While there were slight differences in the peak phosphorylation time of CagA between the strains, the trends across the strains were the same. However, assessment of the five hour time point, which was a time point shown in a previous study to allow detection of high levels of phosphorylation (39), presented more evidence that there were secondary mutations in our strains. While increasing numbers of EPIYA-C motifs corresponded to increasing amounts of phosphorylated CagA, the three isogenic strains that contained only one EPIYA motif (EPIYA-AB^{t}C, EPIYA-AB^{t}D, and the restorant) showed similar levels of phosphorylation (Fig. 19). Unfortunately, these levels of phosphorylation were 2.5 times less than the amount of phosphorylated CagA from the wildtype strain (Fig. 19). This fact combined with the fact that the isogenic strains expressed approximately 70% of the amount of CagA that the wild type strain expressed, immediately suggested that there was something different among the isogenic strains. Additionally, the levels of phosphorylation did not translate into the expected changes in

host cell morphology. Once again, the wild type, EPIYA-AB^{t}C, and restorant strains all behaved differently (Fig. 20) suggesting that there was a second site mutation within at least some of the strains.

Based on the long length of the experiments, the large-scale animal study was regrettably started prior to the completion of the *in vitro* characterization of the isogenic strains. The animal study showed that there was a complete lack of differences in disease progression among Mongolian gerbils infected with the different isogenic strains. Moreover, there was a drastic difference in the pathology induced in animals infected with the wild type strain as compared to the EPIYA-AB^{t}C strain. This fact alone suggests that there is something fundamentally different between these two strains that is not CagA-related. Knowing what we know now, ideally a small pilot experiment of eight to ten weeks duration with fewer numbers of animals should have been conducted before proceeding with the large scale animal study.

Additionally, all *H. pylori* infection groups progressed to gastritis and eventually gastric cancer at the same rate as gerbils infected with the Δ*cagA* strain (Fig. 22), which has been shown previously to cause no gastric cancer in this model (23). This finding led to the re-creation of the ΔEPIYA strain followed by the creation of a new restorant strain. This new restorant strain induced pathology similar to the wild type strain in gerbils six weeks after infection; these animals displayed dysplasia and were progressing to gastric cancer (Fig. 23). This result combined with all the inconsistencies in the *in vitro* work suggests that there was a second site mutation within the original ΔEPIYA strain; therefore, no correlations could be made and there was no reason to further characterize these strains.

Since that time, we have learned that strain 7.13 loses *in vivo* virulence after low number of lab passages (R. Peek, personal communication). Additionally, other work in our lab has shown drastic *in vivo* differences for strains that have been idenitically manipulated to delete the same gene. Knowing this now, it might have been prudent to measure the mutation rate of 7.13 prior to beginning our studies in order to determine whether this strain has a higher mutation rate than other strains of *H. pylori.*

Looking Forward

Since H. pylori strain 7.13 was too genetically unstable to use for these studies, is there a future for this project? I believe that there is and would propose the following possibilities: 1) use strain G27 for *in vitro* characterization and 2) use strain PMSS1 for both *in vitro* and *in vivo* characterization. G27 is our lab's commonly used reference strain and has been shown to be fairly genetically stable. Currently, all of the CagA isogenic strains have been created in the G27 background, and these strains can now be used to determine the effect of the EPIYA motifs on host cell signaling. The EPIYA motif region of these strains has been sequenced to verify the genetic changes. To this end, if I was completing these studies, I would again start by assessing growth kinetics of these strains, as well as analyzing expression of CagA via Western blot analysis. The Western blot analysis should identify any differences in CagA expression that could complicate future experiments. Next, I would look at interaction of the strains with host cells by measuring bacterial localization, adherence, internalization, CagA phosphorylation, and induced host cell morphological changes. Though I would initiate these studies using AGS cells, it might also be interesting to assess another gastric cell

line that is able to form cellular barriers (HGE cells) or a cell line that can be polarized (T84 cells); each have been used to study the dynamics of *H. pylori* infection. Using the protocols in chapter five as a starting point, each assay would need to be optimized for the change in *H. pylori* strain background.

After completion of the basic characterization, modulation of host cell signalling pathways could be assessed. This should initially be assessed through analyzing of changes in host cell morphology, as described in chapter five. However, there are some changes that might be considered. One concern is that we perhaps did not ever achieve the maximal percent of elongated cells in our assay; subsequent smaller experiments showed that the highest percent of elongated cells was observed between 12-18 hours post infection. Additionally, while the percent of elongated cells tapered off, in a single experiment I conducted, the length of the protrusions in elongated cells infected with the EPIYA-AB^{t}D strain still continued to increase even at 18 hours. These results could indicate that the EPIYA-AB^{t}D strain may elicit reactions more slowly, yet its overall effects may be more drastic. Thus, it would be wise to explore additional time points, such as 24 and 36 hours post infection, to assess this possiblity. Of course, if these strains take longer to adhere or if they adhere at much lower levels, then the MOI or time points studied post infection would have to be adjusted accordingly. Alternatively, as noted in chapter five, the trends for phosphorylation of CagA was similar for all isogenic strains, so this may be a measurement of the continued activation of SHP-2. This SHP-2 activation difference would be important if activation of other phosphorylation dependent cellular pathways were analyzed.Once the basic charaterization of these strains is complete, modulation of specific pathways could be assessed. For example, Erk

activation could be assessed via Western blot analysis, as could activation of NF-κB. Localization of β-catenin to the nucleus could be assessed via microscopy or Western blot analysis of nuclear extracts of infected cell lysates. Additionally, since there is a high degree of genetic variablity between strains, these strains could be created within other reference strains, such as J99, 26695, or HPAG-1 to assess the role of the EPIYA motifs *in vitro*. An additional alternative for moving this project forward, could be the use of the parental strain (PMSS1) of the mouse derivative Sydney strain 1 (SS1). PMSS1 has recently been characterized (10, 47), and colonizes mice, but at a lower level than SS1 (10). PMSS1 encodes for a functional CagA protein, whereas SS1 does not express or deliver functional CagA (10, 71). Not only does PMSS1 produce a functional CagA, but this strain has been shown to cause severe pathology in mice, including atrophy, hyperplasia, and metaplasia (10). Since our knowledge about this strain is limited, and it is known to lose the ability to inject CagA into host cells after 1 month *in vivo* (10), it might be wise to measure the spontaneous mutation rate of PMSS1 before conclusively deciding to use it for isogenic strain construction. Ideally, there would be a *H. pylori* strain that infected animals, delivered CagA, causes severe pathology (gastric cancer), and is genetically stable; however in the absence of this ideal strain, PMSS1 may allow us to identify the role of the EPIYA motifs *in vivo*. Provided preliminary results are satisfactory, new primers to amplify the upstream (5') *cagA* region and the downstream (3') *cagA* region would need to be designed in order to create the constructs needed to produce the isogenic strains. Then, *in vitro* characterization of the newly constructed strains could proceed as described above. Once these assays are completed, a small pilot animal study to assess colonization load, histology, and the timeline for

disease progression should be completed. Overall the utilization of this strain could provide another option to explore the role of the EPIYA motifs not only *in vitro* but also in *in vivo* assays.

Importance and the Impact of Future Studies

Gastric cancer is still the second most common cause of cancer morbidity and mortality, and this could be reflective of the high incidence of *H. pylori* infection (20, 60, 68, 86). It could also be a result of the high prevalence of *cagA* in many *H. pylori* strains, or due to the presence of certain CagA polymorphisms that predominate in geographic areas that have high rates of gastric cancer (2, 18, 20, 27, 29, 45, 86). Due to the fact that we do not yet thoroughly understand the process of *H. pylori* induced pathogenesis, including development of gastric cancer, elucidation of virulence factors or virulence factor polymorphisms that impact disease is imperative. Epidemiological studies are traditionally good indicators of trends and serve as a starting point for molecular studies. Unfortunately, *H. pylori* is an organism that shows a high rate of genetic variability, which limits the impact of traditional epidemiological studies. Thus, in order to elucidate the exact role of virulence factors or polymorphisms, it is best to assess differences by creating isogenic strains.

The successful creation of EPIYA isogenic strains will not only answer the question as to what role the EPIYA motifs play in disease manifestations, but will open the door to the assessment of multiple virulence factors. For instance, it has already been demonstrated that CagA is a “master” virulence factor (13, 37, 38), and that there is an association between other virulence factors and disease among the different *cagA* alleles.

This is especially true for different *vacA* alleles (38, 42). Thus, once the EPIYA isogenic strains are created and characterized, they could then be used as the parental strain background to create isogenic strains containing different polymorphic forms of other virulence factors. This would allow us to not only assess the role of different virulence factors in disease, but also their role in disease development within the context of a particular *cagA* allele.

If I were to undertake these assays personally, I would focus first on VacA. As discussed earlier, VacA is polymorphic and different *vacA* alleles impact disease differently (15, 37, 48, 58). Thus, we could create *vacA* isogenic strains within certain *cagA* isogenic backgrounds to more conclusively look at CagA and VacA interaction. Specifically, I would alter the *vacA* alleles within the context of the EPIYA-AB^{t}C, EPIYA-AB^{t}CCC, EPIYA-AB^{t}D and the restorant strains. Experiments with these strains would identify any differences between East Asian (EPIYA-AB^{t}D) and Western (EPIYA-AB^{t}C and EPIYA-AB^{t}CCC) strains. They would also indicate if the number of EPIYA-C motifs influences the effects of the different *vacA* alleles. The restorant strain would provide an important control for genetic manipulation of our strains. These types of studies would allow our lab to not only examine the role of the different *vacA* alleles, but to assess the impact of these alleles within the context of the *cagA* allele. Furthermore, these types of studies may also clarify some of the discrepancies in the literature as to the impact of different *vacA* regions (37, 75).

Additionally, I would likely focus on two other polymorphic virulence factors that have been implicated in disease development, HomB (19, 42, 64) and OipA (16, 61). The *Helicobacter* outer membrane (Hom) proteins are complex because *H. pylori* has two

loci that can encode for a Hom protein; strains can encode for either *homA*, *homB*, *homA/homB*, *homA/homA*, *homB/homB*, or be negative for both *homA* and *homB* (19, 42, 64). The presence of *homB* has been linked to the development of more severe disease, as compared to the presence of *homA* (19, 42, 64). Additionally, a dose response has been identified for strains encoding for *homB/homB* (62, 63). OipA is an outer membrane protein whose expression has been shown to be subject to phase variation due to the number of CT repeats found in the *oipA* signal sequence (6). OipA "on" is used to designate a strain that expresses a functional protein (16, 61). Again, strains that encode for an OipA protein are associated with more severe disease outcomes (16, 61). Moreover, the OipA "on" phenotype is often found in *cagA* positive strains (6). The impact of *homB* and o*ipA* could be assessed individually or in the context of different virulence factors. In other words, besides creating isogenic derivative of these factors within the EPIYA-AB^{t}C, EPIYA-AB^{t}CCC, EPIYA-AB^{t}D and restorant strains to identify difference among the different *cagA* alleles, they could also be assessed within the EPIYA isogenic strains that carried different *vacA* alleles. For instance, I would likely first assess the EPIYA-AB^{t}C, EPIYA-AB^{t}CCC, EPIYA-AB^{t}D and the restorant strains that were s1/i1/m1 and s2/i2/m2. This would help limit the number of strains, especially since there are six different *hom* combinations that could be assessed. This would also seem to be the most likely place to observe differences, since the s1/i1/m1 is the most virulent and the s2/i2/m2 is the least virulent *vacA* allele. If a difference between these populations was observed, isogenic strains within the different *vacA* alleles could then be created and assessed. This system would also allow us to systematically examine these virulence factors in the context of the whole bacterium, and to assess their role in disease

progression, including progression to gastric cancer. This would also be a better system to analyze the role of these different virulence factors in the context of geographical differences, which could really help expand the field of *H. pylori* pathogenesis.

Conclusions

H. pylori is a medically important pathogen that has successfully challenged the preconceived idea that bacteria cannot cause gastric disease. However, more than a few questions remain about this process as well as what host, environmental, and bacterial factors are important for the progression to severe disease. This thesis focused on the bacterial toxins, CagA and VacA, and their role in influencing progression to more severe disease. To that end, we were the first to statistically link a specific *cagA* allele (EPIYA-ABD) to gastric cancer development. We were also able to analyze this large population of clinical isolates for the role of VacA polymorphisms, and found that while the distribution of *vacA* alleles was not directly associated with disease state, it was associated with the distribution of *cagA* alleles and in a three-way association that included the *vacA* allele, the *cagA* allele and disease state. During the course of this work, we identified an amino acid (196) important for the development of gastric cancer within the intermediate region of VacA. Additionally, we identified some associations that were CagA-dependent, such as the association of VacA and disease state in the EPIYA-ABD population and amino acid distribution at position 231 and disease state in the non EPIYA-ABD population. Finally, we were able to optimize techniques that will be used in future studies aimed at characterizing CagA isogenic strains.

While little is currently understood about why some *H. pylori* infected individuals develop gastric cancer while others remain asymptomatic, the data gathered during the course of this thesis will help shed some light on the pathogensis of *H. pylori*-induced disease. Indeed, the elucidation of bacterial factors that are involved in the pathogensis of *H. pylori*-induced disease, such as EPIYA-ABD CagA, is important, because they can serve as a diagnostic marker of infection with a more virulent strain. Understanding any hierarchy of virulence factors is imperative, because more evidence has accumulated that underscores the fact that colonization with *H. pylori* is protective against other illnesses, including asthma (73), tuberculosis (70) and esophageal cancer (reviewed in (17). Thus, suggesting that treatment should only be used for individuals infected with highly virulent strains. Such treatments that only target patients infected with strains that could cause severe disease would be akin to geographically personalized treatment, which I believe is the future for treating *H. pylori*-induced disease. Implementing such location-specific treatments could aid in eliminating a majority of gastric cancer mortality and morbidity worldwide without sacrificing the protective effects provided by infection with *H. pylori*.

References

1. **Achtman, M., T. Azuma, D. E. Berg, Y. Ito, G. Morelli, Z. J. Pan, S. Suerbaum, S. A. Thompson, A. van der Ende, and L. J. van Doorn.** 1999. Recombination and clonal groupings within *Helicobacter pylori* from different geographical regions. Mol Microbiol **32:**459-70.

2. **Ahn, Y. O., B. J. Park, K. Y. Yoo, N. K. Kim, D. S. Heo, J. K. Lee, H. S. Ahn, D. H. Kang, H. Kim, M. S. Lee, and T. S. Park.** 1991. Incidence estimation of stomach cancer among Koreans. J Korean Med Sci **6:**7-14.

3. **Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176-80.

4. **Alm, R. A., and T. J. Trust.** 1999. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. J Mol Med **77:**834-46.

5. **Amieva, M. R., R. Vogelmann, A. Covacci, L. S. Tompkins, W. J. Nelson, and S. Falkow.** 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science **300:**1430-4.

6. **Ando, T., R. M. Peek, D. Pride, S. M. Levine, T. Takata, Y. C. Lee, K. Kusugami, A. van der Ende, E. J. Kuipers, J. G. Kusters, and M. J. Blaser.** 2002. Polymorphisms of *Helicobacter pylori* HP0638 reflect geographic origin and correlate with *cagA* status. J Clin Microbiol **40:**239-46.

7. **Argent, R. H., J. L. Hale, E. M. El-Omar, and J. C. Atherton.** 2008. Differences in *Helicobacter pylori* CagA tyrosine phosphorylation motif patterns between western and East Asian strains, and influences on interleukin-8 secretion. J Med Microbiol **57:**1062-7.

8. **Argent, R. H., R. J. Thomas, F. Aviles-Jimenez, D. P. Letley, M. C. Limb, E. M. El-Omar, and J. C. Atherton.** 2008. Toxigenic *Helicobacter pylori* infection precedes gastric hypochlorhydria in cancer relatives, and *H. pylori* virulence evolves in these families. Clin Cancer Res **14:**2227-35.

9. **Argent, R. H., R. J. Thomas, D. P. Letley, M. G. Rittig, K. R. Hardie, and J. C. Atherton.** 2008. Functional association between the *Helicobacter pylori* virulence factors VacA and CagA. J Med Microbiol **57:**145-50.

10. **Arnold, I. C., J. Y. Lee, M. R. Amieva, A. Roers, R. A. Flavell, T. Sparwasser, and A. Muller.** 2011. Tolerance rather than immunity protects from *Helicobacter pylori*-induced gastric preneoplasia. Gastroenterology **140:**199-209.

11. **Atherton, J. C., P. Cao, R. M. Peek, Jr., M. K. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori.* Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem **270:**17771-7.

12. **Azuma, T., S. Yamazaki, A. Yamakawa, M. Ohtani, A. Muramatsu, H. Suto, Y. Ito, M. Dojo, Y. Yamazaki, M. Kuriyama, Y. Keida, H. Higashi, and M. Hatakeyama.** 2004. Association between diversity in the Src homology 2 domain--containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis **189:**820-7.

13. **Backert, S., N. Tegtmeyer, and M. Selbach.** 2010. The versatility of *Helicobacter pylori* CagA effector protein functions: The master key hypothesis. Helicobacter **15:**163-76.

14. **Bagnoli, F., L. Buti, L. Tompkins, A. Covacci, and M. R. Amieva.** 2005. *Helicobacter pylori* CagA induces a transition from polarized to invasive phenotypes in MDCK cells. Proc Natl Acad Sci U S A **102:**16339-44.

15. **Basso, D., C. F. Zambon, D. P. Letley, A. Stranges, A. Marchet, J. L. Rhead, S. Schiavon, G. Guariso, M. Ceroti, D. Nitti, M. Rugge, M. Plebani, and J. C.**

Atherton. 2008. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology **135:**91-9.

16. **Ben Mansour, K., C. Fendri, M. Zribi, A. Masmoudi, M. Labbene, A. Fillali, N. Ben Mami, T. Najjar, A. Meherzi, T. Sfar, and C. Burucoa.** 2010. Prevalence of *Helicobacter pylori vacA*, *cagA*, *iceA* and *oipA* genotypes in Tunisian patients. Ann Clin Microbiol Antimicrob **9:**10.

17. **Blaser, M. J.** 2008. Disappearing microbiota: *Helicobacter pylori* protection against esophageal adenocarcinoma. Cancer Prev Res (Phila) **1:**308-11.

18. **Blaser, M. J., G. I. Perez-Perez, H. Kleanthous, T. L. Cover, R. M. Peek, P. H. Chyou, G. N. Stemmermann, and A. Nomura.** 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res **55:**2111-5.

19. **Crabtree, J. E.** 1998. Role of cytokines in pathogenesis of *Helicobacter pylori*-induced mucosal damage. Dig Dis Sci **43:**46S-55S.

20. **Crew, K. D., and A. I. Neugut.** 2006. Epidemiology of gastric cancer. World J Gastroenterol **12:**354-62.

21. **El-Etr, S. H., A. Mueller, L. S. Tompkins, S. Falkow, and D. S. Merrell.** 2004. Phosphorylation-independent effects of CagA during interaction between *Helicobacter pylori* and T84 polarized monolayers. J Infect Dis **190:**1516-23.

22. **Farnbacher, M., T. Jahns, D. Willrodt, R. Daniel, R. Haas, A. Goesmann, S. Kurtz, and G. Rieder.** 2010. Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. BMC Genomics **11:**335.

23. **Franco, A. T., E. Johnston, U. Krishna, Y. Yamaoka, D. A. Israel, T. A. Nagy, L. E. Wroblewski, M. B. Piazuelo, P. Correa, and R. M. Peek, Jr.** 2008. Regulation of gastric carcinogenesis by *Helicobacter pylori* virulence factors. Cancer Res **68:**379-87.

24. **Fritz, E. L., T. Slavik, W. Delport, B. Olivier, and S. W. van der Merwe.** 2006. Incidence of *Helicobacter felis* and the effect of coinfection with *Helicobacter pylori* on the gastric mucosa in the African population. J Clin Microbiol **44:**1692-6.

25. **Ghose, C., G. I. Perez-Perez, L. J. van Doorn, M. G. Dominguez-Bello, and M. J. Blaser.** 2005. High frequency of gastric colonization with multiple *Helicobacter pylori* strains in Venezuelan subjects. J Clin Microbiol **43:**2635-41.

26. **Go, M. F., V. Kapur, D. Y. Graham, and J. M. Musser.** 1996. Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. J Bacteriol **178:**3934-8.

27. **Gwack, J., A. Shin, C. S. Kim, K. P. Ko, Y. Kim, J. K. Jun, J. Bae, S. K. Park, Y. C. Hong, D. Kang, S. H. Chang, H. R. Shin, and K. Y. Yoo.** 2006. CagA-producing *Helicobacter pylori* and increased risk of gastric cancer: a nested case-control study in Korea. Br J Cancer **95:**639-41.

28. **Haas, C. E., D. E. Nix, and J. J. Schentag.** 1990. *In vitro* selection of resistant *Helicobacter pylori*. Antimicrob Agents Chemother **34:**1637-41.

29. **Hatakeyama, M.** 2006. *Helicobacter pylori* CagA -- a bacterial intruder conspiring gastric carcinogenesis. Int J Cancer **119:**1217-23.

30. **Hatakeyama, M., and H. Higashi.** 2005. *Helicobacter pylori* CagA: a new paradigm for bacterial carcinogenesis. Cancer Sci **96:**835-43.

31. **Higashi, H., R. Tsutsumi, A. Fujita, S. Yamazaki, M. Asaka, T. Azuma, and M. Hatakeyama.** 2002. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A **99:**14428-33.

32. **Higashi, H., R. Tsutsumi, S. Muto, T. Sugiyama, T. Azuma, M. Asaka, and M. Hatakeyama.** 2002. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. Science **295:**683-6.

33. **Higashi, H., K. Yokoyama, Y. Fujii, S. Ren, H. Yuasa, I. Saadat, N. Murata-Kamiya, T. Azuma, and M. Hatakeyama.** 2005. EPIYA motif is a membrane-targeting signal of *Helicobacter pylori* virulence factor CagA in mammalian cells. J Biol Chem **280:**23130-7.

34. **Hurtado, A., B. Chahal, R. J. Owen, and A. W. Smith.** 1994. Genetic diversity of the *Helicobacter pylori* haemagglutinin/protease (*hap*) gene. FEMS Microbiol Lett **123:**173-8.

35. **Hutchinson, F.** 1996. Mutagenesis, p. 2218-2235. *In* C. R. I. Neidhardt, J. L. Ingraham, C. C. Lin, K. B. Low, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, vol. 2. ASM Press, Washington, DC.

36. **Israel, D. A., N. Salama, U. Krishna, U. M. Rieger, J. C. Atherton, S. Falkow, and R. M. Peek, Jr.** 2001. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. Proc Natl Acad Sci U S A **98:**14625-30.

37. **Jang, S., K. R. Jones, C. H. Olsen, Y. M. Joo, Y. J. Yoo, I. S. Chung, J. H. Cha, and D. S. Merrell.** 2010. Epidemiological link between gastric disease and polymorphisms in VacA and CagA. J Clin Microbiol **48:**559-67.

38. **Jones, K. R., S. Jang, J. Y. Chang, J. Kim, I. S. Chung, C. H. Olsen, D. S. Merrell, and J. H. Cha.** 2010. Polymorphisms in the intermediate region of VacA impact *Helicobacter pylori*-induced disease development. J Clin Microbiol **49:**101-10.

39. **Jones, K. R., Y. M. Joo, S. Jang, Y. J. Yoo, H. S. Lee, I. S. Chung, C. H. Olsen, J. M. Whitmire, D. S. Merrell, and J. H. Cha.** 2009. Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. J Clin Microbiol **47:**959-68.

40. **Jones, K. R., J. M. Whitmire, D. S. Merrell, and** 2010. A tale of two toxins: *Helicobacter pylori* CagA and VacA modulate host pathways that impact disease. Frontiers in Cellular and Infection Microbiology. doi: 10.3389/fmicb.2010.00115

41. **Josenhans, C., K. A. Eaton, T. Thevenot, and S. Suerbaum.** 2000. Switching of flagellar motility in *Helicobacter pylori* by reversible length variation of a short homopolymeric sequence repeat in *fliP*, a gene encoding a basal body protein. Infect Immun **68:**4598-603.

42. **Jung, S. W., M. Sugimoto, D. Y. Graham, and Y. Yamaoka.** 2009. *homB* status of *Helicobacter pylori* as a novel marker to distinguish gastric cancer from duodenal ulcer. J Clin Microbiol **47:**3241-5.

43. **Kang, J., and M. J. Blaser.** 2006. Bacterial populations as perfect gases: genomic integrity and diversification tensions in *Helicobacter pylori*. Nat Rev Microbiol **4:**826-36.

44. **Kansau, I., J. Raymond, E. Bingen, P. Courcoux, N. Kalach, M. Bergeret, N. Braimi, C. Dupont, and A. Labigne.** 1996. Genotyping of *Helicobacter pylori* isolates by sequencing of PCR products and comparison with the RAPD technique. Res Microbiol **147:**661-9.

45. **Kim, J. M., J. S. Kim, H. C. Jung, I. S. Song, and C. Y. Kim.** 2000. Virulence factors of *Helicobacter pylori* in Korean isolates do not influence proinflammatory cytokine gene expression and apoptosis in human gastric epithelial cells, nor do these factors influence the clinical outcome. J Gastroenterol **35:**898-906.

46. **Labigne, A., and P. J. Jenks.** 2001. Mutagenesis, p. 335-344. *In* H. L. Mobley, G. L. Mendz, and S. L. Hazzell (ed.), *Helicobacter pylori:* Physiology and Genetics. ASM Press, Washington, DC.

47. **Lee, A., J. O'Rourke, M. C. De Ungria, B. Robertson, G. Daskalopoulos, and M. F. Dixon.** 1997. A standardized mouse model of *Helicobacter pylori* infection: introducing the Sydney strain. Gastroenterology **112:**1386-97.

48. **Letley, D. P., J. L. Rhead, R. J. Twells, B. Dove, and J. C. Atherton.** 2003. Determinants of non-toxicity in the gastric pathogen *Helicobacter pylori*. J Biol Chem **278:**26734-41.

49. **Maciorkowska, E., I. Roszko, O. Kowalczuk, M. Kaczmarski, L. Chyczewski, and A. Kemona.** 2007. The evaluation of *vacA* gene alleles frequency in *Helicobacter pylori* strains in children and adults in Podlaskie region. Folia Histochem Cytobiol **45:**215-9.

50. **Majewski, S. I., and C. S. Goodwin.** 1988. Restriction endonuclease analysis of the genome of *Campylobacter pylori* with a rapid extraction method: evidence for considerable genomic variation. J Infect Dis **157:**465-71.

51. **Marshall, D. G., W. G. Dundon, S. M. Beesley, and C. J. Smyth.** 1998. *Helicobacter pylori*--a conundrum of genetic diversity. Microbiology **144 (Pt 11):**2925-39.

52. **McClain, M. S., P. Cao, and T. L. Cover.** 2001. Amino-terminal hydrophobic region of *Helicobacter pylori* vacuolating cytotoxin (VacA) mediates transmembrane protein dimerization. Infect Immun **69:**1181-4.

53. **McClain, M. S., P. Cao, H. Iwamoto, A. D. Vinion-Dubiel, G. Szabo, Z. Shao, and T. L. Cover.** 2001. A 12-amino-acid segment, present in type s2 but not type s1 *Helicobacter pylori* VacA proteins, abolishes cytotoxin activity and alters membrane channel formation. J Bacteriol **183:**6499-508.

54. **McClain, M. S., H. Iwamoto, P. Cao, A. D. Vinion-Dubiel, Y. Li, G. Szabo, Z. Shao, and T. L. Cover.** 2003. Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. J Biol Chem **278:**12101-8.

55. **Meining, A., M. Stolte, R. Hatz, N. Lehn, S. Miehlke, A. Morgner, and E. Bayerdorffer.** 1997. Differing degree and distribution of gastritis in *Helicobacter pylori*-associated diseases. Virchows Arch **431:**11-5.

56. **Miehlke, S., A. Hackelsberger, A. Meining, R. Hatz, N. Lehn, P. Malfertheiner, M. Stolte, and E. Bayerdorffer.** 1998. Severe expression of corpus gastritis is characteristic in gastric cancer patients infected with *Helicobacter pylori*. Br J Cancer **78:**263-6.

57. **Miehlke, S., C. Kirsch, K. Agha-Amiri, T. Gunther, N. Lehn, P. Malfertheiner, M. Stolte, G. Ehninger, and E. Bayerdorffer.** 2000. The *Helicobacter pylori vacA* s1, m1 genotype and *cagA* is associated with gastric carcinoma in Germany. Int J Cancer **87:**322-7.

58. **Moss, S. F., S. Legon, A. E. Bishop, J. M. Polak, and J. Calam.** 1992. Effect of *Helicobacter pylori* on gastric somatostatin in duodenal ulcer disease. Lancet **340:**930-2.

59. **Nesic, D., M. C. Miller, Z. T. Quinkert, M. Stein, B. T. Chait, and C. E. Stebbins.** 2010. *Helicobacter pylori* CagA inhibits PAR1-MARK family kinases by mimicking host substrates. Nat Struct Mol Biol **17:**130-2.

60. **Neugut, A. I., M. Hayek, and G. Howe.** 1996. Epidemiology of gastric cancer. Semin Oncol **23:**281-91.

61. **Odenbreit, S., K. Swoboda, I. Barwig, S. Ruhl, T. Boren, S. Koletzko, and R. Haas.** 2009. Outer membrane protein expression profile in *Helicobacter pylori* clinical isolates. Infect Immun **77:**3782-90.

62. **Oleastro, M., R. Cordeiro, J. Ferrand, B. Nunes, P. Lehours, I. Carvalho-Oliveira, A. I. Mendes, D. Penque, L. Monteiro, F. Megraud, and A. Menard.** 2008. Evaluation of the clinical significance of *homB,* a novel candidate marker of

Helicobacter pylori strains associated with peptic ulcer disease. J Infect Dis **198:**1379-87.

63. **Oleastro, M., R. Cordeiro, Y. Yamaoka, D. Queiroz, F. Megraud, L. Monteiro, and A. Menard.** 2009. Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. Gut Pathog **1:**12.

64. **Oleastro, M., L. Monteiro, P. Lehours, F. Megraud, and A. Menard.** 2006. Identification of markers for *Helicobacter pylori* strains isolated from children with peptic ulcer disease by suppressive subtractive hybridization. Infect Immun **74:**4064-74.

65. **Owen, R. J.** 1998. *Helicobacter*--species classification and identification. Br Med Bull **54:**17-30.

66. **Pagliaccia, C., M. de Bernard, P. Lupetti, X. Ji, D. Burroni, T. L. Cover, E. Papini, R. Rappuoli, J. L. Telford, and J. M. Reyrat.** 1998. The m2 form of the *Helicobacter pylori* cytotoxin has cell type-specific vacuolating activity. Proc Natl Acad Sci U S A **95:**10212-7.

67. **Pan, Z. J., D. E. Berg, R. W. van der Hulst, W. W. Su, A. Raudonikiene, S. D. Xiao, J. Dankert, G. N. Tytgat, and A. van der Ende.** 1998. Prevalence of

vacuolating cytotoxin production and distribution of distinct *vacA* alleles in *Helicobacter pylori* from China. J Infect Dis **178:**220-6.

68. **Parkin, D. M., F. Bray, J. Ferlay, and P. Pisani.** 2005. Global cancer statistics, 2002. CA Cancer J Clin **55:**74-108.

69. **Pelz, C., S. Steininger, C. Weiss, F. Coscia, and R. Vogelmann.** 2011. A novel inhibitory domain of *Helicobacter pylori* protein CagA reduces CagA effects on host cell biology. J Biol Chem.

70. **Perry, S., B. C. de Jong, J. V. Solnick, M. de la Luz Sanchez, S. Yang, P. L. Lin, L. M. Hansen, N. Talat, P. C. Hill, R. Hussain, R. A. Adegbola, J. Flynn, D. Canfield, and J. Parsonnet.** 2010. Infection with *Helicobacter pylori* is associated with protection against tuberculosis. PLoS One **5:**e8804.

71. **Philpott, D. J., D. Belaid, P. Troubadour, J. M. Thiberge, J. Tankovic, A. Labigne, and R. L. Ferrero.** 2002. Reduced activation of inflammatory responses in host cells by mouse-adapted *Helicobacter pylori* isolates. Cell Microbiol **4:**285-96.

72. **Pugsley, A. P.** 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiol Rev **57:**50-108.

73. **Reibman, J., M. Marmor, J. Filner, M. E. Fernandez-Beros, L. Rogers, G. I. Perez-Perez, and M. J. Blaser.** 2008. Asthma is inversely associated with *Helicobacter pylori* status in an urban population. PLoS One **3:**e4060.

74. **Ren, S., H. Higashi, H. Lu, T. Azuma, and M. Hatakeyama.** 2006. Structural basis and functional consequence of *Helicobacter pylori* CagA multimerization in cells. J Biol Chem **281:**32344-52.

75. **Rhead, J. L., D. P. Letley, M. Mohammadi, N. Hussein, M. A. Mohagheghi, M. Eshagh Hosseini, and J. C. Atherton.** 2007. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology **133:**926-36.

76. **Roder, D. M.** 2002. The epidemiology of gastric cancer. Gastric Cancer **5 Suppl 1:**5-11.

77. **Saadat, I., H. Higashi, C. Obuse, M. Umeda, N. Murata-Kamiya, Y. Saito, H. Lu, N. Ohnishi, T. Azuma, A. Suzuki, S. Ohno, and M. Hatakeyama.** 2007. *Helicobacter pylori* CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. Nature **447:**330-3.

78. **Suerbaum, S., J. M. Smith, K. Bapumia, G. Morelli, N. H. Smith, E. Kunstmann, I. Dyrek, and M. Achtman.** 1998. Free recombination within *Helicobacter pylori*. Proc Natl Acad Sci U S A **95:**12619-24.

79. **Sugimoto, M., and Y. Yamaoka.** 2009. The association of *vacA* genotype and *Helicobacter pylori*-related disease in Latin American and African populations. Clin Microbiol Infect.

80. **Taylor, D. E., M. Eaton, N. Chang, and S. M. Salama.** 1992. Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. J Bacteriol **174:**6800-6.

81. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539-47.

82. **Tsutsumi, R., H. Higashi, M. Higuchi, M. Okada, and M. Hatakeyama.** 2003. Attenuation of *Helicobacter pylor*i CagA-SHP-2 signaling by interaction between CagA and C-terminal Src kinase. J Biol Chem **278:**3664-70.

83. **Umeda, M., N. Murata-Kamiya, Y. Saito, Y. Ohba, M. Takahashi, and M. Hatakeyama.** 2009. *Helicobacter pylori* CagA causes mitotic impairment and induces chromosomal instability. J Biol Chem **284:**22166-72.

84. **Van Doorn, L. J., C. Figueiredo, F. Megraud, S. Pena, P. Midolo, D. M. Queiroz, F. Carneiro, B. Vanderborght, M. D. Pegado, R. Sanna, W. De Boer, P. M. Schneeberger, P. Correa, E. K. Ng, J. Atherton, M. J. Blaser, and W. G. Quint.** 1999. Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. Gastroenterology **116:**823-30.

85. **Versalovic, J., D. Shortridge, K. Kibler, M. V. Griffy, J. Beyer, R. K. Flamm, S. K. Tanaka, D. Y. Graham, and M. F. Go.** 1996. Mutations in 23S rRNA are associated with clarithromycin resistance in *Helicobacter pylori*. Antimicrob Agents Chemother **40:**477-80.

86. **Yamamoto, S.** 2001. Stomach cancer incidence in the world. Jpn J Clin Oncol **31:**471.

87. **Yamaoka, Y., H. M. El-Zimaity, O. Gutierrez, N. Figura, J. G. Kim, T. Kodama, K. Kashima, and D. Y. Graham.** 1999. Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. Gastroenterology **117:**342-9.

88. **Yamaoka, Y., D. H. Kwon, and D. Y. Graham.** 2000. A M(r) 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. Proc Natl Acad Sci U S A **97:**7533-8.

89. **Yamazaki, S., A. Yamakawa, T. Okuda, M. Ohtani, H. Suto, Y. Ito, Y. Yamazaki, Y. Keida, H. Higashi, M. Hatakeyama, and T. Azuma.** 2005. Distinct diversity of *vacA*, *cagA*, and *cagE* genes of *Helicobacter pylori* associated with peptic ulcer in Japan. J Clin Microbiol **43:**3906-16.